Dependency in Linguistic Description

Igor Mel'čuk and Alain Polguère (eds.)

Contents

Notations, abbreviations and writing conventions	xi
Foreword	xiii
Igor Mel'čuk and Alain Polguère	
1 Dependency syntax: general considerations	xiii
1.1 Definitorial properties of dependency syntactic structures	xiii
1.1.1 Connectedness of the syntactic structure	xiv
1.1.2 Directedness of syntactic relations	xiv
1.1.3 Strict hierarchical organization of the syntactic structure	xiv
1.1.4 "Meaningfulness" of syntactic relations	XV
1.2 Syntactic dependencies, with more Meaning-Text flavor	xvi
1.2.1 Meaning-to-text perspective	xvii
1.2.2 Emphasis on the description of dependencies	xvii
1.2.3 Deep vs. surface distinction	xvii
1.2.4 Ban on linear order in syntactic structures	xviii
2 Presentation of the papers	xviii
2.1 Mel'čuk: a sketch of dependency theory	xix
2.2 Kahane: phrases in Head-Driven Phrase Structure Grammar	XX
2.3 Iordanskaja and Mel'čuk: valence-controlled surface-syntactic relations between a verb and its dependents in French	xx
2.4 Milićević: Serbian second-position clitics in a dependency fran	nework. xxi
3 Making of the volume	xxi
References	xxii
Dependency in Natural Language	1
Igor Mel'čuk	
Introduction	1
PART I: PRELIMINARIES	2
1 Auxiliary notions	2

2	Basic assumptions					
3	Illustrations of sentence structures:					
P	Sen	hantic, Syntactic, and Morphological Structures	5			
$\mathbf{P}_{\mathbf{A}}$	ART I	I: THREE MAJOR TYPES OF LINGUISTIC DEPENDENCY	8			
1	Gen	neral remarks	8			
2	Sen	nantic dependency	10			
	2.1	The concept of semantic dependency	10			
	2.2	The logical and linguistic properties of semantic dependency	10			
3	Mo	rphological dependency	12			
	3.1	The concept of morphological dependency	12			
	3.2	The logical and linguistic properties of morphological dependency	12			
	3.3	The three major subtypes of morphological dependency	15			
4	Syn	tactic dependency	21			
	4.1	General remarks	21			
	4.2	The rationale for syntactic dependency	22			
	4.3	The concept of syntactic dependency	25			
		4.3.1 Criteria A: SSynt-connectedness	25			
		4.3.2 Criteria B: SSynt-dominance	27			
		4.3.3 Criteria C: Labeled SSynt-dependencies	33			
	4.4	The logical and linguistic properties of syntactic dependency	41			
	4.5	Some non-definitorial properties of Governors and Dependents	42			
	4.6	The absolute head of the Synt-Structure of a sentence	44			
	4.7	The three major subtypes of syntactic dependencies.	48			
	4.8	Syntactic dependencies of a language: SSynt-Relations of English	52			
5	Pos	sible combinations of the three types of linguistic dependency	58			
6	Cor	relations between the three types of linguistic dependency	66			
PA	ART I	II: SYNTACTIC DEPENDENCY	67			
1	Cur	rent fallacies concerning syntactic dependency	67			
	1.1	"Double dependency"	67			
		1.1.1 Relative pronouns	67			
		1.1.2 Raisings	71			
		1.1.3 Subordination of coordinate expressions	71			
	1.2	"Mutual dependency"	72			
	1.3	"No dependency"	72			
	1.4	"Insufficient dependency"	73			
		· ·				

2	2 Syntactic dependency in action: Eight illustrative case studies						
	2.1 Russian numeral phrases 7						
	2.2 A Russian "approximation"-marking preposition	76					
	2.3 Determiners as heads?	77					
	2.4 Romance clitics	78					
	2.5 AUX + V phrases, English-style	78					
	2.6 Conjoined nominal phrases N + CONJ + N	80					
	2.7 Russian "exotic" coordination of interrogative/negative pronouns	81					
	2.8 Elliptical constructions	82					
3	Advantages of syntactic dependency	83					
4	Syntactic dependency and syntactic constituency	89					
	4.1 Syntactic constituency	90					
	4.2 Crossing syntactic dependency with syntactic constituency	91					
5	Insufficiency of syntactic dependency: Coordination	92					
A	cknowledgments	95					
Eı	ndnotes	95					
R	eferences	101					

On the Status of Phrases in Head-Driven Phrase Structure Grammar:Illustration by a Fully Lexical Treatment of Extraction111

1	Introduction	111
2	HPSG as a dependency grammar	112
	2.1 Word descriptions	113
	2.2 Combination of words	116
3	Complementizers, relativizers and other transferers	122
	3.1 Transferers	124
	3.2 Relativizers	125
	3.3 <i>That</i> -trace effect and French <i>qui-que</i> alternation	128
	3.4 Transferers and the Locality Principle	130
4	Wh-extraction	133
	4.1 An analysis of extraction based on earlier HPSG analyses	133
	4.2 A more lexical treatment of <i>wh</i> -extraction	136
5	Status of phrases in the syntax-semantics interface of HPSG	142
6	Conclusion	146

Acknowledgments	147
References	147
Establishing an Inventory of Surface-Syntactic Relations: Valence-C	Con-
trolled Surface-Syntactic Dependents of the Verb in French	151
Lidija Iordanskaja and Igor Mel'čuk	
1 Introduction	151
2 A method for establishing an inventory of SSyntRels for a language	155
2.1 The substantive requirement to be satisfied by a SSyntRel	155
2.2 Formal requirements to be satisfied by a SSyntRel.	162
3 SSynt-Relations between a verb and its valence-controlled dependents	169
3.1 SSyntRels whose prototypical dependent is a noun	170
3.2 SSyntRels whose prototypical dependent is an infinitive	198
3.3 SSyntRels whose prototypical dependent is an adjective	205
3.4 SSyntRels whose prototypical dependent is an adverbial expression	213
3.5 SSyntRels whose prototypical dependent is Direct Speech	218
4 The summary: Synoptic tables of French valence-controlled SSyntRels	220
4.1 Subject, Objects, Agent: the prototypical dependent is a noun	220
4.2 Infinitival Objects: the prototypical dependent is an infinitive	222
4.3 Attributive Complements: the prototypical dependent is an adjective	223
4.4 Valence-controlled Circumstantials/Comparates: the prototypical dependent is an adverb/a COMME-phrase	223
5 Conclusion	224
Appendices	224
List of lexemes and constructions considered in the paper	224
List of major clause elements with their possible SSynt-roles	225
Main nominal actantial clause elements: identification tree	228
Acknowledgments	228
Endnotes.	229
References	230
Linear Placement of Serbian Clitics	
in a Syntactic Dependency Framework	235
Jasmina Milićević	
1 Introduction	235

2 What are clitics?

235

3	Bas	Basic facts about Serbian clitics 23					
	3.1	Forms of clitics					
	3.2	Constitution of the clitic cluster	239				
	3.3	Linear placement of the clitic cluster within a clause	244				
		3.3.1 Basic types of hosting constituents	244				
		3.3.2 Fixed placement	245				
		3.3.3 Variable placement	246				
		3.3.4 Special placement: insertion into a host	252				
4	Serł	bian second-position clitics in the Meaning-Text framework	255				
	4.1	The input for the clitic placement operation: SSyntR	256				
	4.2	The output of the clitic placement operation: DMorphR	258				
	4.3	Rules required to carry out the SSyntR-to-DMorphR mapping	259				
		4.3.1 Cliticization rules	260				
		4.3.2 Linearization rules	261				
5	Con	clusion	269				
Acknowledgments							
References 27							

General Index: Notions, Names, Languages

277

Notations, abbreviations and writing conventions

-A	: actant (as in SyntA)	L(w)	: lexeme to which wordform \mathbf{w}
ACC	: accusative (case)		belongs
ADJ	: adjective	L	: a particular language
ADV	: adverb	Morph-	- : morphological (as in MorphS)
AOR	: aorist	MV	: Main Verb
AUX	: auxiliary verb	Ν	: noun
С	: inflectional category	NOM	: nominative (case)
C	: constituency	NUM	: numeral
CLASS	: classifier	OBJ	: object marker
COMPL	: completive (aspect)	PERF	: 1) perfect (tense)
CONJ	: conjunction		2) perfective (aspect)
C(w)	: inflectional category of the	PL	: plural
	wordform w	PREP	: preposition
D	: (syntactic) Dependent	PRES	: present
נ	: dependency	r	: a Surface-Syntactic relation
D-	: deep- (as in D Synt-)	-Rel	: relation (as in DSyntRel)
DAT	: dative (case)	S-	: surface- (as in SSyntRel)
DET	: determiner	-S	: structure (as in SemS)
DirO	: Direct Object	Sem-	: semantic (as in Sem S)
ERG	: ergative (case)	SG	: singular
FEM	: feminine	SUB	: subject marker
G	: (syntactic) Governor	Synt-	: syntactic
g	: a grammeme	Σ	: syntactics (of a linguistic sign)
GEN	: genitive (case)	V	: verb
IMPER	: imperative	W	: a wordform
IndirO	: Indirect Object	w(L)	: wordform which belongs to
INF	: infinitive		the lexeme L
INSTR	: instrumental (case)	'X'	: meaning of X
L	: a lexeme	Ð	: operation of linguistic union

The following writing conventions are adopted throughout the volume:

- semantic quotes (...) for meanings;
- *italics* for linguistic expressions;
- **boldface** for linguistic signs;
- SMALL CAPITALS for names of lexical units;
- *italics 14pts.* for terms on the first mention.

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 $very \leftarrow surprising$ 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 \rightarrow 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** \rightarrow **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 : r_1 for (2a) and 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 \rightarrow MARY and SEND-indirect-objectival \rightarrow 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 \rightarrow 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'.



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 textproduction—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:



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 deepmorphological 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 w_1 and w_2 in a sentence:

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

FOREWORD

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.

Dependency in Natural Language

Igor Mel'čuk

It depends! [the main principle of human science—and of human life]

Introduction

One of the most vital and, at the same time, the most visible characteristics of human speech is a very high degree of organization of utterances. (Nothing astonishing, if we remember that 'information' is, roughly speaking, 'degree of organization'.) More specifically, all the units which constitute the utterance—let us limit ourselves here, for simplicity's sake, to wordforms—are arranged by the speaker in well-specified configurations, according to numerous complicated rules. These rules make up the crucially important component of any language: namely, its syntax. Putting this in a different way, all wordforms within an utterance are always related or linked among themselves. This fact is obvious to any speaker, independently of his educational level or general knowledge.

Thus, in English, we have to say *I love you*, rather than **I you love*, as one does in Russian or French (*Ja tebja ljublju/Je t'aime*), or **Love I you*, or **To-me you love* (still in the sense of 'I love you'!), or **I loves you*, etc. It is clear that the position and the form of the pronouns I (Rus. JA, Fr. JE) and YOU (TY, TU) depend on the verb, while the form of the verb depends on I (JA, JE). To make a long story short, the wordforms in an utterance are linked by dependencies: one wordform depends on another for its linear position and its morphological form. That is how the concept of dependency appears in linguistics.

Just from this it becomes evident to what extent dependency is important for linguistic description. It is one of the most basic concepts of linguistics, situated on the same level of basicness as, say, the signified, the signifier, the syntactics, and the linguistic sign: to speak in a way that guarantees the transmission of information, the speaker has first to select necessary signs (the paradigmatic axis) and then to arrange the signs selected into linear sequence (the syntagmatic axis). The arrangement of signs on the syntagmatic axis—i.e., the signs' temporal sequence—is controlled by dependencies between them. Thus, linguistic dependency merits an in-depth study, which I hope to offer in what follows. Two important provisos:

First, not all the relations between wordforms in utterances are dependencies. For instance, the coreference relation between wordforms **father** and **he** in the sentence *When I saw my father, he was busy repairing the fence* (here, **father** and **he** refer to the same person) is not a dependency. This paper limits itself to dependency relations.

Second, dependency in language is of three major types: semantic *vs.* syntactic *vs.* morphological dependency. This is, however, not easily seen on the surface—hence the widespread confusion of these different types. The failure to distinguish them results in many an incongruous or outright false statement. I will keep the different types of linguistic dependency apart as strictly as possible.

The paper is divided into three parts:

- Part I supplies the introductory information: auxiliary notions, basic assumptions, and detailed illustrations of linguistic representations proposed.
- Part II discusses the three major types of linguistic dependency: semantic, syntactic and morphological. The definitions of these three types of dependency are formulated, after which the properties of each type are described in parallel, their subtypes are specified, and an overview of their fourteen combinations (possible between two wordforms in an utterance) is presented. Criteria for the presence, orientation and specific type of syntactic dependency relation between two given wordforms are proposed.
- Part III concentrates on syntactic dependency. Four current fallacies concerning syntactic dependency are analyzed, and eight case studies are given to illustrate the effect of our criteria for establishing syntactic dependencies. A cursory comparison with constituency representation follows; the cases in which "pure" syntactic dependency proves to be insufficient for the representation of the syntactic structure of a sentence are discussed.¹

PART I: PRELIMINARIES

1 Auxiliary notions

The logical analysis of the concept "dependency in language" requires the following fourteen underlying notions:²

1. *Utterance*: a speech segment that is sufficiently autonomous. An utterance can appear between two major pauses, it constitutes a prosodic unit, and its internal struc-

¹ Two preliminary versions of this paper were published as Mel'čuk 2001 and 2003.

² The formulations that follow are not rigorous definitions, but rather approximate characterizations, which are hopefully sufficient for the purposes of the present paper.

ture is governed by linguistic rules. It is also perceived by speakers as "something that exists in the language." An utterance is a wordform, a phrase, a clause, or a sentence.

2. *Wordform*: a minimal utterance [= not containing other utterances]. In a prototypical case, it is a disambiguated word [= a lexeme] taken in a specific inflectional form; for instance, the verb SPEAK is a lexeme, while **speak**, **speaks**, **spoke**, **spoke**, etc. are its wordforms.³ The wordform is the ultimate unit treated in this paper: only linguistic dependencies between wordforms are considered, but not those between wordform parts [= morphs and other signs of the morphological level] or between wordform configurations [= phrases or clauses].

3. *Phrase*: an utterance consisting of several wordforms (as a limiting case, it can be one wordform).⁴

4. *Clause*: a phrase that is syntactically organized in essential respects as a sentence. A clause can constitute a (simple) sentence by itself or be a constituent of a sentence. In languages of Standard Average European type, a clause always contains at least one finite (\approx tensed) verb.

5. *Sentence*: a maximal utterance, which is a complete communication unit. (Two or more sentences are a sequence of utterances.) The sentence constitutes the upper limit of analysis in this paper: only linguistic dependencies between wordforms within a sentence are considered. (For instance, anaphoric links between wordforms from different sentences are not dependencies and, therefore, are not considered.)

6-8. Semantic predicate, semantic name, argument of a predicate: the notions themselves and the way they are used in linguistics are borrowed from the language of predicate calculus. A (semantic) predicate is a "binding" meaning, which is incomplete without other meanings: it has open slots into which other meanings should be inserted. A meaning that is not a predicate is a (semantic) name. Predicates denote actions, activities, events, processes, states, properties, relations, localizations, quantities, etc.; their linguistic expressions can belong to any part of speech. Semantic names denote objects (including beings), substances, and points in time and space; their linguistic expressions are, mostly, nouns.

A meaning that is inserted into an open slot of a predicate is called its *argument*; the traditional notation for a predicate P and its argument a is P(...; a; ...). Thus, *Leo is sleeping* is represented as SLEEP(LEO). A predicate can have several arguments:

 $P(a_1; a_2; a_3; ...);$

e.g., the verb SEND takes three arguments:

Leo sent a letter to Alan = SEND(LEO; LETTER; ALAN).

The number and the nature of possible arguments of a predicate must be fully specified in its description in one way or another, e.g., by ordering or numbering the arguments,

³ In case of compounding or incorporation, a wordform may represent two or more lexemes. This complication is, however, irrelevant here.

⁴ I allow myself, stretching the terminology a bit, to use the term *phrase* also for the structural representation of an actual phrase. Thus, I will speak of "the ADJ + N phrase," meaning the set of phrases like *intelligent child, expensive houses, former minister, blue sky*, etc. This is simply a convenient abbreviation.

so that, e.g., HIT(LEO; ALAN) \neq HIT(ALAN; LEO). A predicate with its arguments can itself be an argument of another predicate, this phenomenon being recursive:

Leo knows that Alan is in love with Helen = KNOW(LEO; BE-IN-LOVE(ALAN; HELEN)); *I think that Leo knows that Alan is in love with Helen* =

THINK(I; KNOW(LEO; BE-IN-LOVE(ALAN; HELEN))); etc.

9-10. *Inflectional category*: a set of mutually opposed inflectional values, called *grammemes*, such that the selection of one of them is obligatory with lexemes of a particular class (e.g., in English, number for a noun, with grammemes ^(SG) and ^(PL); or tense for a finite verb, with grammemes ^(PRES), ^(PAST), ^(FUT)).

11-13. *Syntactics*: one of the three components of a linguistic sign, in particular of a wordform. Syntactics specifies the cooccurrence of the sign that is not determined by its signified nor by its signifier (i.e., more or less arbitrary cooccurrence). The syntactics of a sign is represented as a set of *syntactic features*, each of which has mutually exclusive values.

14. *Passive syntactic valence* of a lexeme/of a phrase: a set of syntactic roles which the lexeme/the phrase can take in larger constructions (maybe with some inflectional modifications). In other words, the passive syntactic valence of a lexeme/a phrase is its syntactic distribution. Passive syntactic valence is normally defined for major classes of lexemes, known as *parts of speech*. Thus, the passive syntactic valence of the English noun is as follows: 1) the syntactic subject of a finite verb, 2) the Dir(ect) $O(bject) [= DirO]^5$ of a transitive verb, 3) the Indir(ect) O(bject) [= IndirO] of a verb that takes IndirOs (*send Father a letter*), 4) the complement of a copula, 5) the complement of a preposition, 6) the first member of a nominal compositive phrase (*computer program*), 7) an address, 8) a fronted topic, etc.^[1] (Raised numbers in square brackets refer to the endnotes, p. 95*ff*.)

2 Basic assumptions

Assumption 1: levels of sentence representation

A sentence has different *representations* on four levels: semantic, syntactic, morphological, and phonological. (In what follows, the phonological representation will be left out, even if phonological dependency is considered in a number of works: thus, see Durand 1986, Anderson & Ewen 1987, and Árnason 1989.) Each representation reflects a set of such properties of the sentence that are of the same nature. The same is true of any non-minimal part of the sentence —e.g., the clause or the phrase.

A sentence representation is a set of formal objects called *structures*, each of which is responsible for a particular aspect of sentence organization at the given level. Thus:

Sem(antic) R(epresentation	n)= ((SemS(tructure); Sem-Comm(unicative)S; Sem-RhetoricalS; ReferentialS)
D(eep)-Synt(actic) R	= (DSyntS; DSynt-CommS; DSynt-Anaph(orical)S; DSynt-Pros(odic)S
S(urface)-SyntR	= ((SSyntS; SSynt-CommS; SSynt-AnaphS; SSynt-ProsS)
DMorph(ological) R	= ((DMorphS; DMorph-ProsS)
SMorphR	= ((SMorphS; SMorph-ProsS)

⁵ For abbreviations used in this paper, see p. xi of the volume.

Assumption 2: sentence structures

The kernel part of a sentence representation, called its *central*, or *bearing*, structure, appears formally as a labeled graph, whose vertices, or *nodes*, represent linguistic units of the corresponding level, and whose *arcs* represent relations between these units.

It is here that the notion of linguistic dependency comes into play: the major type of relation between linguistic units in a sentence structure is dependency.

Assumption 3: deep vs. surface distinction

On the syntactic and the morphological level the Deep and the Surface sublevels of the sentence structure are distinguished: the former is aimed at meaning and expresses explicitly all relevant semantic contrasts; the latter is aimed at form and expresses explicitly all relevant formal contrasts. (For more on the Deep *vs*. Surface distinction, see Mel'čuk 1988: 59-72.)

3 Illustrations of sentence structures: Semantic, Syntactic, and Morphological Structures

In order to show the reader how linguistic dependencies work, sentence structures that use various types of dependency will be presented—before introducing the corresponding concepts formally. These illustrations will be referred to during subsequent discussions.

Consider the English sentence (1) and its structures at different levels (Figures 1-4):

(1) For decades, cocoa farming has escaped such problems by moving to new areas in the tropics.

The *Semantic Structure* [= SemS] of a sentence is a network whose nodes represent meanings and are labeled with semantemes (roughly, lexical meanings of the language in question); its arcs represent predicate-to-argument relations and are labeled with numbers identifying an argument with respect to its predicate. These arcs correspond to semantic dependencies, see Part II, **2**, p. 10.



Figure 1: The Semantic Structure of sentence (1)

This SemS is simplified; thus, semantic grammemes—number for nouns, tense for verbs—are not indicated.

The *Deep-Syntactic Structure* [= DSyntS] of a sentence is a tree whose nodes are labeled with the full lexemes of the sentence; the arcs of this tree, called *branches*, are labeled with names of universal *Deep-Syntactic Relations* [= DSyntRels]. Their number—across all languages—is 12:

- seven actantial DSyntRels: I, II, ..., VI, plus a special DSyntRel to encode the Direct Speech: II_{dir.speech}, which is, so to speak, a variant of the DSyntRel II;
- two attributive DSyntRels: ATTR(ibutive)_{restr(ictive)} [written, as the default case, ATTR] and ATTR_{qual(ificative)};
- two coordinative DSyntRels: COORD(inative) and QUASI-COORD [the latter is used for the special construction where the following element elaborates on the preceding one, adding more information:

John was born in the USA,-QUASI-COORD \rightarrow in New York,-QUASI-COORD \rightarrow in Manhattan,-QUASI-COORD \rightarrow on 56th Street];

• and appenditive, or extra-structural, DSyntRel: APPEND(itive), which links to the Main Verb [= MV] such "loose" elements as interjections, addresses, sentence adverbials.

DSyntRels are, of course, particular subtypes of syntactic dependency; see Part II, 4, p. 21*ff*.



Figure 2: The Deep-Syntactic Structure of sentence (1)

The diagram in Figure 2 indicates the coreference link between the two occurrences of FARMING (by a dashed bi-directional arrow). This indication belongs to the Deep-Syntactic Anaphoric Structure, mentioned above (and not considered in this paper); a coreference link is not a dependency, and it does not interfere with genuine syntactic dependencies represented in the DSyntS.

The *Surface-Syntactic Structure* [= SSyntS] of a sentence (see Figure 3, next page) is also a tree whose nodes are labeled with all the lexemes of the sentence (including all auxiliary and structural words). The arcs of this tree, also called *branches*, are labeled with names of language-specific *Surface-Syntactic Relations* [= SSyntRels], each of which represents a particular construction of the language (their number, in an

average language, is somewhere between 50 and 100; see a list of SSyntRels of English in Part II, 4.8, p. 52). SSyntRels also are particular subtypes of syntactic dependency.



Figure 3: The Surface-Syntactic Structure of sentence (1)

The *Deep-Morphological Structure* [= DMorphS] of a sentence is the string of lexico-morphological representations of all its wordforms. Its arcs are, so to speak, degenerated: they specify only the strict linear ordering of wordforms (" w_1 immediately precedes w_2 "), so that they need not be indicated explicitly. In sharp contrast to the SemS and the D-/S-SyntSs, the DMorphS of a sentence does not represent morphological dependencies between its wordforms: morphological dependencies are not explicitly presented in MorphSs because they are not universal (see Part II, 3.2, item f, p. 15). They are computed—by means of syntactic rules of the language—on the basis of syntactic dependencies, that is, from the SSyntS, transcoded into grammemes that appear in the DMorphS and expressed in the corresponding wordforms.⁶

⁶ The other structure of the DMorphR of a sentence is the Deep-Morphological Prosodic Structure [= DMorphProsS], which specifies the pauses, i.e., phonological phrases, as well as intonation contours, phrase and sentence stresses, etc. It is here that what are known as *constituents*—in the strict sense of the term—first appear. Cf. P. III, 4.1, p. 90*ff*, as well as Milićević's paper in this volume, Section 4.3.2, p. 261*ff*.

FOR	DECAD)E _{pl}	COCOA	A _{sg}	FARMI	NG _{sg}	
HAVE	nd, pres, sg	g, 3	ESCAF	PE _{ppart}	SUCH		PROBLEM _{pl}
BY MO	OVEger	ТО	NEW	AREA _{pl}	IN	THE	TROPICS _{pl}

Figure 4: Deep-Morphological Structure of sentence (1)

Thus, the SemS of Figure 1 shows semantic dependencies between (the meanings of) the wordforms of sentence (1), while the DSyntS of Figure 2 and the SSyntS of Figure 3 show the Deep- and Surface-syntactic dependencies between the wordforms of (1); but morphological dependencies are not shown explicitly. Given the morphological poverty of English, there is only one case of morphological dependency in (1): the wordform **has** depends morphologically—for the singular and 3rd person—on **farming**. The Russian sentence (2), which is a close translation equivalent of (1), contains many examples of morphological dependency (its major types—agreement, government, and congruence—are considered in Part II, 3.3, p. 15ff):

 (2) V tečenie desjatiletij kul'tura kakao ne znala takix problem blagodarja rasprostraneniju na novye territorii v tropikax.

Here,

- *desjatiletij* [GEN] ^(decades) depends for its case on *v tečenie* ^(during) [government];
- *kul'tura* [NOM] 'farming' depends for its case on [*ne*] *znala* lit. 'has-known' [government];
- *znala* [SG, FEM] 'has-known' depends for its number and gender on *kul'tura* (farming) [agreement];
- *takix* [PL, GEN] 'such⁵ depends for its number and case on *problem* 'problems' [agreement]; etc.

In Russian, unlike English, almost all the wordforms of a sentence may be linked by morphological dependencies.

PART II: THREE MAJOR TYPES OF LINGUISTIC DEPENDENCY

1 General remarks

I will consider three major types of syntagmatic dependency relations between (specific occurrences of) wordforms in a sentence: semantic dependency $[= \text{Sem-}\mathbf{n}]^7$, syntactic dependency $[= \text{Synt-}\mathbf{n}]$, and morphological dependency $[= \text{Morph-}\mathbf{n}]$, as distinguished in Mel'čuk 1964a, 1979: 13, 1981, 1988: 105-149 and developed in Nichols 1986. I will

⁷ In point of fact, Sem-**D** holds between lexical meanings (of wordforms), i.e., between semantemes in the Semantic Structure—rather than between actual wordforms in an actual sentence. However, I have allowed myself this *abus de langage* in order to be able to compare different **D**s between wordforms, doing this in a parallel fashion.

leave aside paradigmatic relations between wordforms, such as synonymy, antonymy or derivation, and syntagmatic relations of a different nature, such as:

— all kinds of lexical correspondences, e.g., between a word and a preposition it requires (INSIST — on, BORROW — from, CENTRAL — to), or between a noun and its classifier (e.g., in Vietnamese an animate noun takes the classifier CON and an inanimate noun, the classifier CÁI, with some exceptions—such as *con sông* (river); in Indonesian, nouns take different classifiers according, roughly, to their semantic class: *tiga helai kemeja* lit. (three sheet shirt) = (three shirts) vs. *tiga ekor ajam* lit. (three tail chicken) = (three chickens) vs. *tiga batang rokok* lit. (three stick cigarette) = (three cigarettes), etc.; let it be emphasized that no morphology is involved in such lexical correspondences);

— the anaphoric relation (coreferentiality: between a pronoun and its antecedent or between two nouns sharing the same referent; lexical identity: between a pronoun of the type of THAT and its antecedent, as in *my hat and that of my friend*);

- the inclusion relation (between a phrase and its constituents);
- the ordering relation (between wordforms, phrases, and clauses);
- the communicative dominance relation (between semantic units in a semantic representation).

I will deal only with direct dependencies, without indicating this explicitly every time.

Dependency is by definition a non-symmetrical relation, of the same type as logical implication: one element implies in some sense the other, but, generally speaking, not vice versa. Therefore, dependency is denoted by an arrow: $w_1 \rightarrow w_2$ means that w_2 depends on w_1 ; w_1 is said to be the governor of w_2 , and w_2 a dependent of w_1 . Other terms used to designate the governor in a dependency relation include: *head, regent, ruler*; here, however, only the term *governor* will be used. The term *head*, extremely popular in the literature, has the following defect: it is natural to speak of the head of a phrase/clause/sentence, but the expression ?the head of this wordform meaning 'the governor of this wordform' seems much less convenient. The concept of head is inherited from phrase-structure syntax and carries with it unnecessary connotations (implying constituency). Moreover, governor of phrase P \neq head of phrase P: P's governor is outside of P, while P's head is inside of P, so that in (3) the head of the phrase P = *abc* is the unit *b*, while P's governor is the unit *d*:

Therefore, in this paper the term *head* is used only in the precise sense 'the Synt-head of a phrase/clause/sentence', never in the sense of the Synt-governor. (Cf. Hudson 1993a: 274-275, on the terminological problem with respect to the expressions *head of a phrase* vs. *head of a wordform*.)

An alternative term for *dependent* is *satellite*.

Because of its intermediate nature—it is "squeezed" between semantics and morphology, Synt-**n** is the most difficult type of linguistic dependency to grasp. Therefore, it will be treated after Sem-**n** and Morph-**n**.

2 Semantic dependency

2.1 *The concept of semantic dependency*

As stated in Part I and illustrated in Figure 1, the meaning of a sentence can be represented using the formalism of the predicate calculus. We say that an argument of a predicate semantically depends on this predicate; for P(a) we write P-sem $\rightarrow a$. As I have said, an argument of a predicate P_1 can be another predicate P_2 with its own arguments $a_{2-1}, a_{2-2}, a_{2-3}, ...$:

$$P_1(P_2(a_{2-1}; a_{2-2}; a_{2-3}; ...))$$

In this case, we write P_1 -sem $\rightarrow P_2$, P_2 -sem $\rightarrow a_{2-1}$, P_2 -sem $\rightarrow a_{2-2}$, P_2 -sem $\rightarrow a_{2-3}$, etc.

The arc between the predicate and its argument carries the number of the argument: $P-1 \rightarrow a_1$, $P-2 \rightarrow a_2$, etc. The meaning of the sentence *Leo sent a letter to Alan* can then be represented (leaving grammemes aside) as



From this, we immediately obtain the definition of Sem- \mathbf{n} between wordforms \mathbf{w}_1 and \mathbf{w}_2 in an utterance.

Definition 1: Semantic dependency

The wordform \mathbf{w}_2 is said to *semantically depend* on the wordform \mathbf{w}_1 in the given utterance if and only if the meaning of \mathbf{w}_1 is (or includes) a predicate and the meaning of \mathbf{w}_2 is an argument of this predicate in this utterance: $(\mathbf{w}_1)(\dots; (\mathbf{w}_2); \dots)$.

I write, as convened above, w_1 -sem $\rightarrow w_2$.

A Sem-dependent of a wordform is called its *Sem-actant*.

2.2 The logical and linguistic properties of semantic dependency

a) Sem-**D** is anti-symmetrical: w_1 -sem $\rightarrow w_2$ entails $\neg(w_1 \leftarrow \text{sem} - w_2)$, i.e., $*w_1 \leftarrow \text{sem} - w_2$. The predicative meaning of a wordform (or any other type of meaning) cannot be an

argument of the predicative meaning of another wordform and, at the same time, have the latter as its own argument.

argument. The anti-reflexivity of the Sem-**n** follows from its anti-symmetry.

c) Sem-**n** is, generally speaking, neither transitive, nor anti-transitive: in most cases, w_1 -sem $\rightarrow w_2$ and w_2 -sem $\rightarrow w_3$ entails neither w_1 -sem $\rightarrow w_3$, nor $\neg (w_1$ -sem $\rightarrow w_3)$. Thus, from *I* saw [= w_1] Alan's [= w_3] wife [= w_2] it does not follow that I saw Alan as well, but it does not follow, either, that I did not (I could).

However, in some cases, i.e., for some predicates, Sem-**D** is transitive: with such predicates, w_1 -sem $\rightarrow w_2$ and w_2 -sem $\rightarrow w_3$ entails w_1 -sem $\rightarrow w_3$. A typical example is the predicate ⁽[to] order⁾ in the sentence *I order* [= w_1] *him* [= w_3] *to go* [= w_2],⁸ which has the following SemS:



Thus, the SemS may contain an undirected circuit (shown in boldface in the above diagram), but not a cycle, i.e., a directed circuit in which all the arrows point in the same direction.

Finally, in some other cases, Sem-**1** is anti-transitive:

 w_1 -sem $\rightarrow w_2$ and w_2 -sem $\rightarrow w_3$ entails $\neg(w_1$ -sem $\rightarrow w_3)$.

Thus, in *I* wrote down $[= \mathbf{w}_1]$ Alan's $[= \mathbf{w}_3]$ address $[= \mathbf{w}_2]$, it is clear that $\neg(\mathbf{w}_1 \text{-sem} \rightarrow \mathbf{w}_3)$. Another example of the same type is *I* heard $[= \mathbf{w}_1]$ that Alan $[= \mathbf{w}_3]$ came $[= \mathbf{w}_2]$ home.

d) Sem-**n**s must be *typed*, or *labeled*: a Sem-**n** arc has to be supplied with a symbol identifying the corresponding argument. In the present approach, this is a purely distinctive number: it does not carry meaning by itself. Thus, an arc $-i \rightarrow$ expresses different semantic roles with different predicates. (The actual semantic role of an argument of the predicate (\mathbf{w}) is specified by the semantic decomposition of (\mathbf{w}) . For instance, $(X kills Y) \approx (X, by acting upon Y, causes that Y dies), which shows that X is the Agent and the Causer, while Y is the Undergoer.) In other approaches, the symbols on Semarcs can be meaningful: e.g., Agent, Perceiver, Beneficiary, etc. Since this does not affect my reasoning in any essential way, I will not deal with this issue here.$

⁸ The English verb ORDER is such that if its meaning takes as its Sem-argument 2 the meaning ^(go), which in turn has Sem-argument 1 ^(he), then the meaning ^(order) has to take ^(he) as Sem-argument 3 (such verbs are known as "subject-to-object raising verbs"). This shows the transitivity of the Sem-**1** under consideration.

e) Sem-**1** does not presuppose the uniqueness of the governor: a wordform can semantically depend simultaneously on many other wordforms, i.e., many different meanings can be predicated about one meaning at the same time:



f) Sem-**D** is universal in the following three respects: it is present in all languages; it appears in all sentences of a language; and it embraces all full wordforms of a sentence (this means that in a sentence, Sem-**D**s always form a connected structure, such that there is a Sem-"path" between any wordform and any other wordform). Cf. the Sem-**D**s in the SemS of Figure 1, p. 5.

3 Morphological dependency

3.1 *The concept of morphological dependency*

In many languages (but by no means in all!), a wordform w_2 in an utterance can take a particular morphological form, or inflect, under the impact of another wordform, w_1 , of this utterance. Thus, in *I am well* vs. *You are well* the verb BE has different forms because of its subject [agreement, cf. 3.3, Definition 3, p. 16]. In German, with the preposition NACH 'after/to' a noun is in the dative (*nach dem Fest* 'after the feast'), but with WEGEN 'because of' it is in the genitive (*wegen des Fest+es* 'because-of the feast') [government, cf. 3.3, Definition 4, p. 18]. Technically, in such cases a grammeme **g** of an inflectional category **C** of w_2 is determined by some properties of w_1 . This leads to the following definition.

Definition 2: Morphological dependency

The wordform w_2 is said to *morphologically depend* on the wordform w_1 in the given utterance if and only if at least one grammeme of w_2 is selected depending on w_1 .

I write w_1 -morph $\rightarrow w_2$.

3.2 The logical and linguistic properties of morphological dependency

a) Morph-**D** is, strictly speaking, anti-symmetrical. In most cases (agreement of an ADJ with an N, government of the case of an N by a V or a PREP, etc.),

However, in some cases, Morph- $\mathbf{1}$ is, so to speak, quasi-symmetrical—or rather reciprocal: a wordform \mathbf{w}_2 can be inflected, for the inflectional category \mathbf{C}_1 , as a func-

tion of the wordform w_1 , and, at the same time, w_1 must be inflected, but for a different category C_2 , as a function of w_2 , so that w_1 -morph $\rightarrow w_2$ entails $w_1 \leftarrow$ -morph $-w_2$; as a result, $w_1 \leftarrow$ -morph $\rightarrow w_2$ obtains, with $C_1 \neq C_2$. This property does not contradict the antisymmetry of dependency in general, since the reciprocity of Morph-**n** is possible only with respect to different Cs.

Examples

(4)	Russian			
	<i>dv</i> + <i>a</i>	stol	+a	'two tables'
	two MASC.NOM	table[MA	ASC] SG.GEN	

The noun *stola* morphologically depends for its singular and genitive case on the numeral *dva*, while *dva* depends on *stola* for its masculine gender; cf., respectively, *pjat' stol+ov* [PL.GEN] 'five tables' and *dvadcat' odin stol+O* [SG.NOM] lit. 'twenty-one table', but dv+e [FEM.NOM] *krovati* 'two beds'.

[Here and below, the grammemes of the wordform \mathbf{w} which are determined by the wordform \mathbf{w}' , as well as their markers, are boldfaced.]

(=)	0	•
(5)	Geo	rgian
		0

a. <i>Gogi+Ø</i>	m +zrdi	+s	me
Gogi NOM	1SG.OBJ bring.up	PRES.3SG.SUB	me-DAT
'Gogi brings me u	p ² .		
VS.			
Gogi+Ø	gv +zrdi	+s	čven
Gogi NOM	1PL.OBJ bring.up	PRES.3SG.SUB	we-DAT
'Gogi brings us up	o ⁾ .		
h <i>Gogi</i> + m	$\sigma_a + m$	+zard +a	me
Gogi FRG	COMPL 1SG OB	\mathbf{J} bring up AOR	3SG SUB me-NOM
Gogi brought me	up ² .	oring.up riore	
VS.			
Gogi+ m	ga + g v	+zard + a	čven
Gogi ERG	COMPL 1PL.OF	BJ bring.up AO	R.3SG.SUB we-NOM
Gogi brought us u	ιp ⁾ .		

[In Georgian, the forms of the nominative and the dative of the 1st/2nd person pronouns are homophonous; a transitive verb, when in the present, requires a nominative construction, with the Subject in the nominative and the Direct Object in the dative, while when it is in the aorist, an ergative construction appears: the Subject is in the ergative and the DirO in the nominative.]

Both Synt-actants depend morphologically on the verb for their case (which is imposed by the syntactic type of the verb and its tense: the present *vs*. the aorist), while the verb depends morphologically on the actants for its person/number (of the Subject and of the DirO).

b) Morph-**D** is anti-reflexive: *
$$\bigcirc_{\circ}^{\text{morph}}$$
,

i.e., the inflection of a wordform cannot be determined by the wordform itself.

However, the inflection of a wordform w can be conditioned by some properties of w itself, which is not a case of Morph- \mathbf{D} .⁹ Thus, in Alutor, in the ergative construction of a transitive verb, the Subject w is in the instrumental if w is not a human proper name, and in the locative if it is:

- (6) Alutor (Chukchee-Kamchatkan family, Russia)
 a. *ənpəŋav* +a Ø +l*ə?uşqiv* +nin +Ø qəlavul+Ø
 old.woman SG.INSTR 3SG.SUB went.to.see 3.OBJ SG.OBJ man SG.NOM ⁽[An] old woman went to see [her] man³.
 - b. *Miti* +nak Ø +lə?uşqiv +nin +Ø qəlavul+Ø Miti SG.LOC 3SG.SUB went.to.see 3.OBJ SG.OBJ man SG.NOM 'Miti went to see [her] man'.

The case of the Subject is determined here by the verb—but conditionally, i.e., according to the indicated property of the subject noun.

Another example of a slightly different kind comes from Arabic:

(7)	Arabic						
	Jā	rafīq+ u	vs.	Jā	rafīq+ a	ab +ī	
	oh	friend NOM		oh	friend ACC	father 1SG	
	⁽ Oh, friend! ⁾		⁽ Oh, friend of my father! ⁾				

The case of the address noun N is determined by the vocative particle $J\bar{A}$, but according to whether or not N has its own nominal Synt-dependent: if such a dependent is absent, N is in the nominative; if the dependent is present, N is in the accusative.

c) Morph-**D** is neither transitive nor anti-transitive. In most cases, Morph-**D** is anti-transitive:

```
w_1-morph\rightarrow w_2 and w_2-morph\rightarrow w_3 entails \neg(w_1-morph\rightarrow w_3).
```

Thus, in Rus. Ja vižu $[= w_1]$ krasivuju $[= w_3]$ knigu $[= w_2]$ 'I see [a] beautiful book' there is no direct Morph-**n** between the verb and the adjective.

There are, however, cases where Morph-**D** is transitive:

 w_1 -morph $\rightarrow w_2$ and w_2 -morph $\rightarrow w_3$ entails w_1 -morph $\rightarrow w_3$.

An example of a transitive Morph-**D** (again, for different inflectional categories and different grammemes) is found in Russian:

(8) Russian Ja zna +l +a ego molod+ym I know PAST FEM he-SG.MASC.ACC young SG.MASC.INSTR ⁽I [a woman] knew him young⁾. vs.

⁹ According to Definition 2, Morph-**D** means the imposition of a grammeme. A wordform cannot impose a grammeme upon itself, but it can, by some of its properties, condition the choice between several grammemes imposed on it by a different wordform.

Ja zna +l +aеë molod+oj know PAST FEM she-SG.FEM.ACC young SG.FEM.INSTR I ⁽I [a woman] knew her young⁾. VS. Ja zna +l+aix molod+ymi they-PL.ACC young PL.INSTR Ι know PAST FEM ⁽I [a woman] knew them young⁾.

Here, $ego/e\ddot{e}/ix$ [= w_2] depends on *znala* ^(knew) [= w_1] for its accusative case, while *molodym/molodoj/molodymi* [= w_3] depends on *ego/eë/ix* for its number and gender, and on *znala* for its instrumental case.¹⁰

d) Similarly to Sem-**n**, Morph-**n** must be also typed (= labeled): if w_1 -morph $\rightarrow w_2$, then in order to fully specify this Morph-**n**, we have to indicate the inflectional category $C(w_2)$ whose grammeme is imposed by w_1 . Thus, the labeling of Morph-**n**s is meaningful rather than purely distinctive, as is the case with Sem-**n**.

e) Morph-**D** does not presuppose the uniqueness of the governor: a wordform can morphologically depend simultaneously on several other wordforms—for different inflectional categories, of course. Cf. (8), where w_3 depends morphologically on w_1 and w_2 at the same time (with transitivity of Morph-**D**); another example of Morph-**D** with multiple governors (without transitivity of Morph-**D**) is (9a), p. 17.

f) Morph-**n** is not universal: in many languages it is not present at all; in a language where it is present it is not present in all sentences; and in a sentence where it is present it is not necessarily present in every wordform. As a result, in a sentence Morph-**n** s do not form, generally speaking, a connected structure: there are wordforms that are not morphologically linked to the rest of the sentence (this is the reason why the MorphS of a sentence is not explicitly specified).

3.3 *The three major subtypes of morphological dependency*

There are three major subtypes of Morph-**D**: *agreement*, *government*, and *congruence* (Mel'čuk 1993 and 2006: 31 ff).¹¹ Here are the corresponding definitions; in all of them the wordform **w**₂ depends morphologically on the wordform **w**₁ according to the inflectional category **C**₂. The wordform **w**₁ is called the *controller*, and the wordform **w**₂, the *target* of the Morph-**D** in question. In the examples below, the controlling grammeme or syntactic feature of the controller is underlined, and the controlled grammeme and its marker in the target are boldfaced.

¹⁰ The Russian verb ZNAT' ^(know) is such that if it has a Morph-dependent w_2 which has a Morphdependent w_3 of its own, then, under specific syntactic conditions (w_2 is a DirO, etc.), w_3 is a Morphdependent of ZNAT' as well. This shows the transitivity of the Morph-**1** under consideration.

¹¹ For a different analysis of the corresponding notions (and a rich bibliography), see Schmidt & Lehfeldt 1995. — Recall that agreement and government were treated for a long time as types of SYNTACTIC dependency, which created confusion.

Definition 3: Agreement

The wordform w_2 is said to *agree* with the wordform w_1 in the inflectional category C_2 if and only if the following two conditions are simultaneously satisfied:

- 1. the wordform w_2 is not a substitute pronoun coreferential with w_1 ;
- 2. a grammeme $\mathbf{g}_2 \in {}^{(\mathbf{w}_2)}$, such that $\mathbf{g}_2 \in \mathbf{C}_2$, is selected depending upon:
 - (a) either the grammeme $g_1 \in {}^{(w_1)}$, such that $g_1 \in C_1$ and C_2 is mirroring for C_1 ;
 - (b) or the value γ_1 of a feature Σ_1 of the syntactics of w_1 , this Σ_1 being one of the following three features of syntactics:
 - agreement class, pronominal person, or pronominal number;
 - (c) or some semantic components of \mathbf{w}_1 or some properties of its referent.

Comments

1. *Substitute*, or anaphoric, *pronouns* are pronouns of the type HE, SHE, IT, THEY, and all the relative pronouns, which replace nouns. Thus, a substitute pronoun is always used instead of a noun, so it is really a PRO-noun. Substitute pronouns must be distinguished from personal pronouns of the type I, YOU, WE, which never replace a noun.

2. An inflectional category C_2 is said to be *mirroring for the inflectional category* C_1 if (roughly) C_2 exists in **L** exclusively to "reflect" C_1 . Thus, adjectival number and adjectival case are mirroring for nominal number and nominal case. (The relation 'to be mirroring for' is, of course, by no means symmetrical: C_1 is not mirroring for C_2 .)

3. The *agreement class* A is (roughly) a subset of lexemes of the same part of speech (essentially, of nouns) such that in any context the following three conditions are simultaneously satisfied:

- if any two wordforms w_i and w_j of A impose on a third wordform w a grammeme of a category C(w), they impose on w the same grammeme g ∈ C(w);
- if in an utterance a wordform w of A simultaneously imposes on different wordforms w_i that morphologically depend on it a grammeme of a category C(w_i), it imposes on all w_i the same grammeme g ∈ C(w_i);
- 3) the grammeme \mathbf{g} is not imposed by anything except the wordforms of \mathbf{A} .

Agreement class is a generalization of *grammatical gender* (as in Indo-European languages) and of *nominal class* (as in Bantu and Daghestanian). Agreement classes are defined in particular languages prior to and independently of the notion of agreement (Mel'čuk 1993: 323-324, 1996b: 206-211, 2006: 47*ff*).

4. Condition 1 of Definition 3 separates agreement from congruence, see below. Condition 2a foresees agreement with grammemes of the controller (e.g., agreement of an ADJ with an N in number and case). Condition 2b foresees agreement with features of the syntactics of the controller: e.g., agreement of an ADJ with an N in gender; or agreement of a V with a pronominal Synt-actant in person and number (the latter being
syntactic features of a pronoun). Condition 2c foresees what is known as semantic agreement (Rus. $Na\check{s}+a$ vrač skazal+a ⁽our[FEM] doctor [a woman] said[FEM]⁾, where, in spite of the fact that the noun VRAČ is masculine, the agreeing adjective and the verb are both in the feminine, because in this utterance VRAČ refers to a woman).

Examples

(9) a. Akhvakh (North-Caucasian, Daghestan, Russia; Boguslavskaja 1991: 11)

In Akhvakh, an adjective or a participle which is used as a restrictive modifier of a noun and, at the same time, has a complement or an actant of its own agrees both with this complement/actant (in nominal class, shown by Roman numerals) and the modified noun (again in nominal class); the first agreement is shown by a prefix, and the second by a suffix:

 $roča + \emptyset$ **b**+exeq +ida +je jaše + \emptyset book[<u>III</u>] SG.NOM **III** bring ADJECT(ivizer) **II** girl[<u>II</u>] SG.NOM lit. 'book-bringing girl' = 'a girl who is bringing a book'

b. Old Georgian

A noun N_2 in the genitive that syntactically depends on another noun N_1 agrees with N_1 in case and thus gets the suffix of the second, "agreeing" case, cf.:

neb+Ø +ita	γmrt+Ø+isa + jta	lit. 'by-[the-]will of-God'
will SG <u>INSTR</u>	God SG GEN INST	`R

c. Kayardild (Australia; Evans 1988: 221-222) All the objects of a verb in Kayardild agree with the verb in tense/mood:

danga+a bargi+da tungal+Ø +i nara+nuni +v chop <u>NON-FUT</u> knife INSTR NON-FUT man NOM tree ACC NON-FUT 'The man just chopped/is chopping the tree with a knife'. vs. danga+a bargi+du tungal+Ø +u nara+ŋuni +wu man NOM chop ACC FUT knife INSTR FUT FUT tree 'The man will chop the tree with a knife'. VS. danga+a bargi+dara tungal+Ø +ina nara+nuni +na ACC PAST knife INSTR PAST man NOM chop PAST tree 'The man (had) chopped the tree with a knife'.

For more examples of "exotic" agreement see Kibrik 1977 and Anderson 1992: 103-118.

Definition 4: Government

The wordform w_2 is said to *be governed* by the wordform w_1 with respect to an inflectional category C_2 if and only if one of the following two conditions is satisfied:

- The grammeme $\mathbf{g}_2 \in {}^{(\mathbf{w}_2)}$, such that $\mathbf{g}_2 \in \mathbf{C}_2$, is selected depending:
- 1. either upon a grammeme $g_1 \in {}^{(w_1)}$, such that $g_1 \in C_1$ and C_2 is not mirroring for C_1 ;
- 2. or upon a value γ_1 of a feature Σ_1 of the syntactics of w_1 , this Σ_1 being neither agreement class, nor pronominal person, nor pronominal number.

Comments

1. Condition 1 foresees government by a grammeme of the controller. These are exotic cases of government: e.g., the comparative that governs the case of the comparand (Rus. *sil'n+ee smerti* [GEN] ^(stronger than-death)) or the tense of the verb governing the case of its actants, see examples (10a-b).

2. Condition 2 foresees government by a feature of the syntactics of the controller; it separates such government from syntactics-induced agreement. These are normal cases of government: e.g., a verb or a preposition governing the grammatical case of an object/complement.

Examples

(10) a. Georgian

A transitive verb governs:

- in the present/imperfect, the nominative of the Subject and the dative of the DirO;
- in the aorist, the ergative of the Subject and the nominative of the DirO;
- in the perfect, the dative of the Subject and the nominative of the DirO.

However, the agreement of the verb does not change: it always agrees with its Subject and with its DirO (in person and number), if the latter is not of the 3rd person:

ceril+0+scer +s Gogi+Ø letter SG DAT Gogi NOM write PRES.3SG.SUB 'Gogi is-writing [a] letter'. VS. +cer +a $ceril + \emptyset + i$ da Gogi +m letter SG NOM COMPL write AOR.3SG.SUB Gogi ERG 'Gogi wrote [a] letter'. VS. $\dot{c}eril+\emptyset$ +i $+\dot{c}er$ +iGogi +s da +u+aletter SG NOM COMPL 3pers write PERF 3SG.SUB Gogi DAT 'Gogi (apparently) has-written [a] letter'.

["COMPL" stands for "completive [aspect];" "3 pers(on)" in the last sentence denotes the grammeme 'FOR THE OTHER' of a special inflectional category of Georgian: the *version*, which signals the person for whose benefit the event in question takes place (this grammeme is obligatorily present in perfect forms). The Georgian perfect implies that the Speaker was not an eye-witness of the event referred to and does not vouch for the truth of what he is saying.]

b. Hindi

A transitive verb governs:

- in the present, the nominative of the Subject and the nominative/dative of the DirO (the dative seems here syntactically optional);
- in the perfect, the ergative of the Subject and the nominative/dative of the DirO.

But, unlike Georgian, the agreement of the verb changes depending on the tense: in the present, the verb agrees with the Subject, but in the perfect either it agrees with the DirO (if the DirO is in the nominative) or it takes the unmarked form of the 3rd person singular masculine (if the DirO is in the dative).

Laṛkā	+Ø	kitāb	+Ø	paṛh +	-tā	hai
boy[masc]-SG	NOM	book[FEM]-SG NOM	read	IMPF.MASC analytical	AUX. <u>PRES</u> .3SG present form
'[The] boy re	ads [a] b	ook'.				
VS.						
Laṛke	+ <i>ne</i>	kitāb	+Ø	paṛh	$+\overline{\iota}$	
boy[masc]-SG	ERG	book[FEM]-SG NOM	read	PERF.FEM.	3SG
'[The] boy re	ad [a] bo	ook ⁾ .				
VS.						
Laŗkõ	+ <i>ne</i>	Sitā -	ŀØ	dekh	$+\overline{\iota}$	
boy[masc]-PL	ERG	Sita[FEM]	NOM	see	PERF.FEM.	3SG
'[The] boys s	aw Sita'					
VS.						
Laṛkiyõ	+ <i>ne</i>	Sitā	+ko	dekh	$+\bar{a}$	
girl[FEM]-PL	ERG	Sita[FEM]	DAT	see	PERF.MAS	C.3SG
'[The] girls s	aw Sita'.					

c. Russian

The infinitive in a special "impossibility" construction governs the dative of its semantic subject:

Mn+e ètu knig +u ne pročest' I **DAT** this book SG.ACC not read-PERF.<u>INF</u> 'I will not be able to read this book'.

Alen+utudanedojtiAlanSG.DATtill.therenotreach.walking-PERF.INF(Alan will not be able to walk till there).

d. Hungarian

When the subordinate conjunction HOGY 'that' syntactically depends on a verb of volition, it requires the imperative of the MV of the subordinate completive clause:

Az +takarta,hogylassan jár +jthis ACCwant-PAST.3SG.SUB.3SG.OBJthatslowlywalk IMPER.2SGlit. ⁽[S/he]this wanted that [you-SG] slowly walk³.

It is impossible to define agreement and government in a simpler way—for instance, following the traditional view that agreement is a correspondence between the inflec-

tional form of a lexeme and the inflectional form of another lexeme, while government is a correspondence between the inflectional form of a lexeme and lexicographic properties of another lexeme. This viewpoint is simply wrong: many types of agreement involve lexicographic properties of the controller (gender, nominal class, animacy), and many types of government are determined by the inflectional form of the controller (cf., among others, examples (10a-b)).

Definition 5: Congruence

The wordform w_2 is said to *be congruent* with the wordform w_1 with respect to the inflectional category C_2 if and only if the following two conditions are simultaneously satisfied:

- 1) $\mathbf{w_2}$ is a substitute pronoun coreferential with $\mathbf{w_1}$;
- 2) a grammeme $\mathbf{g}_2 \in {}^{(\mathbf{w}_2)}$, such that $\mathbf{g}_2 \in \mathbf{C}_2$, is selected depending on \mathbf{w}_1 .

Comments

1. Congruence is, so to speak, a particular case of agreement, namely, "agreement in absentia:" while genuine agreement obtains between an ADJ/a V and the N it combines with syntagmatically, congruence obtains between a substitute pronoun and the N it replaces. Agreement marks semantic and/or syntactic \mathbf{T} s within the borders of a clause, while congruence marks anaphoric links, basically outside the borders of a clause. For congruence, correspondence according to meaning (rather than according to grammatical properties of the controller) is especially typical. Recall that congruence is not a syntactic dependency, but a morphological one; therefore, the absence of its controller from the clause is not a problem.

2. Congruence presupposes the choice of a particular inflectional form of a given lexeme. Thus, in Spanish, the noun **caballo** [MASC] ^chorse, sG³ is replaced with the pronoun él ^(he), the noun **mosca** [FEM] ^(fly, sG³) with the pronoun ella ^(she), caballos ^(horse, PL³) with ellos, and **moscas** ^(fly, PL³) with ellas; this is congruence, since él, ella, ellos and ellas are forms of one lexeme (= ÉL), which is inflected for gender and number. (The same state of affairs obtains in any language in which substitute pronouns grammatically distinguish gender and/or number: Romance, Slavic, Semitic, Bantu languages.) However, the choice between different pronominal lexemes as a function of w₁ to be replaced is not congruence. Thus, in English, general or Alan is replaced by HE, sister or battleship, by SHE, and warning or [a] fly, by IT; but HE, SHE and IT are different lexemes rather than inflectional forms of the same lexeme—because English has no inflectional category of gender. The selection of the appropriate lexeme has to do with lexical correspondences, mentioned in Part II, 1, p. 9, not with congruence, since no Morph-**1** is involved (no grammeme is imposed).

Examples

(11) a. French

Nous étudions un suffixe [MASC.SG] *et deux alternances* [FEM.PL]; *nous traiterons celui-là* [MASC.SG] *immédiatement et celles-ci* [FEM.PL], *au chapitre suivant* 'We study a suffix and two alternations; we will deal with the former right away, and with the latter in the next chapter³.

The wordforms **celui** and **celles** are inflectional forms of the lexeme CELUI 'that-one', so that their choice illustrates congruence. (In contrast, the English wordforms **former** and **latter** belong to two different lexemes, and therefore their use is not related to congruence.)

b. Bushoong (Bantu, Democratic Republic of Congo)

Here, a noun is replaced by different inflectional forms of the same substitute pronoun lexeme -N (s)he, it, they), namely—by the form of the corresponding nominal class:

Ι	aa +n	replaces a noun of the class I;
II	baa +n	replaces a noun of the class II;
III	muu +n	replaces a noun of the class III;
IV	mii +n	replaces a noun of the class IV; etc.

To conclude the discussion of the three subtypes of Morph- \mathbf{n} , let me state three reasons that underlie the intuitive desire of a linguist to distinguish them (cf. also Part III, **6**, p. 66):

1) A morphological reason: under agreement and congruence (which is a particular case of agreement) the target reflects some properties of the controller; under government, this cannot happen.

2) A semantic reason: under agreement the target is prototypically the Sem-governor of the controller, which is its Sem-actant; under government, on the contrary, the target is prototypically the Sem-actant of the controller, which is its Sem-governor; under congruence the target and the controller cannot be linked by a Sem-**n**: they are coreferential.

3) A syntactic reason: under agreement the target can be or not be linked by a direct Synt-**D** to the controller; under government the target is necessarily linked by a direct Synt-**D** to the controller; under congruence the target and the controller cannot be linked by a Synt-**D**.

4 Syntactic dependency

4.1 General remarks

Paraphrasing R. Jakobson, we can say that Sem-**1** is directly related to meaning and therefore it is conceivable or understandable, while Morph-**1** is directly related to (phonological) form and therefore it is perceivable. The Synt-**1** is, however, not directly related either to meaning or to form; it is more abstract, more indirect than Sem-**1** and Morph-**1**, and, as a consequence, more questionable. Even its mere existence needs justification.

Syntactic dependency has been used to talk about the structure of sentences from Antiquity and throughout the Middle Ages to modern times. All respectable pre-20th century grammatical traditions in syntax have been based on it, as has much language teaching. By the 8th century, Arab grammarians (e.g., Sībawaih, who died in 798) already distinguished the governor vs. the dependent in syntax and used this distinction when formulating rules for word order and inflection (Owens 1988: 79-81). One finds dependency trees as a means of describing sentence structure in German syntax books from the 19th century (Weber 1992: 13). In point of fact, constituency representation in syntax, i.e., what became known as phrase-structure, was first introduced—and that, almost exclusively in the domain of the English language!—in the early 20th century. The dependency approach [= **1**-approach] was properly presented for the first time in Tesnière 1959 (the first sketch of Tesnière's theory appearing in Tesnière 1934). This profound treatise made syntactic dependency available for serious theorizing. Yet, due to the dominance of Chomskian Transformational-Generative Grammar—which uses, as its main syntactic tool, the phrase-structure representation (i.e., constituency)—the **1**-approach did not become popular in modern linguistics until the beginning of the 1980s.

Nevertheless, starting in the early 1960s and over a period of about ten years, a number of publications which laid foundations for the **n**-approach had appeared (Hays 1960 [1961], 1964a, b, Lecerf 1960, Fitialov 1962, 1968, Mel'čuk 1962, 1963, 1964a, b, Iordanskaja 1963, 1967, Padučeva 1964, Gaifman 1965, Baumgärtner 1965, 1970, Marcus 1965a, b, Robinson 1970, Heringer 1970). All these studies were more or less inspired by computational applications of linguistics-primarily machine translation and other types of computer text-processing. Gradually, the field grew into real theoretical research, continuing to rely heavily on computer applications (e.g., Kunze & Priess 1967-1971, Sgall et al. 1969, Goralčíková 1973, Machová 1975, Kunze 1975, Hudson 1976, 1980a, b, 1984, Garde 1977, Korhonen 1977, Schubert 1987). And later, several general linguistic theories have emerged, based partially or completely on the **D**approach, including Case Grammar (Fillmore 1968, Anderson 1977), Meaning-Text Theory (Mel'čuk 1974, 1979, 1988, 1997b), Lexical-Functional Grammar (Bresnan 1982), Relational Grammar (Perlmutter 1983), Word Grammar (Hudson 1984, 1990), Functional Generative Description (Sgall et al. 1986, Petkevič 1995), Lexicase Theory (Starosta 1988)-cf. Hudson 1993b: 330-332. Cognitive Grammar (Langacker 1987, 1991, 1997) is also dependency-oriented. One finds a few university manuals which use the **D**-approach (e.g., Matthews 1981, Tarvainen 1981, Weber 1992). The description of German syntax in Engel 1977 [1982, 1994] and the syntactic part of Engel 1988—one of the most authoritative German reference grammars-are developed explicitly within the **D**-approach (see especially Engel 1988: 21-26). A formal **D**-grammar for German syntax (a list of rules and theoretical discussion) is found in Heringer 1996.

NB: When speaking of the **n**-approach in what follows, I mean exclusively a dependency representation of the structure of sentences rather than a dependency grammar, i.e., a logical device consisting of rules that ensure the generation/parsing of sentences. The two notions are logically related, but should be kept distinct. (Cf. Hudson 1993a: 266-269 on the difference between syntactic heads in sentence structure and syntactic heads in grammar rules.)

4.2 *The rationale for syntactic dependency*

The notion of Synt- \mathbf{D} is proper to Syntactic Structure [= SyntS]—a formal object used to depict the formal organization of a sentence in terms of its wordforms and their linear arrangement—as opposed to its meaning, which is the target of the Semantic

Structure [= SemS]. Synt-**n**s are building blocks of a SyntS, and so it will be useful to start with a brief characterization of the latter.

Formal considerations

The SyntS of a sentence is called upon to mediate between its SemS and its Morph(ological)S. The SemS is formalized as an arbitrary *n*-dimensional graph, i.e., a *network*, as we see in Figure 1, p. 5. The MorphS is a 1-dimensional (linear) graph, i.e., a *chain*, cf. Figure 4, p. 8. The SyntS constitutes a convenient bridge between the SemS and the MorphS: under text synthesis, that is, in the transition from meaning to text, the SyntS must be easily produced from the Sem-network and easily converted into the Morphchain; under analysis, that is, in the transition from text to meaning, it must allow the inverse operations to be carried out also with sufficient ease. The simplest formal object that satisfies these requirements is a 2-dimensional (planar) graph—a tree. Networks are relatively easy to arborize, and trees are easy to linearize (text synthesis); vice versa, chains are relatively easy to arborize, and trees are easy to convert to networks (text analysis). In other words, the Synt-tree is the most convenient intermediary between the Sem-network and the Morph-chain. That is how the idea of SyntS as a dependency tree composed of lexemes is formally arrived at. If the SyntS is a tree, then any of its arcs, or branches, represents an anti-reflexive, anti-symmetrical and anti-transitive binary relation between lexemes—i.e., a Synt-n relation. This reasoning leads us to the notion of Synt-dependency as an order relation (see next page) and to the notion of dependency tree as an appropriate formalism for the representation of SyntSs (see the end of 4.4, Part II, p. 41).

What has just been said should not be construed as a proposal to consider the dependency tree as an artifact of the linguist, a figment of his imagination, having no claim to psychological reality. On the contrary, I think that the dependency tree is a reasonably good model of how sentences are organized in the brain of the speakers. The dependency tree is proposed here as an exclusive means for representing the SyntSs of sentences exactly because I believe that my brain is using it when I am writing these lines.

Substantive considerations

Now I will consider the problem of SyntS from another angle. Suppose we want to represent the SyntS of the sentence *Leo knows that Alan is in love with Helen*. There are exactly four types of linguistic means that this sentence uses to express its meaning: lexemes, order of lexemes (i.e., word order), prosody, and inflection. Note that:

- there do not exist other types of linguistic means that could be used to express meanings;
- these four types of linguistic means are used by all languages in all sentences with the notable exception of inflection, which does not exist in quite a few languages and which, even in the languages where it does exist, does not appear in all sentences and on all wordforms;
- each of these means can be used either as a direct expression of meaning, i.e., in a semantic capacity, or without a direct relation to meaning—purely in order to

Linguistic means	used in semantic capacity	used in syntactic capacity
lexical units	full words: <i>for, decades, cocoa, farming, escape, the, when</i> , etc.	empty words—e.g., governed prepositions and conjunctions: [<i>depend</i>] on , [<i>to order</i>] that , etc.
word order	arrangements that mark commu- nicative structure (Theme ~ Rheme, Given ~ New, etc.)	arrangements that mark syntactic constructions: N + N, PREP + N, ADJ + N, V + N [= DirO], etc.
prosody	prosodies that mark question <i>vs</i> . assertion, focus, emphasis,, irony, threat, tenderness, etc.	prosodies that mark borders of lin- ear constituents
inflection	number in nouns; aspect and tense in verbs	case in nouns; person and number in verbs; gender, number and case in adjectives (agreement and gov- ernment inflectional categories)

indicate links between wordforms in the sentence, i.e., in a syntactic capacity, see Table 1 below.

Table 1: Linguistic expressive means and their possible uses

Non-lexical means used in syntactic capacity (shadowed cells in Table 1) should not appear in a SyntS:¹² they are means used to express the SyntS, therefore they cannot be part of it. All of them appear closer to surface, in the DMorphS of the sentence, while the SyntS has to replace them with a formal simple homogeneous device. This device must be able to encode the future linear arrangement of wordforms, i.e., word order, in an explicit, clear and elegant way. Note that syntactic prosody applies to a previously ordered sequence of wordforms, and inflection is absent in many cases; therefore, these two linguistic means are secondary from the viewpoint of the SyntS. The SyntS has to tell us, first of all, where to position a wordform w_2 : before or after another wordforms which have to be positioned on the same side of w_1 . The most economical way to do this is using an anti-reflexive, anti-symmetrical and anti-transitive binary relation between the wordforms of the sentence—an *order relation* (in the logical sense of the term). This is nothing else but a Synt-**n**; thus, we have once again, this time via substantive reasoning, come to the same conception of Synt-**n** relation.

As a bridge between the SemS and the DMorphS of a sentence, the D-/S-SyntS must encode all the relevant semantic contrasts that are expressed on the surface and all the relevant formal contrasts that carry meaning. Therefore, the specific Synt-**D** relations that are introduced for a particular language must be such that they satisfy this requirement.

¹² Lexical means used in syntactic capacity, i.e., what is known as structural, or empty, words, complicate the picture without affecting the essence of my reasoning: they do not appear in the DSyntS, but they are present in the SSyntS—since they are separate wordforms, and the SSyntS is supposed to represent all the wordforms actually found in the sentence. To keep my formulations as simple as possible I leave the lexical means used in syntactic capacity out of the discussion.

Let us emphasize that at the very beginning of the **n**-approach Synt-**n** was not, and is still not always, rigorously distinguished from Sem-**n** and Morph-**n**. Linguists would often talk about dependency *tout court*, aiming at Synt-**n**, but in actual fact taking in a mixture of the three.

Since Synt-**1** is an abstract formal concept, it is not as easy to define as Sem-**1** and Morph-**1**. Three groups of criteria for Synt-**1** have to be introduced; but first, let me emphasize that for simplicity's sake I will be dealing exclusively with Surface-Syntactic [= SSynt-] dependency. The results can be easily generalized to cover Deep-Syntactic dependency as well.

To establish a SSynt-**D** relation between two wordforms in a sentence we need (at least) three groups of criteria:

- A. Criteria for SSynt-connectedness of the two wordforms (i.e., criteria for the PRESENCE of a SSynt-**D** between them).
- B. Criteria for the SSynt-dominance between the two wordforms (i.e., criteria for the ORIENTATION of the SSynt-**1** between them).
- C. Criteria for the specific type of the given SSynt-**D** between the two wordforms (i.e., criteria for the TYPE OF THE SSYNT-RELATION between them; as will be shown, to ensure a proper representation of syntactic structure of utterances, we have to distinguish, in a particular language, many different specific types of Synt-**D**).

These criteria are necessary, but unfortunately not sufficient. Thus, when establishing the specific types of SSynt- \mathbf{n} s in a particular language, the researcher has to invoke relevant linguistic properties of different dependents subsumed under the same SSynt-Rel (Iordanskaja & Mel'čuk 2000). While deciding on the presence and orientation of SSynt-dependencies some heuristic principles have to be recurred to (Part III, 1.1.1, p. 70). And more than anything else, reasoning by analogy remains the most necessary tool: the description, in terms of Synt- \mathbf{n} , of a "dubious" phrase P₁ should correspond to the SSynt-description adopted for the similar phrase P₂ where the situation is clearer or outright obvious.

4.3.1 Criteria A: SSynt-connectedness

First, one has to know whether two particular wordforms in an utterance, w_1 and w_2 , are syntactically directly linked. To answer this question, we need Criteria A1 and A2.

Criterion A1: Linear arrangement of wordforms

Wordforms w_1 and w_2 considered in an utterance of language L have a direct Synt-**D** link between them only if the following condition is satisfied:

The linear position of w_1 or w_2 in this utterance must be specified with respect to the other.

Comments

1. In languages where word order is used semantically—among other things, to express communicative organization (the Rheme/Theme and the Given/New division,

Focalization, Emphasis, Contrast, etc.)—Criterion A1 applies in a limited way: it has to be used only in communicatively neutral expressions, i.e., expressions without extractions, permutations or other communicative transformations of all kinds.

2. "The linear position of the wordform w_1 is specified with respect to the wordform w_2 " means that w_2 either precedes w_1 , or follows it, or else can precede or follow it (optionally or under some conditions). Thus, if in a language, the manner adverbial can indiscriminately precede or follow the verb it modifies, we still formulate the possible positions of the adverbial with respect to the verb, rather than the other way around.

3. The wordform determining the linear position of the other is not necessarily its Synt-governor (cf. Criteria B). Thus, in the phrase PREP \rightarrow N, it is the Synt-governor PREP that is positioned with respect to its Synt-dependent N. Of course, to say that a PREP precedes the N it introduces is logically equivalent to saying that an N introduced by a PREP follows this PREP. However, linguistically, these two statements are not equally acceptable: since a noun can appear without a preposition, while a preposition cannot appear without a noun, it is more natural to specify the place of the preposition with respect to the noun than the other way around. (The etymology itself of the word *preposition* is a witness to this intuition: it is an element that is PRE-posed to the noun.)

4. In some exotic cases Criterion A1 has to be applied with caution. Thus, Serbian has the construction $V_{Aux} \rightarrow V_{non-fin}$, which expresses the past or the future tense of the verb V; here V_{non-fin} is either a Past Participle in the compound past tense, or an Infinitive/a clause with the conjunction DA 'that' in the compound future tense. In this construction it is VAux that is the SSynt-governor: Ja sam-pisao 'I [MASC] have written⁾, Ti si→pisala 'You [FEM] have written⁾, Ja ću→pisati or Ja ću→da pišem 'I will write⁾, etc. However, the linear position of the G(overnor) $[= V_{Aux}]$ is determined without any reference to its D(ependent) [= PPart/Inf/DA-clause]: VAux is a clitic, and Serbian clitics are placed as a cluster, roughly speaking, after the first linear constituent of the clause, whatever the syntactic class and syntactic role of this constituent (some more specific conditions apply; see Milićević's paper in this volume). Nevertheless, before the clitics are placed where they belong, the V_{Aux} clitic-the SSynt-head (= top node) of the clause-serves as the reference point for the linear placement of all the other clause elements, exactly in the same way as any other top node of a clause does. Therefore, even if one of the wordforms w_1 and w_2 is a clitic, Criterion A1 is still fully applicable (but *cum grano salis*).

Criterion A2: Potential prosodic unity

Wordforms w_1 and w_2 considered in a given utterance of language L have a direct Synt-**D** link between them only if one of the following two conditions is satisfied:

- 1. Either w₁ and w₂ can form in L a special prosodic unit—a *phrase* [general case]: e.g., N + V, N + ADJ, V + N, PREP + N, ADV + ADJ, NUM + N, etc.
- Or the wordforms w₁ and w₂ cannot form a phrase, but the wordforms w₁, w₂ and a set of wordforms W can form a phrase such that

 (a) in this phrase, w₁ is the Synt-head and
 - (b) w_2 and W also form a phrase in which w_2 is the Synt-head [special case].

For instance, consider the phrase *escape* $[= w_1]$ *from* $[= w_2]$ *many problems* [= the set **W**], in which *escape* is the Synt-head; *from many problems* is also a phrase, in which *from* is the Synt-head; therefore, we say that *escape* and *from* have a direct syntactic link: *escape*–synt–*from*.¹³ ("X–synt–Y" indicates a non-directed syntactic link between X and Y.)

Of course not every prosodic unit in an actual sentence is a phrase. The concept of phrase needs an elaborate definition, which is outside of my frame here, because it is a concept of the Deep-Morphological, rather than Syntactic, level. I take it to be one of my *indefinibilia*, see Part I, 1, p. 2. However, a phrase must be perceived by speakers as existing in the language, whatever this means.

Criteria A1 and A2 must of course not contradict each other. For instance, in (1), for has to be positioned before decades, and escaped after has, etc.; therefore Criterion A1 indicates the presence of a Synt-**D** in these pairs. Criterion A2 does not contradict this: in (1), for decades is a phrase of English, and so is has escaped (but, e.g., *to new is not); therefore, in the expressions for decades and has escaped the wordforms can be linked by a Synt-**D**. Again in (1), by moving is positioned after escaped, and by before moving (Criterion A1); escaped by moving is a phrase, with escape as the Synt-head, and so is by moving, where the preposition by is the Synt-head; therefore, by Criterion A2b, escaped and by can be linked by a Synt-**D**. Both criteria are again satisfied.

For Criteria A to be satisfied, that is, for there to be w_1 -synt- w_2 , both Criterion A1 and Criterion A2 must be satisfied.

4.3.2 Criteria B: SSynt-dominance

Next, in each pair of wordforms w_1 and w_2 that are syntactically directly linked in the utterance under consideration one of them syntactically dominates the other, i.e., is its Synt-governor. In the phrase w_1 -synt $\rightarrow w_2$ the Synt-governor is the wordform that determines—at least, to a greater extent than the other wordform (its Synt-dependent)—different properties of the phrase according to Criteria B1, B2 and B3.

Criterion B1: The passive SSynt-valence of a phrase

In the phrase w_1 -synt- w_2 , the wordform w_1 is the Synt-governor of w_2 if the following condition is satisfied:

The passive Synt-valence of the whole phrase is determined by the passive Synt-valence of w_1 to a greater extent than by that of w_2 .

Then we have w_1 -synt $\rightarrow w_2$.

To put it differently, the passive Synt-valence of the phrase w_1 -synt $\rightarrow w_2$ is rather that of w_1 than that of w_2 ; the Synt-head of a phrase determines more than any other of

¹³ Here is a more complex case (brought to my attention by N. Pertsov): Rus. [k] domu $[= w_1]$, cvet kryši kotorogo menja razdražaet $[razdražaet = w_2]$ ^c to [the] house [the] color of [the] roof of which irritates me³, where cvet kryši kotorogo menja is W. The wordforms w_1 , w_2 and W constitute a grammatical phrase: domu, cvet kryši kotorogo menja razdražaet, whose Synt-head is domu = w_1 ; w_2 and W also constitute a phrase: cvet kryši kotorogo menja razdražaet, whose Synt-head is razdražaet = w_2 ; therefore, domu and razdražaet are directly linked by a Synt-dependency: domu-synt-razdražaet.

its elements all the external syntactic links of the phrase. (For passive Synt-valence, see Part I, 1, item 14, p. 4.)

Note that Criterion B1 does not require exact distributional equivalence between the Synt-head of a phrase and the whole phrase, as is the case in some similar approaches. For us, it is sufficient if, in the phrase w_1 -synt $\rightarrow w_2$, the wordform w_1 contributes to the passive Synt-valence of w_1 -synt $\rightarrow w_2$ more than w_2 does.

Examples

- (12) a. The passive Synt-valence (= the distribution) of the phrase *for decades* is fully determined by the preposition; therefore, *for*-synt->*decades*.
 - b. Similarly, a phrase like *has escaped* or *does not escape* shows the distribution of, or plays the same Synt-role as, *has/does* (i.e., that of a finite, or tensed, verb) rather than that of the past participle *escaped* or the infinitive *escape*; therefore,

```
has−synt→escaped, does−synt→escape.
```

c. The phrase *Sir Wanner* has the passive SSynt-valence of *Wanner*, not that of *Sir*:

```
I saw Sir Wanner. ~ I saw Wanner. ~ *I saw Sir; therefore, Sir-synt-Wanner.
```

By analogy with such nouns as *Sir* or *Mister* all other nouns possible in this construction (without article) are treated alike:

General—synt–Wanner, Professor—synt–Wanner, President—synt–Wanner, etc.¹⁴

Let us consider now a more difficult case where it is not immediately obvious what element in a construction is the Synt-governor.

(13) French

<i>un</i>	<i>drôle</i>	<i>de garçon</i>	⁽ a strange boy ⁾
a-SG.MASC	strange-SG.MASC	of boy[MASC]-SG	
<i>une</i>	<i>drôle</i>	<i>de voiture</i>	⁽ a strange car ⁾
a-SG.FEM	strange-SG.FEM	of car[FEM]-SG	
<i>ces</i>	<i>drôles</i>	<i>de garçons</i>	⁽ these strange boys ⁾
this-PL.MASC	strange-PL.MASC	of boy[MASC]-PL	
<i>ces</i>	<i>drôles</i>	<i>de voitures</i>	(these strange cars)
this-PL.FEM	strange-PL.FEM	of car[FEM]-PL	

The passive Synt-valence of the phrase *drôle de garçon* (ADJ+*de*+N) is that of a noun and not that of an adjective; what should be taken as the Synt-head of the phrases in (13)?

¹⁴ But in *I saw the Pope John-Paul* (where the article with POPE is necessary) the Synt-**D**s are different: since *I saw the Pope* is perfectly grammatical, we have *the Pope*–synt \rightarrow *John-Paul*.

• Solution I: the noun (GARÇON, VOITURE) is the head. The internal SSynt-structure of the phrase is as follows:

 $v = \frac{v + v}{dr \hat{o}le - synt} - de garçon$

We have then to treat DE not as a preposition, but as a special adjectival marker (homophonous with the preposition DE and depending on the adjective). The adjective agrees in gender and number with its Synt-governor, which is the rule in French.

• Solution II: the adjective (DRÔLE) is the head. The internal SSynt-structure of the phrase is as follows: drôle-synt \rightarrow de-synt \rightarrow garçon. We have then to admit that an adjective of such a type (French has a handful of those: DRÔLE, CHOUETTE, VACHE^[2]) has bizarre SSynt-properties: it can be the head of a noun phrase, while governing a DE-phrase and agreeing with the dependent noun of this DE-phrase, instead of with its own SSynt-governor.

Solution I requires the postulation of a special grammatical element, an adjectival marker DE, which does not exist in French elsewhere, i.e., outside of the construction under analysis. Solution II, on the contrary, requires only the admission of a special character of three French adjectives, which has to be admitted anyway (since even under Solution I, such a construction will be possible only with these adjectives). Moreover, the construction with an adjective that heads an NP while governing a DE-phrase and agreeing with the dependent noun does exist in French independently:

le plus intelligent—synt→*des garçons* 'the most intelligent of the boys' ~ *la plus intelligente*–synt→*des filles* 'the most intelligent of the girls'.

This is an absolutely regular and completely productive construction. Therefore, Solution II has to be preferred. (As we see, the decision is again arrived at by analogy.)

Thus, the examples in (13) show that in more complicated cases one has to proceed with utmost caution. The main tool here is analogy with more normal (= more current, less restricted) constructions. We have to make decisions that will agree with accepted descriptions and try to relegate the eccentricities to restricted sets of phenomena, without allowing these eccentricities to spread on more normal areas of the language.

Criterion B2: The morphological links between the elements of a phrase and its external context

In the phrase w_1 -synt- w_2 in which the passive Synt-valence (of the whole phrase and of its elements) does not allow us to establish the Synt-governor, the wordform w_1 is the Synt-governor of w_2 if the following condition is satisfied:

It is w_1 that controls the inflection of wordforms external to the phrase or w_1 's inflection is controlled by such wordforms.

Then we have w_1 -synt $\rightarrow w_2$.

In the phrase w_1 -synt $\rightarrow w_2$ the wordform w_1 is called the *morphological contact* point.

Examples

- (14) Russian
 - a. The phrase *jubka-štany* lit. 'skirt-pants' does not allow for the application of Criterion B1: both its members are nouns and have the same passive Synt-valence. But Criterion B2 singles out *jubka* as the Synt-governor: *èt+a* [sg.FEM] *jubka-štany byl+a* [sg.FEM] ... 'this skirt-pants was ...', where the external agreement is with *jubka* [FEM.SG], and not with *štany* [PL] (**èt+i jubka-štany byl+i* ...); therefore, the phrase must be represented in the SSyntS as *jubka*-synt→*štany*.
 - b. In the phrase v štate Nebraska ⁽ⁱⁿ [the] state [of] Nebraska⁾, štat is declined regularly (štat+Ø, štat+a, štat+e, ...) in conformity with external context, while Nebraska remains in the nominative (v Nebrask+e, but *v štat+e Nebrask+e); thus, štat is here the morphological contact point, and it is again picked by Criterion B2 as the Synt-governor: štat-synt→Nebraska.
 - c. Similarly, in the phrase *čudo-jabloko* lit. ^(miracle-apple), *jabloko* ^(apple) is the Synt-governor, since it is declined according to the requirements of the external context, while *čudo* is invariable: *čudo-jablok+a*, *čudo-jablok+u*, ..., *čudo-jablok+i*, *čudo-jablok+ami*, ... Thus, we have *čudo*—synt-jabloko.
 - d. In the phrase [*pjat*'] kilogrammov kolbasy '[five] kilos of-sausage', the noun kilogrammov is the Synt-governor, since it is the morphological contact point: [s pjat'ju] kilogramm+ami-synt→kolbasy '[with 5] kilos [of] sausage', [v pjati] kilogramm+ax—synt→kolbasy '[in 5] kilos [of] sausage', etc.
- (15) German

In [*zwei*] *Gläser Wein* lit. ⁽[two] glasses wine⁾, the Synt-governor is *Gläser*, which is the morphological contact point:

a. [*zu diesen zwei*] *Gläser+n Wein* lit. ⁽[to these two] glasses wine⁾, where *Gläsern* is in the dative (*-n*), imposed by the preposition ZU.

b. Dies+e [PL] zwei Gläser Wein sind [PL] notwendig
<sup>(These two glasses [of] wine are necessary⁾,
where Gläser [PL] imposes the plural grammeme on the demonstrative adjective and on the verb.
Therefore, Gläser-synt→Wein.
</sup>

By analogy, the same structure is accepted in the cases where the measure noun remains invariable:

c. *drei Kilo*–synt→*Brot* ⁽three kilos [of] bread⁾ *vierzig Gramm*–synt→*Fleisch* ⁽forty grams [of] meat⁾

(16) In Dutch, the situation is slightly different from that in German: here, the N_{measure} does not inflect under the impact of the external context (it has no case forms), but when in plural, it imposes plural agreement on the verb:

Twee glazen wijn zijn [PL] (**is* [SG]) *noodzakelijk*

'Two glasses [of] wine are necessary';

therefore, in Dutch we also have glazen-synt-wijn.

But in semantically equivalent phrases of Chinese, which has no inflection at all, the Synt-**D**s are different, see (17b).

Criterion B3: The semantic content of a phrase

In the phrase w_1 -synt- w_2 in which neither the passive Synt-valence nor the morphology allows us to establish the Synt-governor, the wordform w_1 is the Synt-governor of w_2 if the following condition is satisfied:

The phrase means 'a kind/an instance of w_1 ' rather than 'a kind/an instance of w_2 '. Then we have w_1 -synt $\rightarrow w_2$.

Examples

- (17) a. In *jam sandwich*, the Synt-governor is *sandwich*, because "*jam sandwich* refers to a kind of sandwich, rather than to a kind of jam" (Hudson 1990: 98).
 - b. In Chinese, where no inflection exists, the phrase shi bàng ròu 'ten pounds [of] meat' consists of morphologically invariable wordforms. Here again, Criterion B3 applies: shi bàng ròu refers to an instance of meat, not to an instance of pounds, so ròu 'meat' is the Synt-governor: shi←synt-bàng←synt-ròu.

One can say (with Zwicky 1993: 295-296) that in a two-word phrase the Synt-governor is the syntactic class determinant, or—if there is no such syntactic determinant—the morphological behavior determinant, or—in case both syntactic and morphological determinants are absent—the semantic content determinant. In one word (Bazell 1949: 11), the Synt-governor is more PROMINENT than its Synt-dependent, namely more prominent syntactically, or else morphologically, or at least semantically.

Most approaches dealing with Synt-**D**s presuppose concord between these properties, i.e., between Criteria B1-B3. In sharp contrast, the Meaning-Text Theory does not require such a concord. Only Criterion B1 is genuinely syntactic; B2 is morphological, and B3 semantic. And we know that the orientations of Sem-D, Synt-D and Morph-D can differ (cf. 5, p. 58ff); therefore, we must expect that these criteria will be in conflict more often than not. In fact, Criteria B1-B3 form a hierarchy: B1 > B2 > B3. Thus, if Criterion B1 is applicable, its indication is sufficient. Only if it is not applicable (because w_1 and w_2 are both of the same part of speech and thus have the same passive Synt-valence), Criterion B2 applies—but only in a language having inflection and only for w_1 and w_2 with different morphological properties. Otherwise, Criterion B3 applies. Therefore, these criteria are never applied together (= simultaneously) and, as a result, they cannot contradict each other. To put it in a slightly different form: I believe that, if we can use passive valence considerations, then the ability of Synt-governors to control the inflectional form of their Synt-dependents or to have their own inflectional form controlled by a Synt-dependent, as well as their ability to be or not to be semantically dominant should not be taken into account when deciding on the Synt-governor status of a wordform. Morphological and semantic properties of heads are, as already stated, freely distributed among Synt-governors and Synt-dependents, so that a consistent combination of these properties cannot be expected. Morphological and semantic properties should be recurred to only if syntactic properties fail.

For Criteria B to be satisfied, at least one of the Criteria B1-B3 must be satisfied, such that other Criteria B higher in the hierarchy are not applicable.

The criteria for the orientation of Synt-**n** (Governor-*vs*.-Dependent problem) are thoroughly discussed in Pittman 1948, Zwicky 1985, 1993, Hudson 1987, 1990: 106-107, and in Corbett *et al.* 1993. For a more rigorous formulation of Criterion B1, see Mel'čuk 1988: 132-135.

Criteria B1-B3 call for the following two important remarks.

First, Criteria B1-B3 are language-specific: if, in the phrase X-synt-Y of language L, these criteria pick X as the Synt-governor, i.e., if we have X-synt-Y, this will not necessarily be the case for a synonymous construction with similar parts of speech in some other language. Thus, Russian and German feature $N_{measure}$ -synt-N, because $N_{measure}$ is the morphological contact point (cf. [v pjati] kilogramm+ax kolbasy in (14d) and [zu diesen zwei] Gläser+n Wein in (15)); yet it does not follow that N syntactically depends on a quantifying $N_{measure}$ in any language: in a language where the $N_{measure}$ does not inflect under the impact of an external wordform and does not itself control the inflection of the quantified N, Criterion B3 picks this N as the Synt-governor: cf. (17b), where we have $N_{measure}$ -synt-N in Chinese.

Second, Criteria B1-B3 are inherently insufficient: there are cases where all the three fail. This must happen where, in a phrase X-synt-Y, both X and Y are of the same part of speech, neither does inflect nor can impose different inflections, and both are semantically "equal." Take, for example, a Russian phrase of the type včera utrom lit. 'yesterday morning' or segodnja popozže lit. (today a-bit-later). Both wordforms in such a phrase are adverbs, both have no morphology, and both denote a stretch of time; which one is the Synt-governor? Note that both are equally omissible: Alen priexal včera 'Alan came yesterday' and Alen priexal utrom 'Alan came in the morning'. In such cases, a more or less arbitrary solution imposes itself: in Russian, the preceding element will be taken as the Synt-governor, so that we have $v\check{c}era$ -synt- $\rightarrow utrom$, segodnja-synt- $\rightarrow popoz$ - $\tilde{z}e$. However, there could be semantic motivation for this solution, after all: 'yesterday' and 'today' are in a sense more important than 'in the morning' and 'later', since 'yesterday¹/today¹ denotes a whole day, of which 'in the morning¹/later¹ is but a part. Then Criterion B3 applies: 'yesterday morning' is a particular moment of yesterday, and 'today later' is a particular moment of today. We can also invoke a syntactic consideration: *yesterday morning* represents a kind of coordination, and in coordinate strings, the subsequent element depends on the preceding one.

An even more problematic case is that of compound numerals in languages where numerals are morphologically invariable themselves and do not govern special inflections of the quantified nouns.¹⁵ Take, for instance, Fr. *soixante-neuf* ⁽⁶⁹⁾. Since both its components are numerals, Criterion B1 is not applicable (*soixante* and *neuf* have the same passive Synt-valence); since almost all French numerals have no morphology and do not affect the morphology of the noun quantified, Criterion B2 is not applicable,

¹⁵ Otherwise, numerals do not create problems. Thus, in Russian, in *šest'desjat tri* ⁽⁶³⁾ the Synt-head is *tri*, because in compound numerals the last numeral is the morphological contact point: *šest'desjat*—*tri stol+a*, but *šest'desjat*—*pjat'* ⁽⁶⁵⁾ *stol+ov* and *šest'desjat*—*odin* ⁽⁶¹⁾ *stol+O*. This means that Criterion B2 applies here and indicates the Synt-governor.

either; finally, their meanings are strictly of the same type (= numbers), so that neither Criterion B3 can be used. The only way open is then to reason by analogy. The compound numeral *soixante et un* lit. ⁽⁶⁰ and 1⁾ (and a few others with 1 as the last digit) would suggest the Synt-dominance *soixante* $\rightarrow et \rightarrow un$; by analogy with regular conjoined strings of the type *Alan* $\rightarrow and \rightarrow Leo$ or *beautiful* $\rightarrow and \rightarrow intelligent$. But then consider two facts that contradict this solution:

• The numeral UN 'one' agrees in gender with the noun quantified: *vingt et un* [MASC] *garçons* 'twenty-one boys' vs. *vingt et une* [FEM] *filles* 'twenty-one girls'; according to Criterion B2, it is UN that must be the Synt-head of the compound numeral.

• Take the ordinals, such as *soixante et unième* (sixty-first) or *soixante-cinquième* (sixty-fifth) (similarly, *soixante et onzième* lit. (60 and 11th) = (71st) et *quatre-vingt-onzième* lit. (80-11th) = (91st)); here the Synt-governor is, according to Criterion B1, clearly the ordinal numeral *unième* (1st), *cinquième* (5th) and *onzième* (11th), i.e., the last numeral in a compound ordinal:

trois-*cent*-*soixante*-*cinquième* '365th', *trois*-*cent*-*soixante*-*et*-*onzième* '371st', etc.

Then, continuing our analogy and taking these two facts into account, we arrive at the same SSyntS in compound cardinals:

 $trois \leftarrow cent \leftarrow soixante \leftarrow cinq$ (365).

And, of course,

 $trois \leftarrow cent \leftarrow soixante \leftarrow et \leftarrow un$ (361).

In a language like German, where some numerals are regularly linked by a conjunction (*und* ⁽and⁾), this gives the following SSynt-structures:

drei - hundert - fünf - und - sechzigster (365th)

lit. 'three hundred five and sixtieth',

where *sechzigster* 'sixtieth' is clearly the Synt-head of the compound ordinal numeral; in a similar way, *drei*—*hundert*—*fünf*—*und*—*sechzig* '365' lit. 'three hundred five and sixty'.

It is possible that elements like 'and' (Fr. *et*, Ger. *und*) that appear within compound numerals should not be considered coordinate conjunctions; then the SyntSs shown above would look less exotic; cf. the Chukchee marker of compound numerals in (24c), p. 44.

4.3.3 Criteria C: Labeled SSynt-dependencies

For each pair of wordforms which are directly linked by a Synt- \mathbb{D} (for w_1 -synt $\rightarrow w_2$), one has to know exactly which specific Synt- \mathbb{D} links them. In order to represent successfully the SyntSs of utterances, we have to use different types of Synt- \mathbb{D} . Thus, I \leftarrow synt-LOVE-synt \rightarrow YOU does not distinguish between 'I love you' and 'You love me'; [JOHN] I \leftarrow synt-SEND-synt \rightarrow YOU can be implemented as *John sends me to you* or *John sends you to me*; etc. In all these and a host of similar cases, different types of Synt- \mathbb{D} s, or different Surface-Syntactic Relations [= SSyntRels] have to be distinguished:

I← r_1 -LOVE- r_2 →YOU, [JOHN] I← r_2 -SEND- r_3 →YOU, ...

This means that SSyntRels must be labeled, the label being meaningful (as is the case with Morph- \mathbf{n}): the label \mathbf{r} of a SSyntRel refers to a family of specific syntactic constructions which implement the SSyntRel \mathbf{r} in the DMorphS of the sentence. Thus, consider the label "subj(ectival)" of a SSyntRel in Russian, i.e., the SSyntRel that appears in phrases of the type:

$$V_{fin}[= w_1] - subj \rightarrow N[= w_2]$$

(*Mal'čik prixodit/Devočka prixodit* 'The boy/The girl comes'; *Mal'čik prišël/Devočka prišla* 'The boy came/The girl came'). The label "**subj**" identifies a set of SSynt-rules that make the finite verb $\mathbf{w_1}$ agree with the noun $\mathbf{w_2}$ in person and number (if the verb is in the present or the future) or in number and gender (if the verb is in the past or the subjunctive); these rules also position $\mathbf{w_2}$ [= N], with respect to $\mathbf{w_1}$ [= V_{fin}]. In other words, the SSyntRel "**subjectival**" is the signified (= Saussure's *signifié*) of every construction in this family. Generally speaking, a SSyntRel is a component of a linguistic sign, whose signifier is the construction in question (an ordered pair of lexemes of particular syntactic classes with particular morphological characteristics).¹⁶

In phrases of the form w_1 -synt $\rightarrow w_2$, the Synt-**D** that links the two wordforms can be labeled **r** (i.e., it can be the SSyntRel **r**) only if it satisfies the following three criteria: C1-C3. If at least one of Criteria C1-C3 is not satisfied, the presumed SSyntRel **r** should be split in two (or more) SSyntRels.

Criteria C are exploited in Iordanskaja & Mel'čuk's paper in this volume: it is shown that these formal SSynt-criteria go hand in hand with substantial requirements concerning SSynt-properties of the dependent members w_j of the phrases under analysis. These w_j must share enough relevant SSynt-properties in order for the phrases w_i -? $\rightarrow w_j$ to be described by one SSyntRel r. In other words,

Criteria C are applied only to phrases $w_i \rightarrow w_j$ such that the common SSynt-properties of w_i make it reasonable to infer that the SSyntRel in question is \mathbf{r} .

If the observable SSynt-properties of the dependent in $\mathbf{w_i}$ -? $\rightarrow \mathbf{w_j}$ phrases do not warrant the same SSyntRel, the question of Criteria C does not even arise. Only if the researcher, based on SSynt-properties of $\mathbf{w_j}$, wants to describe the target set of phrases by the same SSyntRel **r**, he has to use Criteria C in order to see whether they allow him to do so. Thus, the similarity of SSynt-properties of $\mathbf{w_j}$ pushes one to describe the corresponding phrases by the same SSyntRel **r**, while Criteria C allow/disallow this.

Criterion C1: Absence of semantic contrast

Notation: $\mathbf{w}(\mathbf{L})$ is a wordform of lexeme L (\mathbf{w}_i and \mathbf{w}_j can be different or identical); \oplus is the operation of linguistic union, which links signs, in particular—wordforms, according to their syntactics and general rules of L.

A SSyntRel **r** must not describe two different phrases $\mathbf{w_1}(L_1) \oplus \mathbf{w_2}(L_2)$ and $\mathbf{w_3}(L_1) \oplus \mathbf{w_4}(L_2)$, where $L_1 - \mathbf{r} \rightarrow L_2$, which 1) contrast semantically $[{}^{(v_1}(L_1) \oplus \mathbf{w_2}(L_2){}^{)} \neq {}^{(v_3}(L_1) \oplus \mathbf{w_4}(L_2){}^{)}]$ and 2) differ formally only by some syntactic means of expression (i.e., by word order, syntactic prosody, or syntactic grammemes).

Criterion C1 corresponds to what is known in linguistics as the *minimal pair test*, which is used in phonology (two phones cannot be relegated to one phoneme if they are

¹⁶ A SSyntRel is by no means a meaning; but a signified is not necessarily a meaning—it can be, for instance, a command to perform some modifications in the syntactics of a sign (as is the case with voices). I cannot, however, enter here in the discussion of the types of linguistic signifieds.

the only distinguishers of the signifiers of two semantically contrasting word-forms), morphology, and semantics.

Examples

(18) a. In Russian, the construction DESJAT'←r[?]–DOLLAR has two different implementations with different meanings:

desjat' dollarov ⁽¹⁰ dollars⁾ vs. *dollarov desjat'* ^{(approximately 10 dollars⁾. The formal difference between the two phrases is purely syntactic: word order. Therefore, the presumed SSyntRel **r** must be split in two different SSyntRels:}

DESJAT'←quantitative–DOLLAR ⇔ desjat' dollarov DESJAT'←approximate-quantitative–DOLLAR ⇔ dollarov desjat'

b. In English, the construction STARS–**r**[?]→VISIBLE also has two different implementations with different meanings (Quirk *et al.* 1985: 419; cf. Mel'čuk & Pertsov 1988: 136-137):

the visible stars ⁽stars that can be seen in principle⁾ vs.

the stars visible 'stars that can be seen currently'.

The formal difference between the two phrases is again syntactic: word order. Therefore, there are two different SSyntRels as well:

VISIBLE ←modificative −STARS	⇔	the visible stars
-------------------------------------	---	-------------------

VISIBLE←**post-modificative**−STARS ⇔ *the stars visible*

Cf. other pairs with the same semantic/formal contrast: *navigable rivers* ~ *rivers navigable*, *a written word* ~ *a word written*, *a sitting figure* ~ *the people sitting*, etc.

Criterion C2: Syntactic substitutability

The first formalization of the substitutability of syntactic subtrees as a means for establishing SSyntRels was proposed by the German researcher J. Kunze (Kunze 1972: 23; see also Kunze 1975: 5.3, p. 235*ff*): what is known as the *Kunze property*. I start with presenting it here, in order to show that a weaker version of it must be preferred.

Let there be, in **L**, lexemes L(X), L(Y), ... of syntactic classes X, Y, ..., complete SSynt-configurations $\Delta_{(Z)}$ and $\Delta_{(W)}$ (i.e., subtrees having as their top nodes lexemes L(Z) and L(W)), and a SSyntRel **r**.

Definition 6: The Kunze property

A SSyntRel **r** has the Kunze property if and only if for any pair of SSynt-configurations $L(X)-\mathbf{r}\rightarrow \mathbf{\Delta}_{(Z)}$ and $L(Y)-\mathbf{r}\rightarrow \mathbf{\Delta}_{(W)}$, replacing $\mathbf{\Delta}_{(Z)}$ by $\mathbf{\Delta}_{(W)}$ and vice versa does not affect their syntactic well-formedness.

Note that violation of semantic/lexical constraints is not considered as syntactic illformedness. Thus, cf. *inside the car*, but **inside Stuttgart* or *according to Leo*, but **according to the car*; however, the starred phrases are considered as syntactically wellformed (PREP + N being a legitimate phrase of English).

For a SSyntRel that has the Kunze property, any of its potential D(ependent)s can be attached to any of its potential G(overnor)s (= all Ds of a SSyntRel are mutually substitutable in all SSyntSs *salva correctione*). In Mel'čuk 1988: 142 it was required that any

SSyntRel in any \mathbf{L} has the Kunze property.¹⁷ Now, however, I think that the Kunze property is too rigid, since it does not allow for some desirable generalizations. For instance, it does not admit the same SSyntRel for nominal and infinitival SSynt-Subjects, as in the following French sentences:

(19) La course \leftarrow **r**-fatigue lit. 'The running tires'. Courir \leftarrow **r**-fatigue lit. 'To-run tires'.

Since not any verb in French takes an infinitive as its SSynt-Subject (**Pleuvoir m'a surpris* lit. ^{(To-rain has caught me [somewhere]²), the SSyntRel **r** in (19) does not possess the Kunze property: with L(X) = SURPRENDRE, $\Delta_{(Z)} = NP$ (e.g., *La pluie*(**r**-*surprend*) and L(Y) = FATIGUER, $\Delta_{(W)} = Infinitival Phrase ($ *Courir*(**r**-*fatigue* $), the replacement produces the syntactically ill-formed configuration <math>*V_{inf}$ (**r**-SURPRENDRE ^(catch N) [somewhere]²). As a result, using the Kunze property leads to having two different SSynt-Rels for nominal and infinitival SSynt-Subjects (as stated in Kunze 1975: 279). But I think that in (19) the SSyntRel **r** should not be split: all the SSynt-Subjects, whether nominal or infinitival, share a set of important unique properties, and it is preferable to describe all of them by the same SSyntRel.}

Therefore, I propose to use the *quasi-Kunze property*, which is weaker: substituability is required only in one direction and only by at least one particular subtree (which is not a substitute pronoun, since substitute pronouns—see Comment 1, p. 16 constitute a "secondary" syntactic class, being introduced by a transformational rule), rather than in both directions and by any subtree. (The concept of the quasi-Kunze property has been elaborated jointly with L. Iordanskaja; it is introduced in Iordanskaja & Mel'čuk 2000; see in this volume, p. 164*ff*. Another weaker version of the Kunze property was considered in Mel'čuk 1977: 261.)

Definition 7: The Quasi-Kunze property

A SSyntRel **r** has the quasi-Kunze property if and only if there exists in **L** a syntactic class (\approx part of speech) X, which is different from substitute pronouns and such that for any SSynt-configuration L–**r**→ $\Delta_{(Y)}$, replacing $\Delta_{(Y)}$ by $\Delta_{(X)}$ (but not necessarily vice versa!) in any SSyntS does not affect its syntactic well-formedness.

The element $\Delta_{(X)}$ that passes with any G of the SSyntRel r is nothing else but the prototypical D of the SSyntRel r.

The SSyntRel **r** in (19) possesses the quasi-Kunze property, since this **r** has a prototypical D—a prepositionless noun: in French any finite verb admits a nominal SSynt-Subject.¹⁸ As a result, the SSyntRel **r** is allowed: this is the **subjectival** SSyntRel.

¹⁷ A property similar to the Kunze property was used for the identification of SSyntRels in the METATAXIS system (see Schubert 1987: 10: "Interchangeable dependents are grouped in classes and the relations that are definitional for these classes are given names").

¹⁸ The SSynt-Subject of impersonal verbs (PLEUVOIR ⁽[to] rain⁾, NEIGER ⁽[to] snow⁾, etc.)—the impersonal IL—is considered as a particular case of noun (= a pronominal noun, which is not a substitute pronoun). Note that with the Kunze property, SSynt-Subjects in II [= Alan, i.e., a substitute pronoun] *dort* ⁽He is sleeping⁾ and II pleut ⁽It is raining⁾ must be described by two different SSyntRels.

While in Definitions 6 and 7 the G is a particular lexeme, $\Delta_{(Y)}$ is considered up to the syntactic class. Thus, for instance, different prepositions are not distinguished: the SSyntRel **r** in the phrases *insist*-**r** \rightarrow *on*, *supply*-**r** \rightarrow *with* and *compare*-**r** \rightarrow *to* has the quasi-Kunze property, because a PREP+N phrase can be substituted for the D of **r** with any of these verbs, provided the appropriate preposition is chosen according to the verb's Government Pattern.

Criterion C2 can now be formulated as follows:

Criterion C2

Any SSyntRel r must possess the quasi-Kunze property.

Criterion C2 is a particular case of what is known in linguistics as the *substitution criterion*, or *substitution test*, except that here we deal with the substitution of sub-trees which have to hang on the same SSyntRel **r**.

Examples

(20) a. In the Russian phrases $v-\mathbf{r}[?] \rightarrow Pariž$ 'to Paris' and $\check{c}toby-\mathbf{r}[?] \rightarrow \check{c}itat'$ 'inorder-to read' the presumed SSyntRel $\mathbf{r}[?]$ does not possess the quasi-Kunze property:

* $\check{c}toby-\mathbf{r}[?] \rightarrow Pari\check{z}, *v-\mathbf{r}[?] \rightarrow \check{c}itat'$

In other words, Russian has no prototypical D for this SSyntRel: no element can pass with both a preposition and a conjunction; therefore, there are two different SSyntRels:

V–prepositional→PARIŽ and ČTOBY–conjunctional-infinitival→ČITAT'.

b. In the phrases *have*−**r**[?]→*been* and *be*−**r**[?]→*going* the presumed SSyntRel **r** does not possess the quasi-Kunze property:

*have-r[?]→going and *be-r[?]→been; therefore, there are two different SSyntRels: HAVE-perfect-analytical→BEEN and BE-progressive-analytical→GOING.

Criterion C3: Repeatability with the same Synt-governor

Consider the possible number of occurrences of a given SSyntRel \mathbf{r} with the same Governor. In this respect, any \mathbf{r} of language \mathbf{L} can be either non-repeatable or unlimitedly repeatable.

Definition 8: Non-repeatable SSyntRel

A SSyntRel \mathbf{r} is *non-repeatable* if and only if no more than one branch labeled \mathbf{r} can start from any G(overnor).

In a given sentence of **L**, a G of a non-repeatable **r** can have only one D (= one clause element) of the corresponding type. For instance, actantial SSyntRels whose Ds are marked by purely syntactic means (word order, prosody, inflection)—such as the **subj** and the **dir**(ect)-**obj**(ectival) SSyntRels—are obligatorily non-repeatable: otherwise, they would violate Criterion C1, because their Ds would contrast semantically, while differing only in syntactic means. (Only actantial SSyntRels whose Ds are marked by lexical means, that is, by different prepositions—such as the **obl**(ique)-**obj** SSyntRel—can be repeatable.)^[3]

An important warning: In some languages, a clause element can be duplicated by what is called a *resumptive clitic*. Such is, for instance, the D of the dir-obj SSyntRel in Spanish, where we have the construction of the type (21a):

 $\overrightarrow{A \ Alain \ le \leftarrow dir-obj}$ $\overrightarrow{A \ alain \ le \leftarrow dir-obj-veo \ todos \ los \ días \ lit.$ (To-Alan him [I] see every day) [a human DirO in Spanish is introduced by the preposition A (to)].

I do not consider pronominal duplication of a clause element as repeatability: such duplication has a grammaticized character and is "orthogonal" to the genuine cooccurrence of SSyntRels, since the noun and the clitic that duplicates it are necessarily coreferential. In spite of expressions of the type (21a), the dir-obj SSyntRel is considered non-repeatable in Spanish. Similarly, in spite of (21b), the indir-obj SSyntRel is also considered non-repeatable in French:

b. French

indir-obj _______ A mes enfants, je leur←indir-obj-permets tout lit. 'To my children, I permit them everything'.

Here is another example of grammaticized duplication by clitic:

- c. Albanian
 - (i) *Mësuesi u foli fëmijëve* lit. ^{(The-teacher to-them talked to-the-kids⁾, where *fëmijëve* and *u* also are both IndirOs.}
 - (ii) *Njerëzit më panë mua* lit. ^{(The-people me saw me⁾, where *mua* and *më* are both DirOs.}

In all such cases, the corresponding SSyntRel is considered non-repeatable.

Definition 9: (Unlimitedly) repeatable SSyntRel

A SSyntRel \mathbf{r} is (unlimitedly) repeatable if and only if several branches labeled \mathbf{r} can start from a G (with the exclusion of resumptive clitics).

For instance, the **modificative** and the **circumstantial** SSyntRels in English are unlimitedly repeatable; so is the **obl-obj** SSyntRel. For a repeatable SSyntRel **r** the number of branches labeled **r** that can start from a G in any particular case is theoretically unlimited, although in practice this number can be limited either by pragmatic considerations or by the lexicographic properties of concrete Gs, for instance, by their Government Pattern—as is the case with the **obl-obj** SSyntRel; this number cannot be limited by any general syntactic factors.

In other words, a SSyntRel cannot be limitedly repeatable (without being constrained by Government Pattern of the G).

Now I can formulate Criterion C3:

Criterion C3

Any SSyntRel **r** must be unlimitedly repeatable or non-repeatable.

As is always the case, exceptions are possible. Thus, in English, the relative SSynt-Rel is non-repeatable: generally speaking, a noun cannot have more than one relative clause. There is, however, a contradicting phenomenon: two restrictive relative clauses with the same noun are possible under specific conditions in highly colloquial speech, cf. (22):

- (22) a. A student [who comes to my class]₁ [who broke the news to me]₂ left the building.
 - b. We are in the room [I will never forget]₁ [where she kissed me for the first time]₂.

If we decide—in spite of their marginality—to consider such facts, they can be fully and exactly circumscribed. Therefore, they constitute a legitimate exception, which does not prevent us from declaring the relative SSyntRel non-repeatable in English. (This case has been brought to my attention by L. Iomdin.)

Criterion C3 corresponds roughly to the *cooccurrence*, or *iteration*, *test*, used in linguistics on all levels of analysis. Thus, in morphology, an element of a morphological category is either non-repeatable (tense or number in English or French) or unlimitedly repeatable (the causative in Turkish). When we see, for instance, just two possible repetitions—like nominal case suffixes in Basque or Georgian, we speak of two different case categories (semantic case *vs*. syntactic case; governed case *vs*. agreeing case).

Example

(23) In Persian, we find extremely widespread expressions of the following type:

a. *Rāmin+rā*←**r**[?]–*kard*–**r**[?]→*bedār* Ramin DirO made awakening [Noun] lit. ⁽[He/she/it] Ramin made [the] awakening⁾. = ⁽He/she/it awoke Ramin⁾.

These expressions are built on verbal collocations of the type *bedār kard* 'awakening make Y' = 'awake Y' or *dars dad* lit. 'lesson give Y' = 'teach Y', which, although they seem to include a DirO, such as *bedār* or *dars*, behave as transitive verbs and take—as a whole—a genuine DirO (the suffix **-rā** is an unmistakable marker of a DirO with typically transitive verbs meaning 'kill', 'see', 'build', etc.).

The presumed SSyntRel r [dir-obj(ectival)?] in such expressions would be limitedly repeatable—just twice, while no obvious naturally-looking conditions can be formu-

lated; at the same time, this phenomenon can by no means be treated as an exception. Therefore, there are two different SSyntRels:

RĀMIN**←dir-obj**–KARD**–quasi-dir-obj**→BEDĀR.

The nominal element in verbal collocations of the above type is considered to be a Quasi-Direct Object. Here is another similar example (Lazard 1994: 93):

b. Samāvar ateš kardand samovar fire do-PAST.3PL
lit. ⁽[They] samovar fire did⁾. = ⁽They lit the samovar⁾. [⁽fire do Y⁾ = ⁽light (up) Y⁾]

The SSyntS is here as in the previous example:

SAMĀVAR**←dir-obj**−KARD**−quasi-dir-obj**→ATEŠ.

A very similar situation exists in Korean (O'Grady 1991: 236):

c. John+i enehak +ul kongpwu +lul hay+ss +ta SUBJ linguistics ACC study ACC do PAST DECL(arative) lit. 'John linguistics study made'. = 'John studied linguistics'. ['make study Y' = 'study Y']

"SUBJ" stands for "subjective case," known in traditional Korean grammar as nominative; however, this case is never used for naming, but has as its central function the marking of the syntactic Subject, so that it seems preferable to call it the subjective.¹⁹

The SSyntS of (23c) is again very much the same as before: ENEHAK←dir-obj-HAY-quasi-dir-obj→KONGPWU.

For Criteria C to be satisfied, all three criteria C1-C3 must be satisfied.

The SSyntRels of a language form a systematic inventory, just like phonemes or grammemes. Criteria C1-C3 are part of a methodology for establishing SSyntRels' inventories. Note that Criteria C1 and C2 are paradigmatic, while Criterion C3 is syntagmatic.^[4]

Now we are ready for a definition of Synt-**D**.

Definition 10: Syntactic dependency

The wordform w_2 is said to syntactically depend on the wordform w_1 via SSyntRel r in a particular utterance if the three groups of Criteria A, B and C are satisfied for this pair of wordforms and r.

The syntactic dependency **r** between wordforms w_1 and w_2 will be written as w_1 -r \rightarrow w_2 .

¹⁹ Sentence (23c) can be put into the passive; then the DirO ENEHAK becomes the Subject, the Quasi-DirO KONGPWU is a Quasi-Subject, and JOHN (also in the subjective), a proleptic Theme:

4.4 The logical and linguistic properties of syntactic dependency

a) Synt-**D** is anti-symmetrical:

w_1 -synt $\rightarrow w_2$ entails $\neg(w_1 \leftarrow synt - w_2)$, i.e., $w_1 \leftarrow synt - w_2$.

This means that a wordform w_1 cannot be the Synt-governor of another wordform w_2 and simultaneously have w_2 as its own Synt-governor. This follows from our decision to use the dependency tree as the formalism for the representation of Synt-structures. Moreover, since most often w_1 -synt $\rightarrow w_2$ signals that w_2 is linearly positioned with respect to w_1 , it is paradoxical to claim that at the same time w_1 -synt- w_2 , so that w_1 is linearly positioned with respect to w_2 (or vice versa—in case where w_1 is positioned with respect to w_2).

This means that a wordform cannot be linearly positioned with respect to itself. As with Sem-**n**, anti-reflexivity of Synt-**n** follows from its anti-symmetry.

c) Synt-**D** is anti-transitive:

 w_1 -synt $\rightarrow w_2$ and w_2 -synt $\rightarrow w_3$ (in one sentence) entails $\neg(w_1$ -synt $\rightarrow w_3$). Otherwise, the principle of the unique governor—see below, item e)—would be violated. This does not preclude, however, the presence of an indirect Synt-**D** between w_1 and w_3 : w_3 is part of the Synt-subtree hanging from w_1 .

d) Synt-**D**s must be distinctively labeled: to properly represent *Mary loves John*, in the phrases $Mary \leftarrow \mathbf{r_1}$ -love and $John \leftarrow \mathbf{r_2}$ -love the SSyntRels $\mathbf{r_1}$ and $\mathbf{r_2}$ must be different; otherwise the semantic contrast will not be preserved in the SSyntS. (The SSyntS $Mary \leftarrow \mathbf{r}$ -loves- $\mathbf{r} \rightarrow John$ does not show who loves whom.)

e) Synt-**1** presupposes the uniqueness of the governor: a wordform can syntactically depend only on one other wordform (or be independent, as is the top node of a SyntS).

f) Synt-**D** is universal in the following three respects: it is present in all languages; it appears in all sentences of a language; and it embraces all wordforms of a sentence (that is, for a sentence, Synt-**D**s always form a connected structure—like Sem-**D**s, but unlike Morph-**D**s).

The logical properties of Synt-**n** as defined above correspond to the fact that Synt-**n**s between the wordforms of a sentence form a dependency tree: a connected graph in which 1) each node can directly depend only on one other node (= the uniqueness of the Synt-governor) and 2) one and only one node does not depend on anything—the *top node*, or the *root*, of the SSyntS (= the presence of the absolute head). The linear order of the nodes in the SSyntS is of course not explicitly specified; in this way, the **n**-description of the SSyntS consistently separates the SSynt-links between wordforms and the linear order of the latter. (Word order in a sentence is determined by its SSyntS; it is computed by syntactic rules of the language on the basis of Synt-**n**s.)

Examples of Deep-Synt-**n**s and Surface-Synt-**n**s, i.e., DSyntRels and SSyntRels, are given in Figures 2 and 3. For a detailed description of the SSyntRels of English, see Mel'čuk & Pertsov 1987: 85-156 (and 4.8 below), as well as Apresjan *et al.* 1992: 71-

121; the inventories of SSyntRels for Russian are found in Mel'čuk 1974: 221-235, and Apresjan *et al.* 1989, 1992: 204-208; for the inventories of SSyntRels (\approx "dependent types") for German, Danish, Polish, Bangla, Finnish, Hungarian, Japanese, and Esperanto, see Maxwell & Schubert 1989; a list of Synt-**n**s, illustrated in English, is attached to Petkevič 1995. A sketch of syntactic word order rules based on Synt-**n**s for Russian, see Mel'čuk 1967 and 1974: 260-302; see also Sgall *et al.* 1995 (for Czech and German).

4.5 Some non-definitorial properties of Governors and Dependents

Synt-governors and Synt-dependents have three important properties, which, however, cannot be taken as definitorial: some Synt-governors and some Synt-dependents in particular languages do not have them. Nevertheless, these properties are sufficiently characteristic of Synt-governors and Synt-dependents, so that they can be resorted to as convenient heuristic means. These properties are omissibility, cooccurrence control, and incorporability.

Omissibility

This is the most important non-definitorial property that distinguishes Synt-governors and Synt-dependents. Typically, in the configuration w_1 -synt $\rightarrow w_2$, the Synt-dependent w_2 can be omitted without affecting the Synt-correctness of the SSyntS (and without producing an ellipsis), while the Synt-governor w_1 cannot. Such is the case in the constructions ADJ \leftarrow N, N \rightarrow N_{gen}, V \rightarrow PREP + N, X \rightarrow Conj_{coord} + Y, and a few others. (Let it be emphasized that we speak here of omissibility in the Synt-structure, not in the actual sentence.) But this is not always the case:

• The Synt-dependent may be obligatory (= non-omissible): either in some contexts (e.g., the DET in a DET \leftarrow N construction), or always—as in exocentric constructions (e.g., the N in a PREP \rightarrow N construction). Cf., for instance, non-omissible adjectives in phrases like *a man of various talents*.

• The Synt-governor can be omissible: for example, 1) the Russian preposition OKOLO 'about' with a numeral phrase (*okolo trëx tonn* 'about three tons' is syntactically equivalent to *tri tonny* 'three tons') or the English prepositional configuration *from - to*, again with a numeral phrase (*from three to six girls* is syntactically equivalent to *six girls*); 2) the English subordinate conjunction THAT (*John knows that Mary is in town* is syntactically equivalent to *John knows Mary is in town*).

Cooccurrence (= Subcategorization) control

Typically, in the configuration $\mathbf{w_1}$ -synt $\rightarrow \mathbf{w_2}$, it is the Synt-governor $\mathbf{w_1}$ that is subcategorized for by the Synt-governor \mathbf{w} of the whole phrase $(\mathbf{w}$ -synt $\rightarrow \mathbf{w_1})$. To put it differently, the lexicographic description of \mathbf{w} must take into account some properties of $\mathbf{w_1}$, but not of $\mathbf{w_2}$. Thus, if a verb admits a noun as its actant, the lexicographic properties of the noun may be relevant (this verb admits only human nouns, or only mass nouns, etc.); but it is not the case that a verb admits as its actant a noun with a particular determiner—say, only with EVERY, or only with A/AN, etc.^[5] This fact points to N as the Synt-governor in the constructions DET N or ADJ N. Similarly, in the construction CONJ_{subord} $\rightarrow V_{fin}$ (... whether [he] comes; ... that [I] am [here]), it is CONJ_{subord} that determines the subcategorization of the MV in the matrix clause: some verbs take

WHETHER, some others take THAT, etc.; but the verb of the subordinate clause is immaterial in this respect. Consequently, we have $V_{matrix} \rightarrow CONJ_{subord}[\rightarrow V_{fin}]$.

More generally, the Synt-governor w_1 tends to subcategorize for its Synt-dependent w_2 (i.e., w_1 tends to determine the choice of w_2): we say *many*—*books*, but *much*—*noise*, etc.; or else *depend*—*on*, but *borrow*—*from*, etc.

Incorporability

Typically, a language with incorporation manifests two phenomena concerning the orientation of Synt- \mathbf{D} in a configuration $\mathbf{w_1}$ -synt- $\mathbf{w_2}$:

• Internal incorporability. If w_2 can be incorporated into w_1 , and not the other way around, then w_1 is the Synt-governor of w_2 ; if w_2 has its own dependents, they can be incorporated together with it into w_1 or remain stranded in the sentence (as a function of the language and the context). Well-known examples include the incorporation of actants into the verb and of modifying adjectives into the modified noun. Cf., for instance (the incorporated stem is boldfaced):

```
(24) a. Chukchee (Chukchee-Kamchatkan family, Russia)

n∂+tur+qine+te←-synt-kupre+te 'with [a] new net'

ADJ new 3SG SG.INSTR net SG.INSTR

vs.

tur+kupre+te 'with [a] new-net'

new net SG.INSTR
```

[Non-incorporated adjectives in Chukchee have a special prefix **nə**-, marking them as adjectives, and a person/number suffix, here -**qine**; they agree with the modified noun in number and case.]

• External incorporability. If w_1 (or both w_1 and w_2 , but not w_2 alone) can be incorporated into the Synt-governor w of the whole phrase, then w_1 is the Synt-governor of w_2 ; we can thus have $[w+w_1]$ -synt $\rightarrow w_2$, but not $*[w+w_2]$ -synt $\rightarrow w_1$. Again, if w_2 has its own dependents, they can be incorporated with it into w or remain stranded. However, it seems impossible to have a Synt-dependent of w_i incorporated, while w_i itself is not (Allen *et al.* 1984):

b. Sc	b. Southern Tiwa (Kiowa-Tanoan family, New Mexico, USA)				
W	isi	seuan+	$in \leftarrow synt - bi + m\tilde{u} + ban$ ^{(I} saw two men ⁾ .		
tw	0	man	PL 1SG see PAST		
VS.					
W	isi		<i>bi+seuan+mũ+ban</i> lit. 'Two I-man-saw'.		
tw	0		1SG man see PAST		
VS.					
*S	*Seuan+in		$bi + wisi + m\tilde{u} + ban$ (I men two-saw).		
1	man	PL	1SG two see PAST		

Thus, in the phrase *wisi seuanin* 'two men' we have *wisi←seuanin*, because *seuanin* can be incorporated alone into the verb, while *wisi* alone cannot.

(c. Chukchee <i>ŋireqqlikkin</i>	amŋərootken	parol ←synt –li	<i>li</i> + <i>t</i> ⁽ forty-eight gloves)
	forty	eight	extra glo	ove PL.NOM
	vs.			
	ŋireqqlikkin	amŋərootken	parol+lele +	<i>to</i> forty-eight gloves
	forty	eight	extra glove H	PL.DAT

[Incorporation of numerals into the quantified N is obligatory in Chukchee if this N is in an oblique case.]

For a compound numeral, only the marker of compound numerals PAROL, meaning 'extra, added', is incorporated; other components of the numeral remain outside. (We see above that the noun stem *lili* is modified to *lele* in the form with incorporation: this is the effect of vowel harmony under the influence of the incorporated element.) Therefore, the marker *parol* is the Synt-head of the whole numeral, so that the SSynt-dependencies in a Chukchee compound numeral are as follows:

ŋireqqlikkin←amŋərootken←parol ⁽⁴⁸⁾ forty eight extra

From a logical standpoint, incorporability could be a definitorial property of Syntgovernors, if it weren't for the restricted character of incorporation itself: it is far from being universal, since it is not found in a majority of languages. Therefore, it cannot be used as a general criterion for the orientation of Synt-**D**.

Other non-definitorial properties of Synt-governors (listed as early as Pittman 1948) include class size (a Synt-governor belongs, as a rule, to a larger word-class than its dependent), versatility (a Synt-governor appears in a greater variety of syntactic environments), frequency (a particular Synt-governor is less frequent than a particular dependent), as well as some others. However, all of them are violated by many types of Synt-governors, so that they can be used as heuristic considerations only.

4.6 *The absolute head of the Synt-Structure of a sentence*

Since Synt-**n** presupposes uniqueness of the governor (no wordform in the sentence can depend syntactically on more than one other wordform), the SyntS of a sentence must have one absolute head, or a *top node*—a wordform which does not syntactically depend on anything and on which all the other wordforms of the sentence depend (directly or indirectly). Practically, in most versions of the **n**-approach known to me, in a complete clause/a complete sentence this role is filled by the finite, or tensed, verb—the MV (at least in languages that obligatorily have one in each complete clause/sentence, cf. below).²⁰ Thus, in the DSyntS of Figure 2, p. 6, where any form of the MV, including multiword—analytical—forms, is represented by a single node, the top node of the sentence is the verb ESCAPE (in the finite form of the Present Perfect); in the SSyntS of Figure 3, where each wordform, including the auxiliaries, is represented by a

²⁰ This view was held, at least in Europe, as early as in the 13th-14th centuries. Weber (1992: 13) speaks of Siger von Kortrijk, who preached the absolute dominance of the finite verb in a sentence around 1300. Cf. also the following remark by Nicolò Macchiavelli in 1516: "...dicono che chi considera bene le 8 parti de l'orazione ... troverrà che quella che si chiama verbo è la catena e il nervo de la lingua;" quoted in Koch & Krefeld 1991: v. For objections to the status of the Main Verb as the Synt-head of the sentence, see Hewson 1992: 49-51; these objections are (again and again) due to the confusion of different types of **1** (semantic, syntactic, morphological, and communicative).

separate node, the top node is the auxiliary verb HAVE (in the finite form of the Present Indefinite). The choice of the MV as the Synt-head of the sentence is by no means arbitrary: the finite verb is, on Criteria B1-B3, the (direct or indirect) Synt-governor with respect to all other wordforms in the sentence, and in this way it ends up as the absolute head. Let us consider the application of Criteria B1-B3 to the MV of a sentence.

By Criterion B1, the finite verb is the governor of the subject, since the passive Synt-valence of the phrase Subject—synt-MV is determined by the verb: for a phrase to be insertable in the construction *I know that* ... (or any similar context), it has to contain a finite verb; with respect to the phrases Object—synt-MV or Circumstantial—synt-MV the syntactically dominant status of the verb is obvious (and has never been doubted). To this, two arguments can be added:

• In many languages, subjectless sentences exist (Chinese, Japanese, Lezgian): for instance, in the Lezgian sentence *Meqizva* lit. 'Cold-is' = 'It is cold' no Synt-Subject is possible, even a zero one—the Lezgian verb shows no agreement, so that nothing would justify positing a zero dummy subject. Even in languages where the Subject is not omissible, such as English or French, the imperative sentence uses a finite verb, but has no overt Subject. In PRO-drop languages (Spanish, Polish, etc.), sentences without an overt Subject are quite typical (e.g., Sp. *Está muy ocupado* '**He** is very busy' is a current example). Sentences without objects and circumstantials are even more widespread. However, languages that admit full sentences without the MV, or more precisely, without a Synt-predicate (which can be not a verb), are not known—at least, to me. Thus, the presence of the MV, or more generally, of a Synt-predicate, is the necessary and sufficient condition for the existence of a genuine sentence.

• The Sem-valence and the active Synt-valence of the MV determine the syntactic organization of the sentence/the clause. Thus, if the MV is SLEEP, only one Sem-actant is possible and, consequently, the clause allows only the Synt-Subject; with SEE, two Sem-actants and, consequently, a Synt-Subject and a DirO are necessary; KISS involves three Sem-actants (who kisses whom on what part of the body), but there can be two Synt-actants (the Synt-Subject and the DirO: either with the Possessor depending syntactically on the DirO or the body part being not mentioned) or three Synt-actants (the Synt-Subject, the DirO and an Oblique Object): *Alan kissed Helen's hand/Helen* vs. *Alan kissed Helen on the forehead*.

Strictly speaking, we do not need to try Criteria B2 and B3, since Criterion B1 establishes the MV as the top node of a sentence/a clause beyond any doubt; however, I will do this here in order to show that in this case they all agree.

By Criterion B2, it is the finite verb that is the morphological contact point in a subordinate clause (minus the complementizer); for instance:

• In French, after the conjunction QUOIQUE 'although', the MV of the subordinate clause has to be in the subjunctive: $quoiqu'il soit \langle *est \rangle$ malade lit. 'although he should-be ill'.

• In French and English, after the conjunction SI/IF the MV of the subordinate clause has to be in the present, even if it refers to the future: *S'il vient* (**viendra*) *demain* ... /*If he comes* (**will come*) *tomorrow* ...

• If a clause is nominalized in order to be used in the Synt-Structure as a noun, it is its MV that actually undergoes the nominalization: *After John arrived*, ... \Rightarrow *After John's arrival*, ...

And, finally, by Criterion B3 the whole sentence is semantically reducible to its MV rather than to its Synt-Subject; thus, 'John works at IBM' is more an instance of 'work' that an instance of 'John' or of 'IBM'.

However, two complications arise in connection with the Main-Verb-as-the-Synt-Head-of-the-Sentence principle: zero verb forms and verbless sentences.

Zero verb forms

What is the top node of the SyntS of the Russian sentence (25a), which does not contain any overt verb at all?

(25) a. Leo moj drug lit. 'Leo my friend'. = 'Leo is my friend'.

Our first clue is that as soon as this sentence is transposed into the past, the future, the subjunctive or the imperative, a wordform of the verb BYT' (be) obligatorily appears:

'Leo was my friend'.
⁽ Leo will-be my friend ⁾ .
'Leo would be my friend'.
'Leo, be my friend!'

Since the sentence in (25a) stands in an obvious paradigmatic relation to the sentences in (25b), we conclude that the meaning 'PRESENT INDICATIVE' is expressed in (25a) by a zero wordform or, to put it differently, that the verb BYT' has a zero wordform in the present indicative. The SSyntS of (25a) looks then as follows:





 $BYT'_{ind, pres}$ is expressed by an empty (= null) signifier on the SMorph-level only; thus it does not create a problem for the **n**-Synt-structure of a sentence.

See Mel'čuk 1988: 303ff, 1995a: 169ff and 2006: 469ff on zero verb forms in syntax.

Verbless sentences

In quite a few languages, a full sentence does not have to include a finite verb. Thus, in Turkic languages, an equative or locative sentence in the present of the indicative ('John is a doctor/John is Canadian/John is in the room') does not admit a finite verb 'be'; instead, the predicative noun or adjective is supplied with a predicative suffix, which thus marks its Synt-role. In Salishan languages (West Coast, Canada), in particular, in

Lushootseed, all types of full sentences are possible without a finite verb and—unlike Turkic—without any morphological marker of predicativity.

- (26) a. Turkish^[6] Ingiliz+im 'I [am] English'. English 1SG 'You_{PL} [are] soldiers'. Asker+siniz soldier 2PL $Ev + \emptyset + de + sin$ 'You_{SG} [are] in the house'. house SG LOC 2SG *Cocuk+lar+Ø* Ankara+da+dir+lar (Kids [are] in Ankara). PL NOM Ankara LOC 3 kid PL b. Lushootseed (Beck 1997: 98ff; the syntactic predicate is boldfaced) lit. 'Salmon that'. = 'That [is/was] a salmon'. i. s?uladx^w ti?ił salmon that $sq^w ig^w ac$ lit. 'Two this deer'. = 'These deer [are/were] two'. ii. sali? ti?ə? two this deer
 - iii. dx^{w} ?al tə hud tə s + $x^{w}itil$?ə tə biac into the burning the NOMIN(alizer) fall of the meat lit. 'Into the fire [is/was] the fall(ing) of the meat'. = 'The meat falls/fell into the fire'.

This situation is typical of other Salishan languages: any lexeme, whatever its part of speech, can be turned into the syntactic predicate, provided it is rhematic (in Salishan languages, the SyntS of a sentence very closely parallels its communicative structure). In such sentences, the top node of the SSyntS can really be anything. For instance, here is the SSyntS of (26b-iii); it has as its top node the preposition DX^w?AL meaning ^(into):



To take into account languages with verbless sentences, we have to generalize our Main-Verb-as-the-Head-of-the-Sentence principle. This is readily done:

The top node of the SyntS of a sentence is its main, or primary, Synt-Predicate, whatever its surface realization.

In the languages of what Whorf called the Standard Average European type, the Synt-predicate of a full-fledged clause is (almost) invariably a finite verb. However, even these languages have incomplete clauses of different types, in which the top node of the SyntS is not a finite verb, but a noun, an infinitive, or an adverb:

Best wishes to you and your family; Down with Saddam Hussein!; Yours sincerely.

Rus. *Mne eščë domoj idti* lit. 'To-me still home to-go'. = 'I still have to go home'.

- Fr. *Et elle de rire* lit. ⁽And she to laugh⁾. = ⁽She broke out laughing⁾. *Heureusement qu'elle est venue* ! lit. ⁽Luckily that she has come!⁾ = ⁽Thank goodness she has come!⁾
- Ger. "Ich kann wieder Farben unterscheiden", so Charlotte Falk lit. ^{(·}I can again colors distinguish,' so Ch. Falk⁾. = ^{(·}I can distinguish colors again,' says/said Ch. Falk⁾.

Language-specific rules define the admissible top node for each of these minor sentence types.

4.7 *The three major subtypes of syntactic dependencies*

There are three major subtypes of Synt-**n**s, recognized universally: complementation, modification, and coordination (complementation and modification being particular cases of subordination). Specific subtypes of Synt-**n**s, i.e., Deep and Surface SyntRels, introduced above, are distributed between these major subtypes.

Complementation, modification and coordination have been discussed for a long time (cf., e.g., Matthews 1981: 147-167, Lehmann 1985, Zwicky 1993), so here I simply formulate the corresponding definitions (Definitions 11-13). Note that on the SSyntlevel there are several Synt-**n**s to which the distinction of these three major subtypes of Synt-**n** does not apply in a clear-cut way; we have to allow for some SSynt-**n**s that belong to neither of these classes (*is*-**r**_i→*reading* or *from*-**r**_j→*to* [as in *from two to five pounds*]; in such a case, I will speak of *ancillary* SSyntRels). On the DSynt-level, however, the distinction between complementation, modification and coordination creates no problems.

In all of the following definitions, the wordform w_2 depends syntactically on the wordform w_1 in the given sentence: w_1 -synt $\rightarrow w_2$.

Definitions 11-13 are approximate: they cover only the prototypical cases. In order to take into account all possible cases, I would have to add more conditions and thus make the definitions more complex; but in the present context, it does not seem worthwhile.

Definition 11: Complementation

The wordform w_2 is a complement, or a Synt-actant, of the wordform w_1 if w_2 is also a Sem-dependent of w_1 :



Complementation is controlled by the active valence of the Synt-governor. Formally, this means that the complements of the wordform \mathbf{w} must be specified in the lexical entry of $L(\mathbf{w})$, more precisely—in its Government Pattern. In other words, a complement of \mathbf{w} corresponds necessarily to a DSyntA(\mathbf{w}), the inverse being untrue: a DSyntA(\mathbf{w}) can be expressed, as we will see, by a modifier of \mathbf{w} .

Examples (Synt-actants, or complements, are in boldface): *He loves Helen*; *during* [*the*] *meeting*; *worth* [*a*] *trip*; *This must be* [*seen*]; *but* [*not*] *Helen*.

Comment

Definition 11 does not cover, for instance, the construction where a Synt-actant w_2 of the wordform w_1 depends semantically on a different wordform w_3 which also stands in a complementation relation to w_1 : e.g., *He believed* $[= w_1]$ *John* $[= w_2]$ *to be* $[= w_3]$ *sick*. Here, *John* is a DSynt-actant of *believe*, without being its Sem-actant, that is, without depending on *believe* semantically: 'John' is a SemA of 'be'. (On the discrepancy between Sem- and DSynt-actants of the same lexeme, as well as on cases where a Semactant of w is implemented as a modifier of w, see, e.g., Boguslavskij 1985: 10-19 and 1996: 23-43; cf. also Mel'čuk 2004.)

Definition 12: Modification

The wordform w_2 is a modifier, or a Synt-attribute, of the wordform w_1 if w_2 is a Sem-governor of w_1 :



Modification is typically not controlled by the active Synt-valence of the Synt-governor: the modifiers of the wordform \mathbf{w} are, as a rule, not specified in the lexical entry of $L(\mathbf{w})$.

Examples (modifiers are in boldface):

good friend; love *passionately*; *only* him; *not* serious; wrote *in* [Stuttgart]; wrote *when* [he was in Stuttgart].

Comments

1. Definition 12 does not cover, for instance, the construction where a SSynt-modifier w_2 of the wordform w_1 depends semantically on it, since w_2 expresses one of w_1 's Sem- and DSynt-actants: e.g., *French* $[= w_2]$ *participation* $[= w_1]$. Here, *French* depends both syntactically and semantically on *participation*, but in spite of this it is a SSyntmodifier of *participation* rather than its SSynt-actant (it is an adjective, and adjectives are modifiers by definition—because of their specific SSynt-behavior). Another similar example is *a book* $[= w_1]$ *hard* $[= w_2]$ *to find*: the adjective *hard* depends syntactically on *book*, but semantically bears on *find*; however, it is considered as a modifier of *book*.

The opposition "complementation ~ modification" underlies, in an obvious way, the problem of distinguishing between actants (\approx complements, Ger. *Ergänzungen*, Rus. *dopolnenija*²¹) and circumstantials (\approx modifiers, Ger. *Angaben*, Rus. *obstojatel'stva*). This distinction, first introduced explicitly probably in Tesnière 1959, is discussed in Engel 1977: 98-103, 158-179, Somers 1984, 1987, and Helbig 1992: 75-98 (with rich bibliography); see also Mel'čuk 2004: 265-270.

2. Interestingly, as indicated in Bazell 1949: 7-9, some languages strongly prefer complementation, while some others stress modification. Thus, Turkic languages or Japanese formally mark the complements, using morphological government: in particular, they possess developed case systems. On the other hand, Bantu languages formally mark only modification: they actively use agreement and completely lack cases; even in complementation constructions, they make the Synt-governor (say, the verb) agree with its Objects, leaving the latter unmarked. Of course many languages mix the two techniques in different proportions. Thus, Classic and most Slavic languages richly mark both complementation (\approx government) and modification (\approx agreement).

Definition 13: Coordination

The wordform w_2 is a conjunct of the wordform w_1 if and only if the following two conditions are simultaneously satisfied:

- 1) Semantically neither of them depends on the other (= w_1 and w_2 are not directly related semantically), but both are (or at least can be) Sem-dependents of such a semanteme as 'and', 'or', 'but', etc.: ' w_1 '-sem-(and)-sem-(w_2 ').
- 2) Syntactically w_2 depends on $w_1: w_1$ -synt $\rightarrow w_2$

Comments

- 1. The coordination of w_1 and w_2 can be of two types:
- Either direct coordination, where w_1 and w_2 have a direct Synt-**D** between them: w_1 -coord $\rightarrow w_2$

This coordination is called *asyndetic* ('conjunctionless').

Examples: Alan, Leo, Helen; eat, drink, sing, dance; [something] red, [not] white.

²¹ In modern linguistic literature, the terms *arguments* or *terms* (vs. *non-arguments/non-terms*) are also current for the corresponding concept. I prefer avoiding them in linguistics, to reserve their use for logic: arguments of a predicate, predicates *vs*. terms; see Mel'čuk 2004: 5-6.

- Or indirect coordination, where w_1 and w_2 are syntactically linked via a conjunction $\text{CONJ}_{\text{coord}}$:

 w_1 -coord \rightarrow CONJ_{coord}-conjunct \rightarrow w_2 .

Examples: Alan and Helen; either Alan or Leo; eat and drink, but not sing and dance; red, but [not] white.

2. As the two preceding definitions, Definition 13 does not cover all cases of coordination. Thus, on the SSynt-level, formally coordinate structures can be used to express DSynt-subordination. For instance, in Russian we have *izlovčilsja i ukusil* lit. ⁽[he] managed and bit⁾ = ⁽[he] managed to bite⁾, where the conjunct *ukusil* implements the DSynt-A **II** of *izlovčilsja* (example from Boguslavskij 1996: 28-32); a similar English example: *Try and catch the train*.^[7]

The above distinctions between the three major subtypes of Synt-**n**s are reflected in the DSynt-component of the Meaning-Text model by the three-pronged division of the DSyntRels:

actantial (I, II, II _{dir.speech} ,, VI)	= complementation
VS.	
attributive (ATTR, ATTR _{qual} and APPEND)	= modification
VS.	
coordinative (COORD and QUASI-COORD)	= coordination;
see Mel'čuk 1988: 63-65 and 2004: 248ff.	

The inclusion relations between the three major subtypes of Synt-**D** is shown in Figure 5 below:



Figure 5: Major subtypes of syntactic dependency

These subtypes of Synt-**1** were clearly distinguished by L. Bloomfield (1933: 194-198); I slightly change here Bloomfield's formulations, without modifying his main idea:

• In a complementation phrase $w_1 - w_2$ the passive Synt-valence of the whole phrase may be different from those of both of its elements, although it is determined by the passive Synt-valence of one of them, namely that of the Synt-head of the phrase; normally, the head requires the presence of the dependent(s): X - synt-hits-synt-Y.

• In a modification phrase $w_1 - w_2$ the passive Synt-valence of the whole phrase is the same as that of one of its elements, namely—that of its Synt-head; normally, the head does not require the presence of the dependent: *hits*-synt-*severely*.

• In a coordination phrase $w_1 - w_2 - \dots - w_n$ the passive Synt-valence of the whole phrase is the same as that of each of its elements (= *conjuncts*):

read,-synt-write,-synt-(and) count.

The constructions manifesting the SyntRels of the first subtype, i.e., complementation, are called *exocentric*; the constructions manifesting the SyntRels of the second and third subtypes, i.e., modification and coordination, are called *endocentric*.

For the surface SyntS, a fourth major subtype of SSyntRel is needed—to link syntactically-induced wordforms (what are known as structural words, chunks of idioms, parts of compound numerals, etc.), which do not appear in the Deep-SyntS and cannot be covered by the dependencies of the three above-mentioned classes. As proposed at the beginning of this subsection, I will call these SSyntRels *ancillary*, to emphasize their "subservient" character.

4.8 Syntactic dependencies of a language: SSynt-Relations of English

In order to give the reader a better idea about SSyntRels, as they can be used in a description of a language, I cite here a tentative list of SSyntRels of English, taken—with some corrections and additions—from Mel'čuk & Pertsov 1987: 85-160. No claims are laid as to completeness of this list; its purpose is strictly illustrative.

In the examples, the SSynt-dependent is boldfaced; words not participating in the construction illustrated are included in brackets.

For a better surveyability of the list, the SSyntRels are grouped as follows:

- First, they are divided into subordinate and coordinate ones.
- Second, the subordinate SSyntRels are subdivided into two subsets:

— clausal SSyntRels, or those that can hold between (the heads of) noun, verb, adjective, and adverb phrases (they can also appear within these phrases);

— phrasal SSyntRels, or those that can hold only between the elements within phrases (never between phrases).

• Third, inside each subdivision, the line is drawn between valence-controlled SSynt-Rels (that necessarily embody complementation) and non-valence-controlled SSyntRels (that can be either modificative, or coordinative, or ancillary).
A SUBORDINATE SSYNTRELS: 1-50

A.1 Clause-level (= clausal) SSyntRels: 1-21

A.1.1 Valence-controlled SSyntRels: Complementation (1-14)

A.1.1.1 Actantial SSyntRels: 1-11

1. Subjectival

I←subj-am [Russian]; *Smoking*←subj-is [dangerous]. *That*←subj-[Alan comes]-is [clear]. *It*←subj-is [clear that Alan comes].

2. Quasi-Subjectival

[It] is–[clear]–quasi-subj \rightarrow that [Alan comes].

3. Direct-Objectival

[She] sees-dir-obj→me; [to have] written-dir-obj→novels; [Helen] wants-dir-obj→Alan [to tickle her]. prefer-dir-obj→staying [home]; explain-[to him]-dir-obj→that [Alan was absent]; make-dir-obj→it [possible to neutralize the consequences].

4. Quasi-Direct-Objectival

make–[*it possible*]–quasi-dir-obj→*to* [*neutralize the consequences*].

5. Indirect-Objectival

gives-indir-obj \rightarrow Alan [some money]; convince-[Alan]-indir-obj \rightarrow that [he should work less].

6. Oblique-Objectival

depends–**obj**→*on* [*Alan*]; [*much*] *respect*–**obj**–*bj*→*for* [*Alan*]; *translation*–**obj**–*from* Lushootseed [*into* Polish]; *translation*–[*from* Lushootseed]–**obj**–*into* [Polish].

7. Infinitival-Objectival

can-inf-obj→read; want-inf-obj→to [read]; [Helen] wants-[Alan]-inf-obj→to [tickle her]. [Helen] makes-[Alan]-inf-obj→read. [her] desire-inf-obj→to [come].

8. Completive

 $find-[this]-compl \rightarrow easy; consider-[Alan]-compl \rightarrow happy; make-[it]-compl \rightarrow possible; make-[Helen a good]-compl \rightarrow wife.$

9. Copular

be-copul→*easy*; *be*-[*a*]-copul→*teacher*; *be*-copul→*without* [*a hat*]; *seem*-copul→*in* [*a difficult position*].

10. Agentive

written-agent $\rightarrow by$ [Alan]; arrival-agent $\rightarrow of$ [Alan] (NB: the phrase Alan's arrival is described by a different SSyntRel: possessive SSyntRel, No. 24); shooting-agent $\rightarrow of$ [the hunters: 'the hunters shoot']; [a] translation-agent $\rightarrow by$ [Alan]; [I like] for agent-[my son to]-do [well in school] (that is, do-agent $\rightarrow for$ [son]).

11. Patientive

translation-patient $\rightarrow of$ [*this text*]; *shooting*-patient $\rightarrow of$ [*the hunters*: ^{(the hunters are shot⁾].}

A.1.1.2 Copredicative SSyntRels: 12-13

12. Subject-copredicative

[Alan] returned—subj-copr→*rich*.

13. Object-copredicative

[Alan] likes–[Helen]–obj-copr \rightarrow slim. [Alan] hammered–[the coin]–obj-copr \rightarrow flat.

A.1.1.3 Comparative SSyntRel: 14

14. Comparative

older-compar→*than* [*Leo*]; [*Alan loves Helen*] *more*-compar→*than* [*Leo*]; *more*-[*important*]-compar→*than* [*Leo*]; *as*-[*important*]-compar→*as* [*Leo*].

A.1.2 Non-valence-controlled SSyntRels: Modification (15-21)

A.1.2.1 Absolutive SSyntRel: 15

15. Absolute-predicative

[*His first*] attempt–[a]–**abs-pred** \rightarrow *failure*, [*Alan* ...]. [*He went out, his*] anger–**abs-pred** \rightarrow *gone*. [*He went out, his*] gun–**abs-pred** \rightarrow *in* [*his left hand*].

A.1.2.2 Adverbial SSyntRels: 16-19

16. Adverbial

walk-adverb→fast; [will] write-[next]-adverb→week; delve-adverb→deeply; [He] works-adverb→there (in [this office]). [He] went-[out,]-[his]-adverb→gun-abs-pred→in [his left hand]. With←adverb-[her paper finished, Helen]-can [afford this trip].

17. Modificative-adverbial

[As always] *elegant*, ← mod-adverb–[Alan]–walked [away].

18. Appositive-adverbial

[An old] man,←appos-adverb–[Alan]–works [less].

19. Attributive-adverbial

Abroad, ← attr-adverb–[*Alan*]–works [less].

A.1.2.3 Sentential SSyntRels: 20-21

20. Parenthetical Oddly,←parenth-[Alan]-works [less]. As←parenth-[we have known for some time, Alan]-works [less]. To←parenth-[give an example, I]-consider [now nominal suffixes].

21. Adjunctive OK,←adjunct-[1]-agree. Helen,←adjunct-[1]-agree.

A.2 Phrase-level (= phrasal) SSyntRels: 22-50

A.2.1 Any phrase SSyntRel: Modification (22)

22. Restrictive

still restr-*taller*; *most* restr-*frequent*; *not* restr-*here*; [*Alan has*] *just* restr-*arrived*.

A.2.2 Noun Phrase SSyntRels: 23-37

A.2.2.1 Valence-controlled SSyntRel: Complementation (23)

23. Elective

[*the*] *poorest*-elect \rightarrow *among* [*peasants*]; [*the*] *best*-[*ones*]-elect \rightarrow *from* [*these boys*]; [*the*] *most*-[*successful student*]-elect \rightarrow *in* [*our class*].

A.2.2.2 Mixed Type SSyntRels = Valence-controlled/Non-valence-controlled: Modification (24-25)

24. Possessive

Alan's – poss-arrival; *Alan's* – poss-bed; *Alan's* – poss-organs; *Alan's* – poss-pleasure.

25. Compositive

man compos-[-*machine*]-*interaction*; *car* compos-*repair*; *noun* compos-*phrase*; *color* compos-*blind*.

A.2.2.3 Non-valence-controlled SSyntRels: Modification (26-37)

26. Determinative

my determ-*bed*, *a* determ-*bed*, *those* determ-*beds*.

27. Quantitative

three—quant–*beds*; [*three*—num-junct–]*thousand*—quant–*people*.

28. Modificative

comfortable—modif-*beds*, *visible*—modif-*stars*, *French*—modif-*production*.

29. Post-modificative

stars-post-modif-*visible* (vs. *visible stars* above).

30. Descriptive-Modificative

[these] beds,-descr-modif->comfortable [and not expensive], ...

31. Relative

[the] paper-[that I]-relat-read [yesterday]; the girl-[who]-relat-came [first].

32. Descriptive-Relative

[*this*] paper,-[*which* I]-descr-relat-*read* [yesterday]; Alan,-[*who*]-descr-relat-*loves* [*her so much*].

33. Appositive

General ← appos-Wanner; [the] term-appos → 'suffix'.

34. Descriptive-Appositive

[*This*] term-descr-appos→('suffix') [will be considered later]. [You forget about] me,-[your]-descr-appos→mother. Alan-[the]-descr-appos→Powerful; you-descr-appos→children, we-descr-appos→linguists.

35. Sequential

man-sequent *→machine* [*interaction*]; *fifty*-sequent *→to* [*seventy dollars*].

36. Attributive

learners-attr-*with* [*different backgrounds*]; *dress*-attr-*of* [*a beautiful color*]; [*a*] *man*-[*the same*]-attr-*age*; *years*-attr-*of* [*war*], [*the*] *house*-attr-*of* [*Alan*].

37. Descriptive-Attributive

[Professor] Wanner,-descr-attr→from [Stuttgart, was also present].

A.2.3 Prepositional Phrase SSyntRels: 38-39

A.2.3.1 A valence-controlled SSyntRel (Complementation): (38)

38. Prepositional

in-**prepos**-*bed*, *without*-[*three hundred*]-**prepos**-*dollars*.

A.2.3.2 A non-valence-controlled SSyntRel (Ancillary): (39)

39. Prepositional-infinitival

to-prepos-infinit \rightarrow go [to bed].

A.2.4 Verb Phrase (= analytical form) SSyntRels (Ancillary): 40-42

40. Perfect-analytical

has-perf-analyt \rightarrow written, has-perf-analyt \rightarrow been [beaten].

41. Progressive-analytical

was-progr-analyt-*writing*.

42. Passive-analytical

was−pass-analyt*→written*.

A.2.5 Conjunction Phrase valence-controlled SSyntRels (Complementation): 43-46²²

43. Subordinate-Conjunctional

[Suppose] that–[Alan]–subord-conj \rightarrow comes. [so] as–[not]–subord-conj \rightarrow to [irritate Leo].

44. Coordinate-Conjunctional

[Alan] and—coord-conj \rightarrow Helen.

45. Comparative-Conjunctional

than-compar-conj-*Helen*; *as*-compar-conj-*always*; *as*-compar-conj-*for him*.

46. Absolute-Conjunctional

If-[*a*]-**absol-conj** \rightarrow *pronoun*, [*the grammatical subject may* ...]; *while*-**absol-conj** \rightarrow *in* [*bed*].

A.2.6 Word-like Phrase SSyntRels (Ancillary): 47-50

47. Verb-junctive

give-verb-junct $\rightarrow up$, bring-verb-junct $\rightarrow down$, do-verb-junct $\rightarrow in$.

48. Numeral-junctive

fifty—num-junct–*three*; *fifty*—num-junct–*third*.

49. Binary-junctive

if ...-bin-junct→*then* ...; *the*-[*more* ...]-bin-junct→*the* [*more* ...]; *till*-bin-junct→*after*; *from*-[...]-bin-junct→*to* [...]; *either*-[...]-bin-junct→*or* [...].

50. Colligative

[*is*] *dealt*−**collig**→*with* [*stranded prepositions*].

B COORDINATE SSYNTRELS: 51-52

51. Coordinative

Alan-coord→and [Leo]; *Alan*-coord→but [not Leo]; rich,-coord→intelligent-coord→and [beautiful].

²² For an argumentation in favor of the SSyntS Conj→MV, see Hudson 1987: 119-121.

52. Quasi-coordinative

[*He was*] $abroad-quasi-coord \rightarrow without-[a penny]-quasi-coord \rightarrow in a desperate situation.$ [*These moneys we keep hidden* $] under-[a loose board]-quasi-coord \rightarrow under-[the floor]$ $quasi-coord \rightarrow under-[a chamber pot]-quasi-coord \rightarrow under my friend's bed (T. Capote, "A$ Christmas Memory").

Comment

As suggested above (4.7, Comment 1 after Definition 12, p. 50), some of the SSyntRels that belong to the modification class can be valence-controlled, so that their dependents correspond to DSynt-Actants of their governors:

<i>my</i> determ - <i>arrival</i>	\$	I ←I –ARRIVAL
American - modif - participation	¢	AMERICA←I–PARTICIPATION
<i>treat</i> −[someone]− adverb → <i>friendly</i>	÷	TREAT– III →FRIENDLY
income ←compos -tax	¢	INCOME ←II −TAX

Similarly, the coordinative SSyntRel can be valence-controlled:

try-coord
$$\rightarrow$$
 and [*come*] \Leftrightarrow TRY-II \rightarrow COME

In point of fact, the correlation between complementation and modification, as well as between complementation/modification and coordination on the DSynt- and SSynt-levels is complex and cannot be discussed here in depth.

5 Possible combinations of the three types of linguistic dependency

The three types of linguistic syntagmatic dependency that we are studying—semantic, syntactic, and morphological—are logically independent of each other, which means that they can co-occur in all possible combinations. Thus, two wordforms in a sentence can be related by a Sem-**D** with no Synt-**D** or Morph-**D** between them (a); or they can have a Sem-**D** and, at the same time, an inverse Synt-**D**, with still no Morph-**D** (b); or there can be a Synt-**D** with a Morph-**D** having the same orientation, but no Sem-**D** (c); etc.

(a)
$$w_1 \xrightarrow{sem} w_2$$
 (b) $w_1 \xrightarrow{sem} w_2$ (c) $w_1 \xrightarrow{synt} w_2$
morph

All in all, there are fourteen logically possible combinations of direct Sem-**D**, Synt-**D** and Morph-**D** between two wordforms, w_1 and w_2 , of a sentence (cf. Mel'čuk 1964a, 1988: 118-128); all of them are actually found in languages and will be illustrated below.



No syntagmatic dependency whatsoever between two wordforms.

Example

The wordforms *cocoa* and *new* in (1), p. 5.

 W_2



Two wordforms are linked by a Sem-**D**, unsupported by any Synt-**D** or Morph-**D**.

Example

The wordforms *farming* and *problems* in (1) are semantically directly related—*farming* depends on *problems* ('problems are-for farming'), yet there is no Synt-**n** or Morph-**n** between them. Another example could be an expression of the type *He drives me mad*, where we have *mad*-sem→*me*, while syntactically and morphologically *me* and *mad* are not linked.

CASE 3:	$W_1 \longrightarrow W_2$	Two wordforms are linked by a Synt- D , but there
	1	is no Sem- D or Morph- D between them.

Examples

- (27) a. In Japanese, a numeral or a quantitative adverb, while bearing semantically on the SSynt-Subject or the DirO as in 'Five people were injured' or 'He reads many books' (and—for numerals—morphologically depending on it), depends syntactically on the verb, with which it has neither semantic nor morphological links, cf.:
 - i. Sono ziko +de keganin +ga go+nin deta this accident LOC injured.people SUBJ(ective) five CLASS(ifier) emerged lit. ⁽In this accident, injured-people five-ly - synt-emerged⁾. =

⁽In this accident, five people were injured⁾.

ii. Nihongo+no hon+o takusan yomimasita Japanese GEN book ACC many read-PAST
lit. 'Japanese-language books many-ly←synt-read'. = ([I] have read many Japanese books'.

For a detailed characterization of this Floating Quantifier construction and its relations to other numeral constructions in Japanese, see Kim 1995; cf. also Case 8, (31b).

- b. In Russian, a numeral adverb of the type VDVOËM 'being-two', VTROËM 'being-three', etc. is used in a similar construction, where this adverb semantically bears on the SSynt-Subject of the clause, but has with it no syntactic or morphological link: My sideli–[na beregu]–synt→včetverom
- lit. 'We sat [on the-shore] being-four'.

The same holds about most of Floating Quantifiers of different types in various languages.

c. In English, French and many other languages, a measure noun used as a DirO depends syntactically on the verb, but does not have a semantic or morphological link with it (semantically the verb dominates the noun quantified):
 John bought–[five]–synt→kilos [of potatoes].
 Fr. *Jean a acheté–[cinq]–synt→kilos [de pommes de terre*].

Cf. Case 9, example (32b).

d. Any conjoined elements that are morphologically invariable, as, e.g., *Alan*, *Helen*, *Leo*, are linked syntactically without any direct semantic or morphological link between them.

CASE 4:

w₁ w₂

Two wordforms are linked by a Morph-**D** only, without any Sem-**D** or Synt-**D**.

Examples

(28) a. In Tabassaran (Eastern Caucasian, Daghestan, Russia), the MV can agree with the 1st/2nd person Possessor of the Synt-Subject, the Possessor being not directly related to the verb semantically or syntactically, cf.:

 $I\check{c}$ [**w**₁] *mudur* $u\check{c}vuhna$ $he+b+gnu+ji\check{c}$ [**w**₂] Our goat-kid[II] to-you left II left **1PL** ⁽Our kid ran away to your place⁾.

Here the verb *hegnu* 'ran away, fled' agrees in class with *mudur* (class II, the class-marking infix in the verb is **-b**-), but in person and number with $i\check{c}$ 'our'.

The same type of agreement of the MV is characteristic, among others, of Chickasaw (Muscogean, USA), Wichita (Caddoan, USA), Tangut, and Maithili (Indian).

b. In Maasai (Nilotic, Kenya), an infinitive that semantically and syntactically depends on the MV agrees in number with the Synt-Subject of the MV (rather than with its own understood [= semantic] Actor):

Kítaretôoltonjánì/ Iltónjáná \acute{aa} $+m\acute{ok}$ [w_2]enáíshóhelp-PERF.1PL.SUB.3.OBJthe-man-SG.NOM / the-man-PL.NOMINF.PLbrewbeer-SG.NOM'[We]helped the-man/the-men to-brew[PL]the-beer'.

[1) The Synt-Subject w_1 —^(I) in the first sentence and 'we' in the second—does not appear on the surface: Maasai is a Pro-Drop language.

2) The Maasai verbs, both transitive and intransitive, feature the ergative construction—with the DirO in the nominative and the Synt-Subject, which is not shown in the present example, in the oblique.

3) Acute and grave accent symbols stand, respectively, for high and low tones.]

c. In Alutor, a transitive verb of perception on which syntactically depends a DirO clause can show object-agreement either with the Synt-Subject (i) or with the DirO (ii) of this clause (depending on the communicative role of the former and the latter):

i.	Qəmavə-	+nak	na	+laSu	+tkəni	+ $\gamma \partial t [w_2]$			
	Qamav	SG.LOC	3SG.SU	JB see	PRES	5 2SG.OB.	J		
lit.	γ ən you _s (Qamav (Qamav)	+ <i>annə</i> [w _G SG.INSTR sees-you y sees YÓU l	$\begin{bmatrix} 1 \\ 2 \\ 2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0$	+ka G.SUB ha -harness ing the 1	<i>əl?at</i> + arness ing re reinde	ətkə +na PRES 3.0 indeer ³ . = er ³ .	+ <i>wwi</i> BJ PL	<i>qura</i> + <i>wwi</i> reindeer PL	
ii.	<i>Qəmavə</i> Qamav	+ <i>nak</i> SG.LOC	Ø 3SG.SU	+ <i>laSu</i> + JB see	<i>-tkəni</i> PRES	+ <i>nina+w</i> 3.0BJ P	wi [w ₂] L		
lit	γən you _s	+ <i>annə</i> _G SG.INSTR sees-them	Ø R 2S VOU ar	+ka G.SUB ha e-harnes	<i>arness</i>	<i>ətkə +na</i> PRES 3.01 eindeer ⁾ .	- +wwi BJ PL =	<i>qura</i> +wwi [reindeer PL	[w ₁]
	, `		5		0				

'Qamav sees you harnessing the REINDÉER'.

[The Alutor transitive verb enters in an ergative construction, with the Synt-Subject in the locative, if it is a human proper name, and in the instrumental otherwise (cf. (6), p. 14); **na**- is a 3sg subject marker if the DirO is neither of the 3rd person nor 1sg, and $\mathbf{\emptyset}$ - is a 3sg subject marker if the DirO is in the 3rd person or 1sg. A verb of perception can also agree with its DirO clause as a whole, showing 3sg object agreement; this case is, however, irrelevant in the present context.]



Two wordforms are linked by a Sem-**D** and a Synt-**D**, oriented the same way, but no Morph-**D** is present.

This is a typical situation with nominal objects in caseless languages: for instance, the wordforms *escape* and *problems* in (1).



Two wordforms are linked by a Sem-**D** and a Synt-**D**, this time oriented the opposite ways, with no Morph-**D** present.

Examples

- (29) a. An adjective and the modified noun in ADJ + N phrase in a language where adjectives are invariable, cf. *new* and *areas* in (1).
 - b. In Lushootseed, a PREP + NUM phrase syntactically depends on the clause predicate (the nominal phrase boldfaced in the example), which is its Semdependent, and there is no Morph-D between them:
 ?əbs +s+q^wəb+q^wəbaj?[w₂] əlg^wə? ?> [w₁] ti?ə? bə+sali?
 POSS dog dog PL at this two
 - lit. ⁽[They are] dog-dog-possessor+s at $\langle with\text{-respect-to} \rangle$ these two². = ⁽[They] have two dogs² (example of D. Beck).



Two wordforms are linked by a Sem-**D** and a Morph-**D**, oriented the same way, but no Synt-**D** is present.

Example

In a language where Clitic Raising exists, a clitic—in the SSyntS—can semantically and morphologically depend on an infinitive, while there will be no direct Synt-dependency between them, the clitic being a Synt-dependent of a higher verb, which hosts it, cf.:

(30) Sp. *Le* [**w**₂]←synt–quisiera poder enviar [**w**₁] este libro lit. ⁽To-him [I] would-like to-be-able to-send this book⁾.

Semantically, *le* 'to him' depends as an actant on *enviar* 'send'; its dative form is imposed by this verb, so that morphologically *le* also depends on *enviar*.



Two wordforms are linked by a Sem-**D** and a Morph-**D**, oriented the opposite ways, without any direct Synt-**D** between them.

Examples

(31) a. An attributive or copredicative adjective and the Subject/the DirO in a language with adjectival agreement illustrate Case 8. The adjective shows agreement with the Subject/DirO, which is its Sem-dependent, and there is no direct Synt-link between them:

Fr. *Elle* $[\mathbf{w}_2]$ *semblait fatiguée* $[\mathbf{w}_1]$ 'She seemed tired',

where semantically *elle* depends on *fatiguée* [= 'fatigué'('elle')], but morphologically *fatiguée* depends on *elle* for its singular and feminine; syntactically, the two are not directly related.

Similar examples: Fr. *Elle est rentrée heureuse* 'She returned happy', *Il buvait son thé froid/sa tisane froide* 'He drank his tea cold/his herbal tea cold', etc.

For a detailed analysis of the copredicative construction, see Nichols 1978.

b. In Japanese, a numeral w_1 can bear semantically on a noun w_2 and morphologically depend on it (the form of the numeral—more precisely, its classifying suffix—is determined by the noun), while syntactically it depends on the verb (cf. Case 3, example (27a)):

<i>Uma</i> + <i>o</i> horse ACC	$ip + piki \leftarrow synt-tot + ta$ one CLASS(ifier) take PAST	⁽ [He] took one horse ⁾ .
<i>Hagaki+o</i> postcard ACC	<i>ip</i> + <i>pon</i> ← synt − <i>tot</i> + <i>ta</i> one CLASS take PAST	⁽ [He] took one postcard ⁾ .
Kippu +o	<i>iti +mai</i> ←synt–tot +ta	⁽ [He] took one ticket ⁾ .
ticket ACC	one CLASS take PAST	





Two wordforms are linked by a Synt-**D** and a Morph-**D**, oriented the same way, with no Sem-**D** between them.

Examples

- (32) a. In Latin, the construction $ab [\mathbf{w_1}]$ urbe $[\mathbf{w_2}]$ condita lit. (since [the] city founded) = (since the founding of the city), the preposition ab (since) syntactically and morphologically dominates the noun *urbe*, while semantically it bears on (conditio) = (founding).
 - b. A measure noun used as a DirO in a language with cases depends syntactically and morphologically on the verb, but does not have a direct semantic link with it:

Rus. Ivan kupil->tonnu kirpičej 'Ivan bought a ton of bricks'.

Cf. Case 3, example (27c).

CASE 10:
$$W_1 \longrightarrow W_2$$

 $w_1 \longrightarrow W_2$

Two wordforms are linked by a Synt-**D** and a Morph-**D**, oriented the opposite ways, again with no Sem-**D** between them.

Example

A phasic or copular verb syntactically dominates its Synt-Subject, but morphologically depends on it (= agrees with it in person and number), while there is no Sem-**D** between this verb and its Subject, because the Subject semantically depends on the lexical verb, cf.:

(33) The water $[\mathbf{w_2}]$ begins $[\mathbf{w_1}]$ to boil or The water seems to boil,

where *water* syntactically depends on *begin/seem*, morphologically dominates it, and semantically depends on *boil*: 'boil'('water').

CASE 11:
$$W_1 \xrightarrow{sem} W_2$$

morph

Two wordforms are linked by all three types of dependency, oriented all the same way.

Example

A verb and its nominal object in a language with cases, cf. Rus. *problem* [PL.GEN] with respect to (*ne*) *znat*' in (2), p. 8.

Two wordforms are linked by all three types of dependency, with Sem-**D** and Morph-**D** oriented the same way, while Synt-**D** goes in the opposite direction.

Examples

- (34) a. A postnominal modifier and the modified noun in a language having what is known as *izafa* construction. Cf. Persian *daftar+e nav* lit. 'workbook new', where *nav* [w₁] bears semantically on *daftar* [w₂] and imposes on it a special form (= the izafa suffix -e), while being syntactically its dependent.
 - b. A negative particle and the negated verb in a language where the negative particle $[\mathbf{w_1}]$ requires a special form of the verb $[\mathbf{w_2}]$. Thus, in Arabic the particle **lam** (NEG.COMPL(etive).PAST) requires the jussive, while the particle **lan** (NEG.COMPL.FUT) requires the subjunctive (**la** (NEG.INCOMPL(etive).PRES) is neutral in this respect): ja+ktub+u ([he] writes) ~ $la \ ja+ktub+u$ ([He] does not write). VS. IND(icative)

Ø+katab+a	([he] wrote) ~	lam ja+ktub+Ø	([He] did not write).
VS.		JUSS	S(ive)
sa+ja+ktub+u	([he] will write)~	lan ja+ktub+a	⁽ [He] will not write ⁾ .
		SUE	BJ(unctive)

Semantically, the negative particle LAM/LAN bears on the verb and morphologically controls its form; but syntactically, it depends on the verb.



Two wordforms are linked by all three types of dependency, with Sem-**D** and Synt-**D** oriented the same way, while Morph-**D** goes in the opposite direction.

Examples

(35) a. A verb $[w_1]$ and its nominal actant $[w_2]$ in a language with polypersonal agreement of the verb, but no nominal cases, such as, e.g., Abkhaz (West Caucasian, Georgia), where the MV agrees in nominal class and number with the SSynt-Subject, the DirO and IndirO:

Sara	Nadš'a	i+l+əs+teit	aš ^w q ^w ə
Ι	Nadsha	it her I gave	book
'I gav	e Nadsha [a] book ⁾ .	

In Abkhaz, nouns and pronouns have no case inflection themselves, but impose agreement on the verb, whose prefixes cross-reference all three SSynt-actants.

b. Agreement of the participle in an analytical verb form with the preposed DirO in French:

les	fleur	+s	que	je	ť	ai	offert-	+es
the	flower[F	EM] PL	that	Ι	to-you	have	given	PL.FEM
⁽ the	flowers that	at I have	given to	you),			

where *que* (that) semantically and syntactically depends on *offertes* (given), but morphologically controls its gender and number (QUE gets its gender and number from its antecedent, FLEURS, so that, in final analysis, *offertes* is feminine and plural because of *fleurs*; yet, technically speaking, it agrees in gender and number with QUE).

In point of fact, the situation here is more complex, since *que* is an accusative form, imposed by the transitive verb *offrir* ⁽give⁾; so that *que* morphologically depends on

offertes at the same time that offertes morphologically depends on que. This is a case of reciprocal morphological dependency.

Two wordforms are linked by all three types of dependency, of which Synt-**D** and Morph-**D** are oriented the same way, but in the opposite direction with respect to Sem-**D**.

Example

An agreeing adjective and the modified noun in a language with adjectival agreement (Slavic, Romance, Semitic, German, etc.), where the adjective bears semantically on the noun, but syntactically and morphologically depends on it.

Consistent distinction of the three major types of dependency allows for some elegant formulations, for instance:

• The adjective as a part of speech can be characterized in terms of Sem-**D** vs. Synt-**D** (see Beck 2002: 84):

Prototypically, an adjective semantically dominates the noun on which it depends syntactically. (Morph-**1** can go both ways or be absent altogether: cf. Cases 6, 12, and 14.)

Similarly, for the adverb (replacing "noun" with "verb or adjective").

• Taking into account the three types of linguistic dependency, Zwicky (1993) presents the differences between complements and modifiers in a compact form:

Syntactic role	Complement	Modifier
Properties		
Semantic Syntactic Morphological	argument obligatory and unique government target agreement controller	predicate optional and repeatable agreement target government controller

The correlations stated in this table hold only in the most current, prototypical cases; as has been shown above, the syntactic and morphological properties of complements *vs.* modifiers can in principle be inverted. However, the semantic—definitorial—property is stable.

• In the literature, one finds heated debates concerning the split of head-related properties between different sentence elements, which presumably makes the identification of heads difficult and/or dubious: a given element seems to be the head in one respect, but the dependent in another one. However, if one distinguishes the three types of dependency and uses Criteria B1-B3 in the hierarchical way (p. 31), such a split is logically impossible. Thus, a Synt-head must be determined exclusively according to the properties of syntactic heads; it is irrelevant whether or not it has properties of Semheads or Morph-heads (as the 14 combinations above show, in many cases a Synt-head does not have such properties).

6 Correlations between the three types of linguistic dependency

The three types of dependency are linked by the following correlations (these correlations hold only for prototypical cases of morphological agreement and government and are no more than tendencies).

Sem-D vs. Morph-D

- Sem-governors morphologically agree with their Sem-dependents;
- Sem-governors morphologically govern their Sem-dependents.

This is the *Keenan's principle* (Keenan 1974: 298-303 and 1978: 94-98); cf. Zwicky's slogan: 'Functors are agreement targets and government triggers' (1993: 295).

Synt-D vs. Morph-D

• If w_2 morphologically agrees with w_1 , then w_1 and w_2 sometimes are, and sometimes are not, linked by a direct Synt-**D** (there also may be no Sem-**D** between w_1 and w_2).

• If \mathbf{w}_2 is morphologically governed by \mathbf{w}_1 , then \mathbf{w}_1 and \mathbf{w}_2 are always linked by a direct Synt-**D**; however, a Sem-**D** between them can be absent. (In fact, in some constructions, \mathbf{w}_2 morphologically governed by \mathbf{w}_1 depends syntactically not on \mathbf{w}_1 , but on the head of a verbal chain of Synt-**D**s leading to \mathbf{w}_1 : $\mathbf{w}_2 \leftarrow \mathbf{w}_i \rightarrow ... \rightarrow \mathbf{w}_j \rightarrow \mathbf{w}_1$. Such a chain is what is known as a *nucleus*: it consists of auxiliary, modal, phasal and similar verbs: something like $N[=\mathbf{w}_2]$ could have wanted to begin to $V[=\mathbf{w}_1]$; cf. (30).)

As can be seen in our review of theoretically possible cases, in the configuration $\mathbf{w_1}$ -synt $\rightarrow \mathbf{w_2}$, the Morph-**n**s can go both ways: the Synt-governor can be either the controller or the target of a Morph-**n**. The same holds with respect to linear position: rules for positioning can also go both ways, such that in some cases the linear position of the Synt-dependent $\mathbf{w_2}$ must be stated with respect to its Synt-governor $\mathbf{w_1}$ (ADJ-N, N-V, ADV-V, etc.), and in others the linear position of the Synt-governor $\mathbf{w_1}$ must be stated with respect to its Synt-governor $\mathbf{w_1}$ must be stated with respect to its Synt-governor $\mathbf{w_1}$ must be stated with respect to its Synt-governor $\mathbf{w_1}$ must be stated with respect to its Synt-governor $\mathbf{w_1}$ must be stated with respect to its Synt-governor $\mathbf{w_1}$ must be stated with respect to its Synt-governor $\mathbf{w_1}$ must be stated with respect to its Synt-governor $\mathbf{w_1}$ must be stated with respect to its Synt-dependent $\mathbf{w_2}$ (PREP-N, AUX-V, CONJ-V, etc.).

NB: The reference point for linear positioning of the one of two syntactically-linked elements X-synt \rightarrow Y is the element which can appear without the other, the inverse being untrue. Thus, in ADJ \leftarrow synt -N, N can appear without ADJ, but not vice versa, which means that ADJ is positioned with respect to N, its SSynt-governor. Similarly, for PREP-synt \rightarrow N, N can be used without a PREP, but a PREP never appears without its N (except for some special cases, such as stranded prepositions in English); therefore, the position of PREP is stated with respect to N, its SSynt-dependent (cf. 4.3.1, Criterion A1, Comment 3, p. 26).

Sem-/Synt-D vs. Morph-D

Sem-**1** and Synt-**1** are global in that they embrace all the wordforms in a sentence; therefore, they are represented explicitly in the SemS and the D-/S-SyntS of the sentence. On the other hand, Morph-**1** is not global in this sense: it does not necessarily embrace all the wordforms in a sentence (in addition, it is by no means present in all sentences and can be altogether absent from a language); therefore, no special structure is foreseen in which it would be explicitly represented: Morph-**1** s are computed by syntactic rules of the language during the SSyntS \Leftrightarrow DMorphS transition and encoded in the DMorphS via corresponding grammemes.

PART III: SYNTACTIC DEPENDENCY

Among the three major types of linguistic dependency that we are studying, it is the Synt-**D** that attracts the lion's share of attention. It is, beyond any doubt, the most important type of dependency and, at the same time, the most controversial one. I will discuss the Synt-**D** additionally, touching on the following three points:

- Some false dogmas on the subject of Synt- \mathbf{D} (1).
- Analysis of some constructions difficult for the attribution of Synt-**1** (2).
- Advantages of Synt-D (3).

1 Current fallacies concerning syntactic dependency

One finds in the literature references to several problems for the \mathbf{n} -approach as presented above; these can be grouped under four headings: so-called "double dependency," "mutual dependency," "no dependency," and "insufficient dependency." I will consider below examples of each in order to show that these are pseudo-problems, since they stem from the confusion of different types of dependency or from using unlabeled dependencies.

1.1 "Double dependency"

There are three typical cases where many see double syntactic dependency: relative pronouns, raisings, and subordination of coordinate expressions.

1.1.1 Relative pronouns

Analyzing the phrases of the type of *the man whom we saw/the car which we saw*, many linguists—for instance, Tesnière (1959: 560) and Hudson (1990: 117)—say that the relative pronoun syntactically depends both on the MV of the relative clause (here, *saw*) and on its own antecedent (here, *man/car*; in Figure 6 below this second Synt-**n** is shown by a boldfaced branch). Since in this approach the relative pronoun is the Synt-

head of the relative clause, it syntactically governs the MV on which it, at the same time, syntactically depends:



Figure 6: The SSyntS of a relative clause according to Tesnière and Hudson

Were it so, this would be a problem for the **n**-approach, since it would mean the violation of the uniqueness-of-Synt-governor principle, as well as the principle of the absence of cycles in the Synt-structure. This would, in turn, destroy a clear understanding of the substantive nature of Synt-**n**, which is supposed to specify the linear positioning of one of its members with respect to the other—and nothing else.

However, the representation above is simply a result of confusion between different types and/or levels of dependency. I think that in the SSyntS the relative pronoun depends syntactically only on the MV of the relative clause, while standing in an anaphoric relation to its antecedent; and in many languages the relative pronoun also has a Morph-**1** with its antecedent (namely, congruence). This masks the fact that the Synthead of a relative clause is its finite MV, and by no means the relative pronoun: it is only the presence of a finite verb in a clause that licenses the speaker to use this clause as a relative, and it is this use that imposes the pronominalization of the relativized clause element, which thus becomes a marker of relativization. Here is the SSyntS of a relative clause as proposed in the Meaning-Text theory (the dashed bi-directional arrow is part of SSynt-Anaphoric Structure):



Figure 7: The SSyntS of a relative clause in Meaning-Text theory

It is, however, obvious that the relative pronoun has indeed a double syntactic nature: it is both a Synt-dependent of the MV of the relative and, at the same time, the marker of the relative. This leads some researchers to split the relative pronoun into two abstract lexical elements, one of which represents the Synt-head of the relative clause (its MV depends on this element), while the other occupies its legitimate dependent Synt-position with respect of the MV of the relative. Thus, Engel (1977: 234-235 [1988: 292-293]), following the proposal of Tesnière (1959: 561), represents the SSyntS of the German relative clause *der Mann, der Birnen verkauft* (the man who sells pears) by splitting the relative pronoun DER (that) [= (which/who)] into the relative marker part D-

and the pronominal anaphoric part -ER ^(he), obtaining something like *the man that he* [= der] *sells pears* and thus avoiding double dependency:



Figure 8: The SSyntS of a relative clause according to Engel (1977)

Relative clauses with a separate expression of the relative marker and the pronominal anaphoric element (= resumptive pronoun) exist in many languages, for instance, in Arabic, Turkish, Albanian, Persian, Middle High German, Provençal, etc. (see, e.g., Suñer 1998). But this is exactly what proves that there is no need for such a *tour de force* in English, French or German: here, the syntax of the relative clause is different. The double role of the relative pronoun in these languages is reflected on different levels of representation in terms of the three types of dependency plus the separate anaphoric relation. As far as the Synt-**1** is concerned, the relative pronoun does not depend on its antecedent—it depends on the MV of the relative clause. To justify this statement, consider the following facts:

On the one hand, the relative pronoun does not syntactically depend on its antecedent because the antecedent of a relative pronoun and the pronoun itself cannot form a phrase; thus, *[a] man whom and *[a] car which are by no means phrases of English. See Criterion A2, Part II, 4.3, p. 26.

On the other hand, some properties of the relative pronoun clearly point to its syntactic role as a dependent within the relative. The most important in this respect is the fact that relativization may be restricted by the dependent Synt-role of the relative pronoun. For instance, in some languages relativization is possible only if the would-be relative pronoun is the Synt-Subject, or if it is the Synt-Subject or the DirO, or if it is the Synt-Subject, the DirO or the IndirO, etc. Thus, the specific type of the Synt-**D** of the relative pronoun on the MV of the relative clause is crucial. To this, one could add, for instance, that the relative pronoun can be omitted in some languages (as in *the man I saw* or *the man I talk with*) without any effect on the relative; omissibility is a typical feature of Synt-dependents—although it happens to the Synt-heads as well (Part II, 4.5, p. 42). Also, in some languages, the relative clause is marked by a special form of the MV of the relative, without any relative pronoun (Bantu).

But my strongest arguments against the double dependency of a relative pronoun are as follows:

• <u>Deep-Synt-Structure of the Relative Clause</u>. In the DSyntS, the (future) relative clause has no relative pronoun at all—only its nominal source N is allowed to appear there. And this N syntactically depends of course only on the MV of its clause, being anaphorically related to its antecedent (as is the case with all substitute, i.e., anaphoric,

pronouns). When in the transition DSyntS \Rightarrow SSyntS this N is replaced with the corresponding relative pronoun, what could be the reason to add another Synt-**n** between it and its antecedent? I can see none. This consideration can be formalized by the following heuristic principle:

"As-Little-DSynt=SSynt-Restructuring-as-Possible" Principle

When deciding on the SSyntS of a phrase/clause, the researcher should maintain for it the same orientation of syntactic **n**s as in the DSyntS—except in cases of obvious necessity to reverse the dependencies, which have to be explicitly justified. (In other words, the default case must be that a DSynt-**n** remains a SSynt-**n**.)

• <u>SSynt-Structure of the Related Interrogative/Headless Relative Clause</u>. The fullfledged sentence *Who wants a lift?* has the SSyntS with the top node WANT, and this is for me an important argument in favor of establishing the same top node in the corresponding relative [*the boy*] *who wants a lift*—because I adopt another heuristic principle:

"Always-the-same-SSyntS" Principle

When deciding on the SSyntS of a phrase/clause P, the researcher should try to maintain for P always the same SSyntS no matter where this P appears in a larger formation.

Thus, the same phrase *who wants a lift* in a sentence of the type *Who wants a lift has to sign up*—this time, a headless relative—has the SSyntS with the finite verb as its top node. If I have accepted the MV as the head of an independent interrogative clause, I want this clause to have the same SSyntS also when it is used as a relative. If I have accepted the MV as the head of a normal relative clause, I prefer to treat the corresponding headless relative in the same way; and so forth. This means that in English, the finite, or tensed, verb has in its passive valence the role of the head of such phrases (= actually, full-fledged clauses) that can be used as equivalents of noun phrases—under specific conditions, such as the presence of relative pronouns.

The phrase who wants a lift is a partial syntactic equivalent of a noun phrase—it can be, e.g., a SSynt-Subject. The phrase what Alan bought in a sentence of the type What Alan bought is important—again, a headless relative—also has a finite verb as its top node: $Alan \leftarrow bought \rightarrow what$. It is also a partial syntactic equivalent of a noun phrase, since it can be the SSynt-Subject, or the DirO of the MV, or else depend on a preposition:

What $Alan \leftarrow bought \leftarrow is important.$ I like what $Alan \leftarrow bought.$ For what $Alan \leftarrow bought$ [I could pay him \$15]. The phrase *whatever apples Alan bought* (Van Langendonck 1994: 256), which is as well syntactically equivalent to a noun phrase, has a similar SSyntS—in the sense that its top node is the finite verb *bought* and the WH-pronoun depends on it (in this case, indirectly): *whatever*—*apples*—*bought*—*Alan*. (The SSyntS of *whatever apples that Alan bought* is different, its SSynt-head being APPLES:



The situation is the same with indirect-interrogative pronouns, as in *I wonder whom you love* or *He asked what book Alan had brought*. Such a pronoun depends syntactically—directly or indirectly—only on the MV of the completive clause, although there is a direct Sem-**D** between the pronoun and the MV of the matrix clause:

I wonder-sem \rightarrow *whom* ... and *He asked*-sem \rightarrow *what* ...

(for more on the representation of indirect questions of this type, see Kahane & Mel'čuk 1999).

To conclude, let it be mentioned that in the **n**-descriptions of various languages (English, Danish, Esperanto, etc.) for a Machine Translation system (Schubert 1987: 100-102, Maxwell & Schubert 1989), the relative pronoun is treated as a Synt-dependent of the MV of the relative clause.

1.1.2 Raisings

It is also said that in sentences of the type *He keeps talking*, the pronoun *he* depends both on *keeps* and *talking*, because it is the Subject of both (cf. Hudson 1988b: 194*ff*; the construction is even commonly known as Subject Raising). However, if Sem-**D** and Synt-**D** are distinguished, this reasoning does not apply: *he* is the syntactic Subject of *keeps* (HE controls the agreement of KEEP, is positioned with respect to KEEP, undergoes inversion with DO) but it is Sem-actant **1** of *talking* (this allows one to account easily for cooccurrence restrictions: **Something talks*, etc.). At the SSynt-level, there is no direct Synt-link between *he* and *talking*: **he talking* is not a phrase of English. The situation is slightly different with "meteorological" verbs: in *It keeps raining*, the impersonal pronoun *it* is semantically empty. The impersonal IT is inserted in the SSyntS by a special rule, based on the lexical entry for the verb [to] RAIN, so that the question of its semantic compatibility with the verb does not even arise (the verb [to] RAIN has no semantic actant).

1.1.3 Subordination of coordinate expressions

In sentences of the type *Alan reads books, newspapers and magazines*, the elements *newspapers* and *magazines* are often said to have two Synt-governors each: *newspapers* depends on *books* and on the verb *reads*, while *magazines* depends on *and* and again on the verb *reads*. Formally, then, the DSyntS of such a sentence must be as follows:



Figure 9: The DSyntS of the sentence *Alan reads books, newspapers and magazines* (in an approach with "double" syntactic dependencies)

Duplicating Synt-**n**s are shown by bolder branches. This type of representation is discussed in detail on the basis of Russian data in Sannikov 1989: 32-41.

Whichever the advantages of this representation, it reflects again a confusion of different types of **D**s: in point of fact, duplicating branches show Morph-**D**s (if any: *Alan sees Helen and me* $\langle *I \rangle$) and Sem-**D**s; they do not have the same nature as the coordinate Synt-**D**s, which—as all Synt-**D**s—essentially specify linear positioning of wordforms. (One of the disadvantages of a double-dependency representation is immediately clear: it presupposes the repeatability of actantial dependencies, which contradicts the postulate of uniqueness of each actant, widely shared by linguists of all schools of thought.) The intuition that I would like to capture in the case of coordinate, or conjoined, strings on the Surface-Synt-level is not that every element of a conjoined phrase depends in parallel on the same Synt-governor, but rather that a conjoined phrase as a whole depends on its Synt-governor via its Synt-head (= its first element, see 1.3 below).

1.2 "Mutual dependency"

Fairly often, grammarians insist on mutual dependency between the MV of a clause and its SSynt-Subject. They say that even if it is the MV that represents the whole clause, the Subject controls the form of the verb (*The cat is sleeping* vs. *The cats are sleeping*); moreover, the Subject and the MV constitute a communicative unit consisting of a theme/topic and a rheme/comment. Again, such statements are due to confusion between different levels of dependency: the fact that the Subject depends on the MV syntactically does not prevent the MV from depending on its Subject morphologically. In many languages the MV agrees not only with the Subject but with the DirO (and sometimes with the IndirO) as well: cf. (35a) above, p. 64; this, however, does not belie the universally accepted syntactic status of objects as dependents of the MV.

1.3 "No dependency"

While some linguists treat coordination by means of double dependencies (1.1 above), it is also frequently said (Matthews 1981: 196, Hudson 1988a: 314) that there is no Synt-**D** at all within conjoined, or coordinate, expressions: in *Leo and Alan [came]*, as well as in *Leo or Alan [will do it]* nothing is the Synt-head. This viewpoint goes back to Tesnière 1959: 339ff.^[8] Once again, Synt-**D** is being confused with subordination (which is a particular case of Synt-**D**). *Leo and Alan* is a phrase of English, and so is and Alan, while *Leo and is not (the fact that a pause is possible after AND—as, for

instance, in *Leo and*, \parallel *believe me or not*, \parallel *Marga* ...—does not impart to the expression **Leo and* the status of a phrase; it still is not an utterance of English). The phrase *Leo and Alan* has thus the passive Synt-valence of *Leo*, and not that of *and Alan*, the passive Synt-valence of the phrase *and Alan* being determined by *and* rather than by *Alan* (the phrases such as ... *and Alan*, ... *or Alan*, ... *but not Alan* etc. can be only conjoined constituents, and this property comes from the coordinate conjunction); therefore, the Synt-**n** s in *Leo and Alan* are as follows:

LEO-coordinative → AND-conjunctional → ALAN.

In a conjunctionless coordinate phrase such as *Leo, Alan, Helen* the Synt-**D**s are LEO-coordinative->ALAN-coordinative->HELEN.

The Synt-head of a conjoined phrase is, at least in English and similar languages, its first element (independently of the presence/absence of a coordinate conjunction). Note that in a number of languages, the first element in a coordinate string has some special properties. Thus, in some Bantu languages (for instance, Bafia), only the first verb in a coordinate string of verbs (*stood up, drank his coffee, took the book and left*) has a complete morphological marking, including tense; all the following verbs are in a special—*conjunctive*—form, which precludes the expression of tense. In Nias (Indonesia), in a string of coordinated nouns, only the first noun is inflected according to the external context, while all the others remain in the unmarked nominative; etc. For more on Synt-**D** in connection with coordination, see **5** below; a concise review of possible solutions to the problem of **D**-description of coordination is presented in Schubert 1987: 104-119.²³

1.4 "Insufficient dependency"

Many linguists have criticized the **n**-approach for its incapacity to express what they call the multilayer character of syntactic structure. Thus, in *Alan gives an apple to Helen*, APPLE is somehow closer—syntactically, not linearly!—to GIVE than is HELEN; and ALAN has the loosest link to the verb (external argument, as it is known in some approaches). Without going into a discussion of what this syntactic closeness really means, I can point out simply that all such distinctions are easily and naturally expressed via the names of D-/S-SyntRels:



Labeled SSyntRels guarantee the capacity of any **n**-description to state the closeness (or remoteness)—of course, in the structural sense—of any elements of the sentence.

²³ The solution Schubert himself prefers is to take as the head of a conjoined phrase the coordinate conjunction: $\text{LEO} \leftarrow \text{AND} \rightarrow \text{ALAN}$. But this solution is unacceptable for me; see 2.6 below, p. 80.

2 Syntactic dependency in action: Eight illustrative case studies

Let us consider several particular syntactic constructions and show how they are described using Synt-**1**; I will mainly concentrate on the orientation of Synt-**1**, i.e., on the problem of Synt-head. Since in some cases all theoretical arguments advanced in support of this or that viewpoint concerning the treatment of a given phrase in terms of Synt-**1** seem inconclusive, I will try to argue based on the rules necessary to produce the construction in question within the framework of a stratificational multilevel semantically-oriented linguistic model (more specifically, the Meaning-Text Model). My goal is to show that the opposite decision concerning the choice of the Synt-governor entails a complexification of the processing rules.

2.1 Russian numeral phrases

In Russian, a numeral phrase NUM + N shows rather complex behavior:

— if the numeral does not end in ODIN ⁽¹⁾ and is in the nominative or the accusative, the noun is in the genitive and its number depends on the numeral (with DVA ⁽²⁾, TRI ⁽³⁾, and ČETYRE ⁽⁴⁾ or any numeral that ends in these three—23, 32, 44, ..., 1452, etc.—the noun is in the singular, while with all other numerals it is in the plural);

— if the numeral ends in ODIN (1) (e.g., 1231), no matter in what case it is, the number of the noun is singular;

— if the case of the numeral (other than ODIN (1)) is the nominative or the accusative, and if it is (or ends in) DVA, it agrees with the noun in gender; etc.

This complexity engendered much discussion concerning the orientation of Synt- $\mathbf{1}$ in the NUM + N phrase; all logically possible solutions have actually been proposed (NUM \leftarrow N; NUM \rightarrow N; NUM \leftrightarrow N; in the nominative and the accusative it is NUM \rightarrow N, in other cases NUM \leftarrow N; etc.). In actual fact, the orientation of Synt- $\mathbf{1}$ in Russian numeral phrases is always NUM \leftarrow N, since the passive Synt-valence of the phrase is obviously that of N, and not that of NUM. What obscures the picture is again confounding the Synt- $\mathbf{1}$ with variegated Morph- $\mathbf{1}$ s (Mel'čuk 1985: 59-102; for the opposite view—NUM \rightarrow N, i.e., the numeral is the Synt-head,—see Corbett 1993).

To make my point clearer, I will describe the production of two phrases, one with a genuine numeral DVA 'two' and another one with a measure noun KUČA 'heap, a lot':

[*On pročël*] *dva romana* ⁽[He read] two novels². [*On pročël*] *kuču romanov* ⁽[He read] a-lot of-novels².

In the SemS, both phrases have a similar representation, where quantification appears as any semantic modification would:



Or, using an obvious abbreviation:



In the SemS \Rightarrow DSyntS transition, the direct DSynt-link between ČITAT' (read) and ROMANY (novels) is not necessarily preserved. Namely, if the quantifying expression is a noun, it becomes the DSynt-governor of ROMANY:



This is done because the DSyntS is a syntactic structure, so that it is supposed to reflect, in the most faithfull way possible, the syntactic organization of the sentence. And from the purely syntactic viewpoint, the constructions *dva romana* and *kuča romanov* are very different. (Cf. the As-little-DSynt⇒SSynt-restructuring-as-possible Principle, Section 1.1.1 of this part, p. 70.)

Under the DSyntS \Rightarrow SSyntS transition, the NUM DVA remains subordinated to the quantified N, and the N_{measure} KUČA remains the SSynt-governor of the quantified N:



Syntactically, the phrase *čitat' kuču* [romanov] ⁽read a lot of [novels]⁾ behaves exactly as any other pair $V_{(trans)}$ -dir-obj \rightarrow N: the N is positioned and inflected as any regular DirO should. (Among other things, the noun KUČA becomes the Synt-Subject under passivization: Byla pročitana kuča romanov ⁽A-lot of novels was read⁾.) On the other hand, kuča romanov behaves as any other pair N-compl \rightarrow N does. Because of this, for *čitat' kuču romanov*, the Synt-**1** KUČA \rightarrow ROMANY is justified by a considerable economy of rules, which otherwise would have to be doubled: special rules would be needed to describe the treatment of a quantifying modifier (= KUČA) that behaves as a DirO and another set of rules for the treatment of a DirO (= ROMANY) that behaves as an adnominal complement. But for *čitat' dva romana* nothing justifies an inversion of the Synt-**1** DVA \leftarrow ROMANY: the extremely complex rules that compute the inflection of the NUM, of the quantified N and eventually of some depending adjectives ($\dot{e}t+i$ dva romana 'these two novels', dva interesn+yx romana 'two interesting novels', etc.) remain the same, whichever element is taken to be the Synt-head in the phrase NUM + N (see these rules in Mel'čuk 1985: 162-210). Therefore, the overall simplicity of DSynt-rules requires to invert the Sem-**D** between NUM and N, that is, to have between them a Synt-**D** oriented in the opposite direction: 'NUM'-sem \rightarrow 'N', but NUM \leftarrow synt-N (a more complex version of Case 14, p. 65). Ergo, on both the DSynt- and SSynt-levels, we have NUM \leftarrow N.

The treatment proposed for *čitat' kuču romanov* (read a lot of novels) can be easily extended to cover all the constructions including quantifying expressions, in particular the measure phrase, as in Rus. *vypit' tri litra piva* (have drunk three liters of beer), Eng. *have eaten ten dollars of bagels*, Fr. *faire deux heures de sieste* lit. (have two hours of nap), etc. In the SemS, a measure phrase is represented as a modifier of the noun quantified—in the same way as the synonymous expressions *vypit' pivo v količestve trëx litrov* lit. (have drunk beer in quantity of three liters), Eng. *have eaten bagels for ten dollars*, Fr. *faire une sieste de deux heures* lit. (have a nap of two hours). It is on the level of Syntactic Structure (both Deep- and Surface-) that the inversion of dependency—Sem-**D** *vs*. Synt-**D**—takes place.

2.2 A Russian "approximation"-marking preposition

In the Russian phrase *okolo pjati kilogramm* 'about five kilos' the preposition OKOLO (lit. 'close to', here 'approximately') is the Synt-head of the phrase: without it, the numeral phrase has the exact distribution, i.e., the passive Synt-valence, of a noun, but with OKOLO the numeral phrase can only be used as the Synt-Subject or DirO. Thus, the phrase with *okolo* cannot be the complement of a preposition (**dlja okolo pjati kilogramm* 'for about five kilos') or an IndirO (**raven okolo pjati kilogrammam* '[is] equal to about five kilos'). Therefore, on the SSynt-level, we obtain the following structure:

OKOLO→KILOGRAMM→PJAT'.

This representation is buttressed by the complete identity in syntactic behavior of this *okolo* and all other genuine Russian prepositions; thus, all of them, together with the numeral, follow the noun quantified in the **approximate-quantitative** construction: *dlja pjati kilogramm* 'for five kilos' ~ *kilogramm dlja pjati* 'for approximately five kilos' and *okolo pjati kilogramm* 'about five kilos' ~ *kilogramm okolo pjati* 'approximately about five kilos'. (In English *about* 'approximately' does not have the same properties: *for about five kilos* is OK; therefore, its SSynt-status is different: ABOUT (FIVE KILOS.)

A similar construction exists in Latin:

(36) Latin *Circa quingent+os Roman+orum cecid +erunt* around 500 ACC Roman PL.GEN fall-PERF 3PL ^(About 500 Romans fell).

The preposition CIRCA governs the case (namely, the accusative) of QUINGENTI ⁽⁵⁰⁰⁾, as all Latin prepositions do: it is the Synt-head of the phrase *circa quingentos Romanorum*, which is the Synt-Subject; however, CIRCA is omissible without any syntactic effect on the phrase.

In Russian, as in Latin, treating the preposition OKOLO/CIRCA as any other preposition, i.e., as the Synt-governor of the numeral phrase, allows us to avoid writing special syntactic rules to treat these syntactically quite ordinary prepositions, which are only semantically deviant (they manifest a kind of adverbial meaning: semantically, they are monoactantial—in contrast to genuine prepositions, which are biactantial).

2.3 Determiners as heads?

Several linguists argue that in the DET + N phrase the determiner is the Synt-head: thus, in English we should have THE–synt \rightarrow N, ANY–synt \rightarrow N, etc. (Hudson 1984, 1990: 271*ff*, Hewson 1991; cf. also Vennemann 1977: 270, 296). I cannot analyze their argumentation in depth, but within the framework expounded above, such a description is unacceptable, and this, for the following three reasons.

First, the passive Synt-valence of the phrase *the dog* is that of the noun *dog*, not of the article *the*. If in some syntactic positions DOG cannot appear without an article (or any other determiner), this happens because articles and the determiners in general are analytical exponents of grammemes of an inflectional category—namely, of the definiteness of the noun—and in these positions an English noun cannot be used without a marker of its definiteness, just as a Latin noun cannot be used without a case-number suffix. The expression **Dog is barking* is ungrammatical, independently of the fact that its SSyntS is well-formed: the problem here is the incorrect grammatical form of the lexeme DOG, very much like the bad expression **The dogs is barking*, where the SSyntS is also 100% correct, but the grammatical form is not. *Ergo*, the SSyntS is here *the*—synt–*dog*, *any*—synt–*dog*, etc.; cf. the phrase *that* (*stupid*) *John*, which has the distribution of *John* and not that of the determiner *that*, or *Dogs are faithful*, where the noun *dogs* appears without any overt determiner.

Second, it is necessary to reflect the parallelism in the syntactic behavior of such phrases as *the dog*, *this dog* and *Alan's dog*; are we prepared to say that *this* and *Alan's* are the SSynt-governors of *dog*?

Third, analytical exponents of grammemes of a lexeme in most cases syntactically depend on this lexeme—in the SSyntS, since they do not appear at all in the DSyntS (auxiliary verbs that express the grammemes of tense, mode, aspect or voice constitute an important exception, see below, 2.5). For instance, in Tagalog an analytical case marker of an N-that is, ang [NOM], ng [OBL] or sa [DAT]-syntactically depends on the noun, while the preposition (that syntactically dominates N) imposes the choice of the case: in the DSyntS, we have, for instance, PARA 'for'-synt-BABAE 'woman', and in the SSyntS, PARA-synt-BABAE-synt-SA, which gives para sa babae 'for [the] woman' (the preposition PARA requires the dative). In languages where the plural of a noun is expressed by a separate word (Dryer 1989), this plural exponent equally depends syntactically on the noun: Yapese (Austronesian) ea pi - synt-kaarroo neey 'the PL car this' = 'these cars' or Mixe (Mexico) he pi' miš-synt \rightarrow ?aHkš' (the little boy PL') = 'the little boys). Cf. also the Russian particle BY that expresses the subjunctive of a verb on which it depends: Ja poexal-synt $\rightarrow by$ (I would go). In Hawaiian (and other Polynesian languages) all markers of the verb's inflectional grammemes are analytical and syntactically depend on the verb:

ua ← synt-*hele au* ⁽COMPL(etive) go I⁾ = ⁽I went⁾, *ke* ← synt-*kali* – synt → *nei au* ⁽PROGR(essive) wait I⁾ = ⁽I am waiting⁾ [where KE ... NEI is a bilexemic progressive marker] *e* ← synt-*kali* ⁽IMPER(ative) wait⁾ = ⁽Wait!⁾, etc.

Considering an analytical grammeme marker as a Synt-governor in the SSyntS would entail a restructuring of the DSyntS, where this marker does not appear at all; but why engage in something complex when one can easily do with something simple? Since the article is a particular case of an analytical grammeme marker, it should be considered a Synt-dependent of the noun. (The solution DET \leftarrow synt-N is successfully defended in Van Langendonck 1994; for a different treatment of the DET + N phrase in Salishan, that is, DET-synt-N, see Beck 1997: 109-118.)

2.4 *Romance clitics*

Clitics in French (and in other Romance languages, where Clitic Raising exists) pose a special difficulty for a **\mathbf{n}**-description: the clitic does not always syntactically depend on the same wordform on which its source [= the noun the clitic replaces] depends. Thus, compare (37a) with (37b), where the clitic changes Synt-governor vis-à-vis that of its source:

- (37) a. Elle a été fidèle-synt $\rightarrow a$ Pierre ⁽She has been faithful to Peter⁾.
 - b. *Elle lui* synt-*a été fidèle* lit. 'She to-him has been faithful'.

On the SSynt-level, where, under synthesis, clitics first appear (the DSynt-level admits only nominal sources of clitics-to-be), a clitic depends syntactically on its host word, with which it forms a possible prosodic unit, i.e., a phrase (as in Sp. *lo ve* lit. ([(s)he] it sees) or *le da* lit. ([(s)he] to-him/to-her gives) and with respect to which it is linearly positioned. In Romance languages, the host of a clitic in the SSyntS is not necessarily the same element on which the source of the clitic depends in the DSyntS. The "new," i.e., Surface-Syntactic, governor of the clitic is computed by special rules of the DSyntS \Rightarrow SSyntS transition. Roughly speaking, it is the Synt-head of a dependency chain (= a nucleus) on the last element of which the source of the clitic depends.²⁴

2.5 *AUX* + *V* phrases, English-style

What is the orientation of Synt-**n** in the phrase AUX + V in English (and similar languages), $AUX \leftarrow synt-V$ or $AUX - synt \rightarrow V$? An AUX is an analytical exponent of a grammeme (or of several grammemes); according to what was said in Subsection 2.3, it should be considered the dependent of the lexical verb. However, an AUX greatly differs from analytical markers discussed above: while all of them are "lonely particles," an AUX is a verbal lexeme that has its own inflection and accepts its own syntactic dependents. Because of this, it requires a special discussion.

 $^{^{24}}$ Second-position clitics, as found, for instance, in Serbian are even more problematic; see J. Milićević's paper in this volume, p. 235*ff*.

As before, my argumentation is based on the rules necessary to produce the phrase in question from a SemS. Suppose we want to have the sentence *Alan has slept*. Starting with a SemS

(plus the indication of time), the semantic rules of Lexicalization and Arborization construct the DSyntS of the form

The analytical form of the verb is represented, at this level, as one node directly linked to the subject node by the DSyntRel I; thus, all lexical selection constraints (that may exist between the verb and its Synt-Subject) can be easily accounted for. In the SSyntS, the DSynt-node

o SLEEPind, pres, non-progr, perf

is expanded into the subtree

HAVE_{ind, pres} o-perfect-analytical->o SLEEP_{past.participle}

by the following Deep-Syntax rule:

 $L_{(V)ind, pres, perf} o \Leftrightarrow HAVE_{(V)ind, pres} o-perfect-analytical \rightarrow o L_{(V)past.participle}$.

NB: This rule, as well as all similar rules describing analytical verbal forms, needs a set of conditions that foresee possible combinations of several analytically expressed grammemes.

From the purely logical viewpoint, we can take HAVE as the SSynt-head of the phrase AUX + V (as shown in our illustration) or as the SSynt-dependent: for this rule itself it makes no difference. But for all Surface-Syntax rules which have to compute the inflections on HAVE, linearly position AUX and V in the sentence and check the well-formedness of the SSyntS the difference is quite significant. If HAVE is considered to be the SSynt-head, all SSynt-rules that apply to any pair $N \leftarrow subj - V$ will automatically apply to the auxiliary HAVE, as they do to any verb in the role of SSynt-head, including the non-auxiliary HAVE. More specifically, aspects under which an auxiliary (BE, DO and HAVE) must be treated as any full finite verb of English include:

• Agreement with the Subject (*I have written* vs. *He has written* as *I read* vs. *He reads*; including all complex cases of the type *The committee has/have*, etc.).

• Word order, in particular, inversion (Have I written? as Can I write?).

• Being the only verb in the clause ([*I know that*] *he has* as [*I know that*] *he works*; or in tags: *He has not gone, has he*? as *He works, doesn't he*?).

• Carrying the grammemes of mood and tense.

• Parallelisms with the non-auxiliary BE, DO and HAVE (*He was astonished by John* as *He is astonished*; *He does work* as *He does his work*; *He has arrested John* as *He has John arrested*; *He has gone* as *He has to go*).

• Parallelisms with semi-auxiliaries such as GET (*He got robbed*), KEEP (*He keeps going*) or GOING TO (*He is going to read*), which have to be treated in a similar way to the genuine auxiliaries.

On the other hand, there are no idiosyncratic SSynt-properties of English auxiliaries that would require any specific rule to deal with them.

However, if the auxiliary HAVE (or BE, or DO) is not taken to be the SSynt-head of the AUX + V phrase, a bunch of additional rules have to be written to deal with a finite auxiliary verb which is not the SSynt-head of the clause, as well as with a non-finite verb form which is. There is no justification for such useless multiplication of entities; *ergo*, on the SSynt-level, we have AUX \rightarrow V (which corresponds to Criterion B1). For more argumentation in favor of AUX \rightarrow V, see Hudson 1987: 117-118 (English).

2.6 Conjoined nominal phrases N + CONJ + N

According to Criterion B1, in such a phrase as *Alan and Helen* the Synt-head is ALAN: the passive Synt-valence of the phrase is the same as that of the noun ALAN (rather than that of the phrase AND HELEN). Nevertheless, in many syntactic frameworks AND is considered the Synt-head of a coordinate string, ALAN and HELEN being its actants: $ALAN \leftarrow AND \rightarrow HELEN$ (the same description is applied to all coordinate conjunctions: e.g., Schubert 1987: 104ff).

Once again, purely theoretical argumentation proved to be of little use here, so I will reason from the viewpoint of the rules that have to synthesize such strings.

If we take the conjunction to be the SSynt-head of the coordinate string we run into the following four difficulties.

• First, to be able to specify the linear order of conjuncts, which in many cases is significant, both Synt-**1**s starting from CONJ have to carry different labels.

• Second, rules for the conjoined strings without a coordinate conjunction—such as *Alan, Helen, Leo*—must be completely different, or else a dummy artificial conjunction has to be added to the SyntS.

• Third, the linguistic rules that deal with the SSynt-Subject, the DirO, the IndirO, the complement of a preposition, the Prolepsis (= Fronted Topic), and the like—in one word, with an N—have to be repeated for the conjunction! And this will be sufficient only for the conjunction that links two nouns; for conjunctions linking lexemes of other parts of speech still other additional rules are needed. More than that: these rules must be extremely complex, since they have, e.g., to assign grammatical case to the conjunction and then percolate it to the nouns linked by the conjunction, etc.

• Fourth, must the conjunction be the Synt-head at the DSynt-level? Presumably so, since this is closer to its semantic role. Then all the selectional constraints acting between the verb and the conjoined nouns will not be easy to check—again an unnecessary complication.

In case we take the initial (= first) element of the conjoined phrase to be its Synthead, no rules dealing with nouns, verbs, adjectives or adverbs have to be doubled and no additional rules are required (just specific rules for the conjunction and the second conjoined element), nor any dummy added. At the same time, absolutely nothing is lost. Why should then anyone want to add complexity without any visible gain? *Ergo*, on both the DSynt-levels, we have, respectively:

 X_1 -coord \rightarrow CONJ-II \rightarrow X_2

and

$$X_1$$
-coordinative \rightarrow CONJ-conjunctional \rightarrow X_2 .

2.7 Russian "exotic" coordination of interrogative/negative pronouns

In Russian, the interrogative and negative pronouns which bear different Deep-Syntactic relations to the governing verb are allowed to form a coordinate string in the Surface-Syntactic structure (in the Deep-SyntS there are no direct syntactic links between these elements: they are subordinated to the verb "in parallel"):

(38) a. *Kto, komu i čem pomog?* lit. ^(Who, to whom and with what helped?)

b. Nikto, nikomu i ničem ne pomog

lit. Nobody, to nobody and with nothing not helped).

To represent the phrase *kto, komu i čem* on the SSynt-level simply as all other coordinate phrases are represented, that is, as

kto-coord \rightarrow komu-coord \rightarrow i-conjunct \rightarrow čem,

is insufficient. In a regular coordinate construction any Synt-dependent element plays with respect to the Synt-governor of the whole coordinate string the same Synt-role as its Synt-Governor itself. However, in this case, kto [NOM] is the Synt-Subject, but komu [DAT] is an IndirO and *čem* [INSTR] is another IndirO of the verb *pomog* (helped); accordingly, all three pronouns are inflected differently. To account for this, it has been proposed (Sannikov 1989) to use double dependency, namely to add to the SyntS above the indication of the direct Synt-**n** of each pronoun on the verb pomog (see above, Section 1.1.3 of this part, Figure 9, p. 72). But these added Synt-ns do not have the same substantive nature as the Synt-**n**s covering the coordination in this case: the added links are needed only to compute the Morph-ns (under synthesis) and the Sem-ns (under analysis). However, as we have seen, Morph-**D**s and Sem-**D**s can link two wordforms that do not have a direct Synt-**D** between them. Therefore, it is preferable to introduce some special SSyntRels just for this very special construction: coord-subj, coord-dir-obj, coord-indir-obj, etc. Such SSyntRels indicate, in a natural way, the exotic character of this coordinate phrase and encode the actual SSynt-roles of its displaced elements.

A similar method can be used in comparative constructions (Savvina 1976). For instance, in Russian, the two comparative constructions in (39) have to be distinguished in their SSyntSs in the following way:

(39) a. [*Ja ljublju Mašu bol'še*,] *čem*-conjunct-subj \rightarrow *Van*+*ja* [NOM] ^(I) love Masha more than Vanya [does]⁾.

Another possibility to represent the SSyntS of these constructions would be to consider the grammatical case of the SSynt-dependent in such coordinate or comparative strings as semantically meaningful and admit it into the SSyntS of these constructions; this is, however, too technical a point to be discussed here.

2.8 Elliptical constructions

How should one describe in the SSyntS common gappings of the type Alan went to Paris and Leo to Coruña? Since the expression and Leo to Coruña is not a normal phrase of English, it cannot, such as it is, be assigned a well-formed SSyntS. It is a "mutilated" expression, which has lost its top node, in this case a finite verb, and become a "non-canonical conjunct," in Hudson's (1988a: 305) terms. But even if deleted on the surface, the verb elided imposes on its SSynt-actants, where appropriate, government-induced inflection grammemes. Therefore, to represent the expression in question in terms of Synt-**D**, it is proposed to use in its SSyntS a node labeled with the elided verb —in this way the verb's government is specified; the verb is, however, marked as being elided—in this way the verb is invisible to SSynt-to-DMorph-rules and will not appear in the DMorphS of the sentence. The ellipsis of the verb is shown in (40) by shading:





This is how the SSyntS of elliptical expressions—or, more precisely, of non-canonical conjuncts—looks in the Meaning-Text approach: the element to be elided is shown in the SSyntS, but with an explicit indication of its ellipsis (on ellipsis in the SSyntS, see also below, Item 5, p. 88).²⁶

I like her more than-[he]-conjunct \rightarrow does vs. I like her more than-conjunct \rightarrow him. For this construction, we do not need special conjunctional SSyntRels.

 $^{^{25}}$ Note that, for instance, in English the situation is different, because of the impossibility of **I like her more than he*:

²⁶ Alternatively, the representation of ellipsis could be relegated to a later stage—that is, to the DMorphS: the ellipsis would not be shown at all in the SSyntS, but in the DMorphS the elided element simply would not appear; the operation would be carried out by some SSyntS-to-DMorphS rules, which compute the possibility (or the necessity) of this ellipsis based on the SSyntS itself. For the time being, I do not see any logical objections to such a strategy. The reason for which I prefer to keep an explicit indication of ellipsis in the SSyntS of a sentence is my intuitive feeling that a complete sentence and a sentence with ellipsis are syntactically different and, therefore, this difference should be reflected at the SSynt-level. Under such an assumption, the ellipsis is specified by the DSyntS-to-SSyntS rules; SSynt-to-DMorph-rules merely ensure its "mechanical" implementation.

3 Advantages of syntactic dependency

The remarks that follow are very sketchy and superficial: a systematic discussion of the advantages of the **n**-approach and its comparison with the constituency, or \mathbf{c} -, approach would require another paper, perhaps longer than the present one. Still I think that these remarks can be useful, since at least they point in the right direction.

Let me begin with two general considerations. First, in a linguistic description that takes semantics into account seriously, the **n**-approach in syntax imposes itself, since it ensures a much better fit of syntactic structure with semantic structure, where dependencies are universally recognized (most versions of predicate calculus language used in semantics are, in point of fact, **n**-based). A lack of interest in semantics and the postulate that syntax is autonomous are main factors that have lead to the dominance of **c**-representations in syntax. In a theory where the Synt-structure of a sentence is produced (roughly) from the Sem-structure of this sentence, this Sem-structure being written in terms of Sem-**n**s, it is much more natural to see the Synt-structure as being based on Synt-**n**s.

Second, a **D**-representation with labeled SyntRels is formally more powerful than a "pure" **C**-representation—in the sense that the former allows one to present all relevant syntactic details much better than the latter. As a result, linguists have been forced, practically from the beginning, to specify heads of the constituents as opposed to satellites (e.g., Pittman 1948) and the relations between them. But in a **C**-representation, as soon as one starts marking heads and indicating types of SyntRels between heads and satellites, the heavy machinery of constituency—particularly, non-terminal nodes, numerous empty nodes, and artificial ordering of elements in the SyntS—becomes use-less, because redundant: all these pieces of information can be easily computed, if and when needed, from the **D**s specified. Because of this, most modern syntactic theories—such as Perlmutter's Relational Grammar, Bresnan's Lexical-Functional Grammar or Pollard and Sag's Head-Driven Phrase-Structure Grammar—are moving fast in the direction of the **D**-approach.

To these general considerations, one can add a number of specific ones. Namely, there are at least five important linguistic phenomena for the description of which Synt-**D** is really crucial: valence, voice, restricted lexical cooccurrence, word order, and ellipses of all types. (I am not implying that the **c**-approach cannot handle them; but the **D**-approach does it, I think, in a more natural and therefore more economical way.)

1) Valence—or, more precisely, active valence—is a property of lexemes: a lexeme opens slots for other lexemes that it "attracts" as its dependents. Linguistic valence is obviously a metaphor based on valence in chemistry: atoms have valences to link with other atoms and thus form molecules. In much the same way, a lexeme has semantic, syntactic and morphological valences to link with other lexemes. Lexemes L_i that fill the valences of the lexeme L depend on it, exactly in the sense in which dependency has been defined above. Actually, valence and dependency are related in a very direct way; cf. Baumgärtner 1970: 62*ff*, Eichinger & Eroms 1995 and also Agel *et al.* 2003. Active valence is of course not the only source of dependency: there is passive valence of lexemes as well; however, active valence shows the convenience of using Synt-**D** in an especially graphic way.

2) The inflectional category of voice is crucial to the understanding of semantics, syntax, and morphology. Voice grammemes mark the change of the *basic diathesis* of the verb, i.e., the correspondence between its semantic and syntactic actants (Mel'čuk 1997a and 2006: 189), or, to put it differently, between its Sem- and Synt-dependents. No wonder, then, that voice and voice-related categories are much better described in the **D**-approach. In particular, they have been the focus of research within the framework of Perlmutter's Relational Grammar or Foley and Van Valin's Function and Reference Grammar far more than in any **C**-based theory.

3) For a systematic description of *restricted lexical cooccurrence*, or of *collocations*, the apparatus of Lexical Functions is proposed (Žolkovskij & Mel'čuk 1967: 198-218, Mel'čuk 1974: 78*ff*, Mel'čuk 1996a). Each collocation is described as having the structure $\mathbf{f}(x) = y$, where \mathbf{f} is a particular *lexical function*, x is a lexical unit which is the *base* of the collocation, and y, a set of (more or less synonymous) lexical units each of which is the *collocate*; it expresses, contingent on x, the meaning of \mathbf{f} . Here are a few examples:

Intensifier	Support Verb	Realization Verb
Magn(smoker) = inveterate Magn(sleep) = like a log Magn(hot) = burning	Oper ₁ (<i>favor</i>) = <i>do</i> Oper ₂ (<i>exam</i>) = <i>take</i> , <i>sit</i> Oper ₃ (<i>order</i>) = <i>receive</i>	$\begin{aligned} \mathbf{Real}_1(goal) &= achieve\\ \mathbf{Real}_2(exam) &= pass\\ \mathbf{Real}_3(order) &= execute \end{aligned}$

The number of simple standard lexical functions is about 60, and they are universal. Their values, on the contrary, are of course language-dependent: they are specified, for each language and each lexical unit, in a special lexicon. Using them greatly facilitates lexicalization in the transition SemS \Rightarrow DSyntS \Rightarrow SSyntS, when the appropriate collocates have to be selected.

Now, as is easy to see, the lexical-functional dependency between the base lexeme of a collocation and the collocate lexeme is supported by a Synt-**D** between them. Thus, **Magn**(*armed*) = to the teeth, and *armed*-synt→to the teeth. Similarly, **Oper**₁(visit) = pay, and pay-synt→visit, or **Real**₂(exam) = pass, and pass-synt→exam. For each lexical function, a particular Synt-**D** between its base and the collocate is specified. Outside of Synt-**D**, there is no economical way to describe the collocations properly.

4) Synt-**D** is especially convenient for the description of word order. Using Synt-**D**s forces the linguist to separate strictly and consistently the hierarchical (= genuinely syntactic) order from the linear order, which is a surface means for the expression of the former. Thus, the main task of natural language syntax—linearizing a two-dimensional Synt-structure (explicitly formulated in Tesnière 1959: 19-20)—can be solved with much more ease with Synt-**D**s than in any other way. The advantages of Synt-**D**s for the description of word order can be resumed in three points:

• First, word order rules can be easily formulated in terms of positioning a Syntdependent with respect to its Synt-governor (before or after it). Again, Tesnière (1959: 22-25, 32-33) stated this fact explicitly, dividing languages in *centripetal* (where a Synt-dependent precedes its Synt-governor) and *centrifugal* (where a Synt-dependent follows its Synt-governor); both can be so consistently or inconsistently. For some languages, this allows for very compact formulations; e.g., in Japanese all Synt-dependents precede their governors, in Welsh almost all Synt-dependents (the only exception being the article *y*) follow their governors (Hudson 1990: 105):

(41) a. Japanese [a consistently centripetal language]

Itiban takai siraga+de+no sensei+wa kono omosirokunai hon+o kai+ta very tall gray-haired professor this boring book wrote lit. 'Very tall gray-haired professor this boring book wrote'.

b. Welsh [a consistently centrifugal language]

Ysgrifennodd athro tal iawn a gwallt llwyd ganddo y llyfr undonnog hwm wrote professor tall very and hair gray to-him the book boring this lit. ^{(Wrote professor tall very and hair gray to-him the book boring this⁾.}

But even in languages where the linear distribution of Synt-governors vs. Syntdependents is not as clear-cut as in Japanese or Welsh, that is, in "inconsistent" languages, resorting to these notions helps to state the word-order rules. Thus, in Arabic the majority of Synt-dependents follow their governors, with the notable exception of the demonstratives and numerals; in Hungarian, the majority of Synt-dependents precede their governors, with the notable exception of the relative clause; etc. Such facts were discussed already in Trubetzkoy 1939 and practically used in numerous language manuals and descriptive grammars; cf. an outline of word-order typology in Xolodovič 1966. In this vein, a relatively complete description of word order in Synt-**1**-terms (within simple clauses) was proposed for Russian (Mel'čuk 1967, 1974: 260-302).

• Second, Synt-**n** has allowed for the discovery (Hays 1960, Lecerf 1960) of an important property of word order in all languages, called *projectivity*. According to the property of projectivity, in most cases, if we supply a sentence with its SyntS written in terms of Synt-**n** and draw a perpendicular from each wordform to the node that represents it in the SyntS, then:

- no branches of the SyntS intersect;

— no branch intersects with a projection perpendicular.

Let me illustrate the phenomenon of projectivity with sentence (1), associating its surface form with its SSyntS (Figure 10, next page).

As one can easily see, the sentence appears as a projection of the SSyntS such that SSyntS's branches cross neither each other nor the projection perpendiculars; hence the name "projectivity".

An absolute majority of sentences in most languages are projective. Taking this fact into account ensures a more elegant formulation of word-order rules and greatly facilitates the analysis and synthesis of texts: with the exception of particular cases (see immediately below), only projective sentences must be produced from a given SSyntS, and only SSyntSs that guarantee projectivity must be associated with a given sentence.



Figure 10: Sentence (1) and its SSynt-Structure

However, projectivity can be systematically violated in many special cases, for instance:

1. English



The culprit here is the superlative marker of the adjective; cf. the representation in a tree form, where intersections that create non-projectivity are shown:



2. French



lit. ^{(the} girl whose I know the father) = ^{(the} girl whose father I know)

The culprit is the extracted relative pronoun *dont*:



3. Serbian



lit. ^{(Of-faith to-him has [actually, 'is'] teacher answered⁾. =} 'The teacher of faith has answered to him'.

The culprits are two displaced clitics—a dative pronoun mu 'to him' and the auxiliary verb *je* 'is':



4. Latin



lit. 'Last Cumaeian comes now song's age'. =

'Now comes the last epoch of [the] Cumaeian song' [Vergil, *Eclogae*, IV.4].

Non-projectivity is provoked here by the fronting of two adjectival modifiers, *ultima* and *Cumaei*, a wide-spread phenomenon in Latin poetry:



I could indicate more constructions in different languages where non-projectivity is possible or even obligatory; all such cases must be isolated and specified in concrete linguistic descriptions. Now, while within the **D**-approach the representation of nonprojective structures poses no problem whatsoever, the **C**-approach is unable to represent non-projective structures without some additional (and relatively clumsy) machinery, for instance, transformations. In this respect, the **D**-approach is again superior to its rival.

• Third, the **D**-approach is much less rigid (than the **C**-approach) and has the inherent ability to accommodate easily what is known as non-configurationality and long-range dependencies. The perturbations introduced into the word order of a sentence by its Communicative Structure-Frontings, Extractions, Postponings, etc. plus all sorts of "displacements" in such languages as German or Russian-can wreak havoc on a **C**-structure, since even the closest-knit phrases can be torn apart and permuted. On the other hand, **1**-structure, without linearity and contiguity, is totally insensitive to such permutations: they happen in the linearized DMorphS of the sentence and do not at all affect the SSyntS. The reason is obvious-a strict and complete separation of hierarchical (= syntactic) and linear links in the **D**-approach. As a result, the **D**-approach does not know problems in representing discontinuities, which, in the literal sense of the word, simply do not exist in a **D**-structure. Thus, the sentence Which violins are these sonatas easy to play on? is assigned a natural SSynt-structure written in **n**-terms (Figure 11a, below). The linear break of the phrase on which violins is produced by a word-order rule that puts the phrase which violins in the first linear position in the sentence, i.e., extracts it (during the transition SSyntR \Rightarrow DMorphR). Note that such an extraction is not possible for a similar SSyntS of Figure 11b: the result *Which sonatas are these violins easy to play on? is ungrammatical and can be precluded by imposing all the necessary conditions on the extraction rule. The SSyntS of Figure 11b can be only realized with the WHword in situ: These violins are easy to play on WHICH sonatas? (See Hudson 1988b: 199ff and Kahane & Mel'čuk 1999 on the problem of extraction with a **D**-framework.)



Which violins are these sonatas easy to play on? These violins are easy to play on WHICH sonatas?

Figure 11: SSyntSs of sentences with and without extraction

5) As Nichols (1993) has shown, ellipses, i.e., constituent-reducing operations, can be conveniently characterized in terms of Synt-**1**. Thus, four languages studied by Nichols—Russian, Nunggubuyu (Australian), English, and Chechen-Ingush (North-Caucasian)—differ with respect to their preferences in the domain of constituent-reducing: Russian prefers to remove Synt-heads, Nunggubuyu does it more frequently with Synt-dependents, English removes both with equal ease, while Chechen-Ingush does neither (which means that it has few ellipses). Cf. (42):
(42) a. Russian

A Maška emu po morde lit. ⁽And Mashka to-him on the-mug⁾. = ⁽And Mashka gave him a blow in the face⁾,

where the top node—a Synt-head, which is a verb meaning 'hit',—is elided.

b. Nunggubuyu *Anāgugu nā?galimaŋ; nā?galimaŋ* lit. ⁽[He] water fetched-for-him; fetched-for-him⁾. =

⁽He_i fetched water for him_i, he_i did⁾,

where the top node—a verbal Synt-head meaning 'fetch'—is repeated by the narrator for more expressivity, but with its dependent ('water') elided.

c. English

Leo is from Chernigovskij, and Alan from Paris,

where the top node—the Synt-head of the second conjunct clause (the verb BE)—is elided, or, more precisely, factored out;

or

Susan is fond of, while Marga looks askance at, profanity,

where the Synt-dependent of *fond of* is elided/factored out (Russian does not admit this type of dependent removal).

d. In Chechen-Ingush, the answer to the question 'What did he give his son?' must be *Sowyat dennad* lit. '[He a] gift gave', rather than simply **Sowyat* '[a] gift', which is the norm in the above three languages: Chechen-Ingush does not tolerate the removal of Synt-heads. Even the sentence meaning 'Good!/ OK!' must contain the verbal Synt-head: *Dika du!* lit. 'Good is'.

All the five phenomena mentioned above—valence, voice, restricted lexical cooccurrence, word order, and ellipses—are in the focus of a trend in modern linguistics that Hudson (1990) aptly dubbed *lexicalism*: putting at the center of the linguistic description facts about lexical units rather than facts about syntactic constructions, so that the lexicon is at last deservedly given a place of honor in linguistic studies; cf. as well Hudson 1983, 1984 and Mel'čuk 1995b. Stressing the importance of the lexicon goes quite well with the **D**-approach in syntax, because in this approach all the links are established between wordforms and based, in the final analysis, on their lexicographic properties.

4 Syntactic dependency and syntactic constituency

I am not offering here an in-depth comparison of **n**- and **c**-approaches in syntax (cf. Vennemann 1977, Nichols 1978, Hudson 1980a, b, Dahl 1980, Matthews 1981: 71-95, Mel'čuk 1988: 13-17, Sgall & Panevová 1988-89); I will, however, briefly touch upon two topics relevant to such a comparison: the concept of constituent in syntax and hybridization of **n**- and **c**-approaches.

4.1 Syntactic constituency

To compare the **D**-approach in syntax to the **C**-approach (also known as the Phrase-Structure approach), one needs to make precise the concept of syntactic constituent. Let me first take the simplest, or "naive," interpretation of constituent as a linearly ordered string of actual wordforms that shows a prosodic and semantic unity (i.e., a constituent \approx an actual phrase) and consider constituency exclusively as based on contiguity. Such constituents are not syntactic units in the sense that the Synt-structure of a sentence cannot be described in terms of these formations: they are linear, prosodic and morphological IMPLEMENTATIONS of (fragments of) the SSyntS, rather than part of it. The legitimate place of such constituents is in the Deep-Morphological structure of the sentence. (Cf. Langacker 1997 for a convincing discussion of the role constituency plays in language on the semantic and phonological levels, while it has no place on the syntactic level of sentence representation. According to Langacker, syntactic structure must ensure the correspondence between semantic and phonological constituents, especially in cases of unavoidable numerous and variegated mismatches, due to the strictly linear character of human speech, which has to convey utterly "non-linear" meaning. As a result, syntactic structure itself cannot be represented in terms of constituents.)

If, on the other hand, we consider the sophisticated concept of a constituent as a set of lexemes that go together, this set taken before linearization, prosodization and morphologization, then, in order for constituents to be able to carry relevant information about word order and inflection, each constituent has to have both its head and its constituent type specified. This means, more or less, indicating the type of the Synt-relation between the constituent's elements. But no sooner is this done than we have a **D**-representation! Or, to be more precise: a sophisticated *E*-representation carries all **D**information plus some other characteristics of the sentence represented. The question is then whether we need these extra data to be explicitly present in the Synt-structure of the sentence. The answer depends of course on our main theoretical postulates. I, for one, proceed from the postulate that every part of a linguistic representation must be as homogeneous and as compact as possible. In other words, phenomena of different nature should be represented in different components of the representation, so that a multilayered representation must be preferred over a unified one. If this postulate is accepted, then specifying the Synt-heads and the type of SyntRels between the sentence elements makes all other attributes of the **C**-approach redundant and therefore superfluous-particularly, 1) non-terminal nodes and 2) the categorization of constituents in the SyntS.

• Non-terminal nodes indicate the Synt-constituents, but, as I have said, the constituents can be computed from the **n**-representation and are needed only on a closer-to-surface level—on the DMorph-level; therefore, they should not be present in the Synt-structure.

• The categorization of the elements of a Synt-structure, i.e., the syntactic class and other syntactic features of lexemes, should not be part of the SyntS, either: this is not syntactic, but lexicographic information; as such it should remain behind the scenes, in the lexicon's entries for the lexemes involved.

So, if one follows the above postulate—that is, stops specifying non-terminal nodes and keeps lexicographic information in the lexicon, rather than in the SyntS, then nothing remains of the classical \mathbf{c} -approach in syntax.

To avoid misunderstandings, it would probably be worthwhile to formulate the following two provisos concerning the problem of constituency in the **D**-approach.

1. The **n**-approach does not negate the existence of constituents: constituents do of course exist and have the primary importance for any complete linguistic description. (I mean here constituents as real linguistic items—strings of wordforms with an appropriate prosody, not as formal abstract entities that are automatically specified by any dependency tree as projections of complete subtrees.) However, their place is not in the SyntS, but rather, as pointed out above, in the DMorphS of the sentence: in the SSyntS \Rightarrow DMorphS transition, linearization is carried out in terms of constituents that have to be computed from the SSyntS; and prosodization affects the constituents after they have been linearized.

2. The **D**-approach extensively uses standard subtrees (non-linearized and non-morphologized), which specify different constructions that behave identically from the viewpoint of surface syntax. Such are, for instance, Δ NUMP or Δ APPROX—sources of the numeral and numeral-approximate phrases: *three kilos* ~ *about/over three kilos* ~ *more than three kilos* ~ *from three to ten kilos* ~ ... (Mel'čuk & Pertsov 1987: 487-489). Another example is Δ V, or *the verbal nucleus*: a chain consisting of verbs and some special non-verbal elements allowing for particular operations in which it participates as a whole (Kahane & Mel'čuk 1999). However, first, standard subtrees are not constituents; and second, their place is not in the SyntS of a sentence, but in the syntactic rules of the linguistic model, which identifies them in the SyntS and processes them as specified.

4.2 Crossing syntactic dependency with syntactic constituency

For many years, linguists have been talking about the integration of both approaches that is, they have been looking for a hybrid between \mathbf{n} - and \mathbf{c} -representations to be used in syntax (e.g., Baumgärtner 1970 and Vennemann 1977). The incentive for such an integration comes primarily from the problems related to representing coordination in the \mathbf{n} -approach (see Section 5 of this part), as well as to some other linguistic phenomena such as:

• extraction (I know which girl you told my wife Alan was going out with, the extracted component being boldfaced; extraction happens under focusing, relativization, or interrogation);

• analytical forms (verbal and nominal, i.e. AUX \rightarrow V and DET \leftarrow N: has been detected; the book);

• *idioms* and *collocations* (among others, with what is known as *light verbs*: *make headway, pay a visit, launch an attack*, Ger. *zur Aufführung bringen* lit. 'to-the carrying-out bring' = 'carry out'), and the like.

The main idea is to introduce for any of these syntactic constructions a special type of subtree that is allowed to occupy as a whole one node of a dependency tree. In this way, the linguist tries to capture the intuition that such a set of wordforms depends on or governs other wordforms as a unit. The first full-fledged specific proposal for a mixed **D**-/**C**-representation of this type—by means of what is known as *syntactic groups*—was advanced in Gladkij 1966, 1968 and 1985. A similar device is put forward in Lobin 1993: 42ff and 1995 (under the name of *complex elements*). The most recent move in this direction is, as far as I know, Kahane 1997, where the concept of *bubble* is introduced: a subset of nodes of a **D**-tree which is allowed to be treated as a node, while having inside a completely specified **D**-structure of its own, including other bubbles. It is natural that linguists feel the need for some formation of this type; but to what extent such a hybridization, or rather, extension or enrichment, of **D**-trees is welcome remains to be seen.

In particular, using multistructural and multilevel representations (cf. I-3, p. 5) allows for elegant solutions of many problems that otherwise have to be treated via bubble-like entities. Thus, the difficulties of representation related to various extractions can be overcome in a natural way by recourse to the Communicative Structure (Kahane & Mel'čuk 1999). Similarly, the special character of AUX + V or DET + N phrases, as well as of idiomatic phrases (= idioms) such as *with respect to* or *kick the bucket*, is reflected by the fact that in the DSyntS all these phrases are represented each by one single node. Collocations are described, as pointed out in III-3, p. 84, in terms of lexical functions, and such a description makes explicit the specific character of the collocations. For instance, in the DSyntS, a phrase such as *pay a visit* or *do a favor* is represented as

Oper₁ o-**II** \rightarrow o VISIT or **Oper**₁ o-**II** \rightarrow o FAVOR

Here, $Oper_1$ is the symbol of a lexical function which specifies for a deverbal noun the support verb that takes this noun as a DirO and links it to the subject; $Oper_1$'s values, as those of the other LFs, are given in the lexical entries for nouns:

Oper₁(VISIT) = pay [ART ~] **Oper**₁(FAVOR) = do [ART ~]

For the time being, I believe that more progress is needed in the domain of the **D**-approach to syntax before we can determine where and how to use this or that element of the **C**-approach within the **D**-framework. However, what is already clear is that a single **D**-tree is not sufficient to represent all the information that might be necessary at the syntactic level. The linguistic model I propose uses two **D**-trees—namely, the Deep- and Surface-SyntS; in addition, it has recourse to a separate Communicative Structure. And that is not all: in some specific cases, more special machinery is used (groupings, see immediately below).

5 Insufficiency of syntactic dependency: Coordination

If we agree to use two levels of representation for syntactic structures, that is, the DSyntS and SSyntS, plus Communicative Structures on both levels, then a pure **D**-re-

presentation in syntax seems to be sufficient for all syntactic phenomena, except for one type of construction, and that is in the domain of *coordination* (cf. Hudson 1990: 97*ff* and Lobin 1993 on a special place coordination occupies with respect to Synt-**1**; see as well Stassen 2000 on the typology of coordination). The problem arises because the following situation is possible:

A wordform w "relates" either to a whole conjoined phrase or just to its Synt-head alone, such that the two constructions are morphologically, linearly or prosodically distinct and have different meanings; however, within the strict **D**-approach, both types of structure can be shown only by the direct Synt-**D** of w on the Synt-head of the conjoined phrase (the pure **D**-formalism does not allow for the dependency on a phrase as a whole). Consequently, one SSyntS written in terms of Synt-**D**s corresponds in such cases to two different meanings, which is not admissible.

Here are two examples:

- (43) a. The SyntS *old* \leftarrow *men* \rightarrow *and* \rightarrow *women* expresses two meanings, that is
 - either a meaning that corresponds to a joint reading [the adjective bears on both nouns]:

 $(old {men + women}) = (old men) and (old women);$

 — or a meaning that corresponds to a disjoint reading [the adjective bears on one noun only]:

({old men} + women' [the women are not necessarily old].

- b. The SyntS [*He is*] not *tall* and *fat* also expresses two different meanings:
- either '[he is] not {tall + fat}' [joint reading: he is neither tall nor fat];
- or ([he is] {not tall} + {fat}) [disjoint reading: he is not tall and is fat].²⁷

In such cases, different surface implementations that formally distinguish intended meanings are in principle available (depending on the language and particular lexical means used; "||" stands for a pause):

in (43a), old men and women [without a pause] vs. old men || and women;

in (43 b), He is not || tall and fat vs. He is not tall || and fat.

The semantic contrast accompanied by a formal contrast requires that the semantic distinction be maintained in the SSyntS (cf. Criterion C1, p. 34). The only way to do so sticking to pure Synt-**n**, that is, using exclusively **n**-formalism and without admitting multiple Synt-dependencies—seems to be to label differently the SSyntRels involved, i.e., to have in (43a) something like

 $old \leftarrow modif - men \rightarrow and \rightarrow women$ for the disjoint reading and $old \leftarrow coord - modif - men \rightarrow and \rightarrow women$ for the joint reading.

²⁷ I take the negative particle NOT in this example to be a Synt-dependent of TALL rather than of BE; cf. *He is, as everybody knows since the period when ..., not* \rightarrow *tall and fat* vs. *He is* \leftarrow *not* (*isn't*), *as everybody knows since the period when ..., tall and fat*.

However, this solution is no good. First, it is not natural enough linguistically; among other things, it entails doubling all SSyntRels that can link Synt-dependents to conjoined phrases. Second, it is not sufficient formally: it cannot help in the case of more than two conjuncts, such as in *hungry men, and women, and children* 'hungry {men, women and children}' vs. '{hungry {men and women}} and children'. Therefore, a real alternative is to complement the pure **D**-approach with *groupings*—specification, within the SyntS, of the **D**-subtrees relevant in such cases (Mel'čuk 1974: 214-216, 1988: 28-33). For instance:

old←**modif**−*men*→*and*→*women* (without grouping)

stands for the disjoint reading ([{]{old men}} and women⁾). However,

 $old \leftarrow modif - [-men \rightarrow and \rightarrow women]$ (with a grouping indicated by square brackets) corresponds to the joint reading ('old men and old women').

For (43b), we will also write two different SSyntSs:

He is not \leftarrow restr-*tall* \rightarrow *and* \rightarrow *fat* (for the disjoint reading: 'He is {not tall} and fat') and

He is not \leftarrow restr- $[-tall \rightarrow and \rightarrow fat]$ (for the joint reading: 'He is not {tall and fat}').

Note that a grouping is not a constituent in the strict sense: there is no higher node to represent it as a whole (because a grouping is not a projection of a complete subtree), and it does not participate in **n**-links as such (because in a consistent **n**-approach, only single nodes do; this is in contrast to the approach advanced in Kahane 1997, where a configuration of nodes in a dependency tree—a bubble—can be treated as a bona fide node). As we see in the example

 $old \leftarrow modif - [-men \rightarrow and \rightarrow women],$ the branch "modif" leaves the node *men* within the grouping, but not the grouping as such.^[9]

I would like to mention two other syntactic phenomena where groupings in the SSyntS may be necessary.

• The first one is "layered," or recursive, modification:

expensive {*Japanese cars*} vs. *Japanese* {*expensive cars*}

Here, the linear order of adjectives is not arbitrary: it reflects the successive, or stepwise, inclusion of sets of objects on which the modifiers bear, and is thus semantically relevant. (The problem is again created by quasi-coordination, i.e., by co-subordination.) Under the **n**-approach, both expressions have the same SSyntS:



As a result, a semantic difference is lost. It is not, however, clear to me whether this difference should be accounted for in the SSyntS as such (then groupings are needed) or rather in the Syntactic-Communicative or Referential Structure (and then groupings in the SSyntS are avoided). For the time being, I prefer the second solution; if it is adopted, either we have to introduce a special Sem-Communicative opposition, which will represent the order of (sub)set inclusions, or—and this seems more natural—we have to use different referential indications (S. Kahane): in the first, but not the second, reading, the meaning 'Japanese cars' has a direct referent: a subset of (all) Japanese cars is characterized by the property of being expensive; the situation is inverse in the second reading: a subset of (all) expensive cars is characterized as being Japanese.

• The second phenomenon is modification of the type [a] typical American woman $\langle *an \ American \ typical \ woman \rangle$ or [a] former German minister $\langle *a \ German \ former \ minister \rangle$. These cases resemble the syntactic behavior of quantifiers, which also deserves a special study from the viewpoint of **D**-representation. However, it is possible to deal with the adjectives of the TYPICAL or FORMER type on the basis of their special lexicographic properties: they can be linked to their Synt-governor by the same SSyntRel as any other adjective, but their positioning is controlled by their lexicographic features. Therefore, in this case, groupings in the SSyntS are not necessary.

Acknowledgments

This paper was written in May-July 1998 during my stay in Germany under a Humboldt Foundation scholarship and has been reworked several times since. Its final version has profited a great deal from my joint work with L. Iordanskaja (Iordanskaja & Mel'čuk 2000): I have included in the present text several of our common results and findings. The paper has been read and commented upon by D. Beck, I. Bolshakov, N. Bröker, P. Hellwig, R. Hudson, L. Iomdin, L. Iordanskaja, S. Kahane, M. Kracht, H. Lobin, J. Milićević, E. Savvina, L. Wanner, and D. Weiss; N. Pertsov has caught several mistakes and incongruities at the very last moment. The remarks and objections of D. Beck and S. Kahane have been especially deep and constructive and have lead to many substantial changes in the text. I did my best to take into account all the suggestions of my friends and colleagues, and I am happy to extend my heartfelt gratitude to all of them, as well as to the Humboldt Foundation; the usual disclaimers apply, of course.

Endnotes

- [1] (p. 4) The notion of passive Synt(actic)-valence cannot be reduced to that of part of speech. First, passive Synt-valence characterizes not only lexemes, but phrases as well, to which I think the notion of part of speech is not applicable. Second, and more importantly, the passive Synt-valence of a lexeme L is determined, generally speaking, by the part of speech of L only partially: syntactic features of L play here a crucial role. That is, two lexemes of the same part of speech may have different passive Synt-valences because of their syntactic features (≈ subcategorization). Thus, nouns like MONTH, WEEK or DAY may appear in the duration construction with a verb (*work the whole month, travel day after day*, etc.), in which other nouns are impossible; this fact is expressed by the syntactic feature value "temp" assigned to such nouns. For more on syntactic features, see Mel'čuk & Pertsov 1987: 471*ff*.
- ^[2] (p. 29) Here are two more examples (for a detailed analysis of the construction in question, further examples and a bibliography, see Gaatone 1988):

un	vache	de garçon	'an impressive boy'
a-SG.MASC	impressive-SG.MASC	of boy[masc]-SG	
une	vache	de voiture	(an impressive car)
a-SG.FEM	impressive-SG.FEM	of car[FEM]-SG	
се	chouette	de garçon	'this nice boy'
this-SG.MASC	nice-SG.MASC	of boy[masc]-SG	
cette	chouette	de voiture	'this nice car'
this-SG.FEM	nice-SG.FEM	of car[FEM]-SG	

Some French expletive interjections can also appear, along with the abovementioned adjectives, as the SSynt-head of this construction: *cette nom d'un chien de machine* ⁽this darned machine⁾, *cette bon sang de Julie* ⁽this bloody Julie⁾, *ces sacré nom de Danois* ⁽these bloody Danes⁾, *ma nom de Dieu de parole d'honneur* ⁽my damned word of honor⁾, *ce putain de garçon* ⁽this bloody boy⁾, etc., where the SSynt-head is boldfaced. The construction has the SSyntS of the following form:

CE←synt–NOM D'UN CHIEN–synt→DE–synt→MACHINE

Note that the determiner agrees in gender and number with the qualified noun (in this case, MACHINE) rather than with its own Synt-governor—the head of the phrase (NOM D'UN CHIEN), which, unlike an adjective, cannot borrow the gender and number from the noun. A similar English construction (*a bitch of a problem*, '*Ulysses' is murder to read*, etc.) is analyzed in McCawley 1987.

Let it be emphasized that the construction illustrated in (13) is different from such constructions as *ce cochon de Polytte* 'this swine of Polytte', *l'imbécile de ton mari* 'the fool of your husband' or *ce fou de prof* lit. 'this madman of professor', where the head is a noun (it can be a nominalized adjective, but it is anyway a noun). In (13), the head adjective cannot be nominalized: **un drôle, *un chouette*, etc.

- [3] (p. 38) It is sometimes claimed that even actantial SSyntRels can be repeatable. The best-known example is probably the repeatability of the dir-obj SSyntRel in Kinyarwanda: it is said that in this language, a clause can have up to three DirOs of the same MV (Kimenyi 1980: 229; the presumed DirOs are boldfaced); cf.:
 - (i) $Umu+g\acute{o}re$ $\acute{a}+r+\acute{u}ubak+iish+iriz +a$ $\acute{a}b\acute{a}+ana$ umu+gabo inzu Class I woman I PRES build CAUS BENEF CONT II children I man house 'The woman, on behalf of the children, is making the man build the house'.

A detailed analysis of "repeated DirOs" in Kinyarwanda in Gary & Keenan 1977: 87-94 shows that indeed all of them possess the same relevant linguistic properties which set them off with respect to oblique objects: they passivize, reflexivize and relativize, they can be cross-referenced in the verb, etc. And yet, in our framework, all three of them cannot be considered DirOs, because they contrast semantically, that is, they violate our Criterion C1. The presumed **dir-obj** SSyntRel in Kinyarwanda has to be split into three different SSyntRels, which are, so to speak, the subtypes of an abstract SSyntRel: the **dir-obj** SSyntRel, the **caus-dir-obj** SSyntRel and the **benef-dir-obj** SSyntRel. In this way, the commonality of their important properties is explicitly shown. Similarly, in Sanskrit, two objects in the accusative (= "double accusatives") cannot be both DirOs, either:

(ii) *Tā* yajamānam vāc +aya +ti them-ACC sacrificer-ACC name CAUS IND.PRES.3SG ⁽[He] makes [the] sacrificer name them³.

 $T\bar{a}$ is a DirO, but *yajamānam* must be described by a different SSyntRel: the **caus-dir-obj** one.

In other languages (e.g., Latin, German and Serbian) the situation with double accusatives is even clearer:

(iii)

- a. Lat. Quis music+am [ACC] ←?-docuit-?→Epaminond+am [ACC]?
 lit. 'Who taught Epaminondas music?' = 'Who taught music to Epaminondas?' or
 Me [ACC] ←?-rogavit-?→sententi+am [ACC]
 lit. '[He] asked me opinion' = 'He asked me for my opinion'.
- b. Serb. *Ta slika* m+e [ACC] \leftarrow ?-*košta*-?-*hiljad*+u [ACC] *dolara* lit. ^{(This} painting costs me one thousand dollars².
- c. Ger. *Was* [ACC] \leftarrow ?-*fragst* [*du*]-? \rightarrow *mich* [ACC]? 'What are you asking me?'

Neither of these sentences has two DirOs: the two accusatives do not display the same syntactic behavior. Thus, in (iii-b) *me* is omissible, while *hiljadu* is not: *Ta slika košta hiljadu dolara* vs. **Ta slika me košta*. This shows that *me* is here an IndirO, despite its accusative form. In (iii-c), only *mich* is the DirO, *was* being an OblO; etc. For double accusatives in German, see Wilkinson 1983.

For more on multiple objects in Latin, Ancient Greek and Modern Hebrew, see Lazard 1994: 89-96.

An interesting case of double accusatives is found in Korean (O'Grady 1991):

(iv) Kay +ka John+ul son +ul mwul+ess +ta dog SUBJ(ective) ACC hand ACC bite PAST DECL(arative) (The dog bit John's hand).

There can even be multiple double accusatives:

(v) John+i mwune+lul tali+lul kkuthpwupwun+ul cokum+ul SUBJ octopus ACC leg ACC endpart ACC bit ACC cal+la +ta cut PAST DECL
^{(J}John cut the octopus on the end part of the leg a bit²).

However, only the first one in such a chain of accusatives is a genuine DirO (O'Grady 1991: 74-75, 77-78); all the others show special behavior—thus, they cannot be linearly permuted with the DirO, cf. (iv) vs. $*Kay+ka \ son+ul \ John+ul \ mwulessta$. All the accusative nouns that follow the DirO are—from the semantic stand—elaborations introducing additional details (\approx 'The dog bit John, more precisely, on the leg³); syntactically, they seem to be quasi-coordinate conjuncts similar to what we see in English *He lives in the USA, in New York, on Manhattan, on 47th Street, in a luxurious building*.

^[4] (p. 40) Along with Criteria C1-C3, the researcher can use the following heuristic test:

Coordinability with one SSynt-governor

Within a coordinated phrase D_1 -coord $\rightarrow D_2$ which is subordinated as a whole to a SSynt-governor G, each element must in principle be able to bear the same SSyntRel **r** to G:

if, in language L, $G-r \rightarrow D_1$ -coord $\rightarrow D_2$, then it is possible that $G-r \rightarrow D_1$ and $G-r \rightarrow D_2$.

Examples

(i) French

a. Il craint d'être découvert et que l'administration le punisse

lit. 'He fears to be discovered and that the administration punish him'.

b. *Il veut partir et aussi que je parte avec lui* lit. ⁽He wants to leave and also that I leave with him⁾.

c. *Le rendement augmente successivement et par degré* lit. ^{(The yield rises successively and by degrees⁾.}

In (i), the boldfaced phrases stand in the same SSyntRel to the MV.

Unfortunately, this test cannot be raised to the rank of a genuine formal criterion: coordination—at any rate, in many languages—is strongly semantically motivated; therefore, in some cases, syntactically different clause elements can be coordinated, while in some other cases identical clause elements cannot. Here are a few examples.

Coordination of different clause elements (cf. Grevisse 1993: 371)

(ii) French

- a. *Elle vieillissait dans l'aisance et entourée de considération* (She was aging in well-being and surrounded with consideration).
- b. augmentation successive et par degré lit. ⁽[a] rise successive and by degree².
- c. *Je me demande si et sous quelles conditions on pourra régler le problème* ⁽I ask myself whether and under what conditions it will be possible to solve the problem⁾.

(iii) (Defrancq 1998: 118-119)

a. *Je me demande qui travaille et où* 'I ask myself who is-working and where'.
b. *Je me demande qui décide et quoi* 'I ask myself who decides and what'.

(iv)

couper les cheveux très court et de façon à ce qu'ils ne lui tombent pas sur le front 'cut the hair very short and in such a way that it would not fall on his forehead'.

Other examples can be drawn from Russian:

(v)

a. Èto otkrytie bylo sdelano v Anglii i angličaninom

lit. 'This discovery was made in England and by an Englishman'.

- b. Ja govorju s poètom i o poète lit. 'I talk with a poet and about a poet'.
- c. *Nikto, nikomu i nikogda ne pomogaet* lit. ^{(Nobody, to nobody and never helps⁾.}

(Russian coordinate constructions of this exotic type are described in detail in Sannikov 1989: 14-20.)

Impossible coordination of identical clause elements

(vi) French

- a. *Ils étaient cinq et très blonds lit. 'They were five and very blond'.
- b. *des plats français et exquis 'French and exquisite dishes'
- c. **Tout le monde préfère le repos maintenant et partir plus tard* ⁽Everybody prefers the rest now and to leave later⁾.

Consequently, the result of coordination test can serve as an argument in favor of or against a particular solution (especially in less obvious cases); but the test as such cannot be accepted as a rigorous criterion. Cf. the discussion of the role coordination plays in establishing grammatical relations in Sag *et al.* 1985 (*I am neither an authority on this subject nor trying to portray myself as one; Pat was awarded the prize and very upset about it*; and the like) and Hudson 1988a.

^[5] (p. 42) A particular syntactic or communicative role may require a noun in a particular inflectional form, for instance, 'DEF(inite)' or 'INDEF(inite)'; thus, in French, the bold-faced quasi-subject in the construction *Il est venu* 10 étudiants lit. 'It has come 10 students', may be only indefinite. D. Beck pointed out to me another interesting example: in Lushootseed, the negative predicate X^wIS 'be not' requires its actant to be in the subjunctive and have the hypothetical determiner k^wi:

 $x^{w}iS$ $k^{w}i$ $g^{w} + ad + s$ $+S \rightarrow Fard$ be.not DET **SUBJ** 2SG NOM(inalizer) eat lit. (Is-not your eating). = (You did not eat).

Yet, I think, in all such cases the Synt-governor directly requires a particular form of its dependent—rather than the presence of a particular dependent of its dependent.

^[6] (p. 47, (26a)) In Turkish, a zero copula cannot be postulated in the present based on paradigmatic considerations, as we have done for Russian. The main reason is that the forms in (26a) contain the marker of predicativity that precludes the use of the copula: in the past tense of the indicative, both the expression with the marker of predicativity but without copula and the expression with the copula I(*-mek*) ^(be) but without a marker of predicativity are possible (the latter being typical of colloquial speech, while the former is current in the written language):

Cocuk+tu +mandCocuki + di +mkidPAST 1SGkidbe PAST 1SG

both meaning 'I was a kid'. Note that the verb I(-mek) has no present tense.

[7] (p. 51) Languages also differ with respect to the meanings they allow to be coordinated. Cf. numerous examples of coordinate phrases in Latin which should be translated with subordinate phrases in French (Tesnière 1959: 315-316):

Lat. *orare atque obsecrare* ⁽pray and-also implore⁾ ~ Fr. *prier instamment* ⁽pray insistently⁾ Lat. *interdicit atque imperat* ⁽He forbids and-also orders⁾ ~

Fr. Il défend expressément 'He forbids expressly'

Lat. *diuellere ac distrahere* (separate and-also tear-apart) ~

Fr. séparer violamment (separate violently)

Lat. doctrina et ratio 'teaching and method' ~

Fr. *un enseignement méthodique* ⁽a methodical teaching)Lat. *studium et aures* ⁽favor and ears⁾ ~ Fr. *une oreille favorable* ⁽a favorable ear⁾

However, the study of the relationship between coordination and subordination falls outside the scope of this paper.

[8] (p. 72) Tesnière distinguishes Fr. *jonction*, which is coordination, from Fr. *connexion*, which is subordination. In Tesnière's trees (= *stemmas*) conjoined elements are linked horizontally, showing in this way their equal nature. Each of them is then subordinated to the same Synt-governor. Among other examples of how the coordination is represented, one finds in Tesnière 1959: 345 the following complex structure with parallel Synt-ns between the elements of coordinate phrases which themselves are not linked by Synt-ns:

Les maîtres, les pédagogues et les éducateurs donnent, répètent et ressassent des avis, des conseils et des avertissements aux écoliers, aux collégiens et aux lycéens (Teachers, pedagogues and educators give, repeat and trot out opinions, pieces of advice and warnings to the school kids, college students and high-school students).

The structure proposed by Tesnière is as follows (only a part of it is represented here):



MAÎTRES PÉDAGOGUES ÉDUCATEURS

^[9] (p. 94)

Groupings in the D-Syntax

Three remarks concerning relevant aspects of groupings seem in order.

• Dependence on the head of a coordinate string *vs*. dependence on the whole string.

These two cases of dependency are distinguished in the proposed SSyntS with groupings by including into a grouping all "personal" (= disjoint) dependents of its head. Thus, for

(old {{fat men} and women}),

where 'old' bears on the whole conjoined string (= joint reading), but 'fat' on 'men' only (= disjoint reading), we write

• Dependence of the head of a coordinate string vs. dependence of the whole string.

Here again, groupings allow for efficient disambiguation. Let us consider the following French example (Abeillé 1997: 19): *Paul rêvait d'acheter et collectionner des pistolets anglais* ^(P). was dreaming of buying and collecting English pistols⁾. The boldfaced conjoined string of infinitives depends on the verb RÊVER ^(dream) as a whole—it has a shared DirO *pistolets anglais* ^{(English pistols⁾; therefore, the preposition DE that introduces the infinitive need not to be repeated (= joint reading). However, if the two conjoined infinitives do not depend as a whole on RÊVER, the}

preposition has to be repeated: *Paul rêvait de voyager et de collectionner des pistolets anglais* ^(P). was dreaming of traveling and collecting English pistols⁾ $\langle *Paul rêvait de voyager et collectionner des pistolets anglais \rangle$ (= disjoint reading). This difference is readily expressed using groupings: for the joint reading, we write

 $r\hat{e}vait-[\rightarrow acheter \rightarrow et \rightarrow collectionner-] \rightarrow des pistolets anglais, and for the disjoint one,$

rêvait→voyager→et→collectionner→des pistolets anglais.

• Multiple coordinate conjunctions depending on the same Governor.

This is another problematic case for pure dependency. Thus, consider the expressions (i) - (ii), where both conjunctions—AND and OR—syntactically depend on the same noun (MEN, in this case):

(i) {*men and women*} *or children* vs.

(ii) {men or children} and women

These expressions clearly have different meanings; however, in terms of pure dependency, both have the same SSyntS:



We thus see that a pure-dependency SSyntS is unable to preserve the intended meaning in cases of such a type. In order to distinguish (i) and (ii) in the SSyntS, we need groupings—and there is no other way to achieve this goal:

References

Abeillé, Anne 1997. "Fonction ou position objet?". Au gré des langues 11: 8-29.

Agel, Vilmos, Eichinger, Ludwig, Eroms, Hans-Werner, Hellwig, Peter, Heringer, Hans-Jürgen, & Lobin, Henning (eds.) 2003 *Dependency and Valency. An International Handbook of Contemporary Research*, vol. 1, Berlin/New York: W. de Gruyter.

Allen, Barbara, Gardiner, Donna & Frantz, Donald 1984. Noun Incorporation in Southern Tiwa. *International Journal of American Linguistics*, 50(3): 292-311.

Anderson, John 1977. On Case Grammar. Prolegomena to a Theory of Grammatical Relations. London: Croom Helm.

Anderson, John & Ewen, Colin 1987. *Principles of Dependency Phonology*. Cambridge: Cambridge University Press.

Anderson, Stephen 1992. A-Morphous Morphology. Cambridge: Cambridge University Press.

Apresjan, Jurij, Boguslavskij, Igor', Iomdin, Leonid *et al.* 1989. *Lingvističeskoe obespečenie sistemy ÈTAP-2* [Linguistic Software of the ETAP-2 System]. Moscow: Nauka.

Apresjan, Jurij, Boguslavskij, Igor', Iomdin, Leonid *et al.* 1992. *Lingvisitičeskij* processor dlja složnyx informacionnyx sistem [A Linguistic Processor for Complex Information Systems]. Moscow: Nauka.

Árnason, Kristján 1989. "Dependency Phonology". Linguistics 27: 319-339.

Baumgärtner, Klaus 1965. "Spracherklärung mit den Mitteln der Abhängigkeitsstruktur". *Beiträge zur Sprachkunde und Informationsverarbeitung* 5: 31-53.

Baumgärtner, Klaus 1970. "Konstituenz und Dependenz. Zur Integration der beiden grammatischen Prinzipien". In *Vorschläge zu einer strukturellen Grammatik des Deutschen*, H. Steger (ed.), 57-77. Darmstadt.

Bazell, Charles 1949. "Syntactic Relations and Linguistic Typology". *Cahiers F. de Saussure* 8: 5-20.

Beck, David 1997. "Theme, Rheme, and Communicative Structure in Lushootseed and Bella Coola". In *Recent Trends in Meaning-Text Theory*, L. Wanner (ed.), 93-135. Amsterdam/Philadelphia: John Benjamins.

Beck, David 2002. The Typology of Parts of Speech Systems: The Markedness of Adjectives. New York/London: Routledge.

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

Boguslavskaja, Ol'ga 1991. "Razdvoenie grammatičeskix kategorij čisla i soglasovatel'nogo klassa i aspektnye konstrukcii v dagestanskix jazykax" [Doubling of Grammatical Categories of Number and Agreement Class and Aspect Constructions in Daghestanian Languages]. In *Tipologija grammatičeskix kategorij*, 10-12. Leningrad: Nauka.

Boguslavskij, Igor' 1985. Issledovanija po sintaksičeskoj semantike [Studies in Syntactic Semantics]. Moskva: Nauka.

Boguslavskij, Igor' 1996. *Sfera dejstvija leksičeskix edinic* [Scope of Lexical Units]. Moskva: Škola «Jazyki russkoj kul'tury».

Bresnan, Joan (ed.) 1982. *The Mental Representation of Grammatical Relations*. Cambridge, MA/London: MIT Press.

Corbett, Greville 1993. "The Head of Russian Numeral Expressions". In Corbett *et al.* 1993: 11-35.

Corbett, Greville, Fraser, Norman, & McGlashan, Scott (eds.) 1993. *Heads in Grammatical Theory*. Cambridge: Cambridge University Press.

Dahl, Östen 1980. "Some Arguments for Higher Nodes in Syntax: A Reply to Hudson's *Constituency and Dependency*". *Linguistics* 18(5/6): 485-488.

Defrancq, Bart 1998. "Embedded Interrogatives and Coordination in French". *Syntaxis* 1: 113-127.

Dryer, Matthew 1989. "Plural Words". Linguistics 27: 865-895.

Durand, Jacques (ed.) 1986. *Dependency and Non-linear Phonology*. London/Dover, NH: Croom Helm.

Eichinger, Ludwig & Eroms, Hans-Werner (eds.) 1995. Dependenz und Valenz. Hamburg: Buske.

Engel, Ulrich 1977. *Syntax der deutschen Gegenwartssprache*. Berlin: Erich Schmidt. [3rd rev. ed.: 1994.]

Engel, Ulrich 1988. *Deutsche Grammatik*. Heidelberg: Julius Groos & Tokyo: Sansyusya.

Evans, Nicholas 1988. "Odd Topic Marking in Kayardild". In *Complex Sentence Construction in Australian Languages*, P. Austin (ed.), 219-266. Amsterdam/ Philadelphia: John Benjamins.

Fillmore, Charles 1968. "The Case for Case". In *Universals in Linguistic Theory*, E. Bach & R. Harms (eds.), 1-88. New York: Holt, Rinehart and Winston.

Fitialov, Sergej 1962. "O modelirovanii sintaksisa v strukturnoj lingvistike [Modelling Syntax in Structural Linguistics]". In *Problemy strukturnoj lingvistiki*, S. Šaumjan (ed.), 100-114. Moskva: Izdatel stvo Akademii Nauk SSSR.

Fitialov, Sergej 1968. "Ob èkvivalentnosti grammatik NS i grammatik zavisimostej [On the Equivalence of Phrase-Structure Grammars and Dependency Grammars]". In *Problemy strukturnoj lingvistiki-1967*, S. Šaumjan (ed.), 71-102. Moskva: Nauka.

Gaatone, David 1988. "Cette coquine de construction: remarques sur trois structures affectives du français". *Travaux de linguistique* 17: 159-176.

Gaifman, Haim 1965. "Dependency Systems and Phrase Structure Systems". *Information and Control* 8: 304-337.

Garde, Paul 1977. "Ordre linéaire et dépendance syntaxique : contribution à une typologie". *Bulletin de la Société de linguistique de Paris* 72(1): 1-19.

Gary, Judit & Keenan, Edward 1977. "On Collapsing Grammatical Relations in Universal Grammar". In *Grammatical Relations* [= *Syntax and Semantics*, 8], P. Cole & J. Sadock (eds.), 83-120. New York *etc.*: Academic Press.

Gladkij, Aleksej 1966. *Lekcii po matematičeskoj lingvistike dlja studentov NGU* [Lectures on Mathematical Linguistics for Students of the State University of Novosibirsk]. Novosibirsk. [French translation: *Leçons de linguistique mathématique*, fasc. 1, 1970, Paris.]

Gladkij, Aleksej 1968. "Ob opisanii sintaksičeskoj struktury predloženija [Describing the Syntactic Structure of Sentences]". *Computational Linguistics* [Budapest] 7: 21-44.

Gladkij, Aleksej 1985. *Sintaksičeskie struktury estestvennogo jazyka v avtomatizirovannyx sistemax obščenija* [Syntactic Structures of Natural Language in Automated Communication Systems]. Moskva: Nauka. Goralčíková, Alla 1973. "On One Type of Dependency Grammar". In *Functional Generative Grammar in Prague*, W. Klein & A. von Stechow (eds.), 64-81. Kronberg: Skriptor.

Grevisse, Maurice 1993. Le bon usage. Grammaire française [13^e édition]. Paris/ Louvain-la-Neuve: Duculot.

Hays, David 1960. *Basic Principles and Technical Variations in Sentence Structure Determination*. Santa Monica, CA. [Reprinted in: C. Cherry (ed.) 1961. *Information Theory*, 367-374, London: Butterworth.]

Hays, David 1964a. "Connectability Calculations, Syntactic Functions and Russian Syntax". *Mechanical Translation* 8: 32-51. [Reprinted in: D. Hays (ed.) 1966. *Readings in Automatic Language Processing*, 107-125, New York: American Elsevier.]

Hays, David 1964b. "Dependency Theory: A Formalism and some Observations". *Language* 40(4): 511-525.

Helbig, Gerhard 1992. Probleme der Valenz- und Kasustheorie. Tübingen: Max Niemeyer.

Heringer, Hans Jürgen 1970. "Einige Ergebnisse und Probleme der Dependenzgrammatik". *Der Deutschunterricht* 4: 42-98.

Heringer, Hans Jürgen 1996. Deutsche Syntax Dependentiell. Tübingen: Stafenburg.

Hewson, John 1991. "Determiners as Heads". Cognitive Linguistics 2(4): 317-337.

Hewson, John 1992. Review Article of R. Hudson, "English Word Grammar" and R. Langacker, "Concept, Image, and Symbol. The Cognitive Basis of Grammar". *Canadian Journal of Linguistics* 37(1): 41-53.

Hudson, Richard 1976. Arguments for a Non-transformational Grammar. Chicago: University of Chicago Press.

Hudson, Richard 1980a. "Constituency and Dependency". *Linguistics* 18(3/4): 179-198.

Hudson, Richard 1980b. "A Second Attack on Constituency: A Reply to Dahl". *Linguistics* 18(5/6): 489-504.

Hudson, Richard 1983. "Word Grammar". In *Proceedings of the XIIIth International Congress of Linguistics, Aug. 29-Sept. 4, 1982, Tokyo*, Sh. Hattori & K. Inoue (eds.), 89-101. Tokyo.

Hudson, Richard 1984. Word Grammar. Oxford: Oxford University Press.

Hudson, Richard 1987. "Zwicky on Heads". Journal of Linguistics 23: 109-132.

Hudson, Richard 1988a. "Coordination and Grammatical Relations". Journal of Linguistics 24: 303-342.

Hudson, Richard 1988b. "Extraction and Grammatical Relations". Lingua 76: 177-208.

Hudson, Richard 1990. English Word Grammar. Oxford: Basil Blackwell.

Hudson, Richard 1993a. "Do We Have Heads in Our Minds?" In Corbett et al. 1993: 266-291.

Hudson, Richard 1993b. "Recent Developments in Dependency Theory". In Jacobs et al. 1993: 329-338.

Iordanskaja, Lidija 1963. "O nekotoryx svojstvax pravil'noj sintaksičeskoj struktury (na materiale russkogo jazyka) [Some Properties of Well-Formed Syntactic Structures of Sentences [Illustrated with Russian Data]". *Voprosy Jazykoznanija* No. 4: 102-112.

Iordanskaja, Lidija 1967. Avtomatičeskij sintaksičeskij analiz. Tom II: Mežsegmentnyj sintaksičeskij analiz [Automatic Syntactic Analysis. Vol. II: Intersegmental Syntactic Analysis]. Novosibirsk: Nauka.

Iordanskaja, Lidija & Mel'čuk, Igor 2000. "The Notion of Surface-Syntactic Relation Revisited (Valence-Controlled Surface-Syntactic Relations in French)". In *Slovo v tekste i v slovare. Sbornik statej k semidesjatiletiju akademika Ju.D. Apresjana*, L. Iomdin & L. Krysin (eds.), 391-433. Moskva: Jazyki russkoj kul'tury.

Jacobs, Joachim, von Stechow, Arnim, Sternefeld, Wofgang, & Vennemann, Theo (eds.) 1993. *Syntax. An International Handbook of Contemporary Research. Vol. 1.* Berlin/New York: W. de Gruyter.

Kahane, Sylvain 1997. "Bubble Trees and Syntactic Representations". In *Proceedings* of the 5th Meeting of Mathematics of Language (MOL 5), T. Becker & H.-U. Krieger (eds.), 70-76. Saarbrücken.

Kahane, Sylvain & Mel'čuk, Igor 1999. "Synthèse des phrases à extraction en français contemporain". *t.a.l.* 40(2): 25-85.

Keenan, Edward 1974. "The Functional Principle: Generalizing the Notion of 'Subjectof". *Papers from the Tenth Regional Meeting of Chicago Linguistic Society*: 298-310.

Keenan, Edward 1978. "Logical Semantics and Universal Grammar". *Theoretical Linguistics* 5: 83-107.

Kibrik, Aleksandr 1977. "O sootnošenii ponjatija sintaksičeskogo podčinenija s ponjatijami soglasovanija, upravlenija i primykanija [The Correlation between the Notion of Syntactic Subordination and the Notions of Agreement, Government, and Adjunction]". In *Problemy teoretičeskoj i eksperimental noj lingvistiki*, 161-179. Moskva: Izd-vo MGU. [Reprinted in: A. Kibrik, *Očerki po obščim i prikladnym voprosam jazykoznanija (universal noe, tipovoe i specifičeskoe v jazyke)*, 1992, Moskva: Izd-vo MGU, 102-123.]

Kim, Alan 1995. "Word Order at the Noun Phrase Level in Japanese: Quantifier Constructions and Discourse Functions". In *Word Order in Discourse*, P. Downing & M. Noonan (eds.), 199-246. Amsterdam/Philadelphia: John Benjamins.

Kimenyi, Alexander 1980. *Relational Grammar of Kinyarwanda*. Berkeley, CA: UCLA Press.

Koch, Peter & Krefeld, Thomas (eds.) 1991. Connexiones Romanicae. Dependenz und Valenz in romanischen Sprachen. Tübingen: Max Niemeyer.

Korhonen, Jarmo 1977. Studien zu Dependenz, Valenz und Satzmodell. Teil I. Theorie und Praxis der Beschreibung der deutschen Gegenwartssprache. Bern etc.: Peter Lang.

Kunze, Jürgen 1972. "Die Komponenten der Darstellung syntaktischer Strukturen in einer Abhängigkeitsgrammatik". *Prague Bulletin of Mathematical Linguistics* 18: 15-27.

Kunze, Jürgen 1975. Abhängigkeitsgrammatik. Berlin: Akademie-Verlag.

Kunze, Jürgen & Priess, W. 1967-1971. "Versuch eines objektivierten Grammatikmodells". *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 20: 415-448, 21: 421-466, 23: 347-378, 24: 373-402.

Langacker, Ronald 1987. Foundations of Cognitive Grammar, vol. 1. Theoretical Prerequisites. Stanford, CA: Stanford University Press.

Langacker, Ronald 1991. Foundations of Cognitive Grammar, vol. 2. Descriptive Application. Stanford, CA: Stanford University Press.

Langacker, Ronald 1997. "Constituency, Dependency, and Conceptual Grouping". *Cognitive Linguistics* 8(1): 1-32.

Lazard, Gilbert 1994. L'actance. Paris: Les Presses Universitaires de France.

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

Lehmann, Christian 1985. "On Grammatical Relationality". *Folia Linguistica* 19: 67-109.

Lobin, Henning 1993. Koordinationssyntax als prozedurales Phänomen. Tübingen: G. Narr.

Lobin, Henning 1995. "Komplexe Elemente – Indizien aus Nominalphrase und Verbalkomplex". In Eichinger/Eroms (eds.) 1995: 117-133,

McCawley, James 1987. "A Case of Syntactic Mimicry". In *Functionalism in Linguistics*, R. Dirven & V. Fried (eds.), 459-470. Amsterdam/Philadelphia: John Benjamins.

Machová, Svatava 1975. Die Abhängigkeitsgrammatiken. In: *Einführung in die generative Grammatik (Prager Autorengruppe)*, 146-154. Kronberg: Skriptor.

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

Marcus, Solomon 1965b. "Sur une description axiomatique des liens syntaxiques". Zeitschrift für mathematische Logik und Grundlagen der Mathematik 11: 291-296.

Matthews, Peter 1981. Syntax. Cambridge etc.: Cambridge University Press.

Maxwell, Dan & Schubert, Klaus (eds.) 1989. *Metataxis in Practice. Dependency Syntax for Multilingual Machine Translation*. Dordrecht/Providence, RI: Foris.

Mel'čuk, Igor' 1962. "Ob algoritme sintaksičeskogo analiza jazykovyx tekstov (obščie principy i nekotorye itogi) [An Algorithm for Syntactic Analysis of Texts (General Principles and Some Results]". *Mašinnyj perevod i prikladnaja lingvistika* 7: 45-87.

Mel'čuk, Igor' 1963. "Avtomatičeskij analiz tekstov (na materiale russkogo jazyka) [Automatic Text Analysis (for Russian)]". In *Slavjanskoe jazykoznanie*, 477-509. Moskva: Nauka.

Mel'čuk, Igor' 1964a. "Tipy svjazej meždu èlementami teksta i tipologija jazykov [Possible Types of Links between Text Elements and Linguistic Typology]". In *Materialy konferencii «Aktual'nye voprosy sovremennogo jazykoznanija i lingvističeskoe nasledie E.D. Polivanova»*, L.I. Rojzenzon (ed.), tom I, 57-59. Samarkand: Samarkandskij Universitet.

106

Mel'čuk, Igor' 1964b. Avtomatičeskij sintaksičeskij analiz. Tom I: Obščie principy. Vnutrisegmentnyj sintaksičeskij analiz [Automatic Syntactic Analysis. Vol. I: General Principles. Intrasegmental Syntactic Analysis]. Novosibirsk: Nauka.

Mel'čuk, Igor 1967. "L'ordre des mots dans la synthèse automatique du texte russe". *T.A. Informations* 2: 65-84.

Mel'čuk, Igor' 1974. *Opyt teorii lingvističeskix modelej "Smysl ⇔ Tekst"* [Outline of a Theory of 'Meaning-Text' Linguistic Models]. Moskva: Nauka. [Reprinted: 1999, Moskva: Škola «Jazyki russkoj kul'tury».]

Mel'čuk, Igor' 1977. "O tipax poverxnostno-sintaksičeskix otnošenij (tri kriterija različenija) [On Possible Types of Surface-Syntactic Relations (Three Criteria for Distinguishing Them)]". *Russian Linguistics* 3(3-4): 245-270.

Mel'čuk, Igor 1979. Studies in Dependency Syntax. Ann Arbor, MI: Karoma.

Mel'čuk, Igor 1981. "Types de dépendance syntagmatique entre les mots-formes d'une phrase". *Bulletin de la Société de linguistique de Paris* 76(1): 1-59.

Mel'čuk, Igor 1985. *Poverxnostnyj sintaksis russkix čislovyx vyraženij* [Surface Syntax of Russian Numeral Expressions]. Wien: Wiener Slawistischer Almanach.

Mel'čuk, Igor 1988. Dependency Syntax. Albany, NY: SUNY Press.

Mel'čuk, Igor 1993. "Agreement, Government, Congruence". *Lingvisticae Investigationes* 17: 307-373.

Mel'čuk, Igor 1995a. *The Russian Language in the Meaning-Text Perspective*. Moscow: Škola «Jazyki russkoj kul'tury» & Vienna: Wiener Slawistischer Almanach.

Mel'čuk, Igor 1995b. "The Future of the Lexicon in Linguistic Description and the Explanatory Combinatorial Dictionary". In *Linguistics in the Morning Calm 3 (Selected Papers from SiICOL-1992)*, Ik-Hwan Lee (ed.), 181-270. Seoul: Hanshin.

Mel'čuk, Igor 1996a. "Lexical Functions: A Tool for the Description of Lexical Relations in the Lexicon". In *Lexical Functions in Lexicography and Natural Language Processing*, L. Wanner (ed.), 37-102. Amsterdam/Philadelphia: John Benjamins.

Mel'čuk, Igor 1996b. *Cours de morphologie générale. Vol. 3. Troisième partie : Moyens morphologiques. Quatrième partie : Syntactiques morphologiques.* Montréal: Les Presses de l'Université de Montréal & Paris: CNRS.

Mel'čuk, Igor 1997a. "Cas grammaticaux, construction verbale de base et voix en massaï : vers une meilleure analyse de concepts". *Bulletin de la Société de linguistique de Paris* 92(1): 49-113.

Mel'čuk, Igor 1997b. Vers une linguistique Sens-Texte. Leçon inaugurale. Paris: Collège de France.

Mel'čuk, Igor 2001. "Language: Dependency". In *International Encyclopedia of the Social and Behavioral Sciences*, N.J. Smelser & P.B. Baltes (eds.), vol. 12, Amsterdam *etc.*: Elsevier, 8336-8344.

Mel'čuk, Igor 2003. "Levels of Dependency in Linguistic Description: Concepts and Problems". In Agel *et al.* 2003: 188-229.

Mel'čuk, Igor 2004. "Actants in Semantics and Syntax, I/II". *Linguistics*, 42(1): 1-66, 42(2): 247-291.

Mel'čuk, Igor 2006. *Aspects of the Theory of Morphology*. Berlin/New York: Mouton de Gruyter.

Mel'čuk, Igor & Pertsov, Nikolaj 1987. Surface Syntax of English. A Formal Model within the Meaning-Text Framework. Amsterdam/Philadelphia: John Benjamins.

Nichols, Johanna 1978. "Double Dependency?". *Papers from the Fourteenth Regional Meeting of Chicago Linguistic Society*: 326-339.

Nichols, Johanna 1986. "Head-marking and Dependent-marking Grammar". *Language* 62(1): 56-119.

Nichols, Johanna 1993. "Heads in Discourse: Structural Versus Functional Centricity". In Corbett *et al.* 1993: 164-185.

O'Grady, William 1991. *Categories and Case. The Sentence Structure of Korean*. Amsterdam/Philadelphia: John Benjamins.

Owens, Jonathan 1988. The Foundations of Grammar: An Introduction to Mediaeval Arabic Grammatical Theory. Amsterdam/Philadelphia: John Benjamins.

Padučeva, Elena 1964. "O sposobax predstavlenija sintaksičeskoj struktury predloženija [Techniques for Representing the Syntactic Structure of a Sentence]". *Voprosy Jazykoznanija* No. 2: 99-113.

Perlmutter, David (ed.) 1983. *Studies in Relational Grammar I*. Chicago/London: The University of Chicago Press.

Petkevič, Vladimír 1995. "A New Formal Specification of Underlying Structures". *Theoretical Linguistics* 21(1): 7-61.

Pittman, Richard 1948. "Nuclear Structures in Linguistics". Language 24(2): 287-202.

Quirk, Randolph, Greenbaum, Sidney, Leech, Geoffrey, & Svartvik, Jan 1985. *A Comprehensive Grammar of English Language*. London/New York: Longman.

Robinson, Jane 1970. "Dependency Structures and Transformational Rules". *Language* 46(2): 259-285.

Sag, Ivan, Gazdar, Gerald, Wasow, Thomas, & Weisler, Steven 1985. "Coordination and How to Distinguish Categories". *Natural Language and Linguistic Theory* 3(2): 117-171.

Sannikov, Vladimir 1989. *Russkie sočiniteľ nye konstrukcii: semantika, pragmatika, sintaksis* [Russian Coordinate Constructions: Semantics, Pragmatics, Syntax]. Moskva: Nauka.

Savvina, Elena 1976. "Fragment modeli russkogo poverxnostnogo sintaksisa. III. Sravnitel'nye konstrukcii (sravnitel'nye i sojuznye sintagmy) [A Fragment of a Model of Russian Surface Syntax. III. Comparative constructions (comparative and conjunctional phrases)]". *Naučno-texničeskaja informacija*, serija 2, No. 1: 38-43.

Schmidt, Peter & Lehfeldt, Werner 1995. Kongruenz, Rektion, Adjunktion. Semantische und historische Untersuchungen zur allgemeinen Morphosyntax und zu den Wortfügungen (slovosočetanija) im Russischen. München: Sagner. Schubert, Klaus 1987. *Metataxis. Contrastive Dependency Syntax for Machine Translation.* Dordrecht/Providence, RI: Foris. [Rev.: B. Sigurd, *Studia Linguistica*, 1988, 42(2): 181-184.]

Sgall, Petr, Nebeský, Ladislav, Goralčíková, Alla, & Hajičová, Eva 1969. *A Functional Approach to Syntax in Generative Description of Language*. New York: American Elsevier.

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

Sgall, Petr & Panevová, Jarmila 1988-89. "Dependency Syntax—A Challenge". *Theoretical Linguistics* 15(1): 73-86.

Sgall, Petr, Pfeiffer, Oskar, Dressler, Wolfgang, & Puček, M. 1995. "Experimental Research on Systemic Ordering". *Theoretical Linguistics* 21(2/3): 197-239.

Somers, Harold 1984. "On the Validity of the Complement-Adjunct Distinction in Valency Grammar". *Linguistics* 22(4): 507-530.

Somers, Harold 1987. *Valency and Case in Computational Linguistics*. Edinburgh: Edinburgh University Press.

Starosta, Stanley 1988. The Case for Lexicase. London: Pinter.

Stassen, Leon 2000. "AND-Languages and WITH-Languages". *Linguistic Typology* 4(1): 1-54.

Suñer, Margarita 1998. "Resumptive Restrictive Relatives: A Crosslinguistic Perspective". *Language* 74(2): 335-364.

Tarvainen, Kalevi 1981. Einführung in die Dependenzgrammatik. Tübingen: Max Niemeyer.

Tesnière, Lucien 1934. "Comment construire une syntaxe". Bulletin de la Faculté des lettres, Université de Strasbourg 7: 219-229.

Tesnière, Lucien 1959. Éléments de syntaxe structurale. Paris: Klincksieck.

Trubetzkoy, Nikolaj 1939. "Le rapport entre le déterminé, le déterminant et le défini". In *Mélanges de linguistique offerts à Charles Bally*, 75-82. [Reprinted in: E. Hamp *et al.* (eds.), *Readings in linguistics*, 1966, II, Chicago: University of Chicago Press, 133-138.]

Van Langendonck, Willy 1994. "Determiners as Heads?" Cognitive Linguistics 5(3): 243-259.

Vennemann, Theo 1977. "Konstituenz und Dependenz in einigen neueren Grammatiktheorien". *Sprachwissenschaft* 2(3): 259-301.

Weber, Heinz 1992. Dependenzgrammatik. Ein Arbeitsbuch. Tübingen: G. Narr.

Wilkinson, Edwin 1983. "Indirect Object Advancement in German". *Proceedings of the Ninth Annual Meeting of the Berkeley Linguistics Society*, 281-291.

Xolodovič, Aleksandr 1966. "K tipologii porjadka slov [On Word Order Typology]". *Naučnye doklady vysšej školy — Filologičeskie nauki* 3. [Reprinted in: A.A. Xolodovič, 1979. *Problemy grammatičeskoj teorii*, Leningrad: Nauka, 255-268.] Zwicky, Arnold 1985. "Heads". Journal of Linguistics 21(1): 1-30.

Zwicky, Arnold 1993. "Heads, Bases and Functors". In Corbett et al. 1993: 292-315.

Žolkovskij, Aleksandr & Mel'čuk, Igor' 1967. "O semantičeskom sinteze [On Semantic Synthesis]". *Problemy kibernetiki* 19: 177-238. [There is a translation into French: *T.A. Informations* 2: 1-85, 1970.]

On the Status of Phrases in Head-Driven Phrase Structure Grammar: Illustration by a Fully Lexical Treatment of Extraction

Sylvain Kahane

1 Introduction

The main goal of this paper is a better understanding of the way linguistic units combine to make sentences and particularly of the role of phrases in this process. We will compare two families of approaches that have often been opposed (Hudson 1980, Mel'čuk 1988): dependency grammars (DG) and phrase structure grammars (PSG). DGs and PSGs differ by the status they give to phrases: DGs models syntax only in terms of combination of words and dependencies between words, while PSGs models syntax in terms of phrases, namely grouping of words. Nevertheless, most of contemporary PSGs—and among them the Head-driven PSG (HPSG; Pollard & Sag 1987, 1994; Sag & Wasow 1999) and all the grammars based on the X-bar Theory (Jackendoff 1977)—are very near to DGs: specifying a head in each phrase makes a PSG more or less equivalent to a DG (Lecerf 1961, Padučeva 1964, Gaifman 1965). However, some differences remain that we will explore. An important question arises immediately: what exactly is the role played by phrases in PSGs if PSGs are strongly equivalent to grammars that do not consider phrases?

To try to understand the status of phrases in PSG and particularly in HPSG (probably the most developed formalism for PSGs), we will focus on extraction-a phenomenon whose treatment in previous HPSG presentations (Pollard & Sag 1994, Sag 1997, Bouma et al. 2001) does not fit into the dependency grammar framework. In these HPSG accounts of extraction, wh-words are considered simply as markers and the essence of the construction is attributed to the phrase corresponding to the whole (relative or interrogative) clause. We do not deny that extraction phenomena involve particular constructions (in the sense of Goldberg 1995), but we do not think that these constructions justify using specific phrase descriptions (schemata, in HPSG terms) for each of them. It is still possible (and even simpler and more economical) to associate the description of the particularities of these constructions to their lexical markers, that is, not to use phrasal descriptions (and even not to consider phrases) and to consider more complex lexical descriptions. We will show how to implement in the HPSG formalism a dependency analysis of extraction, that is, an analysis that does not use phrasal descriptions but only combination of words. We will see that this analysis is descriptively equivalent to traditional analyses (and incorporates all refinements of such analyses) while appearing significantly more elegant.

KAHANE

It is important to emphasize that this study focuses on the syntax-semantics interface. Phrases are components of syntax proper, that is, the ordering of words and their grouping in phrases (see Gerdes & Kahane 2007 for a new characterization of phrases based on these ideas). Not considering phrasal descriptions in the syntax-semantics interface is not just a game. We defend a lexicalist approach, where the structure of a sentence (and more generally of a whole text) is completely determined by the combination of structural descriptions associated to the minimal morphological, syntactic and semantic units, that is, morphemes, lexemes and lexical units—as well as prosodic units, not considered here (Kahane 2007). In what follows, for simplicity's sake, we will not distinguish morphemes and lexical units, calling both *words*.

Our purpose is not to oppose HPSG and DG but, on the contrary, to show what both formalisms or theoretical frameworks can bring to each other. Our description of extraction does not fit "pure" DG, since it involves simultaneous combination of more than two words. We will show how the HPSG formalism allows us to model such phenomena in an elegant way. As a result, the HPSG presented here can be seen as a possible formalization of an extended DG.

The rest of the paper is structured as follows. Section 2 focuses on the core of the grammar (combinations of words with their actants and modifiers) and shows how to adjust an HPSG in order to simulate a DG. Section 3 discusses the treatment of complementizers and relativizers and the interpretation of the SLASH feature in DG terms. Section 4 is the main part of the paper; we propose a new description of extraction based on a double identity of *wh*-words and involving only word combinations. Section 5 is devoted to the status of phrases in HPSG; it also shows how the SLASH feature is linked to the underlying phrase structure adopted by HPSG.

This paper is addressed to people working in HPSG as well as those working in other frameworks, dependency grammars or not. For this reason, the presentation is more or less autonomous and does not really presuppose familiarity with HPSG or with DGs.

2 HPSG as a dependency grammar

In dependency grammars (Hays 1964, Mel'čuk 1988, Hudson 1990), a sentence is built directly by combining the words of the sentence, without resorting to strings of words such as phrases. The main conjecture of DGs is that combining the words of a sentence can be reduced to combining the words pairwise. Moreover, DGs assume that the combination of words is asymmetric: one word is the governor and the other is the dependent. The governor is roughly what is defined as syntactic head in PSG; the difference is that DGs consider only relations between words: a governor governs words, while a head heads a phrase. But if we take two words that form a phrase, the word that is the head of the phrase is the word that governs the other word. For instance, it is equivalent to say that *apples* is the head of the phrase big apples and that *apples* governs big in big apples.

In many respects, HPSG can be envisaged as a DG. Given that in most cases the description of a phrase is reduced to the description of its head, the combination of the

head of a phrase with a subphrase can be seen as the combination of two words: the head of the phrase and the head of the subphrase.

In HPSG, linguistic signs such as words are represented by feature structures. In this paper, focusing on the syntax-semantics interface, we do not mention features concerning phonology or word order. In order to be more understandable for HPSG users, we keep usual HPSG notations whenever possible.

Section 2.1 discusses the interpretation of HPSG word descriptions in terms of DGs; Section 2.2 links phrase descriptions and combination of words in a DG.

2.1 Word descriptions

The structural description of a word in a sentence is a typed-feature structure.¹ All words are of type *word*, which is a subtype of the type *grammatical-structure* (abbreviated to *gram-struc*). The relation 'isa' is the relation of inheritance on types: for instance, a feature structure of type *word* inherits all the descriptions of its supertypes and is then the unification of the description attached to the type *word* and the descriptions of its supertypes, that is, *gram-struc* and *feature-struc*, see (2).

(1)
$$gram-struc: \begin{bmatrix} HEAD \ cat \\ VAL \ bag(cat) \\ DEP \ bag(cat) \end{bmatrix}$$
 is a feature-struc
(2) $word: \begin{bmatrix} VAL \ /ebag \\ DEP \ ebag \end{bmatrix}$ is a gram-struc
(3) $word = \begin{bmatrix} HEAD \ cat \\ VAL \ /ebag \\ DEP \ ebag \end{bmatrix}$

The value of the HEAD feature of *word* is the description of the word itself; this description is a feature structure of type *cat*, described in (4).

The VAL(ence) feature contains the valence (= subcategorization frame, or government pattern) of the word, i.e, a set of slots to be filled by the syntactic actants of the word.

Terminological remark: We use the term *actant* proposed by Tesnière (1959) rather than the term *complement*. In particular, actants include the subject. Contrary to Chomskyan tradition (partially adopted by HPSG), we do not particularly distinguish the subject from other actants. Of course the subject is the most salient actant and it can have contrastive properties that oppose it to other actants. But we believe that these properties do not justify a special status for the subject any more than the properties of the direct object justify a special status for it among complements. The case of infinitives

¹ As said previously, we will not discuss a very important question concerning the units of the syntaxsemantics interface. Clearly, words are not these units. Words are the units of syntax proper, while the units of the syntax-semantics interface are lexical and grammatical units (Kahane 2007), which can correspond to a group of words (idioms), as well as to a part of a word (*went* = GO \oplus past).

is special: they do not realize their subject as their own syntactic actant but it can be realized somewhere else (for instance, as the subject of a raising verb). Therefore, the "subject" of an infinitive can be put in a special feature (we can call it, e.g., SUBJ or RAISING) whose behavior resembles that of SLASH rather than that of VAL. The same feature will be used for adjectives due to the parallelism between the constructions *Peter seems to sleep* and *Peter seems asleep*.

The value of VAL is a bag of feature structures of type *cat*(egory). A bag is a set whose elements can be repeated, that is, a list whose elements are not ordered. (HPSG uses lists traditionally. As we do not need order on the values of VAL and DEP, we think it is simpler to consider them as bags, but this is not a serious point.) The empty bag is noted *ebag*. The default value of VAL is empty. A default value is preceded by the symbol "/". It can be changed by a further assignment of the value.

The DEP(endence) feature of a word contains the bag of its dependents.² The DEP of a word is empty as long as this word has not combined with other words. When it combines with one of its actants, this actant fills an actant slot, that is, unifies with an element of VAL, which is then removed from VAL and put in DEP (see 2.2).

A *cat* is described as follows:

(4) $cat: \begin{bmatrix} FCT & syntrel \\ POS & pos \\ INFL & [...] \\ CONT & cont \end{bmatrix}$ is a feature-struc

The feature FCT indicates the syntactic function of the word. Its value is a subtype of the subtype *syntrel: root, subj, dobj, iobj, obl, mod*, etc.³

The feature POS indicates the part of speech of the word. Its value is a subtype of the atomic type *pos*: *verb-pos*, *noun-pos*, *adj-pos*, *adv-pos*, etc.

The feature INFL contains information about the inflection of the word, such as mood and tense for verbs or number for nouns. This topic is not in the focus of this paper, and INFL will be mentioned only when necessary.

The feature CONT specifies the semantic content of the word.⁴ For the time being we just mention that CONT contains the list of the word's semantic arguments as the value of the attribute ARGS:

(5) cont: [ARGS list(cont)]

² Bouma *et al.* 2001 introduces a feature DEPS (= dependents) for the dependents of a word. Contrary to their feature, our DEP feature is only filled when an element is removed from VAL and it does not contain slashed elements. Moreover, we prefer to use the term *dependence* (rather than *dependents*), which is more parallel to *valence*.

³ We make a distinction between syntactic functions and syntactic relations, depending on whether we are speaking about the syntactic function of a word (with respect to its governor) or the syntactic relation between two words. The edges of a dependency tree are labeled by syntactic relations.

⁴ In fact, we put in CONT only the semantic content of the radical. The semantic content brought by the inflection is put in INFL.

Note that we treat CONT as a head feature (see (12), p. 117), that is, the CONT value of a phrase is equal to the CONT value of its head word. But through ARGS, the CONT value of all words in a sentence points at other CONT values, giving us a graph, which is always connected. This graph is the semantic structure of sentences considered by Meaning-Text Theory (Žolkovskij & Mel'čuk 1967, Mel'čuk 1988); for semantic structures, see Mel'čuk in this volume, p. 5*ff*. In the HPSG tradition, the CONT value of a phrase is the set of the CONT values of its daughters. Nothing requires us to pass up and accumulate all the CONT values, just as the percolation of syntactic information is constrained by the Locality Principle (see Section 3.4).

We propose now some subtypes of the type *word*. Words are subdivided into *act*(ant) for *noun* and *verb* and *modfr* (= modifier) for *adj*(ective) and *adv*(erb).⁵ Modifiers are generally adjoined to their governor, which is their first semantic argument. They must contain in their description a description of their governor, which is the value of MOD.⁶ We also instantiate the feature MOD of actants by the atomic type *none* that cannot unify with a non-empty value. This ensures that actants cannot be used to modify another word (see *mod-val-computation* in Subsection 2.2).⁷

(6) a.
$$act$$
: $\begin{bmatrix} HEAD \begin{bmatrix} MOD & none \end{bmatrix} \end{bmatrix}$ is a word
b. $noun$: $\begin{bmatrix} HEAD \begin{bmatrix} POS & noun-pos \end{bmatrix} \end{bmatrix}$ is a act
c. $verb$: $\begin{bmatrix} HEAD \begin{bmatrix} POS & verb-pos \end{bmatrix} \end{bmatrix}$ is a act
d. $transf$ - $verb$: $\begin{bmatrix} HEAD \begin{bmatrix} CONT \begin{bmatrix} ARGS & (1, 2) \end{bmatrix} \end{bmatrix}$
VAL $\{ \begin{bmatrix} FCT & subj \\ POS & noun-pos \end{bmatrix}, \begin{bmatrix} FCT & dobj \\ POS & noun-pos \end{bmatrix} \}$ is a $verb$
(7) a. $modifr$: $\begin{bmatrix} HEAD \begin{bmatrix} FCT & mod \\ CONT \begin{bmatrix} ARGS & (1, ..., \rangle \end{bmatrix} \end{bmatrix}$ is a $word$
 $mod \begin{bmatrix} CONT \begin{bmatrix} 1 \end{bmatrix} \end{bmatrix}$ is a $word$
b. adj : $\begin{bmatrix} HEAD \begin{bmatrix} POS & adj-pos \\ MOD \begin{bmatrix} POS & noun-pos \end{bmatrix} \end{bmatrix}$ is a $modfr$
c. adv : $\begin{bmatrix} HEAD \begin{bmatrix} POS & adj-pos \\ MOD \begin{bmatrix} POS & noun-pos \end{bmatrix} \end{bmatrix}$ is a $modfr$

⁵ Words of type *verb* are necessarily finite verbs. Our treatment of non-finite verbs will be explained in Section 3.1.

⁶ Warning: it is not the value of HEAD but the value of MOD that is the head of the potential phrase containing the modifiers.

⁷ We leave the question of determiners aside. They could be put in the VAL or INFL features of the noun, or in a special feature—see the SPR feature of Sag & Wasow (1999).

The word descriptions can be seen as subtrees of the dependency tree. These subtrees can be represented as proposed in Nasr 1995; see also Kahane 2001. In such a representation, see (8), a black node corresponds to the current node (= the HEAD value), while a white node corresponds to a requirement, either a syntactic actant (= an element of VAL value) or the governor (= the MOD value). In our diagrams we use the labels N, V, Adv and Adj for feature structures *cat* with [POS *noun-pos*], [POS *verbpos*], etc.

(8) a. Dependency interpretation of *transf-verb*



b.	Dependency	interpretation	of adj
		1	

MOD	\longrightarrow	N Q
HEAD	\longrightarrow	Adj

2.2 Combination of words

In this subsection we will see how word descriptions can be combined to produce a sentence and its representation (namely, its syntactic dependency tree). The exposition will be exemplified with sentence (9a) and its dependency tree (9b).

(9) a. Peter often eats red beans.

b. **Dependency tree of (9a)**



The words of sentence (9a) have the following (partial) descriptions:

(10) a.
$$eats$$
: $\begin{bmatrix} HEAD [CONT \ eat-rel] \end{bmatrix}$ is $transf-verb$
b. $Peter$: $\begin{bmatrix} HEAD [CONT \ Peter-rel] \end{bmatrix}$ is $noun$
c. $beans$: $\begin{bmatrix} HEAD [CONT \ bean-rel] \end{bmatrix}$ is $noun$
d. red : $\begin{bmatrix} HEAD [CONT \ red-rel] \end{bmatrix}$ is adj
e. $often$: $\begin{bmatrix} HEAD [CONT \ often-rel] \end{bmatrix}$ is adv

HPSG comes from the phrase structure grammar tradition, which in turns comes from the Immediate Constituent Analysis (Bloomfield 1933). It does not combine words directly, but sets them into phrases. A new phrase is obtained by combination of two subphrases (or daughters) called the HEAD-DAUGHTER (= HDTR) and the NON-HEAD-DAUGHTER (= NHDTR). Phrases are subtypes of *gram-struc* like words. Compared to words, phrases have two additional features, one for each daughter:

(11) phrase :
$$\begin{bmatrix} HDTR & gram-struc \\ NHDTR & \begin{bmatrix} gram-struc \\ VAL & ebag \end{bmatrix}$$
 is a gram-struc

The core of a phrase description, as that of a word description, is given in the HEAD feature. HPSG is a Head-driven PSG, due to the fact that, in most cases, the description of a phrase comes directly from the description of its HEAD-DTR:⁸

(12) Head Computation (Head Feature Principle)

$$hd-phrase: \begin{bmatrix} HEAD \\ HDTR \\ HDTR \end{bmatrix} is a phrase$$

Given that the description of a phrase comes only from its head word, a phrase can be identified with its head word. To put it differently, the combination of two phrases can be interpreted as the combination of the two head words: inside the resulting phrase, one of these two head words becomes the HEAD-DAUGHTER (that is, the governor) and the other one the NON-HEAD-DAUGHTER (that is, the dependent). Let's emphasize that only the HEAD values of a phrase and of its head word are identical; the other values (VAL and DEP) change when the word takes new dependents. In other terms, what is called a phrase in HPSG is the description of a word with the indication of the dependents it has combined with. In dependency terms, HPSG phrases are not real linguistic entities, but rather computational objects used to compute the dependency tree and to ensure that each word is saturated (that is, its whole valence is saturated).

Note that in HPSG it is essential that the NHDTR has an empty VALENCE when it combines with the HDTR—see (11). In dependency terms, this means that the dependency tree is built bottom up: a word can combine with its governor only if it has finished combining with its dependents.

In order to be more understandable for HPSG users we have kept the usual HPSG notations. In a more dependency style, we would prefer to use NODE instead of HEAD, GOVERNOR instead of HEAD-DAUGHTER and DEPENDENT instead of NON-HEAD-DAUGHTER. But there is another reason to maintain the usual HPSG notations: as said previously, HEAD does not correspond to the description of a single node but rather of a larger part of the dependency tree which generally contains only one node but can contain two or three nodes. This is the case for modifiers, which contain,

 $^{^{8}}$ A boxed number (such as 1) indicates that two features share the same value. If these two features receive descriptions separately, these descriptions must be unified to give their common value. If the descriptions cannot be unified, the analysis fails.

in their HEAD value, the feature MOD describing their (potential) governor. We will see another case with complementizers in Section 3.

There are two major types of *phrase* structures (or combinations of words) according to whether the NHDTR (= dependent) is an actant or a modifier. These two types differ in the computation of the VALENCE: an actant NHDTR must be removed from the valence of the HDTR, while a modifier does not change its valence. We call them *hd-act-phrase* and *hd-mod-phrase*.⁹

(13) Valence Computation

a.
$$hd$$
-act-phrase : $\begin{bmatrix} VAL & [] \\ HDTR & [VAL[] \oplus \{[2]\}] \\ NHDTR & [HEAD[2] \end{bmatrix}$ isa hd -phraseb. hd -mod-phrase : $\begin{bmatrix} VAL & [] \\ HDTR & [HEAD[2] \\ VAL & [] \\ NHDTR & [HEAD[2] \\ VAL & [] \end{bmatrix}$ isa hd -phrase

Notice that *hd-mod-phrase* cannot be used with a nominal or verbal NHDTR, because it has an empty MOD value that cannot unify with the HEAD value of HDTR. On the contrary, an adjective or an adverb can express an actant, like the adjective *French* in *the French production* or the adverb *here* in *Peter came here*.

To conclude, we show the computation of the dependency tree with the DEP feature. The computation is quite easy: the NHDTR (= the dependent) must be added to the DEP bag of the HDTR (= the governor). We introduce the following additional constraint to hd-phrase.

(14) Dependency Computation

 $hd-phrase \Rightarrow \begin{bmatrix} DEP & [] \oplus \{2\} \\ HDTR & [DEP []] \\ NHDTR & [HEAD 2] \end{bmatrix}$

The feature FCT of each element of DEP is instantiated either directly in the description of the word when it is a modifier, or via the valence of its governor when it is an actant. This is due to the fact that an actant necessarily unifies with an element of the valence of its governor—see (13a).

⁹ The symbol \oplus designates the union of bags. It is preferred to the \cup symbol, which could be confused with unification. Union of bags differs from set union by the fact that all the occurrences of an element are counted: $\{1,2\}\oplus\{1,3\}=\{1,1,2,3\}$. Note that in (13), \square refers to a bag of *cat* and \square to a *cat*, due to the definition of VAL and HEAD features—see (1).

Finally, after the unification of all the inherited constraints, we obtain for our two types of phrases the following structures:

$$(15) \text{ a. } hd\text{-}act\text{-}phrase = \begin{bmatrix} \text{HEAD} & 1\\ \text{VAL} & 2\\ \text{DEP} & \exists \oplus \{\mathcal{A}\} \\ \text{HDTR} & \begin{bmatrix} \text{HEAD} & 1\\ \text{VAL} & 2 \oplus \{\mathcal{A}\} \\ \text{DEP} & \exists \end{bmatrix} \\ \text{NHDTR} & \begin{bmatrix} \text{HEAD} & 4\\ \text{VAL} & ebag \end{bmatrix} \end{bmatrix}$$
$$(15) \text{ b. } hd\text{-}mod\text{-}phrase = \begin{bmatrix} \text{HEAD} & 1\\ \text{VAL} & 2\\ \text{DEP} & \exists \oplus \{\mathcal{A}\} \\ \text{HDTR} & \begin{bmatrix} \text{HEAD} & 1\\ \text{VAL} & ebag \end{bmatrix} \end{bmatrix}$$
$$(15) \text{ b. } hd\text{-}mod\text{-}phrase = \begin{bmatrix} \text{HEAD} & 1\\ \text{VAL} & 2\\ \text{DEP} & \exists \oplus \{\mathcal{A}\} \\ \text{HDTR} & \begin{bmatrix} \text{HEAD} & 1\\ \text{VAL} & 2\\ \text{DEP} & \exists \end{bmatrix} \\ \text{NHDTR} & \begin{bmatrix} \text{HEAD} & 1\\ \text{VAL} & 2\\ \text{DEP} & \exists \end{bmatrix} \end{bmatrix}$$

These two "phrases" can be interpreted in terms of combinations of words—where the word descriptions are represented as in (16):

(16) a. Dependency interpretation of hd-act-phrase (15a)



b. Dependency interpretation of hd-mod-phrase (15b)



As we can see, both ways of combining are based on the unification of a white node of one structure with a black node of the other. In fact, white and black colors of a node correspond to VALENCE and DEPENDENCE. From this point of view, the potential governor (the MOD value) might be included in the valence.¹⁰ Both word combination operations *hd-act-phrase* and *hd-mod-phrase* can be subsumed under a single word combination operation, illustrated in (17): a white node can be identified with a black node, which results in a black node.

(17) Word Combination Operation



A dependency grammar based on the above principles has been proposed by Nasr (1995) and extended by Kahane (2001). Such a grammar is inspired by Tree Adjoining Grammars (TAG; Joshi 1987), where each word is associated with an elementary structure, the greater structure being obtained by simple combinations of these elementary structures. But unlike TAG, which uses two different operations of combinations—substitution and adjoining, the proposed dependency grammar uses a single operation.

A unified treatment of actants and modifiers, comparable to our single operation (17), has been proposed by van Noord & Bouma (1994) (and improved in Bouma *et al.* 2001). The idea is to add all the modifiers of a word in its VAL bag. This can be done by the recursive application of the following lexical rule:¹¹

(18) Modifier Introduction Lexical Rule



Now, the two ways of the VALENCE computation of (13) are subsumed under only *hd-act-phrase* (13a). In case of head-modifier combination, an element [MOD [...]] is removed from the VAL of the HDTR.

It is time for an example. In (19), we give the analysis of sentence (9a) *Peter often eats red beans*. In this figure, an HDTR is placed vertically under its mother phrase. We use the abbreviation N for the feature structure [HEAD [POS *noun-pos*]] (and similarly V, Adj and Adv).

¹⁰ In the dependency tradition it is usual to include the governor in the valence. Mel'čuk (1988: 112, 145) distinguishes the active valence, that is, the specification of the potential dependents, from the passive valence, that is, the specification of the potential governor.

¹¹ Bouma *et al.* 2001 introduces a special feature for the extended valence, which they call DEPS. Moreover, they propose a way to compute various possible values of DEPS without using lexical rules. Although their proposal is more elegant, it remains that a given word can have numerous elementary structures, which is not convenient either from a cognitive or computational point of view. Our dependency grammar avoids this problem.

(19) Analysis of sentence (9a)



The fact that the root of a sentence must be a finite verb is ensured by imposing the condition that a sentence must be a saturated phrase of subtype *sentence*:

b. *hd-act-sentence* isa *hd-act-phrase* + *sentence* c. *hd-mod-sentence* isa *hd-mod-phrase* + *sentence*

Note that the order in which a word combines with its various dependents does not matter. For instance, in our example, the verb *eats* combines with the adverbial *often*, its subject *Peter* and then its direct object (*red*) *beans*. In Chomskyan tradition, the subject is the last dependent of the verb to combine with it. This is justified by internal considerations of the theory relative to binding. Although traditional HPSGs continue to ensure that the subject is the last dependent to combine with the verb (putting it in a specific feature, SUBJ or SPR, that must be discharged after the rest of the valence), nothing in HPSG requires this, since binding is based on an obliqueness order of the syntactic relations (Pollard & Sag 1994: 248).¹²

¹² More exactly, Pollard & Sag (1994) do not use syntactic relations explicitly, but order the actants in VALENCE (called there the SUBCAT list) and use this order as obliqueness order.

Another reason to order the combination of a verb with its dependents is to avoid spurious ambiguities. Indeed, our grammar allows three phrase structures for the sentence (9a):

(21) a. ((Peter (often eats)) (red beans))
b. (Peter ((often eats) (red beans)))
c. (Peter (often (eats (red beans)))

But from the viewpoint of dependency grammar this does not matter because these three phrase structures give us the same dependency tree—presented in (9b).¹³

We will now see how the analysis (19) can be interpreted in terms of dependency grammar. The computation of the dependency tree comes from the DEP values of the maximal projection of each word. The maximal projection of a word is the highest phrase it heads. The computation of the dependency tree can be represented by using the representation we have proposed for word descriptions in (8) and for phrase descriptions in (16). The diagram in (22)—next page—is a copy of (19) where each feature structure is replaced with its representation in terms of dependency. The HPSG analysis in (19) now appears as a particular computation of the dependency tree, namely a bottom-up computation. The diagram in (22) can be read in this way: for instance, at the bottom of the figure we see the structures associated to *often* and *eats*, and just above the result of the combination of these two structures with the *hd-mod-phrase* schema. In these two lexical structures, the features \exists have different values, but after the combination these two features share the same value, which is the unification of the two previous values.

3 Complementizers, relativizers and other transferers

Tesnière (1959), who is above all known to be the father of modern dependency theory, proposes an original classification of words that we will adopt here. Tesnière postulated that there are only four main parts of speech—verb, noun, adjective and adverb—with the following connections:

- verb actants are nouns;
- verb modifiers are adverbs;
- noun dependents (actants or modifiers) are adjectives;
- adjective and adverb dependents are adverbs.

¹³ A third reason to order the combination of a verb with its dependents concerns coordination. The argument is that the subject can be easily factorized, like in *Peter eats red beans and drinks water*. But this argument is not reliable, because it is also possible to factorize the direct object: *Peter likes, and Bill hates, red beans*. Moreover, the contrast between subject and object factorizations can be captured in a different way.



When a word of a part of speech other than adjective needs to modify a noun, it must be "transferred" into an adjective, that is, it must be introduced in the syntactic structure by a special word which masks it and makes it appear as an adjective for its governor. The transfer of nouns (into adjectives and adverbs) is carried out by prepositions.¹⁴

In Section 3.1, we will see how Tesnière's transfer can be formulated in HPSG. Section 3.2 will study relativizers and introduce the SLASH feature and the constraints on the computation of this feature. Section 3.3 proposes a simple solution of the *that*-trace effect in English and the *qui-que* alternation in French. Section 3.4 deals with the representation of transferers and the Locality Principle.

(22) Dependency interpretation of (19)

¹⁴ Tesnière's terms for *transfer* and *transferer* are Fr. *translation* and *translatif* (in French, *translation* only means a translatory movement). The theory of transfer is implicit in the theory of the three ranks of Jespersen (1924).

3.1 Transferers

Transferers receive the common POS *t*(*ransfere*)*r-pos*. Alongside POS, we will introduce a new feature UPPOS (up part of speech) oriented upwards, that is, towards the governor. Plain words—verbs, nouns, adjectives and adverbs—have the same value for both POS and UPPOS:

(23) *plain-word* :
$$\begin{bmatrix} UPPOS \\ POS \end{bmatrix}$$
 is a word is a word

The UPPOS of a transferer is the part of speech into which it transfers its actant: its value is of type *main-pos*, which is a supertype for *noun-pos*, *verb-pos*, *adj-pos* and *adv-pos*. The actant of a transferer will be called the *transferee*.

(24) transferer:
$$\begin{bmatrix} \text{HEAD} \begin{bmatrix} \text{UPPOS main-pos} \\ \text{POS} & \text{tr-pos} \end{bmatrix} \\ \text{VAL} \left\{ \begin{bmatrix} \text{FCT transf} \end{bmatrix} \right\}$$
 is a word

For instance, complementizers such as *that*, *if* or *whether* are transferers of verb into noun:

(25) a. verb-transferer :
$$\left[V_{AL} \left\{ \left[UPPOS \ verb-pos \right] \right\} \right]$$
 is a transferer
b. complementizer : $\left[H_{EAD} \left[UPPOS \ noun-pos \right] \right]$ is a verb-transferer

From a theoretical point of view, our feature UPPOS attributes a common property to a class of words, predicting that they are a natural class—that is, they should behave similarly in some respects. For instance, the subject of the verb *annoy* can be a noun or a *that*-clause, that is, a UPPOS *noun-pos* (*Peter's departure <= That Peter leaves> annoys me*).

(26) *annoy*:
$$\begin{bmatrix} FCT & subj \\ UPPOS & noun-pos \end{bmatrix}, \begin{bmatrix} FCT & dobj \\ UPPOS & noun-pos \end{bmatrix}$$
 is a verb

Note that among verbs only finite verbs are *plain-word* and have UPPOS *verb-pos*: infinitives and gerunds are UPPOS *noun-pos* and participles are UPPOS *adj-pos*, because they respectively alternate with nouns (27a-b) and adjectives (27c-e) and never with finite verbs:¹⁵

¹⁵ As proposed by Tesnière, the copula *to be* in (27e) should be not considered as a plain word, but rather as a transferer of adjective into verb.
(27) a. Peter wants to drink / wants a beer.

- b. Peter left without sleeping / without his cap.
- c. the cap forgotten by Peter / the red cap
- d. the guy sleeping with his cap / the sleepy guy
- e. Peter is sleeping with his cap / is asleep.

The noun-transferers are the prepositions. Prepositions can also take infinitives and gerunds as dependents. This is not surprising because these are verbs transferred into nouns.

3.2 Relativizers

Relativizers are transferers of verbs into adjectives. In other words, a relativizer introduces a clause that can modify a noun, that is, a relative clause:

(28) a. the guy that came yesterdayb. the guy (that) you knowc. the guy (that) you are talking to

Relativizers are obviously more complicated than complementizers because the noun modified by a relative clause needs to play a semantic role in this clause: it shares its semantic content with a slashed noun. The usual assumption is that the antecedent noun is coindexed with the gap (= the slashed noun); however, we do not think that indices are useful here. The use of indices comes from the Chomskyan tradition, where the gap is filled by a trace and traces are considered as some sort of pronouns without phonological realization. Although, following Sag & Fodor 1994, traces have been abandoned in HPSG's analyses of extraction, coindexation has survived. Here we defend a more radical analysis, coming from the Meaning-Text framework (Žolkovskij & Mel'čuk 1967, Kahane & Mel'čuk 1999): the antecedent noun *is* the semantic argument of the governor of the gap. Formally, it means that the CONT values of the antecedent noun and the slashed noun are identical. So relativizers have the following description, where the BIND value must be discharged with a slashed element as explained below.¹⁶

(29) relativizer: $\begin{bmatrix} UPPOS & adj-pos \\ FCT & mod \\ MOD & N : [i] \end{bmatrix}$ is a verb-transferer BIND $\{N : [i]\}$

Slashed elements belong to the set of potential syntactic dependents of a word that are not realized as dependents of this word. For this reason, they are put outside of the VALENCE in a special bag called SLASH. The fact that VAL \oplus SLASH is equal to the bag of potential dependents can be expressed by an equation as in Bouma *et al.* 2001 or ensured by a lexical rule as traditionally done and proposed here. The lexical rule (30) states that each word description with a non-empty VAL can give a new word description where one of the elements of VAL has been moved into the special feature SLASH,

¹⁶ We use the abbreviated notation N : i for the feature structure $\begin{bmatrix} POS & noun-pos \\ CONT & i \end{bmatrix}$.

which is discharged by a different process than the valence computation in (13). This lexical rule can be applied recursively.¹⁷

(30) Slash Introduction Lexical Rule

 $\begin{bmatrix} word \\ VAL & [] \oplus \{ \Im \} \\ SLASH 2 \end{bmatrix} \longrightarrow \begin{bmatrix} word \\ VAL & [] \\ SLASH 2 \oplus \{ \Im \} \end{bmatrix}$

In (30), contrary to (15a), an element is removed from VAL without being added to DEP. In fact, a slashed element never appears in the DEP bag of its syntactic governor. We consider that the link between a slashed element and its governor, which we call a *quasi-dependency*, is not a true syntactic dependency (Kahane 2001). This is justified by two reasons. First, unlike true dependencies, quasi-dependencies do not obey the well-formedness rules of the dependency tree since they can form cycles (see (41), p. 131). Second, unlike true dependencies, quasi-dependencies do not intervene in the word order rules due to the fact that a slashed element is not realized in the projection of its governor. In our figures, a quasi-dependency is represented by a dashed arrow. Figure (31) proposes a dependency representation of the Slash Introduction Lexical Rule.

(31) Dependency interpretation of the lexical rule (30)



We will now study some constraints on extraction. Two types of constraints must be considered (a third type concerns only the *wh*-extraction and the pied-piping): the elements put into SLASH and the path followed by a slashed element.

Constraints on slashed elements. Not all elements of VALENCE can be slashed. For instance, in French, in contrast to English, the transferee of a preposition can never be slashed:

(32) **le type que tu parles à* the guy that you are-talking to

Therefore, we introduce a subtype *gaprel* of *syntrel* such that only a syntactic relation that is a subtype of *gaprel* allows for its dependent node to be slashed. For instance,

¹⁷ Combined with the Modifier Introduction Lexical Rule (18), the Slash Introduction Lexical Rule allows for modifiers to be slashed and extraction of modifiers to be handled. In the absence of the Modifier Introduction Lexical Rule, a second Slash Introduction Lexical Rule would be necessary.

the relation transf(er), between a transferer and its transferee, will be a subtype of *gaprel* in English but not in French. This constraint on the type of slashed elements is captured by allowing the Slash Introduction Lexical Rule (30) to apply only to an element whose syntactic function is a subtype of *gaprel*. Note that we do not introduce a particular treatment for subjects and, contrary to Pollard & Sag (1994), we consider that subjects can be slashed and that the relation *subj* is a subtype of *gaprel*. We will return to subject extraction in the next subsection.

Constraints on the path of a slashed element. The SLASH value of a phrase is the union of the SLASH values of its daughters. Such a feature, which gathers the values of both daughters, is called a *foot feature* and contrasts with head features, such as HEAD, which only takes the value of the head daughter. If the HDTR has a non-empty BIND value and this value can be unified with an element of the SLASH value of the NHDTR, the BIND value is unloaded from the SLASH. Both properties are expressed by the following constraint:

(33) Slash Computation

$$phrase \implies \begin{bmatrix} SLASH \] \oplus \ 2 \\ HDTR \ \begin{bmatrix} SLASH \ 1 \\ BIND \ 3 \end{bmatrix} \\ NHDTR \ \begin{bmatrix} SLASH \ 2 \\ \oplus \ 3 \end{bmatrix}$$

Nevertheless, there are well-known constraints on the path a slashed element can follow. These constraints have been first brought to the fore by Ross (1967) in terms of islands, that is, in terms of types of phrases that do not allow for a slashed element to escape.¹⁸ Here, following Kaplan & Zaenen's (1989) LFG description of extraction, we prefer to use syntactic relations to control the path of slashed elements. For this, we define a subtype *liftrel* of *syntrel* such that only a syntactic relation that is a subtype of *liftrel* allows for a slashed element to lift (see Hudson 1990 for a DG description of extraction in terms of lifting, called there raising, and Kahane *et al.* 1998 for a polynomial processing). To put it differently, a word which does not have a syntactic function of type *liftrel* cannot allow for a slashed element to lift and hence must have an empty SLASH bag. This gives us the following constraint on *gram-struc* (where *nebag* is a supertype for all non empty bags):

(34) Island Constraint

$$\begin{bmatrix} gram-struc \\ SLASH & nebag \end{bmatrix} \implies \begin{bmatrix} HEAD & FCT & liftrel \end{bmatrix}$$

For instance, the subject island in (35a) is accounted for by the fact that the syntactic relation *subj* is not a subtype of *liftrel* (and therefore the slashed element cannot lift from the *that*-clause to the verb *annoys*). The possibility of extraction in the impersonal construction (35b) comes from the fact that here the *that*-clause is no longer subject (the

¹⁸ Ross' description was formulated of course in terms of transformations and not in terms of slashed elements, but the ideas remain the same.

subject is the dummy *it*) but has another syntactic function we call *quasi-subject* and *quasi-subject* is a subtype of *liftrel*.

(35) a. *Who does that Peter loves annoy you? (⇐ That Peter loves her annoys you)b. Who does it annoy you that Peter loves?

Moreover, lexical constraints can be easily added to the constraint (34): a particular lexical unit can require one of its actants to have an empty SLASH.

3.3 That-trace effect and French qui-que alternation

Following our hypothesis that a finite verb cannot be an actant of a verb without being transferred into a noun, we consider that, in *that*-less sentential clauses and in *that*-less relative clauses (as in (28)), there is a verb-transferer without phonological realization we call the *null transferer*.¹⁹ We will show how to encode the difference of behavior between *that* and the null transferer (thus solving the problem of the *that*-trace effect) and how to describe the *qui-que* alternation in French in a similar way.

As we will see, verb-transferers are sensitive to whether the subject of the verb they transfer is slashed. For this reason, we introduce on finite verbs a feature SUBJ that takes the value '+' when the subject is in VAL and the value '-' when it is in SLASH (the value must be assigned when applying the lexical rule (30) transferring VAL elements into SLASH).²⁰

We can now easily encode the contrast between *that* and the null verb-transferer. First, the null complementizer cannot be used as subject (**Peter leaves annoys me* vs. *That Peter leaves annoys me*). Second, the null relativizer does not allow for the subject of the verb it transfers to be slashed (**The man lives here is a student* vs. *The man that lives here is a student*). Third, like the null relativizer, the *that*-complementizer does not allow for the subject of the verb it transfers to be slashed (**Who did Peter claim that left?* vs. *Who did Peter claim left?*). This last phenomenon is known as *that*-trace effect.²¹ Note that the null relativizer is possible with a subject extraction as long as it is not the subject of the verb it transfers (*The man Peter claims left is a student*). The following constraints are sufficient to get the right properties:

¹⁹ Our analysis contrasts with earlier HPSG analyses, which tend to consider that a verb can subcategorize a clause and that *that* can be present but without major effect (see, for instance, Sag & Wasow 1999: 259).

 $^{^{20}}$ Instead of introducing a special feature to control whether or not the subject is slashed, Bouma *et al.* 2001, following a recent trend in HPSG, proposes to retain a copy of the argument structure of a word (it is the value of ARG-ST) in its HEAD description. In such a case, it is possible to verify whether the subject (the first element of the list ARG-ST) is slashed or not—slashed elements being subtypes of a special type (*gap-ss*). Nevertheless, preserving the argument structure in the HEAD description of a word is a clear violation of the Locality Principle (see Section 3.4). It is preferable to single out the information on non-head daughters that is useful rather than to allow for the whole of the information on the dependence to be accessible (especially since, by recursivity, not only the dependence of a word but its whole projection is accessible).

²¹ See Pollard & Sag 1994: 171*ff* for a more sophisticated analysis, where subjects are never slashed, and for a comparison with the GB treatment.

(36) a. *that-complementizer* : $VAL \{ [SUBJ +] \}$ is a *complementizer*

b. null-complementizer : [FCT ¬subj] is a complementizer

c. that-relativizer isa relativizer

d. *null-relativizer* : $\left[VAL \left\{ \left[SUBJ + \right] \right\} \right]$ is a *relativizer*

French also has a verb-transferer which can be both a complementizer and a relativizer; this is a qu-word.²² It has two alternating forms: qui and que. The traditional analysis says that qui is the subject relativizer in (37a), que is the object relativizer in (37b), and only que is a complementizer, as in (37c). But this analysis fails in case of subject extraction in a subordinated clause as in (37d), where the relativizer is que and the complementizer is qui.

(37) a. *le type qui vient*

'the guy that comes'

- b. *le type que tu regardes* 'the guy that you are watching'
- c. *Tu penses que Pierre vient.* 'You think that Peter comes'
- d. *le type que tu penses qui vient* 'the guy that you think (that) comes'

The solution (adapted from Moreau 1971) is very simple.²³ The words *qui* and *que* are two forms of the same verb-transferer, and this transferer has the marked form *qui* when the verb it transfers has a slashed subject and the non-marked form *que* otherwise:

This gives us the right forms of the transferer. In (37a), the relativizer is *qui* because it transfers a verb with a slashed subject (which is discharged by the relativizer). In (37b), the relativizer is *que* because it transfers a verb with a canonical subject. In (37c), the complementizer is *que* because it transfers a verb with a canonical subject. And in (37d), the complementizer is *qui* because the verb it transfers (*vient* 'comes') has a slashed subject and the relativizer is *que* because the verb it transfers (*pense* 'think') has a canonical subject. The whole analysis of (37d) is given in (39); we use the abbreviated notation [↑]Adj or [↑]N for a verb transferer of UPPOS *adj-pos* or *noun-pos*.

²² French *qu*-words (like English *wh*-words, German *w*-words, Italian *ch*-words, etc.) come from the Indo-European root k^w . Contrary to English, French has no verb-transferers of another family (such as English *that* or German *d*-words coming from the Indo-European root to); French *qu*-words form a less homogeneous family than English *wh*-words, containing both pure verb-transferers and pronouns. See, for instance, Haspelmath 1997 or Le Goffic 2002.

 $^{^{23}}$ The analysis of Moreau (1971), also sketched by Gross (1968: 124), has been adopted by Kayne (1974-75), to whom it is generally attributed. The contribution of Kayne is to defend the thesis that the relativizer and the complementizer *que* are the same word.



(39) Analysis of the relative clause (37d)

In (41), next page, we give an interpretation of (39) in terms of DG. The representation of the valence of *penses* 'think' is justified in the next subsection. The BIND feature of the relativizer is represented by a dashed arrow that must be unified with one of the dashed arrows of the SLASH bag. This arrow points to the antecedent noun, that is, the governor of the relative clause. This noun will be a semantic argument of the verb introducing the SLASH value (here, *vient* 'comes'); this semantic argument has no syntactic counterpart in the clause.

3.4 *Transferers and the Locality Principle*

Whether a complementizer or a transferer in general is the head of the transferred group or not has been debated for a long time (see, for instance, Pollard & Sag 1994: 44). The reason is that many transferers have no meaning and only play a syntactic role namely, the transfer of the transferee. For meaningless transferers, the meaning of the combination of the transferer and the transferee evidently comes from the transferee. Moreover, in some cases, government-related information passes through the transferer: for instance, among verbs that take a *that*-clause as complement, some require that the subordinated verb carries the indicative mood (as *know* in (40a)), while others require that it carries the subjunctive mood (as *demand* in (40b)).

(40) a. *I know that Peter leaves today.*b. *I demand that Peter leave immediately.*

(41) Dependency interpretation of (39)



KAHANE

HPSG incorporates the Locality Principle (Pollard & Sag 1987: 143-44 and 1994: 23), inherited from the Chomskyan tradition, which ensures that the HDTR cannot impose conditions on the internal structure of the NHDTR. This principle is implemented by the fact that VAL elements are only *cat* and not the whole *gram-struc*, that is, that VAL specifies only the HEAD of the NHDTR and not the whole NHDTR.

Therefore, if we want to pass information through the transferer to the transferee, we need to raise some features of the transferee into the HEAD of the phrase obtained by the transferer-transferee combination. One solution, adopted for a long time by HPSG, is to introduce a special combination schema (the traditional *head-marker-phrase*). Another solution, more or less equivalent, is to use the standard *head-actant-phrase* but to copy the whole transferee's description into the transferer's description. For that, we add a feature TRANSF into the HEAD value of a transferer, whose value is the transferee:

(42) transferer
$$\Rightarrow \begin{bmatrix} \text{HEAD} \begin{bmatrix} \text{TRANSF} \\ \end{bmatrix} \end{bmatrix}$$

Г

This amounts more or less to treating the transferer and the transferee as co-heads.²⁴ Note that in Pollard & Sag 1994 the main verb of the *that*-clause is treated as its head, while in Sag & Wasow 1999 it is the complementizer *that*. Nevertheless, both analyses are similar, due to the fact that in both cases information from the two daughters is raised to the mother. Tesnière himself also considered the transferee and the transferer more or less as co-heads. In his dependency trees, a transferer and its transferee are represented side by side in a bubble he calls the *transfer nucleus* (Fr. *nucléus translatif*), which occupies a node in the dependency tree (Tesnière 1959: 46, 56).

٦

Now we can have the following description for *demand*:

(43) demand :
$$\begin{vmatrix} V_{AL} & \{ \begin{bmatrix} FCT & subj \\ POS & noun-pos \end{bmatrix} , \begin{bmatrix} FCT & dobj \\ UPPOS & noun-pos \\ TRANSF \begin{bmatrix} POS & verb-pos \\ INFL & [MOOD & subj-mood] \end{bmatrix} \end{vmatrix}$$
 is a verb

In DG's terms, in conformity with (43), the direct object valence slot of the verb *demand* is filled not with a single node, but with a piece of dependency tree composed of the complementizer and its transferee (see the representation for Fr. *penses* 'think' in Figure (41) above). Repeating the transferee description in the transferer HEAD bypasses the Locality Principle. If each word includes its dependents' descriptions in its

²⁴ Coordination is another phenomenon where several daughters, the conjuncts, are treated as coheads. Coordination cannot be adequately dealt with in a pure dependency approach, although some propositions have been made—see, for instance, Mel'čuk's paper (Part III, Section 1.3) in this volume. Even in very simple examples such as *Peter and Mary are walking* we see that the verb agrees with none of the words *Peter*, *Mary* or *and* but with the result of their combination. In this case, it seems better to clearly introduce a non-lexical node for the coordination as it is done in HPSG. Such analyses have been proposed in dependency frameworks by Petkevič (1995) and Kahane (1997). But contrary to phrase structure interpretations, such a non-lexical node corresponds only to the list of words concerned by the coordination, that is, conjuncts and coordinating conjunctions, and not to the whole coordination phrase.

own description, by recursivity, each word will have in its description the description of all the words it dominates, that is, the description of the whole dependency subtree it roots.²⁵ But, as the repetition of a dependent's description is confined to transferers, we avoid this problem and the Locality Principle is preserved. In fact, the problem is that locality cannot be limited to the direct dependents (and to the governor by way of MOD). It is necessary to allow a deeper control. That is the case for a verb such as *demand*, which constrains a dependent of one of its dependents.

4 Wh-extraction

We traditionally call *wh*-extraction an extraction involving a *wh*-word in English (or a similar word in another language). *Wh*-extraction contrasts with *that*-extraction with respect to pied-piping:²⁶

(44) a. the person [to whom] they dedicated the buildingb. *the person [to that] they dedicated the building

Earlier HPSG analyses of *wh*-word extraction are particularly problematic from the viewpoint of DGs given that these analyses resort to phrasal descriptions. Section 4.1 presents a description of extraction based on HPSG analyses, but adapted to the notions and notations introduced in the previous sections. Section 4.2 proposes a new description entirely in terms of combinations of words, that is, without phrasal descriptions.

4.1 An analysis of extraction based on earlier HPSG analyses

In earlier HPSG analyses (Pollard & Sag 1994; Sag 1997; Ginzburg & Sag 2000; Bouma *et al.* 2001), *wh*-words have rather simple descriptions: they are only distinguished from other pronouns by the fact they duplicate their semantic content in a special feature: REL for relative *wh*-words and QUE for interrogative *wh*-words. To be precise, in traditional HPSG analyses, it is the index rather than the whole semantic content which is duplicated, following the Chomskyan tradition, which, for a long time, did not really consider the meaning. The features REL and QUE are empty for other words.

(45) a. wh-rel-word :
$$\begin{bmatrix} \text{HEAD} \begin{bmatrix} \text{CONT} & \\ \end{bmatrix} \end{bmatrix}$$
 is a word
b. wh-int-word :
$$\begin{bmatrix} \text{HEAD} \begin{bmatrix} \text{CONT} & \\ \end{bmatrix} \begin{bmatrix} \text{variable} \end{bmatrix} \end{bmatrix}$$
 is a word
QUE $\begin{bmatrix} \end{bmatrix}$
c. word $\Rightarrow \begin{bmatrix} \text{REL} / none \\ \text{QUE} / none \end{bmatrix}$

²⁵ It is possible to preserve some form of Locality Principle by only repeating a part of the dependents' descriptions and to limit the depth of the dependency subtree copied in the governor's description.

 $^{^{26}}$ In the transformational analysis, the pied-piping is the fact that the *wh*-word which moves to the beginning of the clause carries along with it other words, much as the Pied Piper of Hamelin attracted rats and children to follow him.

KAHANE

Interrogative *wh*-words have a very particular semantic content we call *variable*. We will not discuss this meaning here (see Kahane & Mel'čuk 1999 for one approach).²⁷ Nevertheless, we consider that it is this particular meaning which licenses interrogative clauses in some syntactic positions (the same meaning is introduced by the complementizers *if* and *whether*). On the contrary, relative *wh*-words have no proper semantic content, because their CONT value is the CONT value of the antecedent noun of the relative clause. The identification of the two values is realized by the percolation of REL. In the same way, the semantic value of an interrogative *wh*-word is lifted via QUE in order to license an interrogative clause:

(46) **Pied-Piping Computation**



The operation +, which appears above, is defined in (46) as follows: none+none = none, none+[...] = [...]. The combination of two non-empty values fails. In other words, only one non-empty value is allowed. We do not consider languages that authorize several extraposed *wh*-words.

All of the complexity of *wh*-extraction is concentrated in the final step of the analysis, the computation of the structure of the whole clause. Recent analyses of extraction in HPSG are based on phrasal descriptions of *wh-relative-clause* and *wh-interrogative-clause*. These two phrase descriptions inherit from two families of constraints: *head-filler-phrase* and *clause*. The *clause* description constrains the HDTR to be a clause, that is, a saturated verb.²⁸ The *head-filler-phrase* constrains the NHDTR to be combined with an element of the SLASH of the HDTR, that is, to fill a gap. The *wh-rel-cl*, which inherits from both *clause* and *hd-fill-phrase*, constrains the NHDTR to have a non-empty REL that will be equal to the semantic content of the antecedent noun.

²⁷ The so-called interrogative *wh*-words are not inherently interrogative and can be found in non-interrogative contexts:

⁽i) a. She told me to whom they dedicated the building.

b. I know to whom they dedicated the building.

c. To whom they dedicated the building does not matter.

An interrogative *wh*-word introduces an unspecified meaning similar to a free variable (in the mathematical sense). It is the context—the rising intonation in direct questions or the meaning of the governor verb in indirect questions—that requires to instantiate this variable and gives the interrogative interpretation.

 $^{^{28}}$ Ginzburg & Sag (2000) consider different types of clauses for relatives and interrogatives. We do not retain this idea, which is an attempt to preserve the description of an interrogative clause as a *hd-phrase*. See our solution below.

$$(47) \text{ a. } hd-fill-phrase : \begin{bmatrix} SLASH [] \oplus [2] \\ HDTR [SLASH [] \oplus [3]] \\ NHDTR [HEAD [3] \\ SLASH [2] \end{bmatrix} \text{ isa } phrase$$
$$b. clause : \begin{bmatrix} VAL & ebag \\ HDTR [HEAD V \\ VAL & ebag \end{bmatrix} \text{ isa } phrase$$
$$c. wh-rel-cl : \begin{bmatrix} HEAD & \begin{bmatrix} UPPOS & adj-pos \\ FCT & mod \\ MOD & N : [i] \\ NHDTR [REL [i] \end{bmatrix} \text{ isa } clause + hd-fill-phrase$$

The *head-filler-phrase* (abbreviated in *hd-fill-phrase*) is usually considered as a *headed-phrase* (defined in (12), p. 117). From this it follows, given the *clause* constraint, that *wh*-clauses are considered as clauses, that is, as phrases headed by a verb. This is not compatible with the Chomskyan Complementizer Phrase analysis of *wh*-clauses nor with Tesnière's transfer theory, which predicts that an interrogative clause, which is a verb actant, must be UPPOS N and that a relative clause, which modifies a noun, must be UPPOS Adj. The analysis in Sag 1997 can be adapted to our theoretical framework by considering that *head-filler-phrase* is not a *headed-phrase* but a verb-transferred phrase:²⁹

(48) *hd-fill-phrase*
$$\Rightarrow \begin{bmatrix} \text{HEAD} [\text{TRANSF}] \\ \text{HDTR} [\text{HEAD}] \end{bmatrix}$$

This alternative analysis is particularly adapted for the treatment of interrogative clauses.

We can now propose that the semantic content of a *wh-int-clause* comes from the *wh*-word:

The presence of a *variable* meaning allows for the right distribution of the interrogative clause. Moreover, as extensively argued by Kahane & Mel'čuk (1999: 37-40), the semantic content of the *wh*-word is the semantic argument of the verb governing the interrogative clause.

²⁹ In the *head-filler-phrase*—see (47a)—the two subphrases are called HDTR and NHDTR, but these notations are not very significant here, because no one actually is a head daughter.

The variant of Sag's (1997) analysis we have proposed brings us back to the analysis of relatives presented in Pollard & Sag 1994: 212, where *wh-rel-clause* is obtained by the combination of the *hd-fill-clause* with a null relativizer (with a substantial difference, given that the relativizer is not considered as a head in Pollard & Sag 1994). This analysis does not discharge a slashed element in the *hd-fill-phrase*, but instead discharges it with the BIND feature of the null relativizer (which is the same null relativizer as in *that*-less relatives). Such an analysis, inspired by Chomskyan analyses (see Sag 1997: 433, for a criticism), is not well motivated for languages that, unlike English, do not allow for the deletion of the relativizer.

Let us come to our point: what are the drawbacks of the above analysis?

First, this analysis uses complex *ad hoc* phrasal descriptions. Much information is added via *wh-rel-clause* and *wh-int-clause*, but we do not know from which words it comes and why. Independently of the fact that such an approach is incompatible with a DG, we can suppose we have lost generalizations somewhere.

Second, we can observe that the semantic content of a *wh*-clause does not come from the HDTR but from the *wh*-word: the semantic content is lifted via the features REL or QUE and the distributional properties of the clause (the fact that the relative clause can modify a noun or the interrogative clause can be the object of an interrogative verb) are no longer controlled by the HDTR (the main verb of the clause).

In fact, consider the combination of the filler daughter and the head daughter, that is the rest of the relative clause. This combination depends:

- for the filler daughter, both on its head, which will govern the gap, and on the presence of a *wh*-word in this phrase (which controls the distribution of the clause);
- for the head daughter, both on its head—that is, the main verb of the clause—and on the word that governs the gap—that is, the word that has introduced the content of the slash feature.

We will show that this phrase description involves four words and the simultaneous combination of two couples of words: the *wh*-word with the main verb of the relative/ interrogative clause and the governor of the gap with the head of the filler phrase.

4.2 A more lexical treatment of wh-extraction

The treatment of *wh*-extraction we propose is again based on an original idea of Tesnière, which has been successfully implemented in a dependency unification grammar based on Meaning-Text Theory (Kahane 1996, 1997, 2001), in TAG (Kahane *et al.* 2000) and in DTG, a TAG-related formalism (Candito & Kahane 1998).³⁰ Tesnière

³⁰ Questions similar to those studied here arise in Lexicalized TAG. LTAG is a fully lexicalized grammar which does not use phrasal descriptions: all of the information must be attached to elementary structures (= word descriptions). In TAG traditional analyses, the ability of relative clauses to modify a noun (given by the *wh-rel-clause* constraint in HPSG) is marked on the main verb of the relative clause (Kroch & Joshi 1986; Joshi 1987; XTAG 1995), significantly increasing the number of elementary structures associated with a verb entry. A simpler and more powerful analysis treating *wh*-words as transferers, comparable to the analysis proposed here, is possible (Kahane *et al.* 2000).

(1959: 61) postulates that a *wh*-word has two roles, a transferer role and a pronominal role, and it behaves as two words simultaneously. In Tesnière's dependency trees, the *wh*-word occupies two nodes. For instance, in (50b), the French relative pronoun *qui* occupies two nodes called *qu*- and *qui*.³¹ The *qu*- node of the pronoun *qui* corresponds to the transferer role and is common to all French *qu*-words (pronouns and pure transferers). In the same way, an English *wh*-word occupies two nodes; all *wh*-words have a common part *wh*- which behaves as the pure transferer *that*.

(50) a. la personne à qui je veux parler the person to whom I want to.speak
b. Dependency tree à la Tesnière for (50a)



In order to introduce two descriptions of nodes in the description of a *wh*-pronoun, we add a feature SUPERHEAD (SRHD), whose value is a *word*. The pronominal part of the *wh*-word is associated to the core of word description (= the main HEAD feature), while the transferer part of the *wh*-word is associated to the superhead (= the HEAD feature of SRHD). By default, a word has an empty SRHD.

(51) a. wh-word : $\begin{bmatrix} pronoun \\ SRHD & verb-transferrer \end{bmatrix}$ is a word

b. word \Rightarrow [SRHD /none]

There are four subtypes of *wh-word*: those corresponding to relative pronouns, to interrogative pronouns in indirect questions, to interrogative pronouns in direct questions and to pronouns in so-called headless relatives.³² Only the two first subtypes are considered here:

³¹ This relative pronoun qui is a different word from the verb-transferer que/qui presented in Section 3.2. This qui is [+human] and can be used only after a preposition.

³² Headless relatives are traditionally described as relatives with an empty antecedent. There are good reasons to consider that they are more closely related to indirect interrogatives than relatives (Le Goffic 1993).



It can be noted that the SUPERHEAD of *rel-wh-word* and that of *int-wh-word* are similar to *relativizer* and *complementizer*, respectively. The semantic cohesion between the relative or interrogative clause and the rest of the sentence is ensured by the *wh*-word and the meaning i which appears in the two-head descriptions of the word and hence will make the connection between the two parts of the sentence.

We will now show how the above description of *wh*-words can be interpreted in terms of dependency grammars. The two parts of a *wh*-word (the main HEAD and the HEAD of SRHD) correspond to two nodes in the dependency tree, as proposed by Tesnière. We give in Figures (53) graphical representations of *int-wh-word* and *rel-wh-word*. The dashed line linking two nodes indicates that these two nodes are two parts of the same word and must be merged at the morphological level.

Note that our analysis (and therefore Tesnière's description of *wh*-words) does not imply that we consider a *wh*-word to be the merger of two words: a *wh*-word is well and truly a single word. We simply consider that a *wh*-word has two roles—transferer and pronoun—and it is why it occupies two positions in the dependency tree. Even these two positions are not very significant: what is important in our analysis is the fact that a *wh*-word is governed in two different ways; it is only because we impose on the dependency structure to be a tree (and each node of the structure to have at most one governor) that we need to associate the *wh*-word to two nodes. Note also that the constraints on extraction can be in a large part imputed to the double role of the *wh*-word: it is because a single word must both transfer the clause and fill a gap that the way between the head of the clause and the gap is not free. Ideally, the gap should be governed by the head of the transferred clause, allowing for the *wh*-word to fill its two roles directly, which would be not so far from the truth (as argued in Kahane 2001).

(53) a. Dependency interpretation of *int-wh-word*





Our analysis works in the following way. The feature SUPERHEAD merges the features QUE and REL of the earlier analyses.³³ Like them, it is a foot feature:

(54) Superhead Computation

 $hd-phrase \Longrightarrow \begin{bmatrix} SRHD & 1 + 2 \\ HDTR & SRHD & 1 \end{bmatrix}$ $NHDTR \begin{bmatrix} SRHD & 1 \end{bmatrix}$

Only certain syntactic relations allow pied-piping. To control that, we define a subtype of *syntrel* called *whrel*. A word which does not have a syntactic function of type *whrel* cannot allow for SUPERHEAD to percolate and hence must have an empty SUPERHEAD. This gives us the following constraint on *gram-struc*:

(55) Pied-Piping Constraint

 $\begin{bmatrix} gram-struc \\ SRHD & verb-transferrer \end{bmatrix} \implies \begin{bmatrix} HEAD & [FCT whrel] \end{bmatrix}$

To combine the filler phrase with the rest of the clause, we use a *filler-gap-phrase*. When this phrase is a *hd-phrase*, it gives the traditional *hd-fill-phrase*, which is still needed for topicalization. Like the traditional *hd-fill-phrase*, the *fill-gap-phrase* constrains the NHDTR to be combined with an element of HDTR|SLASH and thus to fill a gap.³⁴

(56) a. *fill-gap-phrase*: $\begin{bmatrix} SLASH & [] \oplus [2] \\ HDTR & [SLASH & [] \oplus \{ \mathbb{3} \}] \\ NHDTR & \begin{bmatrix} HEAD & \mathbb{3} \\ SLASH & [2] \end{bmatrix} \end{bmatrix}$ is a *phrase*

b. *hd-fill-phrase* isa *fill-gap-phrase* + *hd-phrase*

b. Dependency interpretation of *rel-wh-word*

³³ Pollard & Sag (1994: 159) justify introducing two distinct features QUE and REL instead of a single *wh*-feature by the following contrast between interrogative and relative clauses:

⁽i) a. This is the farmer pictures of whom appeared in "Newsweek."

b. *Pictures of whom appeared in "Newsweek"?

Whatever the reasons of this contrast, it can still be handled in our analysis given that *rel-wh-word* and *int-wh-word* have distinct SUPERHEAD values.

³⁴ We continue to take the filler phrase as NHDTR, but it is clear that it is once again not very significant.

We introduce another subtype of *filler-gap-phrase* that we call *superhead-filler-phrase*, which is not headed by the HDTR but by the SUPERHEAD, see (57a). The SUPERHEAD takes the HDTR as transferee. Therefore, *srhd-fill-phrase* inherits from a Valence Computation schema comparable to the standard *hd-act-phrase* in (13a), this one involving SUPERHEAD and HDTR, rather than HDTR and NHDTR, see (57b).

(57) a. Alternative Head Computation (Head from Superhead)

 srhd-fill-phrase :
 HEAD
 I

 NHDTR
 Image: SRHD
 none

 Image: SRHD
 Image: SRHD
 Image: SRHD

 Image: SRHD
 Image: SRHD

b. Alternative Valence Computation (Superhead-Head Combination)

$$srhd-fill-phrase \Longrightarrow \begin{bmatrix} VAL & [] \\ HDTR & HEAD & [] \\ VAL & ebag \end{bmatrix}$$
$$NHDTR \begin{bmatrix} SRHD & [VAL & [] \oplus \{2]\} \end{bmatrix}$$

Let us explain what happens: the two parts of a *wh*-word act alternatively as heads. At the beginning, for the pied-piping, only the pronominal head is recognized and the SUPERHEAD is simply percolated (as a SLASH feature). At the last step, the SUPER-HEAD becomes the head and the extracted phrase (= filler phrase) acts as a verb-transferer. The extracted phrase in a relative clause becomes equivalent to the relativizer (29), p. 125:

$$(58) \begin{bmatrix} \text{HEAD } \boxed{2} \\ \text{SRHD } \begin{bmatrix} \text{HEAD } \boxed{1} \end{bmatrix} \equiv \begin{bmatrix} \text{HEAD } \boxed{1} \\ \text{BIND } \boxed{2} \end{bmatrix}$$

The unification of the three constraints of srhd-fill-phrase gives us (59).

$$(59) srhd-fill-phrase = \begin{cases} HEAD & 1 \\ VAL & 6 \\ SLASH & 4 \oplus 5 \\ SRHD none \\ HDTR & \begin{bmatrix} HEAD & 2 \\ VAL & ebag \\ SLASH & 4 \oplus \{3\} \end{bmatrix}$$
$$HEAD & 3 \\ VAL & ebag \\ SLASH & 5 \\ SRHD & \begin{bmatrix} HEAD & 3 \\ VAL & ebag \\ SLASH & 5 \\ SRHD & \begin{bmatrix} HEAD & 1 \\ VAL & 6 \oplus \{2\} \end{bmatrix} \end{bmatrix}$$

Given that SUPERHEAD is a verb-transferer, we obtain the equivalent of the *clause* constraint (47b), that is, that $\square = ebag$ and $\square = [UPPOS verb-pos]$. We give only the last step of the computation involving the *srhd-fill-phrase*.

From the DG point of view, the *srhd-fill-phrase* corresponds to two simultaneous combinations of words: the NHDTR's SUPERHEAD with the HDTR and the NHDTR with an element of HDTR|SLASH:



Let's exemplify the *srhd-fill-phrase* with the computation of the French relative clause in (61). Its analysis is given in (62) and its dependency interpretation, in (63).

(61) *la personne à la mère de qui je veux que tu parles* the person to the mother of whom I want that you speak 'the person whose mother I want you to speak to'

To conclude, let us recapitulate our treatment of constraints on extraction and piedpiping. We have defined three subtypes of syntactic relations: *gaprel*, *slashrel* and *whrel*; *gaprels* allow for the dependent to be a gap, that is, the dependent to be slashed; *slashrels* allow the extraction, that is, a non-empty SLASH to percolate; and *whrels* allow the pied-piping, that is, a non-empty SUPERHEAD to percolate. For the major syntactic relations of English, we obtain the inheritance graph (64).

(62) An example of srhd-fill-phrase

(à la mère de qui je veux que tu parles)	
HEAD	$ \begin{bmatrix} MOD & N : [i] \\ UPPOS & adj-pos \\ TRANSF [2] \end{bmatrix} $
VAL	ebag
SLASH	ebag
SRHD	none
HDTR	(je veux que tu parles) HEAD [] VAL ebag SLASH []
NHDTR	$\begin{bmatrix} (\dot{a} \ la \ mere \ de \ qui) \\ HEAD \end{bmatrix} \begin{bmatrix} POS & tr-pos \\ TRANSF \ [N] \end{bmatrix}$ $VAL \ ebag$ $SRHD \begin{bmatrix} (qu-) \\ HEAD \end{bmatrix}$



(64) Syntactic relations' inheritance graph for English



This completes the presentation of our alternative treatment of extraction in HPSG. We hope that we have convinced the reader that this description using only the combination of word descriptions and no phrasal descriptions works at least as well as earlier approaches.

5 Status of phrases in the syntax-semantics interface of HPSG

In this paper, we have argued that, with no substantial modifications, the syntax-semantics interface of HPSG can be viewed as a true dependency grammar, that is, a grammar where all of the information about sentences comes from the information on words and their combination.³⁵ We have shown that even puzzling phenomena such as extraction can be described exclusively in terms of combinations of words. In such a grammar, no linguistic information is attached to objects other than words and in particular, no linguistic information is attached to objects such as phrases. (It is no longer the case in the syntax-phonology interface, where linearization and grouping for prosody are handled.)

Adopting our view of syntax, it can be said that syntactic phrases are simply the result of particular word combination procedures. Indeed, HPSG is based on a bottomup process: HPSG allows for a word to combine with a dependent word only when this dependent has combined with all of its own dependents (and its VAL is empty). In other words, the underlying tree of an HPSG analysis is built from the leaves to the root. Such a procedure imposes the restriction that a word can combine with its governor only when its projection has been processed. Consequently, we claim that syntactic phrases are simply the result of a particular way of processing and that HPSG contains in itself this way of processing, which might be avoided.

A syntactic "representation" is always linked to processing. The syntactic dependency tree stores the information on how words combine with each other; a syntactic dependency testifies to the combination of two words and indicates which one is the HDTR (= governor) and which one is the NHDTR (= dependent). Nevertheless, a dependency tree is less procedural than a phrase structure tree: it does not store the information on the order in which words have been combined during the process. And we think that this last information is not relevant from a linguistic point of view.

We will now see that, even with a formalism such as HPSG, it is possible to use a strategy other than the bottom-up process. For instance, dealing with a sentence such as *Peter wants to eat beans*, we can begin with combining *wants* and *to eat*. For that, we must allow for a word (here *wants*) to combine with a dependent (here *to eat*) whose valence is not filled. This can be done with a schema such as (65).³⁶

(65)
$$hd$$
-act-phrase :
$$\begin{bmatrix} VAL & \boxed{1} \oplus \boxed{3} \\ HDTR & \begin{bmatrix} VAL & \boxed{1} \oplus \{2\} \end{bmatrix} \\ NHDTR & \begin{bmatrix} HEAD & \boxed{2} \\ VAL & \boxed{3} \end{bmatrix} \end{bmatrix}$$
 is a hd -phrase

Here we do not require the VAL(ENCE) of NHDTR to be empty; this valence is added to the valence of HDTR. This idea comes directly from categorial grammars (see the Forward Partial Combination rule in Ades & Steedman 1982: 527). The SLASH feature has been introduced to allow such a combination; we distinguish the VAL coming from HDTR from the VAL coming from NHDTR, which is put in SLASH. Such a

³⁵ As said previously (Footnote 1, p. 113), it would be better to handle lexical units and grammatical morphemes rather than words and to obtain word descriptions by combinations of lexeme and grammatical morpheme descriptions. This is another point, which cannot be discussed here; we think that the architecture of HPSG is not completely convenient, due to the fact that grammatical morphemes are dealt with by lexical rules and therefore are not considered as signs (and are not assigned to a feature structure combining by unification).

 $^{^{36}}$ It is not the first time one proposes to introduce such a schema in HPSG. See, for instance, the handling of auxiliaries by Hinrichs & Nakazawa (1994).

schema should be constrained in the same way as SLASH introduction and SLASH percolation are constrained. It could be useful for dealing with phenomena such as gapping (*Peter wants to eat beans and Mary, spaghetti*) or word order in languages such as German, where verbs can form a cluster, the order of their dependents being free ($da\beta$ *niemand Bohnen essen will* 'that nobody beans eat wants' vs. $da\beta$ Bohnen niemand essen will).³⁷

The use of phrases in syntactic structures has been argued for based on word order configurations. But in non-configurational languages such as Russian or German, observable groupings do not correspond to syntactic phrases (= projections of heads) and another structure must be invoked. For example, in the description of German, a structure other than phrase structure, the topological model, is used to compute word order (see Kathol 1995 for a formalization in HPSG and Donohue & Sag 1999 for an application to Warlpiri). We do not deny that it could be useful to introduce particular groupings to compute word order. But in any case, the introduction of a particular type of grouping must be assumed only for the treatment of word order and prosody.

We will conclude by showing that even the SLASH feature, which seems to be indispensable for the HPSG analysis of extraction, results from the underlying process presupposed by HPSG, which consists of combining words according to the traditional phrase structure. Consider the topicalized sentence in (66). In HPSG, following the traditional X-bar Theory, the sentence is obtained by a combination of words giving us the phrase structure indicated in (66a).³⁸

(66) a. [Sandy (I know [you said (Peter hates)])]b. [([(Sandy) I know] you said) Peter hates]

But it is also possible to obtain this sentence by a combination of words giving us the grouping in (66b). In this case, we must first combine *Sandy* and *I know*. We can allow that, but as *Sandy* is not a syntactic dependent of *know*, it will be added in a special feature we call VIS(ITOR).³⁹ A new visitor is introduced with the *head-visitor-phrase*. In this schema, NHDTR|HEAD does not combine with HDTR|VAL but is simply added to the VIS(ITOR) bag:



³⁷ Word order is actually never free. It can depend on factors such as information structure and be not constrained by purely syntactic properties such as the height of the complements *Peter* and *Bohnen* 'beans' in the dependency tree (contrary to Dutch, which is a closely related language).

³⁸ X-bar Theory encodes, via the phrase structure and the c-command, the syntactic relations between words. But X-bar phrase structure is not useful for that in HPSG, where the syntactic relations between words are encoded in VAL and MOD. We hope that this point, extensively developed in Section 2, is absolutely clear to the reader.

³⁹ Hudson (1990, 2000) uses a similar relation to link the main verb with an extracted phrase.

Next, the visitor is lifted to the verb governing the gap and fills the gap using the following schema *generalized-head-actant-phrase*:

(68)
$$g$$
-hd-fill-phrase :
$$\begin{bmatrix} VAL & 1 \oplus 3 \\ VIS & 4 \\ HDTR & VAL & 1 \oplus \{2\} \\ VIS & 4 \oplus 5 \end{bmatrix}$$
 is a hd-phrase NHDTR
$$\begin{bmatrix} HEAD & 2 \\ VAL & 3 \oplus 5 \end{bmatrix}$$

This schema generalizes *hd-act-phrase* (65). (It is equal to it when $\square = ebag$ and $\square = ebag$.) It allows us to lift visitors (if $\square \neq ebag$) and fill gaps (if $\square \neq ebag$). When Sandy I know and you said are combined, $\square = \{Sandy\}$ and $\square = ebag$; when Sandy I know you said and Peter hates are combined, $\square = ebag$ and $\square = \{Sandy\}$. The major lines of the VISITOR-like analysis of (66) are given in (69), next page.

The analysis in (69) has the advantage of allowing us to parse the sentence incrementally (that is, to build its dependency tree and its semantic structure as soon as possible) and therefore it is better motivated cognitively. Contrary to the SLASH analysis, no lexical rule (or an equivalent device) is needed to introduce the VISITOR. Finally, this analysis is compatible with the SUPERHEAD analysis presented in Section 4.2; it simply need to replace the *head-visitor-phrase* by a *superhead-visitor-phrase* where NHDTR|SRHD takes HDTR as actant (and NHDTR becomes a VISITOR).

SLASH and VISITOR are linked to two different word groupings corresponding to different processes in the combination of words: in one case, the gap is lifted up to the filler by way of SLASH, while in the other, the filler is lowered down to the gap by the way of VISITOR. Kahane (2001) proposes a formalization that subsumes these two analyses. This more generic formalization is less procedural and does not presuppose any particular word grouping. Such a formalism is associative, in the mathematical sense of the term: it allows for the dependency tree to be built in whatever direction, while in the traditional HPSG the dependency tree is necessary built bottom-up, that is, from the leaves to the root.⁴⁰ The implementation of an associative analysis necessitates a serious revision of the foundations of HPSG.

 $^{^{40}}$ Rather than lift an element from the verb introducing the gap to the main verb, Kahane (2001) treats the extraction path—the string of verbs (or verb-equivalent forms)—as a syntactic unit (called a *verbal nucleus*).



6 Conclusion

For HPSG users, the main result of this paper is certainly the demonstration that, in the same way as the *head-daughter-phrase* can be interpreted as the combination of two words, the traditional *head-filler-phrase* (the schema of combination of a filler phrase and a clause with a gap) can be interpreted as a schema involving two simultaneous combinations of words where the two daughters act as heads, each one in a different combination. The filler phrase, by the way of the *wh*-word, controls the distribution of the clause (allowing it to be a relative or an interrogative clause and to modify a noun or to occupy an actantial position of an interrogative verb) and is thus the true head of the clause, but, simultaneously, it fills a gap in the clause and is thus governed by the clause. We have solved this problem by giving a double role to the *wh*-word, a pronominal role and a verb-transferer role. Contrary to earlier analyses of extraction in HPSG, we do not introduce special schemata for each type of extraction and we do not allow for direct introduction of linguistic information in a phrasal schema.

Our study proves that phrasal descriptions are not needed in the analysis of extraction and moreover that the recourse to phrasal descriptions masks what combinations of words are involved in extraction. From a theoretical point of view, we think that purely linguistic data might be introduced in word descriptions and that phrasal schemata might only encode how subphrases combine together.⁴¹ This gives us a fully lexicalist version of the syntax-semantics interface of HPSG, where the phrase structure simply stores the way in which words have combined. As a consequence, phrases have lost any linguistic status. They are simply a result of a particular procedure of word combination, so that another procedure would give us another grouping of words. Of course, it is possible to give a theoretical status to the phrase structure considered by traditional HPSG and inherited from X-bar Theory: a phrase is the maximal projection of a word, that is, a word with all the phrases headed by its dependents. It is even easy to introduce other types of phrases by considering partial projection that involves only a part of the dependents (excluding, for example, the subject). What we say is that such a phrase structure is simply a by-product of the dependency structure and does not play any particular role other than a procedural one.

For DG users, the main result is a proposal of an implementation of a DG in the welldefined formalism of HPSG. To be exact, the grammar proposed here is not a genuine DG—it allows for a dependency structure that is not a true tree, which results of the double role of the *wh*-word. Such an analysis needs an extension of traditional DG formalisms and it is interesting to see that the formalism of HPSG allows elegant solutions.

For all, we hope that this paper proposes a better understanding of both HPSG and DG and particularly the way they model extraction.

Acknowledgments

I thank Mark Dras, Igor Mel'čuk, Piet Mertens, Alain Polguère and Jesse Tseng for many valuable comments on earlier versions of this paper.

References

Ades, Anthony & Steedman, Mark 1982. "On the Order of Words". *Linguistics and Philosophy* 4: 517-558.

Bloomfield, Leonard 1933. Language. New York.

Bouma, Gosse, Malouf, Robert & Sag, Ivan 2001. "Satisfying Constraints on Extraction and Adjunction". *Natural Language and Linguistic Theories* 19:1-65.

Candito, Marie-Hélène & Kahane, Sylvain 1998. "Defining DTG Derivations to Get Semantic Graphs". *Proceedings TAG+4 Workshop*, 25-28. Philadelphia.

Donohue, Cathryn & Sag, Ivan 1999. "Domains in Warlpiri". HPSG 1999, 101-106. Edinburgh.

Gaifman, Haim 1965. "Dependency Systems and Phrase-Structure Systems", *Information and Control* 18: 304-337 [= Technical Report, Rand Corporation, 1961, RM-2315].

⁴¹ It is not exactly the case that all constructions, in the sense of Goldberg 1995, are lexically marked and in some case some linguistic material can not be easily associated to a particular word. It is the case of *that*-less relatives in English. But even in such cases, it seems preferable to introduce a "lexical" sign without phonological realization rather than introduce a specific phrasal schema.

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.

Ginzburg, Jonathan & Sag, Ivan 2000. *Interrogative Investigations: The Form, Meaning, Use of English Interrogatives*. Stanford/Chicago: CSLI Publications/University of Chicago Press.

Goldberg, Adele 1995. A Construction Grammar Approach to Argument Structure. Chicago: University of Chicago Press.

Gross, Maurice 1968. *Grammaire transformationnelle du français. Syntaxe du verbe.* Paris: Larousse.

Jespersen, Otto 1924. *The Philosophy of Grammar*. Chicago: University of Chicago Press.

Joshi, Aravind 1987. "Introduction to Tree Adjoining Grammar". In *The Mathematics of Language*, A. Manaster-Ramer (ed.), 87-114. Amsterdam/Philadelphia: Benjamins.

Haspelmath, Martin 1997. Indefinite Pronouns. Oxford: Clarendon Press.

Hays, David 1964. "Dependency Theory: A Formalism and Some Observations". *Language* 40 (4): 511-525.

Hinrichs, Erhard & Nakazawa, Tsuneko 1994. "Linearizing AUXs in German Verbal Complexes". In *German In Head-Driven Phrase Structure Grammar*, J. Nerbonne, K. Netter, C. Pollard (eds.), 11-38. Stanford: CSLI Publications.

Hudson, Richard 1980. "Constituency and Dependency". Linguistics 18: 179-198.

Hudson, Richard 1990. English Word Grammar. Oxford: Blackwell.

Hudson, Richard 2000. "Discontinuity". In *Dependency Grammars*, *T.A.L.* 41 (1), S. Kahane (ed.). Paris: Hermès.

Jackendoff, Ray 1977. X Syntax: A Study of Phrase Structure. Cambridge, MA: MIT Press.

Kahane, Sylvain 1996. "If HPSG Were a Dependency Grammar ...", *Proceedings TALN'96*, 45-49. Marseille.

Kahane, Sylvain 1997. "Bubble Trees and Syntactic Representations". In *Proc. 5th Meeting of the Mathematics of Language (MOL5)*, T. Becker & H.-U. Krieger (eds), 70-76. Saarbrücken: DFKI.

Kahane, Sylvain 2001. "A Fully Lexicalized Grammar for French Based on Meaning-Text Theory". In *Computational Linguistics and Intelligent Text Processing, Proceedings of CICLing 2001*, A. Gelbukh (ed.), 18-31. Lecture Notes in Computer Science 2004, Berlin: Springer.

Kahane, Sylvain 2007. "A Formalism for Machine Translation in MTT, Including Syntactic Restructurings". In *Proceedings of the Third Meaning-Text Conference*, Klagenfurt. [To appear in *Wiener Slawistisher Almanach*, Sonderband 69, T. Reuther & L. Wanner (eds.), Vienna: Wiener Slawistisher Almanach.]

Kahane, Sylvain, Candito, Marie-Hélène & Kercadio, Yannick de 2000. "An Alternative Description of Extractions in TAG". *Proceedings of TAG+5 Workshop*, 123-28. Paris.

Kahane, Sylvain & Mel'čuk, Igor 1999. "La synthèse sémantique ou la correspondance entre graphes sémantiques et arbres syntaxiques. Le cas des phrases à extraction en français contemporain". *T.A.L.* 40 (2): 25-85.

Kahane, Sylvain, Nasr, Alexis & Rambow, Owen 1998. "Pseudo-Projectivity: A Polynomially Parsable Non-projective Dependency Grammar". *Proceedings COLING- ACL'98*, 646-652. Montreal.

Kaplan, Ronald & Zaenen, Annie 1989. "Long Distance Dependencies, Constituent Structure and Functional Uncertainty". In *Alternative Conceptions of Phrase Structure*, A. Baltin & A. Kroch (eds), 17-42. Chicago: University of Chicago Press.

Kathol, Andreas 1995. *Linearization-Based German Syntax*, Ph.D. thesis, The Ohio State University.

Kayne, Richard 1974-75. "French Relative que", Recherches Linguistiques, II: 40-61, III: 27-92.

Kroch, Anthony & Joshi, Aravind 1986. "Analysing Extraposition in a TAG", In *Discontinuous Constituents*, G. Huck & A. Ojeda (eds.), *Syntax and Semantics* 20, 107-149. Orlando: Academic Press.

Le Goffic, Pierre 1993. Grammaire de la phrase française. Paris: Hachette.

Le Goffic, Pierre 2002. "Marqueurs d'interrogation/indéfinition/subordination: essai de vue d'ensemble". *Verbum* 24(4): 315-340.

Lecerf, Yves 1961. "Une représentation algébrique de la structure des phrases dans diverses langues natuelles". *Comptes rendus de l'Académie des sciences de Paris* 252, 232-34.

Mel'čuk, Igor 1988. *Dependency Syntax: Theory and Practice*. Albany: State University of New York Press.

Moreau, Marie-Lyne 1971. "L'homme que je crois qui est venu ; qui, que : relatifs et conjonctions". *Langue Française* 11: 77-90.

Nasr, Alexis 1995. "A Formalism and a Parser of Lexicalized Dependency Grammars". *4th Int. Workshop on Parsing Technologies.* State University of New York Press.

Padučeva, Elena 1964. "O sposobax predstavlenija sintaksičeskoj struktury predloženija" [On methods of representing the syntactic structure of a sentence]. *Voprosy jazykoznanija* 2: 99-113.

Petkevič, Vladimir 1995. "A New Formal Specification of Underlying Representations". *Theoretical Linguistics* 21: 7-61.

Pollard, Carl & Sag, Ivan 1987. *Information-Based Syntax and Semantics*. Stanford: CSLI Publications.

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

Ross, John 1967. Constraints on Variables in Syntax. Ph.D. Thesis, MIT [printed as Infinite Syntax!, Dordrecht: Reidel, 1985].

Sag, Ivan 1997. "English Relative Clause Constructions". *Journal of Linguistics* 33: 431-484.

Sag, Ivan & Dean Fodor, Janet 1994. "Extractions Without Traces". In *Proceedings of the West Coast Conference on Formal Linguistics*, 13, 365-384. Stanford.

Sag, Ivan & Wasow, Thomas 1999. *Syntactic Theory: A Formal Introduction*. Stanford: CSLI Publications.

Tesnière, Lucien 1959. Éléments de syntaxe structurale. Paris: Klincksieck.

van Noord, Gertjan & Bouma, Gosse 1994. "Adjuncts and the Processing of Lexical Rules". *COLING*'94, 250-256. Kyoto.

XTAG Research Group 1995. "A Lexicalized TAG for English". Technical report IRCS 95-03. University of Pennsylvania (On-line updated version).

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

Establishing an Inventory of Surface-Syntactic Relations: Valence-Controlled Surface-Syntactic Dependents of the Verb in French

Lidija Iordanskaja and Igor Mel'čuk

1 Introduction

The goal of the present paper¹ is to sketch a general method for establishing an inventory of labeled Surface-Syntactic Relations [= SSyntRels²] for a language **L**. More specifically, such a method must allow the researcher to determine the set of SSyntRels in **L** and distribute relevant SSynt-constructions of **L** among them, that is, decide by which SSyntRel a given SSynt-construction must be described. As a test case, we have chosen a particular subset of SSyntRels in French—all SSyntRels controlled by the active valence of the verb, i.e., by its Government Pattern. (The *active valence* of a verb is the set of all its complements, including the Subject—that is, of all its SSynt-Actants; the *passive valence* of a verb is the set of all constructions in which it can be used as a dependent.)

The research is carried out within the general framework of Meaning-Text theory (see, e.g., Mel'čuk 1974, 1988: 43-91, 1997). We take for granted the levels of linguistic representation assumed by the theory and the corresponding notions. (See, however, Mel'čuk's text in this volume, p. 2*ff*.)

From a formal viewpoint, a SSyntRel **r** is a particular type of syntactic dependency relation holding between two lexemes in a SSynt-Structure [= SSyntS], a G(overnor) and its D(ependent): $G-r \rightarrow D$. A SSyntS formed by SSyntRels is a tree; it respects the following two principles: the unicity of the Synt-Governor (= no node can have more than one G) and the presence of the unique Synt-head, or *top node*, of the whole SSyntS (= there is one and only one node on which all other nodes depend, directly or indirectly).

From a linguistic viewpoint, the name \mathbf{r} of a SSyntRel specifies a family of SSyntconstructions of \mathbf{L} that feature sufficiently similar linguistic properties—i.e., they show "family resemblances." A SSynt-construction represents a set of binary phrases (a binary phrase being, roughly, two wordforms linked by a direct SSynt-dependency) and is specified by indicating 1) its members—a pair of part-of-speech symbols, perhaps

¹ A shorter version of this paper was published as Iordanskaja & Mel'čuk 2000.

² For abbreviations used in this paper, see p. xi.

with some additional syntactic features (such as "copula," "transitive," "anteposed," etc.), supplied with corresponding inflectional markings, 2) their mutual linear order and 3) the direction of SSynt-dependency between them. Thus, the French construction "adjective — noun," showing obligatory agreement of the adjective in $\mathbf{g}(\text{ender})$ and $\mathbf{n}(\text{umber})$ with the noun, appears as follows:

$$\text{ADJ}_{(\text{antepos})\mathbf{g}, \mathbf{n}} + \text{N}_{(\mathbf{g})\mathbf{n}}$$

Subscripts without parentheses stand for inflectional values; subscripts in parentheses represent syntactic features, specified in the corresponding lexical entry.

This construction represents such phrases as *différents éléments* 'different elements' and *belles fillettes* 'beautiful little.girls'. Note that $N_{(g)n} + ADJ_{(postpos)g,n}$ —équation différentielle 'differential equation', gouvernement américain 'American government', etc.—is a different SSynt-construction of French, because of different word order.

The set of SSyntRels for L must meet at least the following three formal requirements, which have to be satisfied in the SSyntS:

1. Ensure the preservation of all semantic contrasts which appear on the semantic level of utterance representation and are formally expressed on the surface, but which cannot be taken care of by other entities of the syntactic level (for instance, by the syntactic-communicative structure).

2. Ensure the appropriate substitutability (within a SSyntS) of the constructions described by a SSyntRel.

3. Ensure the appropriate combinability of the SSyntRels that share the same SSynt-Governor.

These requirements are, *mutatis mutandis*, the same as those underlying all *-emic* units, or X-emes, in natural language—e.g., phonemes and morphemes. Therefore, the family of constructions covered by one SSyntRel is, so to speak, a "construction**eme**." The elements of a constructioneme—concrete constructions—behave as all other allo-Xs do: they do not contrast semantically (requirement 1), and show the same regularities in their substitutability (requirement 2) and combinability (requirement 3).³

The above requirements can be formalized as three criteria for postulating different SSyntRels in **L**. Such criteria are needed, in particular, because, in contrast to Deep-Syntactic Relations [= DSyntRels], the SSyntRels are not universal: the set of SSynt-Rels has to be established empirically, individually for each **L**, very much like the set of phonemes or of inflectional values, i.e., grammemes (e.g., tenses or grammatical cases). Consequently, one needs some principles on the basis on which different SSyntRels should be distinguished in **L**.⁴

 $^{^3}$ When speaking of the properties of constructions, we mean, in point of fact, the properties of their lexical realizations: two constructions contrast semantically if it is possible to fill them with the same lexemes and obtain semantically different phrases. For more, see 2.2, Criterion 1, p. 162.

⁴ For one of the earliest theoretical characterizations of the notion of SSyntRel, see Apresjan *et al.* 1978: 255-265; it was developed and enriched in Apresjan *et al.* 1984-1985, No. 155: 4-11.

This paper is limited to a subset of the SSyntRels of French: it considers only the SSyntRels that are necessary for the description of the SSyntS of all verbal constructions in Modern French that involve the surface realizations of the verb's DSynt-actants [= DSyntAs]. In other words, we will be dealing only with valence-controlled SSynt-Ds of a verb. Any type of valence-controlled SSynt-**D** and the corresponding SSyntRel must be foreseen in the lexicographic description of the verb, i.e., in its Government Pattern. Thus, as stated at the beginning of the paper, the SSyntRels introduced below all correspond to the active valence of French verbs.

In the framework of dependency syntax, lists of labeled SSyntRels have been proposed for various languages: Russian (cf. Mel'čuk 1962: 47-87, 1963: 490-493, 1964: 20-24 and 1974: 221-235); English (Mel'čuk & Pertsov 1987: 85-156, Apresjan *et al.* 1989: 71-121); German, Polish, Danish, Bangla, Finnish, Hungarian, Japanese, and Esperanto (Maxwell & Schubert 1989); and French (Apresjan *et al.* 1984-1985, Isaac 1986, Candito 1999). If we feel the need to return to the task, this is because we find the following three drawbacks in previous attempts:

First, the decisions made with regard to specific SSyntRels were not systematically justified. Now we would like to supply the reasons for postulating this or that SSyntRel according to relevant linguistic properties of phrases under consideration and to the formal criteria for the differentiation of SSyntRels (introduced in Mel'čuk 1979: 99-143 and 1988: 141-144 and reworked in this volume, p. 25*ff*).

Second, SSyntRels were established for each language involved more or less in isolation from typological considerations and with no particular regard for the general theory of surface syntax. Now the time seems ripe for taking into account modern developments of syntactic theory and thus making the inventory of SSyntRels we are proposing theoretically "clean" and linguistically substantiated.

Third, in the Meaning-Text approach, verb-to-actant SSyntRels used to be too "semantic;" they were tailored to fit the DSyntRels (we mean, e.g., such SSyntRels as "1st completive," "2nd completive," etc., which were in one-to-one correspondence with DSyntRels). On the other hand, several other approaches known to us seem too formal: the SSyntRels are defined mostly by the distribution classes of their Ds. Now we would like to reconsider previous solutions under the angle of strict separation of levels—such that the SSyntRels be isolated based mainly on syntactic considerations, striking a necessary balance between the two extremes, excessive semanticism and excessive formalism.

The present paper belongs to the domain of linguistic research into what is known as grammatical relations and clause elements: the dependent member of a grammatical relation (in point of fact, a syntactic relation) is, roughly speaking, a clause element, so that the problem of distinguishing grammatical relations bears directly on clause elements and vice versa (for more, see below, Section 2.1). The corresponding literature is too vast to allow even a cursory review. We will limit ourselves to indicating just four publications of general character: Plank 1984, Dryer 1986, Hudson 1992, and Van Valin & LaPolla 1997: 242-309, where a rich bibliography can be found.

Speaking more specifically about French, the topic of establishing/distinguishing clause elements is here also well studied. Again, it is impossible for us to undertake a

survey of existing works. As our principal source of data we have made extensive use of the classic reference book Grevisse 1993 [13th edition] and a monograph by P. Le Goffic (1993). Also, we have drawn some additional data from Gross 1968 and 1975, Boons *et al.* 1976, Kayne 1977, Apresjan *et al.* 1984-1985, and Candito 1999. (Candito 1999 is especially close to our own study: the author proposes a set of SSyntRels for French and Italian, supplying for them syntactic justifications.)

Now we can state the goal of the paper in more precise terms:

To propose a method that would ensure a partition of the set of the syntactic constructions of French such that each subset obtained in this way

1) is sufficiently homogeneous—that is, all the constructions that belong to it share a sufficient number of relevant linguistic properties; and

2) is saturated—that is, it contains all and only the constructions that can be brought together in accordance with some pre-established formal criteria (see below, 2.2, p. 162*ff*, Criteria 1-3).

We want such a method to be satisfactory from the viewpoint of general linguistic theory. More precisely:

• The method must produce a set of SSyntRels/clause elements for **L** such that it would be typologically valid. For instance, consider the hierarchy of major nominal clause elements established by Keenan & Comrie (1977), see below, p. 156; it is preferable to obtain such SSyntRels/clause elements that satisfy this hierarchy. To meet the requirement of typological plausibility, we need a good selection of relevant properties.

• The method must (more or less) correspond to the universally accepted methods for establishing linguistic units at all levels. Criteria 1-3 constitute an attempt to ensure such universality.

Given its methodological orientation, our paper does not present new facts about French syntax. Our main tool—linguistic properties of clause elements—is not new, either; everyone is using more or less the same set of properties. What is particular to our approach is its systematicity. We have put systematicity above completeness or even factual truth. Our linguistic data are of course incomplete, as is the list of properties we exploit. As a result, some of the solutions we propose might turn out wrong. Yet we believe that a coherent general picture of what an inventory of SSyntRels in a language must look like is, at least for the time being, quite crucial; one cannot successfully work out a host of details until the organization of the whole domain is made sufficiently clear.

The structure of this paper is straightforward: Section 2 proposes a method for establishing an inventory of SSyntRels for a language and then Section 3 gives a partial list of SSyntRels for French, which have been isolated following this method; Section 4 contains a synoptic overview of the SSyntRels proposed in the form of tables and a condensed discussion of the hierarchy of SSyntRels; Section 5 offers a short conclusion.

2 A method for establishing an inventory of SSyntRels for a language

When introducing SSyntRels for a language, a SSyntRel to be postulated must satisfy two types of requirements: a substantive, or linguistic, requirement; and a set of formal, or logical, requirements.

2.1 *The substantive requirement to be satisfied by a SSyntRel*

The relevant linguistic properties of SSyntRels, that is, of their Ds, are linked to the fact that SSyntRels are subject to three types of constraints.

• First, a SSyntRel is associated to the semantic role its D fulfills—via the corresponding DSynt-relation (which, in its turn, is related to a semantic role). Therefore, SSyntRels must be such as to allow our linguistic model to compute them in a convenient, straightforward way from the DSyntS.

• Second, SSyntRels are combined within the SSyntS, where they show a particular behavior with respect to each other (omissibility, cooccurrence, paraphrastic relations). Therefore, SSyntRels must be such as to allow for convenient verification of the well-formedness of the SSyntS.

• Third, SSyntRels are aimed at linearization, prosodization (punctuation, in written texts), and morphologization of the SSyntS. That is, a SSyntRel must ensure convenient computation of word order for its D, along with the appropriate syntactically-induced prosody/punctuation and syntactically-induced inflections (= agreement and government, cf. Mel'čuk 1993 and 2006: 31*ff*). As a result, the SSyntRels are constrained from "below" (= by the DSyntS), from "the side" (= by the SSyntS), and from "above" (= by the Deep-Morphological Structure [= DMorphS]). Therefore, all possible Ds of a SSyntRel must have identical or very similar linguistic properties with respect to these three types of constraints.

Each SSyntRel is thus characterized by some specific properties of its dependent member D. In conformity with what has just been said and following Keenan's work (1976), we distinguish three types of such properties:

1) Syntactic-semantic properties (Keenan's "semantic interpretation properties"): properties of Ds from the viewpoint of the correspondence to a deeper level of representation—in our case, to the DSynt-level (and further to the Sem-level).

2) Purely syntactic properties (Keenan's "behavior/control properties"): properties of Ds from the viewpoint of the SSyntS.

3) Syntactic-morphological properties (Keenan's "coding properties"): properties of Ds from the viewpoint of their expression in the DMorphS.

All these properties can be readily described as values of a set of parameters, relevant for a given language; cf. below.

Now, as we have already said, the D of a particular SSyntRel is nothing else but what is known in the traditional grammar as a clause element (cf., e.g., Quirk *et al.* 1991: 59*ff*): the D of the **subjectival** SSyntRel is the Subject, the D of the **direct-objectival** SSyntRel is the Direct Object [= DirO], and so on.⁵ Thus, establishing different SSyntRels boils down to defining different clause elements. For this task, we can avail ourselves of two interesting results obtained in syntactic typology:

• A clause element is defined by a bundle of linguistic properties of the above-mentioned types (Mel'čuk & Savvina 1974 [1978], Keenan 1976, Borg & Comrie 1984;⁶ cf. also Quirk *et al.* 1991: 723*ff*, where such definitions are supplied for main clause elements in English, Lazard 1994b: 101*ff*, with a list of relevant properties of subjects, mainly, but not exclusively, in French, as well as Bresnan 2001: 5-10, 94*ff*, 302*ff* and Dalrymple 2001: 8-44). Therefore, the labeled SSyntRel that corresponds to a particular clause element must be defined by the same bundle of properties. Consequently, SSynt-Rels will be "multifactor" (Keenan 1976: 323), or cluster, concepts.

• Clause elements form a hierarchy such that an element of a higher stand has some relevant properties which no lower element has—but not vice versa (Keenan & Comrie 1977: 66, 1979):⁷

It is natural to require that the corresponding labeled SSyntRels of L form the same hierarchy. (At the end of the paper, we will briefly return to this hierarchy, known as Keenan-Comrie hierarchy: see p. 222.)

Taking into account these results, we have chosen for French the following sixteen linguistic properties, or parameters, relevant for the description of the valence-controlled SSyntRels (and, of course, of the corresponding clause elements) in the constructions that have a verb as the SSynt-head.

Relevant Properties of the Dependents of Valence-Controlled SSyntRels in French

The properties on the list are parameters admitting some pre-established values; most of them admit just two values: yes (= the property in question is present) or no. For each property P we indicate the syntactic classes (roughly, parts of speech) to which P is applicable, except for the cases when P is applicable to all syntactic classes: the default case.

⁵ The main difference between a system of SSyntRels in our approach and a system of traditional clause elements is that the former has to distinguish and, as a rule, does distinguish many more different SSyntRels than the latter distinguishes different clause elements.

⁶ "... each identified grammatical relation represents a clustering of syntactic properties in the language, sufficient to justify the internal cohesion of the grammatical relation and to set it off from other grammatical relations" (Borg & Comrie 1984: 109).

⁷ Let it be emphasized that this hierarchy is not a hierarchy of concepts or classes, where each lower element shows all the properties of all higher elements, but not vice versa. Therefore, it is by no means an inheritance hierarchy.

Syntactic-semantic properties of Ds

1. Corresponding to a particular DSynt-actant (of the Governor).

The DSyntA(ctant)s of a verb correspond to its semantic actants and are numbered in the order of growing obliqueness. Roughly speaking, DSyntA I corresponds, for instance, to the semantic Causer, the Actor or the Perceiver, DSyntA II—to the Patient, DSyntA III—to the Beneficiary or the Instrument, etc. (For more on Deep- and Surface-Syntactic actants, see Mel'čuk 2004.)

Purely syntactic properties of Ds

2. Being obligatorily present in any full-fledged clause (of course alongside the Main Verb [= MV]).

To avoid a misunderstanding, let us emphasize that we really mean EVERY clause; thus, for instance, a Direct Object, even if it is obligatory with some verbs, is not found in every clause of French (because not every French clause contains a transitive verb that requires a Direct Object).

3. Being the dependent of the MV only (that is, being unable to depend on anything but a finite verb).

4. Being implicated in SSynt-promotion/demotion (applicable to nouns, infinitives and subordinate clauses).

5. Being the target of cliticization of a specific type.

6. Being the target of relativization (applicable to nouns).

7. Being the target of clefting (applicable to nouns and infinitives).

8. Being the controller/target of reflexivization (applicable to nouns).

9. Being the controller of the SSynt-role of the Secondary Actor in the causative FAIRE-construction, namely imposing on it the SSynt-role of IndirO. (For more on the representation of the causative FAIRE-construction in French, see 3.1, No. 3, property 2, on p. 177.)

10. Being the controller of an actantial-attributive complement (applicable to nouns).

Syntactic-morphological properties of Ds

11. Being the target of non-specific morphological or lexical government.

We speak of non-specific government when a particular case form or a particular preposition is imposed on the D of a SSyntRel **r** by **r** itself—rather than by the lexical entry of its G, i.e., by the G's Government Pattern. Cf., for instance, case forms of the clitics le_{ACC} vs. lui_{DAT} and of the relative pronouns qui_{NOM} vs. que_{ACC} as a function of the subordinating SSyntRel (subjectival, direct-objectival and indirect-objectival); or else the selection of PAR 'by' for the D of the agentive SSyntRel. (For non-specific government, see Mel'čuk 1993: 321-322 and 2006: 45-46.)

12. Being the target of morphological agreement (applicable to adjectives: *Alain le trouve intelligent* vs. *Alain la trouve intelligente* ⁽Alain finds him/her intelligent⁾).

13. Being the controller of agreement of the MV.

14. Linear position with respect to G and/or to other $D_i(G)$ (not applicable to clitics and relative pronouns, whose ordering does not depend on the subordinating SSyntRel).

15. Impossibility of left dislocation.

In what follows, we deal only with "pure" left dislocation—namely, prosodic separation from the rest of the clause without use of a resumptive clitic. Note that the dislocated phrase must carry a mounting contour, contrary to fronted phrases characterized by a contrastive descending contour: **Rodin**, vous aimez ? lit. 'Rodin, do you like?', with a dislocated phrase, vs. **Douze ans elle avait**, pas treize lit. 'Twelve years she had, not thirteen', with a fronted phrase. (Thanks to C. Blanche-Benveniste for drawing our attention to this distinction.)

16. Particular prosody/punctuation (applicable to Direct Speech).⁸

These properties logically correspond to distinctive features used in phonology (to identify the allophones of the same phoneme and to oppose different phonemes), as well as in morphology, syntax and semantics.

The above list was established empirically and by consulting relevant publications. It clearly could be developed into a logical calculus of possibilities, which would help us for a better coverage. At present, it is far from complete. Thus, we are aware that it lacks, for instance, the following properties:

• Typical question (*Qui*? 'Who?', À qui ? 'To whom?', *Combien* ? 'How much?', etc.).

• Pseudo-Clefting of the type *Ce que* CLAUSE¹, *c'est que* CLAUSE²: *Ce qu'Alain cherche, c'est que tout soit en ordre* ^{(What Alain is looking for is that everything be in order⁾ (Candito 1999: 264).}

• Other types of pronominalization (*ce, cela, là-dessus*, etc.; Blanche-Benveniste 1975 and Eynde & Mertens 2003).

• Reaction to negation (*Alain ne lit pas de romans* 'Alain does not read novels', etc.; Abeillé 1997b: 23).

⁸ Using distributional and transformational properties to characterize syntactic entities is by no means a novel idea. For instance, such an approach has been developed for French and applied for a detailed description of thousands of verbs by M. Gross and his collaborators (Gross 1968, 1975, Boons *et al.* 1976, Guillet & Leclère 1992). They use about a hundred properties to specify the government patterns and isolate useful semantico-syntactic classes of verbs. Since our goals are essentially different (we are interested in SSynt-roles of Dependents, rather than in classes of Governors), our set of relevant properties is also different. However, a few intersections occur: thus, "Antéposition des compléments prépositionnels" of Boons *et al.* 1976: 200-201 corresponds to our "Impossibility of left dislocation."

• Control of anaphora (*On les présentera les uns aux autres* lit. ⁽Someone them will introduce to each other⁾ vs. **On leur présentera les uns les autres* lit. ⁽Someone to them will introduce each other⁾; Abeillé 1997a: 25).

• Possible/impossible extraction (*le garçon dont le père est en prison* 'the boy whose father is in jail' ~ *le garçon dont Alain connaît le père* lit. 'the boy whose Alain knows the father' vs. **le garçon dont Alain plaît au père* lit. 'the boy whose Alain is-likable to the father'; Abeillé 1997a: 23).

- Right dislocation.
- Mutual ordering of codependents of the same governor.
- Possibility/impossibility of inserting a parenthetical.

Many additional relevant syntactic properties of the dependents of the French verb are found, in particular, in Gross 1968.

As the next important step, we have to introduce the notion of *prototypical* D of a given SSyntRel. Since we are working with four major parts of speech—N(ouns), V(erbs), A(djectives), and ADV(erbs), we will group the SSyntRels into four major prototypical-dependent classes: those whose prototypical D is a N, those with the prototypical D being a V in the infinitive, those having an A as their prototypical D, and, finally, those where the prototypical D is an ADV. To this, we add a fifth class with the D being a full utterance (for Direct Speech). As we will see below, within each class, the SSyntRels feature many similar properties, so that our grouping is justified. Now, what exactly is the prototypical D of a SSyntRel?

Definition 1: Prototypical dependent of a given SSyntRel

A *prototypical* D of a SSyntRel **r** is a D of such a syntactic class (\approx part of speech) that this D can be used with any G(overnor) possible for **r**.

Thus, the prototypical D of the subjectival SSyntRel is a (prepositionless) noun. Although a D of this SSyntRel can be not only a noun, but also an infinitive and a subordinate clause (*Fumer nuit à la santé* lit. 'To-smoke harms to the health'; *Qu'Alain ne soit pas là nous inquiète* lit. 'That Alain should not be here bothers us'), any MV in French admits an N (or an impersonal pronoun) as its Subject.

The above 16 linguistic properties that characterize French SSyntRels constitute the maximal set; for a particular class of SSyntRels only a subset of these properties may be relevant—in conformity with the prototypical D of this class. Thus, if the prototypical D of a SSyntRel is an A, the property "Being the target of reflexivization" is simply not applicable: adjectives in French cannot be reflexivized. Therefore, when introducing a SSyntRel **r**, we characterize it according to the properties which are relevant for its prototypical D. We indicate first those properties—or combinations thereof—whose specified values accrue only to (all Ds of) this **r** and which thus distinguish it from all the other SSyntRels of the same prototypical-dependent class: these are *defining* properties of **r**. Afterwards, we supply other relevant properties, which are, however, not

exclusive to (the Ds of) \mathbf{r} within the same prototypical-dependent class: *descriptive* properties. (Lazard 1994b: 68-77 proposes a similar, although different, distinction between "prime" and "secondary" properties.)

The subdivision of SSyntRels into these five prototypical-dependent classes is similar to the subdivision of phonemes into vowels, consonants, and glides; or to the subdivision of inflectional categories into nominal, verbal, adjectival, and adverbial ones; or else to the subdivision of lexemes into semantic classes "action," "event," "substance," etc.

Now we can formulate the substantive requirement a SSyntRel must satisfy:

Similarity of Relevant Linguistic Properties

Any D of a SSyntRel \mathbf{r} 1) must share at least some defining properties with the prototypical D' of \mathbf{r} and 2) must not contradict any property of D'.

A D of **r** does not contradict a property P of **r** if either it has the same value of P as the prototypical D' of **r** or P is not applicable to this D. Thus, if two Ds of a SSyntRel belong to different parts of speech, some of the properties of **r** may be simply inapplicable to one of them and so there is no contradiction.

Let us illustrate the above requirement with an example. The prototypical Direct Object, an N, has the following four defining properties:

1) It can be promoted to the Subject status by passivization (Property 4).

2) It can be replaced with a clitic in the accusative (Property 5).

3) It controls the SSynt-role of the Secondary Actor in the causative FAIRE-construction: it does not admit the expression of the Secondary Actor as Direct Object (Property 9).

4) When preceding the MV—that is, when replaced with a clitic or a relative pronoun, it controls the number-gender agreement of the past participles in compound verb forms (Property 13).

Based on these properties, we can consider as Direct Objects not only nouns, but some infinitives as well (introduced by a preposition or prepositionless). Let us consider the infinitive in sentence (1):

(1) [*Tout le monde*] *propose*−dir-obj→*de partir* ⁽Everybody proposes to leave⁾.

This infinitive has all four defining properties of the prototypical DirO: 1) **Partir** a été proposé par tout le monde 'To leave was proposed by everybody'; 2) **Partir**, tout le monde le propose lit. 'To leave, everybody proposes it'; 3) Cela lui $\langle *le \rangle$ fait proposer de **partir** demain lit. 'It makes to-him propose to leave tomorrow'; 4) **Partir**, tout le monde l'a **proposé** [MASC, SG] lit. 'To leave, everybody has proposed it'. At the same time, this infinitive does not contradict any property of the prototypical DirO: for a given property, it either features the same value or the property is inapplicable to it. In French, an
infinitive, whatever its syntactic role, cannot be the target/controller of reflexivization, but this does not prevent us from considering some infinitives as DirOs, along with nouns, which do control reflexivization and can be reflexivized.

The requirement of similarity of relevant properties calls for two important provisos. When checking the values of relevant properties for a given SSyntRel, one must always bear in mind the following two important facts about natural languages.

• If we say something true about a language, it is true everywhere—except for some particular cases, which must be explicitly identified. This happens, for instance, in all types of phraseologized expressions, where general rules and properties of the language can be suspended. Thus, in the notorious idiom kick the bucket the DirO bucket cannot be promoted to Subject-which is one of defining properties of DirOs (*The bucket was kicked, although the verb KICK has the passive). However, being part of a phraseme and as a result losing many of its relevant properties does not in principle prevent a DirO from being a DirO. Thus, in the French phrasemes, more precisely, collocations, faire *pipi* (make peepee), or *faire dodo* (be gone beddy-bye), PIPI and DODO have almost none of the properties of DirOs - and yet in the causative FAIRE-construction they behave as DirOs, requiring that the SSynt-expression of the Secondary Actor be an IndirO: Alain *lui* (**le*) *fait faire pipi/dodo* (Alain makes to-him go peepee/go beddy-bye) (Morin 1980: 206). We observe quite a similar behavior of DirOs in many even more opaque phrasemes: *lui* (**le*) *faire plier bagage* lit. (make to-him pack up and go), *lui* (**le*) *faire* rendre gorge lit. 'make to-him give back ill-gotten gains', lui (*le) faire tenir compte de cela lit. 'make to-him take this into account', etc. (Gaatone 1993: 40-41). In all such cases, the boldfaced noun has to be considered a DirO in order to ensure that the Secondary Actor in the causative construction is realized as an IndirO—in spite of the fact that inside the phraseme such a DirO lacks other typical properties of DirOs.

In the same vein, we find individual cases of impossible cliticization which run counter to our general statements. All such particular cases must of course be described in the lexical entries of the corresponding Governors, but we can safely ignore them in the process of establishing the set of SSyntRels of **L**.

• In many cases, a given clause element seems to lack a relevant property which it should have. This happens because other factors of a completely different nature intervene. Thus, cliticization of a concrete phrase can be precluded by its semantic and communicative features (for instance, by its non-referentiality). The same holds for left dislocation. Again, such cases should be ignored when introducing SSyntRels of L.

In cases in which, from the viewpoint of the above properties, the Ds of two presumed SSyntRels $\mathbf{r_1}$ and $\mathbf{r_2}$ are sufficiently similar although not identical, we use three formal criteria (stated in Section 2.2 below)—in order to see whether $\mathbf{r_1}$ and $\mathbf{r_2}$ can be collapsed into one SSyntRel \mathbf{r} or should be kept apart. For SSynt-constructions whose Ds manifest very dissimilar properties the question of a common SSyntRel does not even rise and the criteria need not to be recurred to.⁹

⁹ Such is, for instance, the case for the synonymous French constructions implemented by the phrases *l'aide canadienne* 'the Canadian aid' and *l'aide du Canada* 'the aid of Canada'. An agreeing adjectival modifier and a prepositional phrase are so dissimilar in their SSynt-properties that there is not the slightest temptation to describe them with the same SSyntRel. (Adjectival agreement and the presence of a preposition are among the most important syntactic properties in French.)

2.2 Formal requirements to be satisfied by a SSyntRel

Determining the set of SSynt-Relations in a language requires the use of three types of formal syntactic criteria: A. for SSynt-connectedness, B. for SSynt-dominance, and C. for the type of SSynt-Relations. All these criteria are described and discussed in Mel'čuk's text in this volume: p. 25*ff*, so that the reader can be referred there. However, since the present paper aims specifically at establishing the types of some SSyntRels for French and therefore Criteria C1-C3 are extensively used, we reproduce their formulations in the present paper (as Criteria 1-3) and illustrate them with French data.

The application of Criteria 1-3 presupposes that the presence and the direction of SSynt-dependency between two lexemes under consideration is already established. Criteria 1-3 help the researcher to decide, for any pair of binary phrases of language L (and of course for the pair of corresponding constructions) whether both can be described by the same SSyntRel **r**. These criteria formulate the requirements any SSyntRel must satisfy; if a presumed **r** does not satisfy at least one of them, it has to be split into **r**' and **r**''. If Criteria 1-3 do not require such a split, this by no means entails the acceptance of **r**: **r** still can be undesirable because it does not satisfy the requirement of similarity of the relevant properties of the dependent members of the phrases in question. Thus, Criteria 1-3 state only necessary but not sufficient conditions for grouping several SSyntconstructions under the same SSyntRel.

As mentioned above, these criteria are a specific variation of the three basic criteria used in linguistics for all -emes: semantic contrast (= Minimal Pair Test), substitutability, and combinability.

Criterion 1: Absence of semantic contrast

Notations: w(L) is a wordform of lexeme L (w_i and w_j can be different or identical); \oplus is the operation of linguistic union, which links signs, in particular—wordforms, according to their syntactics (i.e., their combinatorial properties) and general rules of L.

A SSyntRel must not describe two different phrases

 $\mathbf{w_1}(L_1) \oplus \mathbf{w_2}(L_2)$ and $\mathbf{w_3}(L_1) \oplus \mathbf{w_4}(L_2)$, where $L_1 \rightarrow L_2$,

which 1) contrast semantically $[{}^{(}\mathbf{w_1}(L_1) \oplus \mathbf{w_2}(L_2){}^{)} \neq {}^{(}\mathbf{w_3}(L_1) \oplus \mathbf{w_4}(L_2){}^{)}]$ and 2) differ formally only by some syntactic means of expression (i.e., by word order, by syntactic prosody, or by syntactic grammemes).

Examples

1. Two French phrases $Alain \leftarrow aime \dots$ 'Alain loves ...' and ... $aime \rightarrow Alain$ '... loves Alain' contrast semantically and differ only by word order; therefore, they must be described by different SSyntRels (the subjectival SSyntRel in the first one, and the direct-objectival in the second). Cf. also the discussion of the quotative-objectival SSyntRel below, p. 219.

2. Consider the phrases *semble* \rightarrow *fatigué* and *fatigué*, \leftarrow *semble* in sentences (2a-b):

(2) a. *Ivre, Alain semble fatigué* <sup>(Drunk, Alain seems tired).
b. *Fatigué, Alain semble ivre* ^{(Tired, Alain seems drunk).}
</sup>

Can both be described by the same SSyntRel \mathbf{r} ? No, because \mathbf{r} would contradict Criterion 1: there is a semantic contrast—(2a) is not synonymous with (2b); and this contrast is expressed by purely syntactic means (word order and prosody).

Criterion 1 corresponds to what is known in linguistics as the *minimal pair test*, which is used in phonology (= two phones cannot be relegated to one phoneme if they are the only distinguishers of the signifiers of two semantically contrasting wordforms), morphology, and semantics.

Criterion 2: SSynt-substitutability

The first formalization of the SSynt-substitutability of syntactic subtrees as a means for establishing SSyntRels was proposed by J. Kunze: the so-called *Kunze property*; see this volume, p. 35. In actual fact, we need a weaker version of it: the *quasi-Kunze property*. Both notions are formulated and explained in Mel'čuk's text (p. 36); however, we will partly repeat the discussion here (sometimes verbatim), since we propose to substantiate our decision to use the quasi-Kunze property with French data.

Let there be, in **L**, lexemes $L_{(X)}$, $L_{(Y)}$, ... of syntactic classes X, Y, ..., complete SSynt-configurations $\Delta_{(Z)}$ and $\Delta_{(W)}$ (i.e., subtrees having as their top nodes lexemes $L_{(Z)}$ and $L_{(W)}$), and a SSyntRel **r**.

Definition 2: Kunze property¹⁰

A SSyntRel **r** has the *Kunze property* if and only if for any pair of SSynt-configurations $L_{(X)}$ -**r** \rightarrow **\Delta_{(Z)}** and $L_{(Y)}$ -**r** \rightarrow **\Delta_{(W)}**, replacing $\Delta_{(Z)}$ by $\Delta_{(W)}$ and vice versa does not affect their syntactic well-formedness.

This means that for a SSyntRel that has the Kunze property any of its potential Ds can be attached to any of its potential Gs (= all Ds of a SSyntRel are mutually substitutable in all SSyntSs *salva correctione*). We think, however, that the Kunze property is too rigid, since it does not allow for some desirable generalizations. For instance, it does not admit the same SSyntRel for nominal and infinitival Subjects, as in (3):

(3) a. La course←r-fatigue lit. 'The running tires'.
b. Courir←r—fatigue lit. 'To-run tires'.

b. *Courtr*←**r**—*Jatigue* lit. 10-run ures'.

Since far from any finite verb in French can take an infinitive as its Subject (**Pleuvoir m'a surpris* lit. 'To rain has caught me [somewhere]'), the SSyntRel **r** in the phrases of (3) does not have the Kunze property: with $L_{(X)} = SURPRENDRE$ (catch N [somewhere])', $\Delta_{(Z)} = Noun$ Phrase [e.g., *La pluie* (**r**-*surprend*] and $L_{(Y)} = FATIGUER$, $\Delta_{(W)} = Infinitive Phrase$ [e.g., *Courir* (**r**-*fatigue*], the replacement produces the syntactically ill-formed configuration V_{inf} (**r**-SURPRENDRE. As a result, using the Kunze property leads to having in (3) two different SSyntRels: one for nominal and the other for infini-

¹⁰ A violation of semantic/lexical constraints is not considered as syntactic ill-formedness. Thus, cf. *inside the car* vs. **inside Stuttgart* or *according to Leo* vs. **according to the car*; however, the starred phrases are considered as syntactically well-formed (PREP + N being a legitimate phrase of English).

tival Subjects (as stated in Kunze 1975: 279). But we think that this \mathbf{r} should not be split: all the Subjects, whether nominal or infinitival, share a set of unique linguistic properties, and we prefer to describe all of them by the same SSyntRel.

Therefore, we propose to use the *quasi-Kunze property*, which is weaker: substitutability is required only in one direction and only at least by one particular subtree.

Definition 3: The quasi-Kunze property

A SSyntRel **r** has the *quasi-Kunze property* if and only if there exists in **L** a syntactic class (\approx part of speech) X which is different from substitute pronouns and such that for any SSynt-configuration L–**r**→ $\Delta_{(Y)}$, replacing $\Delta_{(Y)}$ by $\Delta_{(X)}$ (but not necessarily vice versa!) in any SSyntS does not affect the syntactic well-formedness.

The element $\Delta_{(X)}$ that passes with any Governor of the SSyntRel **r** is nothing else but the prototypical D of the SSyntRel **r**, which has been introduced above, Section 2.1, Definition 1, p. 159.

The SSyntRel **r** in (3) has the quasi-Kunze property, since **r** has a prototypical D: a prepositionless noun/(impersonal) pronoun—because in French any finite verb admits a nominal Subject.¹¹ As a result, the same SSyntRel **r** is allowed in (3a) and (3b): this is the subjectival SSyntRel.

Note that, while the G is a particular lexeme, the D is considered up to the syntactic class. Thus, for instance, different prepositions are not distinguished: the SSyntRel **r** in *insister*–**r**→*sur* (insist on), *dépendre*–**r**→*de* (depend on) and *comparer*–**r**→*avec* (compare with) has the quasi-Kunze property (because a PREP + N phrase can be substituted for its D with any of these verbs, provided the appropriate preposition is chosen according to the verb's Government Pattern).

Now we can formulate **Criterion 2**:

Any SSyntRel **r** must have the quasi-Kunze property.

Or, to put it differently:

Any SSyntRel must have a prototypical D.

Examples

1. Two French phrases *pouvoir respirer* 'be-able to-breathe' and *couper le bâton* 'cut the stick' cannot be described by the same SSyntRel \mathbf{r} , since French has no element that could be used as the D of this \mathbf{r} with any modal and any transitive verb; that is, such a

¹¹ The Subject of meteorological verbs (PLEUVOIR ⁽[to] rain⁾, NEIGER ⁽[to] snow⁾, etc.)—the impersonal IL—is considered as a particular case of noun (= a pronominal noun which is not a substitute pronoun). Note that with the Kunze property, Subjects in II = Alain dort ⁽He is sleeping⁾ and II pleut ⁽It is raining⁾ must be described by two different SSyntRels.

SSyntRel would have no prototypical D (thus, **pouvoir* \rightarrow N and **couper* \rightarrow V_{inf}).¹² Consequently, these two phrases require two different SSyntRels. For *pouvoir* \rightarrow *respirer* we propose the **infinitival-objectival** SSyntRel, while for *couper*–[*le*] \rightarrow *bâton*] of course the **dir-obj** SSyntRel should be used.

2. On the other hand, Criterion 2 does not forbid to use the dir-obj SSyntRel for the infinitives with such verbs as INTERDIRE 'forbid' or PRÉFÉRER 'prefer', cf. (4):

(4) a. interdire-dir-obj→de partir 'forbid to leave' préférer-dir-obj→partir 'prefer to leave'

Here, a substitution of the infinitive by a prototypical—nominal—DirO is possible:

b. *interdire*–[*le*]–**dir-obj**→*départ* ⁽forbid the departure⁾ *préférer*–[*le*]–**dir-obj**→*départ* ⁽prefer the departure⁾

With the Kunze property, the description shown in (4a) would not be allowed because of (4c):

c. *couper*-dir-obj-[*le*] \rightarrow *bâton* 'cut the stick' but **couper*-dir-obj \rightarrow *de partir* 'cut to leave'

In other words, since there are many French transitive verbs that do not take infinitives as DirOs, as COUPER above, the Kunze property disallows us to treat an infinitive as DirO with every verb.

3. Nor does Criterion 2 forbid the use of the same cop(ular)-attr(ibutive)-compl(etive) SSyntRel to describe the phrases $\hat{e}tre \rightarrow avec$ [sa mère] 'be with his/her mother' and sembler \rightarrow malade 'seem ill': although *sembler \rightarrow avec [sa mère] 'seem with his/her mother' is impossible, it suffices that the adjective, which is the prototypical D of the cop-attrcompl SSyntRel, passes with both Gs ($\hat{e}tre \rightarrow malade$ 'be ill' and sembler $\rightarrow malade$ 'seem ill'). Again, should we use the Kunze property, we would have to split the cop-attr-compl SSyntRel into at least two different SSyntRels.

NB: Criterion 2 should not be applied to the phrases that realize the SSyntconstructions which do not have direct DSynt-correlates because they are obtained from underlying DSynt-constructions by special rules (so to speak, transformations; for instance, to the construction of the type *Il est venu trois étudiants* lit. ^{(II} thas come three students⁾ = ^{(There} came three students⁾.

Criterion 2 corresponds to what is known in linguistics as the *substitution test*. Linguistic units A and B that are mutually substitutable (optional distribution) or, at least, unilaterally substitutable (B can always be substituted by A, but not vice versa: inclusive distribution) are included—under some additional conditions— in one unit of a higher level. Thus, in phonology, two phones must be relegated to the same phoneme

¹² The phrases [*II*] *le peut* lit. ⁽[He] can it⁾ and [*II*] *le coupe* ⁽[He] cuts it⁾ do not constitute a counterexample: the two *le* seen here (they belong to two different lexemes) are substitute pronouns, which are explicitly excluded from consideration.

if one can be substituted by the other in any context salva significatione. In morphology, the notions of gender and grammatical case are often defined as substitution classes (cf., e.g., Gladkij 1983). In this paper, we deal with substitution of SSynt-subtrees which depend, in a given SSyntS, on the same G via the same SSyntRel r. Let it be emphasized that in principle mutual substitutability is neither necessary nor sufficient property of units belonging to the same -eme. On the one hand, allophones or allomophs are, in a majority of cases, in complementary distribution-i.e., they are never substitutable; on the other hand, absolute synonyms, which are fully mutually substitutable, do not belong to one -eme. In our case, Criterion 2, as it is formulated, is a necessary condition for the unification of some constructions under one SSyntRel: it can forbid such unification. However, it is not sufficient: it does not impose the unification. For us, the possibility of substitution entails no more than the possibility of unification of several constructions, since a substantial difference in their linguistic properties may prevent us from unifying them under one SSyntRel. For instance, Le Goffic (1993: 169) mentions the fact that, with a phasal verb, the infinitive alternates, more or less freely, with an obvious DirO: commencer à travailler 'begin to work' \approx commencer le travail (begin the work); he concludes that this infinitive is a DirO. We, however, do not believe that such an alternation is a decisive argument: it is typologically well-known that the same DSyntA can be expressed by different SSynt-elements. Thus, consider *l'aide canadienne* (the Canadian aid) \approx *l'aide du Canada* (the aid of Canada); here the adjective and the DE-phrase express both the same DSyntA (= I) of AIDE '[the] aid'. Criterion 2 does not forbid subsuming these two constructions under the same SSyntRel, since the substitution of such a denominal adjective by the DE-phrase is always possible. However, the defining properties of the Ds in both constructions do not warrant such unification (cf. Footnote 9, p. 161).

Criterion 3: Repeatability with the same SSynt-Governor

In order to formulate Criterion 3 we need to define non-repeatable and repeatable SSyntRels.

A SSyntRel **r** is *non-repeatable* if and only if no more than one branch labeled **r** can start from any G.

In other words, in L, a G of a non-repeatable r can have, in a given SSyntS, only one D (= one clause element) of the corresponding type.

For instance, actantial SSyntRels whose Ds are marked by purely syntactic means (word order, prosody, inflection)—such as the **subj** and the **dir-obj** SSyntRels in French or in English—are obligatorily non-repeatable: otherwise, they would violate Criterion 1, because their Ds would contrast semantically, while differing only in syntactic means. (Actantial SSyntRels whose Ds are marked by lexical means, that is, by different prepositions—such as the **oblique-objectival** SSyntRel—can be repeatable.)

In some languages, a clause element can be duplicated by a pronoun; as a rule, this pronoun is what is called a *resumptive clitic*. Such is, for instance, the D of the **dir-obj** SSyntRel in Spanish, where we have the construction of the type (5a):

167

(5) a. Spanish

dir-obj _____ A Alain le←dir-obj-veo lit. 'To Alain him [I] see'

We do not consider pronominal duplication of a clause element as repeatability, since such duplication has a (more or less) grammaticized character and is orthogonal to the genuine cooccurrence of SSyntRels. In spite of the expressions of the type (5a), the **dir-obj** SSyntRel is considered non-repeatable in Spanish. Similarly, in spite of (5b), the indir-obj SSyntRel is also non-repeatable in French:

b. French

indir-obj ______ A mes enfants, je leur - indir-obj-permet tout lit. 'To my children, I permit them everything'.^[1]

A SSyntRel **r** is *unlimitedly repeatable*—or, for short, *repeatable*—if and only if several branches labeled **r** can start from a G such that their possible number is theoretically unlimited.

In practice, this number is limited by pragmatic considerations (interpretability by the addressee, etc.) or by the properties of particular Gs, as a rule—by their Government Pattern, as is the case with the **obl-obj** SSyntRel, cf. below, p. 196, Item 1. Thus, the **modificative** and the **circumstantial** SSyntRels in French and English are unlimitedly repeatable; so is the **obl(ique)-obj(ectival)** SSyntRel (although the actual number of possible OblOs is obviously controlled by the Government Pattern of the G).

Criterion 3 runs as follows:

Any **r** of language **L** must be either non-repeatable or unlimitedly repeatable.

An equivalent formulation:

No SSyntRel can be limitedly repeatable, i.e., if it is repeatable, its repeatability cannot be constrained to a particular number by general syntactic factors.

Examples

- 1. To illustrate limited repeatability, consider the French sentence (6):
- (6) Ils lisent tous ce roman très jeunes 'They read all this novel very young'.

If we try to describe the phrases *lisent* \rightarrow *tous* and *lisent* \rightarrow *jeunes* by the same SSyntRel **r**, it would be repeatable exactly two times, since no more non-actantial adjectives can be added to the construction in question as attributes of the MV, no matter what particular verb we take. Thus, **r** is neither non-repeatable, nor unlimitedly repeatable, which is not allowed by Criterion 3. Therefore, we need here two different SSyntRels.

2. Another example:

(7) Il est venu trois étudiants lit. 'It has come three students'.

One could think that both nominal Ds of the MV (the impersonal *il* and the noun *étudiants*) are Subjects, so that the phrases *il est venu* and *est venu trois étudiants* can be described by the same **subj** SSyntRel. (Some actually say so, treating *il* and *trois étudiants* as two Subjects in the same clause.) Criteria 1 and 2 do not prevent us from doing so: they are not applicable. (More specifically, Criterion 1 is not applicable, because the phrases *il est venu* ... and ... *est venu trois étudiants* cannot be contrasted semantically; and Criterion 2 is not applicable because the construction in question has been produced by a "transformational" rule, see above, p. 165.) However, Criterion 3 is not satisfied: the **subj** SSyntRel would be limitedly repeatable—again exactly two times, and this, not because of the Government Pattern of the MV. Therefore, we have to use here two different SSyntRels:

Il←subj-*est*-[*venu*]-quasi-subj-[*trois*]→*étudiants*.

This decision agrees quite well with our linguistic intuition, which is based on the following two considerations, a specific and a general one:

• The SSynt-properties of *il* and *trois étudiants* in this construction are very different; for details, see the properties of the **subj** SSyntRel (p. 170) vs. the properties of the **quasi-subj** SSyntRel (p. 174).

• One of the tenets of syntactic typology is the uniqueness of the Subject in a clause; we would like to retain this feature.

Criterion 3 corresponds roughly to the *cooccurrence test*, used in linguistics on all levels of analysis. Thus, two phones cannot be included in the same phoneme if one of them contradicts the general conditions for phonemic cooccurrence in **L**. In morphology, an element of a morphological category is either non-repeatable (tense or number in English or French) or unlimitedly repeatable (the causative in Turkish). When we see, for instance, just two possible repetitions—like nominal case suffixes in Basque or Georgian, we speak of two different case categories (semantic case *vs.* syntactic case; governed case *vs.* agreeing case; see Mel'čuk 2006: 110*ff*).

Criteria 1 and 2 are paradigmatic, while Criterion 3 is syntagmatic.

To sum up the discussion in Section 2:

A SSyntRel is postulated for a particular family of SSynt-constructions $G_i \rightarrow D_j$ if and only if all these constructions

1) satisfy the requirement of the similarity of the relevant properties of their Ds and

2) satisfy Criteria 1-3.

The similarity of relevant properties, on the one hand, and the Criteria 1-3, on the other, play different roles:

— A sufficient similarity of properties of the constructions A and B argues for using the same SSyntRel \mathbf{r} to describe them; more precisely, it recommends uniting A and B under the same \mathbf{r} .

— Criteria 1-3 argue for not using the same SSyntRel \mathbf{r} to describe A and B; more precisely, they forbid uniting A and B under the same \mathbf{r} .

3 SSynt-Relations between a verb and its valence-controlled dependents

The theoretical principles discussed in Section 2 have been systematically applied to French data in the domain of valence-controlled verb dependents. In this section we present the results: a partial list of 16 SSyntRels of French, with some explanations and justifications.

A SSyntRel is designated by an adjective derived from the name of the SSynt-role of its D, that is, the name of the corresponding clause element. For instance, we call the SSyntRels linking *John* to *have* and *children* to *have* **subjectival** and **direct-objectival**:

 $John \leftarrow subj(ectival) - has - dir(ect) - obj(ectival) \rightarrow children,$

because John is the Subject, and children, the Direct Object of the MV HAVE.

As stated above, the SSyntRels considered are grouped into five classes, according, roughly speaking, to the part of speech of their prototypical D:

Class I: the prototypical D is an N (without or with a preposition);

Class II: the prototypical D is a V in the infinitive (without or with a preposition);

Class III: the prototypical D is an A;

Class IV: the prototypical D is an adverbial (= an ADV, a prepositional phrase or a phrase introduced by the comparative conjunction COMME ^(as));

Class V: the prototypical D is a clause.

Each SSyntRel **r** is described in four steps:

• Relevant linguistic properties of \mathbf{r} . (When a property is not applicable to all types of Ds of a given \mathbf{r} , but only to some of them, we indicate its value for the prototypical D.)

• Formal types of the Ds of **r**.

• Justification of \mathbf{r} : in non-obvious situations, we try to show that \mathbf{r} cannot be collapsed with a different \mathbf{r}' .

• Comments on difficult cases and interesting features of specific constructions (if any).

169

3.1 SSyntRels whose prototypical dependent is a noun

SUBJECTIVAL, OBJECTIVAL and AGENTIVE SSynt-Relations (1 - 7)

The order of the presentation of the SSyntRels in this group is in accordance with the syntactic hierarchy of the corresponding clause elements (see below, 4.1, p. 222).

1. Subjectival SSyntRel: G-subj→D

The prototypical D is a prepositionless noun. Any French verb can take an N (or an impersonal pronoun) as its Subject: cf. *Me promener là-bas me tente* 'To go for a walk there tempts me' ~ *Que je me promène là-bas me tente* 'That I go for a walk there tempts me' ~ *Une promenade là-bas me tente* 'A walk there tempts me'; etc.

Properties

The dependent member of the subj SSyntRel is the Subject, which is the most privileged clause element directly depending on a verb in L. In French, the SSynt-privileges of the Subject are the following seven properties that accrue to Subjects only and thus may be considered defining properties:

1) Only the Subject corresponds to DSyntA I of the MV.

Exception

The Quasi-Subject, which is obtained by the Impersonalization DSynt-rule, also corresponds to DSyntA I of the MV, see below, No. 2, property 3, p. 174.

Note that:

— A Subject may correspond to no DSyntA at all, as the formal, or dummy, Sur-face Subject, i.e., the impersonal IT, or the Surface Subject of an idiomatic expression of the type *Quelle mouche l'a piqué*? lit. 'What fly bit him?' \approx 'What makes him so irritated?'. Cf. as well *Dans cette famille, la culotte a été portée par madame* lit. 'In this family, the pants have been worn by the lady', where the Surface Subject CULOTTE is part of the idiom *porter la culotte* 'wear the pants' \approx 'dominate'. This idiom is represented at the DSynt-level as one node; therefore, CULOTTE does not correspond to any DSyntA.

— The Agentive phrase that depends on the infinitive in the construction with FAIRE, SE FAIRE and SE VOIR corresponds to its DSyntA I: *Il a fait écrire la lettre par la secrétaire* lit. 'He has made write the letter by the secretary', or *Il s'est fait renverser par un camion* lit. 'He has made himself hit by a truck', No. 6, property 2, p. 193.

2) Only the Subject is obligatorily present in (the SSyntS of) any full-fledged clause.¹³ (Note that we consider here only clauses with a finite verb form as their SSynt-predicate; such clauses as *Voilà Alain* ^(Here is Alain) or *Heureusement* [qu'il est $l\dot{a}$] lit. ^(Happily that he's here³ do not have a Subject.)

3) Only the Subject can depend on nothing but the MV. This means that in no situation a clause element different from the finite verb—for instance, an infinitive or a participle—can govern a Subject.

¹³ We believe that in imperative sentences the Subject is present in the SSyntS (it controls the person and number of the verb), but does not appear on the surface.

Exception

The Quasi-Subject can also depend on nothing but the MV.

4) Only the Subject can be the target of demotion by passivization (to an Agentive phrase).

Let it be emphasized that this approximate formulation is used here for simplicity's sake. In the Meaning-Text theory, a Subject, i.e., an element of the SSyntS, cannot be "moved." What we mean here is as follows: "In French, only DSyntA I, which is to be implemented as a Subject, can be demoted by passivization, that is, become DSyntA II."

5) Only the Subject can be the target of a particular non-specific morphological government—namely, if it is replaced by a clitic or a relative pronoun, the latter is in the nominative: II_{NOM} [= le vin] est bon 'It [= the wine] is good', [le pain] qui_{NOM} nous nourrit '[the bread] which feeds us', C'est partir qui_{NOM} m'inquiète 'It is to leave that bothers me', C'est que Helen soit là qui_{NOM} nous intéresse lit. 'It is that Helen should be here that interests us'.

6) Only the Subject controls the agreement of the finite MV: *Vous* êtes chez vous 'You are at home'; *Alain et Helen sont chez eux* 'Alain and Helen are at home'; *Travailler trop et boire beaucoup sont mauvais pour la santé* 'To work too much and drink a lot are bad for your health'.¹⁴

Exception

With the Subject CE 'this' and the MV ÊTRE 'be', it is the Copular-Attributive Complement that controls the agreement of the MV, cf. *Ce sont mes amis* lit. 'This are my friends'.

7) Among the valence-controlled dependents, only the Subject normally precedes the Main Verb.

Exception

A Subject may follow the MV in a set of well-defined constructions, such as interrogative inversion, inversion with the introduction of Direct Speech, Subject inversion in a relative clause, etc. Cf., e.g., *Estelle arrivée*? 'Has she arrived?'; ... sans que se modifie le chiffre total des ventes lit. 'without that modifies itself the total figure of sales'; *Je me demande quand partira Alain* lit. 'I ask myself when will-leave Alain'; *« Me voilà ! » dit Alain* ('I am here!', said Alain'; C'est là où Alain veut aller 'It is there where Alain wants to-go'~ C'est là où veut aller Alain lit. 'It is there where wants to-go Alain'; etc.

In addition, the Subject has seven more relevant (descriptive) properties:

8) The Subject can be the target of cliticization by a personal clitic: *Est-elle venue* ? 'Has she come?'.

NB: In accordance with the proposal of Morin 1985: 794*ff*, we do not consider as the Subject the clitic that participates in the interrogative/exclamatory inversion in the presence of an explicit anteposed nominal Subject. Thus, in *Alain a-t-il mangé* ? lit. ^(Alain has he eaten?), ALAIN is the Subject, but IL—or, more precisely, *-t-il/-t-elle*—is taken to be an interrogative/exclamatory marker that agrees with the

¹⁴ "The MV agrees with the Subject" means that the form of the MV is determined as a function of the Subject, not that their relevant grammemes coincide. Thus, in *Travailler deux jours et qu'on me paye le double me convient* [= SG] *parfaitement* lit. 'To work two days and that they pay me the double suits me perfectly', the MV agrees with the conjoined Subject, just as it does in *La plupart sont* [= PL] *heureux* 'The majority are happy', although the number of the Subject and the number of the MV are different.

Subject or with the DirO;¹⁵ it depends on the MV via a special auxiliary SSyntRel. (If a nominal Subject is postposed to the MV, this marker cannot be used: \hat{A} quelle heure a(*-t-il) mangé Alain? lit. (At what hour has he eaten Alain?) On the contrary, in *A*-t-il mangé? (Has he eaten?) or *Pleut-il*? (Does it rain?), the clitic IL (he)(it) is of course the Subject.

9) The Subject can be the target of relativization: *J'aime ce livre, qui décrit les voy-ages d'Amundsen* (I like this book, which describes Amundsen's travels).

10) The Subject can be the target of clefting: C'est **Helen** qui veut partir 'It is Helen who wants to leave'; C'est (de) partir le plus vite possible qui m'intéresse lit. 'It is to leave as soon as possible that interests me'.

11) The Subject can be the controller of reflexivization: *Alain se rase* ⁽Alain shaves himself⁹.

Exception

In the Impersonalization construction, it is the Quasi-Subject that controls reflexivization, rather than the Subject, No. 2, property 8, p. 174.

12) The Subject does not control the SSynt-role of the Secondary Actor in the causative FAIRE-construction: since the Subject is present in any clause, it cannot be distinctive in this respect, so that this property is not applicable to Subjects.

13) The Subject can be the controller of an actantial-attributive complement: *Cette nouvelle est considérée comme intéressante* lit. ^{(This} piece of news is considered as interesting).

14) The Subject cannot be the target of pure left dislocation: **Alain, est venu hier* lit. ⁽Alain, has come yesterday⁾.

Formal types of Subject

1. A prepositionless noun in the broad sense:

- a. A noun (including the subclass of stressed pronouns): *Alain travaille* 'Alain is working'; *Celui de mon père est plus grand* 'That of my father is bigger'.
- b. A SSynt-equivalent of a noun: a substantivized adjective (*Le plus sage est de tout oublier* 'The wisest [thing to do] is to forget everything'), a headless relative (*Qui veut y aller doit faire une demande* lit. 'Who wants to go there should make an application') or a quantitative phrase of the type *Beaucoup de livres ont été perdus* 'Many books have been lost'.
- c. The impersonal clitic IL or a personal clitic in the nominative: *Il pleut* 'It rains'; *Elle lit* 'She reads'.

¹⁵ The reasons for this decision are based on the following examples (from Morin 1985), which we quote without discussion:

⁽i) *Ces conclusions, ne voilà-t-il* (*-*t-elles*) *pas qu'elles sont caduques maintenant*? 'These conclusions, is it not the case that they are obsolete now?'

⁽ii) Ne la voilà-t-elle $\langle *-t-il \rangle$ pas prisonnière de ses mensonges ? 'Isn't she a prisoner of her lies?'

⁽iii) Cela la gêne-t-il ?/ Cela la gêne-t-elle ? 'Does this bother her?'

⁽iii) Cela est-il vrai ? 'Is this true?'

⁽iv) Cela te gêne-t-il ?/ *Cela la gêne-t-elle ? 'Does this bother you [SG, MASC/FEM]?'

⁽v) Comment Marie le trouve-t-il ? 'How does Mary find him?'

2. An infinitive (with or without preposition): *Courir fatigue Alain* 'To run tires Alain'; *De voir ça m'a bouleversé* 'To see this has upset me³; *Se plaindre passe pour un signe de faiblesse* 'To complain is considered as a sign of weakness³.

3. A subordinate clause:

- a. A QUE-clause (with the MV in the subjunctive): *Que Helen soit arrivée étonne Alain* lit. ^{(That Helen should have arrived amazes Alain).}
- b. A clause with an interrogative pronoun: *Pourquoi Alain a dit cela reste un grand mystère* ⁽Why Alain has said this remains a great mystery⁾.

4. Direct Speech: « *On doit partir le plus tôt possible* » *est la consigne que tout le monde comprend* ^{("}We have to leave as soon as possible"⁾ is the slogan that everyone understands⁾.

Comments

1. Nominal Subjects and non-nominal Subjects show two types of differences:

• Different behavior with respect to word order: thus, inversion rules do not apply to non-nominal Subjects (= formal types 2 - 4).

• A nominal and a non-nominal Subjects are not easily coordinated with each other, even where semantics allows this coordination: ^{??}Courir et le travail physique fatiguent Alain 'To run and physical work tire Alain'. Such coordination, however, does not seem to be completely ungrammatical, and it is better in some cases: [?]Manger trop de fraises et la sieste sont mes deux péchés 'To eat too many strawberries and the siesta are my two sins'. Coordination of two different non-nominal constructions is possible without problems: Travailler deux jours et qu'on me paye le double m'arrange 'To work two days and be paid the double suits me'.

Facts of this type are irrelevant for SSyntRels as such; they must be taken care of in DSynt-rules (during the DSyntS \Rightarrow SSyntS transition).

2. An impersonal verb such as PLEUVOIR ⁽[to] rain⁾, which governs the impersonal clitic IL ⁽it⁾ as its SSynt-Subject, may in turn depend on a chain of modal, phasic etc. verbs: *Il commence à pleuvoir* ⁽It begins to rain⁾, *Il risque de pleuvoir* ⁽It risks to rain⁾, *Il risque de commencer à pleuvoir* ⁽It risks to begin to rain⁾, *Il continue de tonner* ⁽It continues to thunder⁾, *Il a cessé de neiger* ⁽It ceased to snow⁾, etc.; the impersonal verb is then one of possible Objects of the last verb in the chain. As a result, the impersonal IL must percolate through the chain to its head, of which it becomes the SSynt-Subject: *Il*-subj-commence-obl-infin-obj-à-preposit-pleuvoir.

This is a well-known phenomenon, related to the existence of a verbo-nominal dependency chain of the above type, called *nucleus*. (This formal notion was introduced by S. Kahane: Kahane 1997, 2000, Kahane & Mel'čuk 1999: 61*ff*.) A nucleus is thus transparent to the subjectival government of its dependent full verb.

2. Quasi-subjectival SSyntRel: G-quasi-subj→D

This SSyntRel has no prototypical D: the constructions covered by it do not have direct DSynt-correlates, being produced by a lexically restricted DSynt-rule of Impersonalization (cf. NB after the formulation of Criterion 2, p. 165).

Properties

The dependent member of the **quasi-subj** SSyntRel is the Quasi-Subject: an element that corresponds—in an impersonal construction—to DSyntA I of the MV, but is not the Subject, the Subject being the impersonal clitic IL ^(it) or the demonstrative pronoun CELA/ÇA ^(this) (*Il est raconté beaucoup d'histoires bizarres* lit. ^{(II is [being]} told many strange stories⁾; *Cela m'étonne qu'Alain soit venu* lit. ^{(This} amazes me that Alain should have come³). The Quasi-Subject is introduced, together with IL or CELA, in the SSyntS by the Impersonalization rule—under the control of communicative information. With some verbs the application of this rule is obligatory: *Il semble qu'Alain soit venu* ^{(II} seems that Alain has come³ ~ *Qu'Alain soit venu semble.

The Quasi-Subject has only two out of seven defining properties of the genuine Subject: the correspondence to DSyntA I of the MV and the ability to depend only on the MV; therefore, it can by no means be considered a Subject. In the French grammatical tradition, the Quasi-Subject is called *sujet réel* ⁽real subject)</sup>, while the impersonal IL is *sujet apparent* ⁽apparent subject)</sup>. These terms show the failure to distinguish the semantic and the syntactic levels: from the viewpoint of syntax, it is rather IL which is the "real" subject, while our Quasi-Subject is an "apparent" subject.

The Quasi-Subject has two defining properties:

1) The Quasi-Subject can be the target of clefting: *C'est des étudiants qu'il est venu* lit. ⁽It is students that it has come⁾.

2) The Quasi-Subject is not obligatorily present in every full-fledged clause.

Its descriptive properties are:

3) The Quasi-Subject corresponds to DSyntA I of the MV, being the product of DSynt-rule of Impersonalization.

4) The Quasi-Subject can depend on nothing but the MV.^[2]

5) The Quasi-Subject cannot be promoted/demoted.

6) A nominal Quasi-Subject can be the target of cliticization and is substitutable by the clitic EN: *Il* **en** [= bâtiments de ce type] a été construit en France en 1970 lit. 'It has been built thereof [= buildings of this type] in France in 1970', *Il* **en** [= des coups] pleuvait lit. 'It rained thereof [= blows]', *Il* **en** [= des camions] arrive lit. 'It is arriving thereof [= trucks]'.

7) The Quasi-Subject cannot be the target of relativization: *toutes ces histoires* *qu'il *a été racontées* lit. ^{(all these stories that it has been told).}

8) The Quasi-Subject can be the controller of reflexivization: *Tous les ans, à la même période, il se baigne dans le Gange des miliers de fidèles* lit. 'Every year, at the same period, there bathe themselves in the Ganges thousands of [the] faithful³.

9) The Quasi-Subject does not control the role of the Secondary Actor in the causative FAIRE-construction: the Quasi-Subject is incompatible with the causative construction, so that this property is not applicable to it.

10) The Quasi-Subject cannot be the controller of an actantial-attributive complement: *Il est considéré de telles théories comme fort intéressantes lit. (It is considered such theories as very interesting). (However, as C. Blanche-Benveniste points out, with some verbs this construction is possible: Il a été choisi quelques exemples comme représentatifs de l'ensemble lit. 'It has been chosen a few examples as representative of the set 2 .)

11) The Quasi-Subject cannot be the target of non-specific morphological government. (Non-specific morphological government manifests itself in French only with personal clitics, which distinguish cases; however, the Quasi-Subject is not cliticizable by personal clitics.)

12) The Quasi-Subject cannot be the controller of the agreement of the MV.

13) The Quasi-Subject always follows the MV.

14) The Quasi-Subject cannot be the target of pure left dislocation: *Des étudiants, il est venu lit. 'Students, it has come'.

NB: A nominal Quasi-Subject has an additional important property: it must be indefinite or-for mass nouns-partitive (Il tombait de la neige lit. 'It was falling of the snow). Typical exceptions to this condition include phrases with possessive or demonstrative pronouns: Il pourrait encore venir ton frère, son copain et sa femme lit. 'It could still come your brother, his pal and his wife'; S'il survenait encore ce même accident, voici ce qu'il faudrait faire lit. (If it happens again the same accident, here is what is to be done³.

Formal types of Quasi-Subject

1. An indefinite or partitive prepositionless noun (including indefinite and interrogative pronouns): Il pleuvait des coups lit. 'It was raining blows' = 'Blows were raining'; Il risque d'être venu des gens lit. (It risks to have come people) = (It is possible that some people have come'; Il est tombé de la neige lit. 'It has fallen of the snow'; Il est entré quelqu'un lit. (It has entered somebody); Il est entré qui ? ~ Qui est-il entré ? lit. (It has entered who?' = 'Who has entered?'.

2. The clitic EN.

3. An infinitive of a normal (non-impersonal) verb with the preposition DE: Cela fatigue Alain de courir lit. (This tires Alain to run); Il n'est pas difficile de savoir quand Alain partira 'It is not difficult to know when Alain will leave'.

4. A subordinate QUE-clause: Il semble qu'Alain est venu 'It seems that Alain has come'; Il/Cela m'étonne qu'Alain soit venu lit. (It/This amazes me that Alain should have come).

5. Direct Speech: *Il a été annoncé: « Restez tranquilles ! » ⁽It has been announced: 'Remain calm!').*

Comments

1. Although the constructions listed above do not have a common prototypical D, this does not prevent us from subsuming all of them under the same SSyntRel. Criterion 2 is simply not applicable in this case: the constructions in question do not have direct DSynt-correlates, being produced by a lexically restricted DSynt-rule of Impersonalization.

2. G. Lazard (1994a) considers only nominal expressions, including EN, as possible Quasi-Subjects (*actant H*, in his terms). Actually, he describes these as something intermediate between Subjects and Direct Objects, with which they share several properties (Lazard 1994a: 9), but, in any rate, as a separate SSynt-role. We do the same thing, with the difference that we include under the label of **quasi-subjectival** SSyntRel three other constructions, which have not been considered by Lazard.

3. Nominal and non-nominal Quasi-Subjects show the following important difference: the presence of a nominal Subject with a given verb does by no means entail the possibility of the Impersonalization rule, which will produce a Quasi-Subject; only some existential verbs and all the verbs in the passive admit it. However, the presence of a non-nominal Subject with a verb automatically entails the possibility of the Impersonalization; as indicated above, in some cases, this is even obligatory.

4. Another important difference cuts across nominal and non-nominal Quasi-Subjects: some of them correlate with genuine Subjects possible on the surface while some others do not. For instance: *Il pleuvait des coups* 'It rained blows' ~ *Des coups pleuvaient* 'Blows rained' or *Il fatigue Alain de courir* 'It tires Alain to run' ~ *Courir fatigue Alain* 'To run tires Alain'; but *Il semble qu'Alain est venu* 'It seems that Alain has come' ~ **Qu'Alain est venu semble*.

3. Direct-objectival SSyntRel: G-dir-obj→D

The prototypical D is a prepositionless noun: cf. *Alain propose de partir* 'Alain proposes to leave' ~ *Alain propose le départ* 'Alain proposes the departure'; *Alain sait quoi lui répondre* lit. 'Alain knows what to answer to him' ~ *Alain sait la réponse* 'Alain knows the answer'; etc.

Properties

The dependent member of the dir-obj SSyntRel is the Direct Object [= DirO]. It is the second most privileged clause element depending on a verb in \mathbf{L} . In French, the SSynt-privileges of the DirO are the following four defining properties:

1) Only the DirO can be the target of promotion (to Subject—by passivization).

Exception

In the Accusativus-Cum-Infinitivo construction, the promotion of the DirO is not accepted by some speakers in some contexts: [?]*Helen a été vue sortir de l'immeuble* ⁽Helen has been seen to go out of the building[?] (Abeillé 1997b: 15).

If the verb governing a DirO is not passivizable as such, this property is taken to be not applicable. Thus, it is not applicable, for instance, in the case of *avoir ce livre* ⁽have this book⁾, since AVOIR (in this sense) does not have the passive. The same holds for *comporter trois parties* ⁽contain three parts⁾. (Many transitive verbs that do not passivize inside phrasemes are quoted in Gaatone 1993: 42-48: *donner lieu* ⁽give rise⁾ ~ **Le lieu a été donné*; *prendre froid* ⁽catch cold⁾ ~ **Le froid a été pris*; *faire le désespoir* (*d'Alain*) ⁽make (Alain's) despair⁾ ~ **Le désespoir a été fait*; *foutre le camp* ⁽flee, get out of here⁾ ~ **Le camp a été foutu*; etc. On unpassivizable French transitive verbs, see also Leclère 1993.) But the case of *Alain a commencé à travailler* ⁽Alain has begun to work⁾ vs. *(*À*) *travailler a été commencé par Alain* ⁽To work has been begun by Alain⁾ is different: here this property is applicable and not satisfied, since COMMENCER—in this sense!—has a passive: *Le travail a été commencé* ⁽The work has been begun⁾. (As a result, *à* V_{inf} with COMMENCER is not considered to be a DirO, cf. Comment 6 below, p. 185.)

2) The DirO imposes the IndirO realization of the Secondary Actor in the causative FAIRE-construction (Kayne 1977: 202-211, Morin 1980, Tasmowski-de Ryck 1984).

In order to explain this property, we have to say a few words about the syntactic representation of the French causative construction FAIRE + V (*Alain fait dormir Helen* lit. ^{(Alain makes sleep Helen⁾, *Alain fait lire un roman à Helen* lit. ^{(Alain makes read a novel to Helen⁾ = ^{(Alain makes Helen read a novel⁾). At the DSynt-level, it is represented as follows:}}}



The causative FAIRE is considered to have three DSyntAs: its DSyntA I is the Primary Actor, or the Causer; the DSyntA II is the Secondary Actor, i.e., the demoted Actor of the lexical verb V (*Helen* in our examples); and the DSyntA III is the lexical verb itself (see also Comment 1, p. 203).

The SSynt-role of the corresponding clause elements depends on the transitivity of V and the presence of a DirO with this V.

• If V is intransitive or, being transitive, has no DirO, the Secondary Actor is realized as the DirO of FAIRE; cf.:

- (8) a. *Helen sort de sa chambre* [= OblO of SORTIR] ⁽Helen goes out of her room⁾. vs.
 - b. Alain fait sortir Helen [= DirO of FAIRE] de sa chambre [= OblO of SORTIR]
 lit. 'Alain makes go out Helen from her room'. ~
 Alain la (*lui) fait sortir de sa chambre
 lit. 'Alain makes her go out from her room'.

c. *Helen lit chaque soir* 'Helen reads every night'.
vs.
d. *Alain fait lire Helen* [= DirO of FAIRE] *chaque soir* lit. 'Alain makes read Helen every night'. ~

Alain la (*lui) fait lire chaque soir lit. 'Alain makes read her every night'.

• If V is transitive and has an expressed DirO, the Secondary Actor is realized as an Indirect Object of FAIRE; cf.:

(9) a. *Alain fait lire le roman* [=DirO of LIRE] à *Helen* [= IndirO of FAIRE] (**Helen*) lit. ⁽Alain makes read the novel to Helen⁾. ~

Alain **lui** $\langle *la \rangle$ fait lire le roman lit. ⁽Alain makes to her read the novel⁾. = ⁽Alain makes her read the novel⁾.

Cf. also:

(10) J'ai fait comprendre à Alain [= IndirO of FAIRE] (*Alain) que c'était inutile [= DirO of COMPRENDRE],

lit. ^{(I} have made understand to Alain that this was useless). \sim

Je lui $\langle l \rangle$ *ai fait comprendre que c'était inutile*, lit. ^{(I} have made understand to him that this was useless⁾.

Thus, when we say that a DirO is the controller of the SSynt-role of the Secondary Actor, we mean the following:

If the DirO of the lexical verb is present, the Secondary Actor in a causative construction must be realized as an IndirO of FAIRE, and in its absence, as a DirO of the latter.^[3]

3) Only the DirO can be the target of particular non-specific morphological government—the clitic/the relative pronoun replacing the corresponding noun is in the accusative: Je le_{ACC} [= le vin] bois ^{(I} drink it [= the wine]⁾, [le pain] $que_{ACC} j$ ai acheté ^{([the bread]} that I have bought⁾.

Exception

The Pseudo-DirO can also be replaced by an accusative clitic, see p. 190.

4) Only the DirO is the controller of the agreement of the past participle in compound forms of the MV in case the DirO precedes the participle (the DirO is the relative pronoun or a clitic): *les lettres que j'ai écrites* (the letters that I have written), *Je les* [= *les lettres*] *ai écrites* (I have written them [= the letters]).

Exceptions

1. The past participle of inherently impersonal verbs does not agree with the DirO: *les lettres qu'il m'a fallu* $\langle *fallues \rangle$ lit. (the letters that it needed me), *les chaleurs qu'il a fait* $\langle *faites \rangle$ lit. (the heats that it made).

2. The DirO implemented by the clitic EN does not control the agreement of the past participle, so that we have *Des lettres, j'en ai écrit (*écrites) dans ma vie* lit. ⁽Letters, I have written thereof in my life⁾.

In addition, the DirO has ten descriptive properties:

5) The DirO corresponds exclusively to DSyntA II of the governing verb.

Exception

A DirO in an idiom [= full phraseme] of the type *faire* **l'amour** [*avec qqn*] ^{(make love [with someone]⁾ or *kick the bucket* (cf. p. 161), where—under sentence production—the governing verb appears first only in the SSyntS (in the DSyntS, the whole idiom is represented as one node) and therefore has no DSyntAs.}

In order to avoid cluttering our presentation with non-relevant details, we will not discuss a complication related to the DSynt-representation of constructions with an "internal" DirO, as in *vivre sa vie* 'live one's life' or *suer la sueur de tes nuits* [Ver-laine] 'sweat the sweat of your nights', where the verb receives—by a semantic rule—the DSyntA II that does not correspond to any of its SemAs.

6) The DirO is not obligatorily present in every clause.

7) The DirO does not obligatorily depend on the MV. (It can, unlike the Subject, depend on an infinitive or a present participle.)

8) The DirO can be the target of personal cliticization: Je les [= les conditions] considère 'I consider them [= the conditions]'. Moreover, an indefinite or partitive DirO can be replaced with the clitic EN: J'ai trouvé des livres 'I have found books' ~ J'en ai trouvé lit. 'I have found thereof'; Je mange du pain blanc 'I eat white bread' ~ J'en mange lit. 'I eat thereof'. A partitive DirO can be also replaced with the relative pronoun DONT: Le pain, dont j'ai mangé à plusieurs reprises, est excellent lit. 'The bread, of which I have eaten several times, is excellent'.

9) The DirO can be the target of relativization: *J'aime ce livre, que j'ai acheté à Paris* ^{(I} love this book, which I have bought in Paris³.

10) The DirO can be the target of clefting: C'est ce livre que je veux lire 'It is this book that I want to read'; C'est travailler la nuit que je déteste 'It is to work at night that I hate'.

11) The DirO can be the controller/the target of reflexivization: *Helen a forcé Alain à se raser* 'Helen has forced Alain to shave himself'; *Helen a fait se raser Alain* 'Helen has made Alain to shave himself'.

12) The DirO can be the controller of an actantial-attributive complement: *Il considère cette nouvelle comme intéressante* ⁽He considers this piece of news as interesting⁾.

13) The DirO follows the governing verb.

Exception

A DirO expressed by a quantifying pronoun TOUT 'everything' or RIEN 'nothing' can precede the governing verb: *J'ai tout fait* 'I have done everything'; *Il ne veut rien faire* 'He does not want to do anything'.

14) The DirO cannot be the target of pure left dislocation: *Les romans, Alain lit 'The novels, Alain reads'; ??Ce film, je préfère de loin à un roman lit. 'This film, I strongly prefer to a novel'; ??Partir tôt, je préfère lit. 'To leave early, I prefer'. (Recall that a dislocated phrase must carry a mounting contour, as indicated in 2.1, Property 15, p. 158.)

Exception

In colloquial speech a few verbs admit pure left dislocation of the DirO (without the resumptive clitic), cf. *Ce film*, *j'ai aimé* lit. ^{(This film}, I have loved²; *Les mille-feuilles*, *Alain adore* lit. ^{(The mille feuilles, Alain adores³; *Nager le matin*, *j'ai beaucoup aimé* ^{(To swim in the morning, I have loved a lot².}}

On some SSynt-properties of the DirO in French, see Gaatone 1998.

Formal types of DirO

- 1. A prepositionless noun in the broad sense:
 - a. A noun (including the subclass of stressed pronouns): *J'ai trouvé un livre* ⁽I have found a book⁾; *Je trouve cela dangereux* ⁽I find this dangerous⁾; *Je vois celui de mon père* ⁽I see that of my father⁾.
 - b. A SSynt-equivalent of a noun, such as a headless relative (*J'embrasse qui je veux* ^{(I} kiss who I want⁾; *Alain sait qui sa femme voit* ^{(Alain knows who his wife is seeing⁾; *Alain chassera qui apparaîtra* ^(Alain will chase who will appear), a subordinate clause of the type *Helen a invité tu ne devineras jamais qui* ^{(Helen has invited you will never guess who), ¹⁶ a quantitative phrase (*Alain lit beaucoup de livres* ^{(Alain reads many books³), etc.}}}
 - c. A personal clitic in the accusative: *Alain la trouvait partout* ⁽Alain was finding it everywhere⁾; *Alain ne me quitte jamais* ⁽Alain never leaves me⁾.

2. A noun without article introduced by the preposition DE—with a negated transitive verb: *Alain n'a pas trouvé de livre* ⁽Alain has not found a book⁾.

3. The clitic EN: *Alain en* [= *des légumes frais*] *trouve partout* lit. ⁽Alain finds thereof [= fresh vegetables] everywhere³.

- 4. An infinitive:
 - a. Without preposition or with the preposition À or DE: Alain préfère partir 'Alain prefers to leave'; Alain propose de partir 'Alain proposes to leave', Tous les étudiants apprennent à parler chinois 'All students learn to speak Chinese'.
 - b. Governing an interrogative pronoun: *Alain sait combien lui donner/à qui parler/ qui inviter* 'Alain knows how much to give him/to whom to speak/whom to invite'.
- 5. A subordinate clause:
 - a. Including a relative/interrogative pronoun: *Alain sait quand son ami part* 'Alain knows when his friend is leaving'.
 - b. Introduced by the conjunction QUE: *Alain sait que Helen est là* ⁽Alain knows that Helen is here³.
 - c. Introduced by the interrogative conjunction SI: *Alain veut savoir si nous travaillons ce vendredi* ⁽Alain wants to know whether we work this Friday⁾.

¹⁶ We owe this type of example to Y.-Ch. Morin.

Comments

1. A prepositionless noun that corresponds to DSyntA II of the governing verb is not necessarily a DirO: it may be an Oblique Object [= OblO]. This is, for instance, the case of the noun that designates the topic of the discussion with the verbs PARLER 'speak' and CAUSER 'chat': *parler politique* 'talk politics' or *causer argent* lit. 'chat money' (cf. *talk shop*), cf. No. 7 below, Item 5, p. 196. The noun in question does not have three of the four defining properties of a DirO: it does not forbid the SSynt-implementation of the Secondary Actor in the causative construction as a DirO (*Alain la/lui fait parler politique*) and, since it does not allow either cliticization (**la parler*) or relativization (**politique que nous parlions*), it cannot be the target of non-specific morphological government and does not control the agreement of the past participle in compound verb forms (**politique qu'on a parlée*). It does not passivize, either (**Politique a été parlée*), but this is because the verbs PARLER and CAUSER do not have a genuine passive form.

2. Another controversial case is a prepositionless quantitative phrase that corresponds to DSyntA II of the governing verb of measure and designates the value of a parameter: *coûter 300 euros* (cost 300 euros)/*toute une fortune* (a whole fortune)/*une somme rondelette* (a nice little sum), *mesurer 4 mètres* (measure 4 meters) = (be 4 meters long), etc. The traditional French grammar (e.g., Grevisse 1993: 1336) does not consider this phrase as a DirO; we agree, since this clause element violates two of DirO's defining properties:

— It allows the Secondary Actor in the causative FAIRE-construction to be in the accusative if it is cliticized, which a genuine DirO does not allow: *Ce régime la/lui* [= *Helen*] *fera peser 45 kilos* 'This diet will make her weigh 45 kilos' [measure verbs do not readily admit the causative construction, so that this example is a bit far-fetched] vs. *Ce régime lui* (**la*) *fera perdre 5 kilos* 'This diet will make her lose 5 kilos'.

— It does not control the agreement of the past participle in compound verb forms: 300 euros que cette robe m'a coûté(*s) '300 euros that this dress has cost me', la somme que cette robe m'a coûté(*e) 'the sum that this dress has cost me'.

As for two other defining properties of the DirO, the picture is as follows:

• This construction does not passivize (*300 euros est/sont coûté(s) par cette robe), but then the verbs involved do not have passive forms.

• Although the accusative cliticization in this construction is not readily done (because of the inherent non-referentiality of the noun), it is possible: e.g., 300 euros ! Cette robe les coûte lit. '300 euros! This dress costs them'; Le sac les pèse, ses 15 kilos lit. 'The bag weighs them, its 15 kilos'. In addition, the quantitative phrase admits of relativization via the accusative que: les 300 euros que cette robe m'a coûté/les grosses sommes que ces voitures m'ont coûté 'the 300 euros that this dress has cost me/the big sums that these cars have cost me' (but note the lack of agreement in the past participle!). Because of this, the quantitative phrase with measure verbs is fairly close to DirO, so that the distinction is really tenuous (especially if we take into account the fact that the agreement of the past participle with these verbs would be purely orthographic). However, in conformity with what has been said above we declare the quantitative phrase a special clause element, which we call *Pseudo-Direct Object* [= Pseudo-DirO], and the corresponding SSyntRel is pseudo-dir(ect)-obj(ectival); see No. 5, p. 190ff.

A similar, but actually different case is represented by sentence (11):

(11) Nous avons payé cette robe 300 euros lit. 'We have paid this dress 300 euros'.

On the one hand, since the **dir-obj** SSyntRel is non-repeatable in French, the quantitative phrase 300 euros cannot be a DirO—the DirO here is cette robe ^(this dress). On the other hand, unlike the Pseudo-DirO, the phrase 300 euros expresses DSyntA **III** of PAYER ^(pay) and does not admit cliticization. We consider the phrase 300 euros in (11) to be an ObIO: it is substitutable by a prototypical ObIO—a noun introduced by a preposition, cf. Nous avons payé cette robe **avec les 300 euros** d'Alain ^{(We} have paid this dress with Alain's 300 euros^{).17}

3. The prepositionless noun which depends on the verb FALLOIR (Il faut des livres (Books are needed), on the meteorological verb FAIRE (do) (Il fait une chaleur ! lit. (It does a heat!' = 'It is very hot') or on the idiom IL Y AVOIR (Il y a des livres 'There are books) is considered a DirO, although this is a controversial DirO. Namely, the defining properties 1, 2 and 4 of the DirO are not applicable here: the corresponding verbs do not have the passive form, they cannot be used in the causative construction, and their past participles are invariable. However, the defining property 3 (= the accusative form of the clitic or of the relative pronoun) is satisfied: Il me les [= les médicaments] faut 'I need them [= the drugs]⁾, la chaleur **qu**'il fait lit. 'the heat that it does⁾, and les inondations qu'il y a eu lit. 'the floods that it there has had'. Moreover, the noun in question does not contradict the descriptive properties of the DirO: thus, it does not admit the pure left dislocation: *Ces outils, il me faut 'These tools, I need' or *Des inondations, il y a eu partout 'The floods, it there has everywhere', etc. We take all this to be sufficient for its status as a DirO, because we do not have much of a choice: the only other SSynt-Rel whose D admits the accusative form is the Pseudo-DirO (see No. 5, p. 190), but we cannot present our suspicious clause element as a Pseudo-DirO for the following two reasons:

• Control of attributive complements. The DirO controls attributive complements, actantial (*On considère cette théorie fausse* 'They consider this theory wrong'), as well as non-actantial (*La soupe, je la mange très chaude* 'The soup, I eat it very hot'), while the Pseudo-DirO does not (**La somme, cette robe la coûte rondelette* lit. 'The sum, this dress costs it tidy'). In this respect, the element in question is closer to the DirO: it controls non-actantial attributive complements, although of a slightly different form—they must be preceded by an indefinite article, cf.: *Des outils de constructions, il m'en faut des pas trop chers* lit. 'Construction tools, I need thereof not too expensive'; *Des inondations, il y en a eu des meurtrières* lit. 'Floods, it thereof there had been deadly ones'; *Des chaleurs, il en a fait des vraiment accablantes* lit. 'Heats, it thereof has done really oppressive'.

• Treatment of the infinitive and the QUE-clause with FALLOIR (*Il faut partir* lit. 'It isnecessary to leave'; *Il faut qu'il parte* lit. 'It is-necessary that he leave'). Considering the suspicious clause element as a DirO, we open the way for this infinitive/QUE-clause to be described also as DirOs: they are substitutable by a prototypical nominal D, and many other French transitive verbs admit an infinitive/a QUE-clause as a DirO. Otherwise (i.e., if we decide that this element is a Pseudo-DirO), we meet with a serious

¹⁷ In *Nous avons payé 300 euros pour cette robe* ⁽We have paid 300 euros for this dress⁾ the phrase *300 euros* is a DirO—here it corresponds to DSyntA II of PAYER and satisfies all the defining properties of DirOs.

183

problem: how to treat the QUE-clause with FALLOIR? We do not have a SSyntRel whose prototypical D is a completive clause; and it seems not natural to subsume it under the **pseudo-dir-obj** SSyntRel, since no other verb taking a Pseudo-DirO admits a QUE-clause.

4. The infinitive in a construction of the type *préférer partir, apprendre à parler, proposer de partir, permettre/défendre de partir* and *savoir quand partir* is considered a DirO since it satisfies all four defining properties of DirOs:

1) Passivization (in many cases, additional Impersonalization is required, as is typical of infinitives):

(12)

a. <i>Tout le monde préfère partir</i> 'Everybody prefers to leave ³ .	~	<i>Partir</i> a été préféré par tout le monde 'To leave has been preferred by every- body'.
b. <i>Ils apprennent à parler chinois</i> 'They learn to speak Chinese ⁾ .	~	<i>Il est appris par eux à parler chinois</i> ⁽ It is learned by them to speak Chinese ⁾ .
c. <i>Il a proposé de partir</i> 'He has proposed to leave'.	~	[?] Partir a été proposé par lui ⁽ To leave has been proposed by him [?] ./ Il a été proposé par lui de partir ⁽ It has been proposed by him to leave [?] .
d. <i>Tout le monde sait quand partir</i> (Everybody knows when to leave).	~	[?] <i>Il est su de tout le monde quand partir</i> (It is known to everybody when to leave).

2) Control of the SSynt-role of the Secondary Actor in the causative FAIRE-construction (Morin 1980: 206): the Secondary Actor cannot be realized as a DirO. Here are some relevant examples:

- (13) a. [?]J'ai fait apprendre à parler chinois à Alain (*J'ai fait apprendre à parler chinois Alain)
 (I have made Alain learn to speak Chinese).
 - b. [?]*Cela fait proposer à Alain de partir* (**Cela fait proposer Alain de partir*) 'It makes Alain propose to leave[?].
 - c. *J'ai fait promettre à Alain de partir (*J'ai fait promettre Alain de partir)* 'I have made Alain promise to leave³.

However, for the infinitive with the verb PRÉFÉRER this property is not distinctive: if the Secondary Actor is a noun, the causative FAIRE-construction is impossible; but if it is expressed by a clitic both realizations (the accusative and the dative one) are possible, although judged awkward:

d. *Cela fait préférer partir à Alain. ~ *Cela fait préférer partir Alain (It makes Alain prefer to leave).
vs.
?Cela lui fait préférer partir. ~ ?Cela le fait préférer partir. This makes the infinitive that depends on PRÉFÉRER "less" of a DirO.

3) Accusative cliticization (the infinitive is substitutable by the accusative "neuter" clitic LE):

(14)	a. <i>Tout le monde préfère partir.</i>	~	Tout le monde le préfère.
	b. Ils apprennent à parler chinois.	~	Ils l' apprennent.
	c. Il a proposé de partir .	~	Il l' a proposé.
	d. Tout le monde sait quand partir .	~	Tout le monde le sait.

4) Control of the agreement of the past participle in compound verbal forms (masculine singular):

(15) Partir, tout le monde l'a préféré; Parler chinois, nos étudiants l'ont appris.

Moreover, this infinitive has the descriptive properties of the prototypical DirOs that are applicable to the verb: it corresponds to DSyntA II of the governing verb, follows the governing verb, admits clefting and cannot be the target of pure left dislocation.

Criterion 2 does not forbid us to consider the infinitive in these constructions a DirO since it is substitutable by the prototypical DirO—a noun: *Il préfère le départ* ⁽He prefers the departure⁾, *Il propose le départ* ⁽He proposes the departure⁾, *Il sait l'heure du départ* ⁽He knows the time of the departure⁾.

5. However, Criterion 2 precludes treating the infinitival complement of the modal verbs POUVOIR 'can' and DEVOIR 'must' as a DirO, since it is not substitutable by a noun: *Alain peut/doit partir* 'Alain can/must leave' ~ **Alain peut/doit le départ* 'Alain can/must the departure').

This result can be (and sometimes is) questioned: would it not be more consistent to consider the infinitives with POUVOIR and DEVOIR as DirOs? We do not think so, because these infinitives do not share with the DirO enough defining properties. The first, the second and the fourth properties are not applicable: POUVOIR and DEVOIR have no passive forms; they do not readily admit the causative construction (*Alain lui/l'a fait pouvoir partir); and they do not participate in constructions in which agreement of the past participle can be observed. The third property is, so to speak, only partially satisfied: the infinitive with POUVOIR/DEVOIR can be cliticized via the accusative "neuter" clitic LE (Alain peut/doit partir 'Alain can/must leave'. ~ Alain le peut/?le doit 'Alain can/must it), but it is obsolete with DEVOIR. Thus, the SSynt-properties of the infinitive governed by POUVOIR/DEVOIR do not contradict Criterion 2. What is crucial is the fact that POUVOIR/DEVOIR do not accept nominal objects. Their DSyntA II inherently designates an action, a state, etc.; therefore, their prototypical D must be verbal rather than nominal, while the prototypical DirO is nominal-the name of an entity. As a result, the construction POUVOIR/DEVOIR- $\mathbf{r} \rightarrow V_{inf}$ is described by the direct-infinitival-objectival SSyntRel (3.2, No. 8, p. 198).

The situation with the modal verb VOULOIR 'want' is different: with it, the infinitive is substitutable by a noun (*Je veux ce départ* 'I want this departure'), so that Criterion 2 does not forbid us to treat it as a DirO. The infinitive that depends on VOULOIR admits as well limited passivization (*Partir immédiatement est voulu par tous* 'To-leave immediately is wanted by everybody') and the accusative cliticization—with the resulting control of past participle agreement (*Alain a voulu partir* 'Alain has wanted to-leave' ~ *Alain I'a voulu* 'Alain has wanted it'). True, with respect to the FAIRE-construction, VOULOIR behaves like the semantically close PRÉFÉRER (see above)—that is, for VOULOIR this property is not distinctive. However, since the other three defining properties of DirOs are satisfied, the dependency of an infinitive on VOULOIR is described by the direct-objectival SSyntRel—like that with PRÉFÉRER (as well as with DÉSIRER 'desire' and SOUHAITER 'wish').

6. Not all infinitives that correspond to DSyntA II of the governing verb and are introduced by the preposition À or DE are DirOs: they can also be Oblique Objects as well as Direct-Infinitival Objects and Oblique-Infinitival Objects (see below, 3.2, Nos. 8-10, p. 198*ff*). For instance, the infinitive with a phasal verb such as COMMENCER 'begin', CONTINUER 'continue), or CESSER 'cease' is not a DirO (*contra* Le Goffic 1993: 169). This is so because such an infinitive violates the defining properties of DirOs: it does not admit passivization (*(À) *travailler a été commencé*, **II a été commencé* à *travailler*),¹⁸ does not impose the realization of the Secondary Actor in the causative FAIRE-construction as IndirO (*Alain Ia* (**lui*) fait commencer à travailler 'Alain makes her begin to work'), and disallows cliticization with the "neuter" LE (**Alain I'*[= à travailler]a commencé).

7. The direct reflexive clitic pronoun SE—even in genuinely reflexive verbs, such as SE LAVER 'wash oneself' or SE RASER 'shave oneself'—is by no means a DirO, in sharp contrast to other personal clitics (*me*, *te*, *le*, ...): it does not correspond to DSyntA II of the governing verb, since this verb in the direct-reflexive voice has only DSyntA I. (In the DSyntS, a verbal form of this type is represented by one node and is supplied with the grammeme "dir-refl.") Moreover, the clitic SE contradicts the only defining property of DirOs which is applicable to it: namely, SE does not impose the IndirO realization on the Secondary Actor in the causative construction (*Elle fait se raser* (*à) *Alain tous les jours/Elle le* (**lui*) *fait se raser tous les jours* 'She makes Alain/him shave every day'). In the SSyntS, the lexeme SE is linked to its verb by a special auxiliary SSyntRel; its behavior is different from that of normal object clitics; cf., for instance:

(16) French

a. <i>*Helen fait le</i> [= <i>le roman</i>] <i>lire à Alain</i> 'Helen makes Alain read it [= the novel]'.	~	Helen le [= le roman] fait lire à Alain.
VS.		
b. Helen fait se raser Alain	~	*Helen se [= Alain]
'Helen makes Alain shave himself'.		fait raser Alain.

These facts fit nicely with a general typological consideration: a direct-reflexive verb like SE LAVER is by definition intransitive and cannot have a DirO. The inherent reflex-

185

¹⁸ Phasal verbs themselves have passive forms: *Le travail a été commencé* ⁽The work has been begun⁾, cf. p. 177.

ive SE, as in *se lever* 'rise', *se taire* 'keep silent', etc., has even less reasons to be considered a DirO.

8. The quantitative adverb in such phrases as *Il a mangé beaucoup (trop)* (*plus qu'elle*) ⁽He has eaten much (too much, more than she)⁾ is treated as a Circumstantial—rather than a DirO; for explanations and a justification, see 3.4, No. 14, p. 216.

9. In other approaches, other properties may be taken as defining for the DirO. Thus, M. Gross (1968: 27) defines DirO as the clause element that can be replaced by one of the clitics LE, LA, LES. This definition includes among DirOs, for instance, the attributive adjective with a copula: *Alain est intéressant* ⁽Alain is interesting⁾ ~ *Alain l'est*, etc. Abeillé (1997b: 22-25) proposes the following two defining properties for the DirO in French: possibility of the quantitative EN (*Alain en a mangé trois* lit. ⁽Alain thereof has eaten three⁾) and possibility of the preposition DE with negation (*Alain n'a pas mangé de steak* lit. ⁽Alain has not eaten of steak⁾). We, however, cannot use them as defining properties:

• On the one hand, they *a priori* exclude from DirOs what we believe should be considered as such: the infinitive and the completive clause that satisfy our defining properties.

• On the other hand, they include among DirOs what we believe should not be considered as such: the Quasi-Subject (otherwise, one is forced to admit the existence of a DirO with typically intransitive verbs, which is typologically implausible).

However, Abeillé's properties isolate the central subclass of DirOs: nominal referential DirOs. Thus, only such a DirO, as shown in Abeillé (1997b: 29-30), precludes the otherwise possible inversion of the Subject in a relative clause: **C'est là où lui rendront un hommage éclatant les chefs d'État* lit. 'It is there where will pay to him a strong homage the chiefs of State⁾ vs. *C'est là où lui rendront hommage les chefs d'État* lit. 'It is there where will pay homage to him the chiefs of State⁾ and *C'est là où souhaite aller Alain* 'It is there where wants to go Alain⁾.

4. Indirect-objectival SSyntRel: G-indir-obj→D

The prototypical D is a noun introduced by the preposition À.

Properties

The dependent member of the indir-obj SSyntRel is the Indirect Object [= IndirO]. It is the third most privileged clause element depending on a verb in L. In French, the SSynt-privilege of the IndirO is the following defining property:

1) The IndirO can be the target of non-specific morphological government—that is, if it is replaced with a clitic, this clitic is in the dative: $Je \ lui_{DAT} \ donne \ ce \ livre \ ^{(I)}$ give this book to him³.

Other relevant properties are descriptive:

2) The IndirO corresponds to DSyntA II or III of the governing verb: *Helen plaît* \dot{a} *Alain* [= DSyntA II] lit. 'Helen pleases Alain', *Helen donne une poire* \dot{a} *Alain* [= DSyntA III] 'Helen gives a pear to Alain'.

As is the case with the DirO, we will not discuss here in detail two complications related to the DSynt-representation of the following constructions with IndirOs:

• The IndirO expresses a raised Possessor of the Subject or of the DirO: La gorge lui brûlait lit. 'The throat was burning to-him'; Alain lui a touché l'épaule lit. 'Alain has touched the shoulder to-her'; Se fâcher avec Helen a gâché la vie à Alain lit. 'To quarrel with Helen has spoiled the life to Alain'.

• The IndirO expresses a Beneficiary: *Alain a acheté une glace à Helen* lit. ⁽Alain has bought an ice-cream to Helen⁾.

In both cases the governing verb receives—as a result of the application of a special semantic rule—an additional DSyntA (= III), which does not correspond to any of its SemAs.

3) The IndirO is not obligatorily present in every clause.

- 4) The IndirO does not obligatorily depend on the MV.
- 5) The IndirO cannot be promoted/demoted.

Exception

With the verbs OBÉIR 'obey', DÉSOBÉIR 'disobey' and PARDONNER 'pardon' the IndirO—which corresponds to DSyntA II of the verb—can be promoted to the Subject by passivization: Alain obéit à Helen lit. 'Alain obeys to Helen' ~ Helen est obéie d'Alain 'Helen is obeyed by Alain'; Alain pardonne à Helen lit. 'Alain pardons to Helen' ~ Helen est pardonnée par Alain 'Helen is pardoned by Alain'.

6) The IndirO can be the target of personal cliticization: Alain lui [= a Helen] envoie un cadeau 'Alain sends a present to-her'.

7) The IndirO can be the target of relativization: *le garçon à qui j'ai envoyé ce livre* ⁽the boy to whom I have sent this book⁾.

8) The IndirO can be the target of clefting: *C'est à Alain que je donne mon texte à lire/que je fais lire mon texte* lit. ^{(It is to Alain that I give my text to read/that I make read my text²).}

9) The IndirO can be the controller/the target of reflexivization: *Helen a ordonné* **à** *Alain de se* [= 'Alain', which is the controller] *raser* 'Helen has ordered Alain to shave himself'; *Alain se* [= 'à Alain', which is the target] *parle* 'Alain talks to himself'.

10) The IndirO cannot be the controller of the Secondary Actor in the causative FAIRE-construction (that is, the presence of an IndirO of the lexical verb does not impose the realization of the Secondary Actor as an IndirO of FAIRE: *Alain le* $\langle *lui \rangle$ [= *le chien*] *fait obéir à son maître* ⁽Alain makes him [= the dog] obey his master⁾.

11) The IndirO cannot be the controller of an actantial-attributive complement.

12) The IndirO cannot be the controller of the agreement of the MV.

13) The IndirO follows the governing verb.

14) The IndirO can be the target of left dislocation: \hat{A} mes enfants, j'envoie des livres 'To my children, I send books'.

Formal types of IndirO

- a. A noun (including the subclass of stressed pronouns) with the preposition À: *Alain a donné son livre à Helen* ⁽Alain has given his book to Helen⁾.
- b. A phrase equivalent to a noun with preposition À: *Alain donne son livre à qui le veut* ⁽Alain gives his book to who wants it⁾.
- c. A personal clitic in the dative: *Alain lui parle souvent* 'Alain often speaks to him/ her'.

The indir-obj SSyntRel is non-repeatable.

Justification

The IndirO plays quite a special role in French syntax:

• The IndirO can express the Beneficiary, which is represented at the DSynt-level as DSyntA III of the governing verb: *Alain a fait un bon repas* **à** *Helen* lit. ^{(Alain has made a good meal to Helen⁾ ~ *Alain lui a fait un bon repas* lit. ^{(Alain has made a good meal to her⁾.}}

• The IndirO can express the raised Possessor in the French possessive construction of the type *Alain a lavé la tête à Helen* lit. ⁽Alain has washed the head to Helen⁾ ~ *Alain lui a lavé la tête* lit. ⁽Alain has washed the head to her⁾, where the IndirO is a surface-syntactic realization of DSyntA I of DSyntA II of the governing verb (HELEN is DSyntA I of TÊTE ⁽head⁾).

• The IndirO can express the Secondary Actor of the causative FAIRE-construction with a transitive verb having an expressed DirO: *Alain fait lire le roman à Helen* lit. 'Alain makes read the novel to Helen' ~ *Alain lui fait lire le roman* lit. 'Alain makes read the novel to her'.

• The clitic IndirO can express DSyntA II of the verbs that govern the adjuncts DES-SUS 'on ...' and APRÈS 'behind ...': On lui a tiré dessus lit. 'They to-him have fired on' = 'They have fired at him' or Alain lui court après lit. 'Alain to-her is-running after' = '... is courting her'. (If DSyntA II is not cliticized, it is implemented as an ObIO: On a tiré sur Alain 'They have fired at Alain' and Alain court après Helen lit. 'Alain is-running after Helen'.)

The IndirO in French is typical of animate nouns only; its very existence is one of the manifestations of the syntactic feature of animacy.

These considerations enhance our decision to introduce the IndirO as a clause element different from the Oblique Object (No. 7, p. 194; in Justification 2, p. 196, it is shown that the IndirO and the OblO have different relevant properties). The traditional French grammar does not make this distinction, subsuming all prepositional nominal Ds of a verb which are its actants under the name of "complément d'objet indirect," so that this term is much broader that our Indirect Object. (In general typological studies, the IndirO is carefully distinguished from OblOs; cf., e.g., Comrie 1975: 4 and passim, Keenan & Comrie 1977, 1979. For a discussion of the opposition "DirO vs. IndirO" in a vast typological perspective, see Dryer 1986.)

Comments

1. In the Meaning-Text description of the SSynt-structure of French sentences, a nonsubject pronominal clitic depends syntactically on its host—rather than on the clause element on which its source depends;¹⁹ see, for instance, Figure 1 below (the clitic and its source are encircled). The clitic is, so to speak, transferred from its genuine, i.e. DSynt-, Governor (= ENVOYER 'send') to its host (= the auxiliary AVOIR 'have')—retaining the same subordinating SSyntRel, in our case, the **indir-obj**; we do not posit here a new SSyntRel. This solution is adopted because the linear position of French clitics in general does not depend on the specific subordinating SSyntRel—it is determined by the nature of the clitic itself (a property which is orthogonal to the properties of SSyntRels). At the same time, the grammatical case of the clitic—the accusative *vs*. the dative—must be computed from (the name of) the subordinating SSyntRel, since the Government Pattern of the new G (i.e., of the host) should not contain the necessary information: thus, AVOIR by no means itself governs the dative. This fact constitutes another argument in favor of distinguishing the IndirO *vs*. the ObIO.



Figure 1: A clitic and its host

2. Criterion 3 does not allow us to consider the so-called Dativus Ethicus of the type *Ne me fais pas de bêtises !* lit. ^(Don't do to-me stupidities!) as a particular case of IndirO, since these two clause elements can be combined:

¹⁹ The host of a clitic is the clause element that determines its linear position, that is, the wordform to which this clitic attaches linearly and prosodically (cf. Zwicky 1977).

(17) a. Il te [= Dat. Ethic.] recite trois poèmes en trois minutes à qui veut l'écouter [= IndirO]

lit. 'He recites to-you three poems in three minutes to who wants listen to him'.

b. *Il te* [= Dat. Ethic.] *lui* [= IndirO] *fabriquera une table en vingt minutes* lit. 'He manufactures to-you to-him [= for him] a table in twenty minutes' (Leclère 1979: 134).

Uniting the Dativus Ethicus and the IndirO under the same SSyntRel would create an \mathbf{r} which would be limitedly repeatable (exactly two times, and that, without intervention of the Government Pattern of the verb).

3. Similarly to what has been said about the direct-reflexive SE, the indirect-reflexive SE—as in S'ACHETER *une maison* 'buy oneself a house'—is not an IndirO: it does not correspond to any DSyntA of the governing verb. A verbal form of the type *s'acheter* [qqch.] 'buy [something] to-oneself', *se préparer* [qqch.] 'prepare [something] to-oneself' or *s'imaginer* [qqch.] lit. 'imagine [something] to-oneself', is the form of the indirect-reflexive voice. This verbal form is represented in the DSyntS as one node supplied with the grammeme ('INDIR-REFL'). In the SSyntS, the indirect-reflexive SE is linked to the verb by the same auxiliary SSyntRel as the direct-reflexive SE (there is no need to distinguish the two SE, since they do not have different case forms).

5. Pseudo-direct-objectival SSyntRel: G-pseudo-dir-obj→D

The prototypical D is a prepositionless noun.

Properties

The dependent member of the **pseudo-dir-obj** SSyntRel is the Pseudo-Direct Object [= Pseudo-DirO]—the expression of the value of a parameter or of a property. It appears with verbs of measure—PESER 'weigh', COÛTER 'cost', etc.—and a few verbs of the type SENTIR 'smell [intrans.]': *Ça coûte une fortune/300 euros* 'This costs a fortune/300 euros'; *Ça sent le hareng* 'This smells of herring'.

It is the fourth most privileged clause element depending on a verb in **L**. In French, the SSynt-privileges of the Pseudo-DirO—with respect to the Agent and to the Oblique Object—consist in that it shares more linguistically relevant properties with the DirO than the Agent or the Oblique Object. Namely, the Pseudo-DirO has the same type of cliticization and the same case government as the DirO.

The Pseudo-DirO is defined by the following combination of properties:

1) The Pseudo-DirO can be the target of personal cliticization: *Le poisson, la caisse le sent encore* lit. 'The fish, the box still smells of it'; *Cette grosse somme, ma robe la coûte* lit. 'This big sum, my dress costs it'; *La table les mesure, ses deux mètres* lit. 'The table measures them, its two meters'.

2) The Pseudo-DirO can be the target of non-specific morphological government the clitic/the relative pronoun replacing the corresponding noun is in the accusative: *Cette robe* la [= *cette somme*] *coûte* ^{(This dress costs it [= this sum]⁾; [*la somme*] *que cette robe coûte* ⁽[the sum] that this dress costs⁾.} 3) The Pseudo-DirO cannot be promoted or demoted.

4) The Pseudo-DirO cannot be the controller of the past participle in compound forms.

The first two properties are shared with the DirO, but the second two oppose the Pseudo-DirO to the DirO.

The Pseudo-DirO's descriptive properties are:

5) The Pseudo-DirO corresponds to DSyntA II of the governing verb.

6) The Pseudo-DirO is not obligatorily present in every clause.

7) The Pseudo-DirO does not depend exclusively on the MV.

8) The Pseudo-DirO can be the target of relativization: *Je trouverai les 300 euros que cette robe coûte* 'I'll find the 300 euros that this dress costs'; *On ne mangera pas ce hareng pourri que la caisse sent encore* 'We will not eat this rotten herring of which the box still smells'.

9) The Pseudo-DirO can be the target of clefting: *C'est 300 dollars que cette robe coûte*[, *pas 300 francs*] lit. ^{(It is 300 dollars that this dress costs[, not 300 francs]); *C'est le hareng que le pain sent* lit. ^{(It is the herring that the bread smells of).}}

10) The Pseudo-DirO cannot be the controller/the target of reflexivization.

11) The Pseudo-DirO does not impose the IndirO realization of the Secondary Actor in the causative FAIRE-construction: $\zeta a \ le \ \langle *lui \rangle \ [= le \ pain] \ fera \ sentir \ le \ hareng \ {}^{\circ}$ This will make it [= the bread] smell of herring³.

12) The Pseudo-DirO cannot be the controller of an actantial-attributive complement.

13) The Pseudo-DirO follows the governing verb.

14) The Pseudo-DirO cannot be the target of pure left dislocation: *300 euros, la robe coûte '300 euros, the dress costs'; *Le hareng, la caisse sent 'The herring, the box smells of'. (300 euros, la robe coûte can be grammatical—with a descending contour on 300 euros; in this case, 300 euros is not dislocated, but fronted.)

Formal types of Pseudo-DirO

- a. A noun (particularly, a Num + N phrase): *Il pesait 60 kilos* 'He weighed 60 kg'; *Ça* sent le brûlé lit. 'It smells of [something] burnt' (*rien* 'of nothing)/*le hareng* 'of herring').
- b. A phrase equivalent to a noun: *Ça va me coûter exactement ce que je veux* 'It will cost me exactly what I want'.
- c. A personal clitic in the accusative: *Cette robe la* [= *cette somme*] *coûte* ⁽This dress costs it [= this sum]⁾.

191

Justification

The **pseudo-dir-obj** SSyntRel cannot be subsumed under the **dir-obj** or the **obl-obj** SSynt-Rels, which both have prepositionless nouns among their Ds, for the following reasons:

• A Pseudo-DirO contradicts two defining properties of a DirO—it does not control the Secondary Actor in the causative construction and the agreement of the participle in compound forms.

• A Pseudo-DirO is not substitutable by the prototypical ObIO, i.e., a prepositional phrase (Criterion 2). Moreover, a Pseudo-DirO differs from an ObIO with respect to cliticization: unlike an ObIO, a Pseudo-DirO is replaceable by an accusative clitic.

Neither can the **pseudo-dir-obj** SSyntRel be subsumed under the **pred-attr-compl** SSyntRel (see below, 3.3, No. 13, p. 212: *coûter cher* lit. 'cost expensive', *sentir bon* 'smell good'), which is semantically very close, since Criteria 2 and 3 do not allow us to collapse them:

• They do not have the same prototype (the prototypical D is N for the Pseudo-DirO and A for the PredAttrCo).

• A Pseudo-DirO and a PredAttrCo can appear as co-subordinates of the same verb: *La cuisine sent bon le pain frais* lit. ^(The kitchen smells good—fresh bread). Therefore, the presumed "united" SSyntRel would not feature an admissible value of repeatability: it would be limitedly repeatable.

On the other hand, the **pseudo-dir-obj** SSyntRel covers two semantically different cases: real measures and perceptional parameters, which are distinguished, for instance, by different interrogative words: *Ça mesure combien* ? 'This measures how-much?' [-2 mètres '2 meters'] vs. *Ça sent quoi* ? 'This smells of-what?' [-Le hareng 'The herring']. This can be an indication that we deal here with two different SSyntRels.

The pseudo-dir-obj SSyntRel is non-repeatable.

6. Agentive SSyntRel: G-agentive→D

The prototypical D is a noun introduced by the preposition PAR or DE.

Properties

The dependent member of the **agentive** SSyntRel is the Agent. It is the fifth most privileged clause element depending on a verb: *Helen a été reçue par son frère* ⁽Helen was received by her brother'; *Alain fait lire le roman par Helen* lit. ⁽Alain makes read the novel by Helen'; *Aimée de tout le monde, Helen* ... ⁽Loved by everybody, Helen ...).

NB: The term *Agent* should not be construed as a semantic entity; the corresponding clause element does not necessarily denote people nor semantic agents. Thus, *par la cérémonie* is an Agent in *Sa fureur a été changée en excessive amabilité par la cérémonie du matin* ⁽His fury has been changed into excessive amiability by the morning ceremony⁾.

In French, the SSynt-privileges of the Agent are the following two defining properties:

1) The Agent corresponds to DSyntA II of the governing verb in the passive form (it corresponds to the Subject of the active form of the verb) or to DSyntA I of the governing verb in the infinitive in a number of constructions (*faire lire le roman par Helen* [= Agent of LIRE] lit. 'make read the novel by Helen'; *se faire voler par un mendiant* lit. 'make oneself rob by a beggar', *se voir refuser un contrat par le gouvernement* lit. 'see oneself refuse a contract by the government', etc.).

2) The Agent is the target of non-specific lexical government: it is always introduced by the preposition PAR 'by' (with the exception of a handful of verbs which take an Agent with the preposition DE).

NB: When depending on a noun, the Agent corresponds to its DSyntA I and has more means of expression (*une traduction par Alain* 'a translation by Alain', *l'arrivée d'Alain* 'Alain's arrival', *une conversation entre amis* 'a conversation among friends', etc.).

The Agent has the following descriptive properties:

3) The Agent is not obligatorily present in every clause.

4) The Agent does not obligatorily depend on the MV.

5) The Agent cannot be promoted/demoted.

6) The Agent can be the target of cliticization, but only if it is introduced by the preposition DE (the phrase PAR + N is not cliticizable in principle); the replacing clitic is EN: *Il plaisait aux femmes; en fait, il en était adoré* lit. ^{(Women liked him; actually, he was adored thereby⁾.}

7) The Agent can be the target of relativization: *La femme par qui* Alain est gâté est très belle 'The woman by whom Alain is spoilt is very beautiful'; *Les étudiants dont ce prof est tellement aimé* ... 'The students by whom this professor is loved so much ...'.

8) The Agent can be the target of clefting: *C'est par Alain que Helen est gâtée* ⁽It is by Alain that Helen is [being] spoilt⁾.

9) The Agent cannot be the controller/the target of reflexivization.

10) The Agent cannot be the controller of the Secondary Actor in the causative FAIRE-construction (strictly speaking, this property is not applicable).

11) The Agent cannot be the controller of an actantial-attributive complement.

12) The Agent cannot be the controller of the agreement of the MV.

13) The Agent follows the governing verb.

14) The Agent can be—in informal speech—the target of left dislocation: [?]*Par Alain*, *ce travail sera fait vite et très bien* lit. ^{(By} Alain, this job will be done fast and very well⁾.

Formal types of Agent

- 1. A noun (in the broad sense) introduced by the preposition PAR or DE:
 - a. A noun (including the subclass of stressed pronouns): *Aimée de tous ses amis/de nous, Helen* ... ⁽Loved by all her friends/by us, Helen ... ⁾.
 - b. A phrase equivalent to a noun (e.g., a headless relative): *Elle se laisse séduire par qui le veut* (She lets herself to be seduced by who wants it).
- 2. The clitic EN: Il en était adoré lit. 'He was adored thereby'.

For considerations that justify the introduction of the **agentive** SSyntRel as different from the **oblique-objectival** SSyntRel, see below, No. 7, Justification 3, p. 197.

The agentive SSyntRel is non-repeatable.

7. Oblique-objectival SSyntRel: G-obl-obj→D

The prototypical D is a noun with a preposition.

Properties

The dependent member of the **obl-obj** SSyntRel is the Oblique Object [= OblO]. It is the least privileged clause element, which does not have defining properties: it is character-ized rather negatively, by the opposition to other nominal objects.

The OblO's descriptive properties:

1) The OblO can correspond to any DSyntA of the governing verb, except for DSyntA I: *insister sur* N [= DSyntA II] ^(insist on N), *inviter* N \dot{a} V-er [= DSyntA III] ^(invite N to V-inf), *louer* N \dot{a} N *pour* Num *euros* [= DSyntA IV] *pour* Num *mois* [= DSyntA V] ^{(rent out N to N for Num euros for Num months).}

2) The OblO is not obligatorily present in every clause.

3) The OblO does not obligatorily depend on the MV.

4) The OblO cannot be promoted/demoted.

5) If the OblO is a noun introduced by the preposition \dot{A} or DE, it can be the target of cliticization by Y (but not by **lui**!) or EN: J'y [= \dot{a} ce projet] renonce lit. 'I renounce thereto [= this project]'; Ma décision en [= de ta présence] dépend 'My decision depends thereon [= your presence]'. With a few verbs, the OblO introduced by the preposition SUR can also be cliticized by Y: J'y [= sur ta présence] compte beaucoup 'I count much thereon [= your presence]'.

6) The OblO can be the target of relativization: *le principe sur lequel j'insiste* 'the principle on which I insist'; *le principe dont ma décision dépend* 'the principle on which my decision depends'; *Ces dix centimètres dont Alain la dépasse sont très importants* 'These 10 cm by which Alain exceeds her are very important'.

7) The OblO can be the target of clefting: C'est sur ce point que j'insiste 'It is on this point that I insist'; C'est à construire la maison qu'Alain m'a aidé 'It is to build the

house that Alain has helped me³; *C'est de dix centimètres qu'Alain la dépasse* 'It is by 10 cm that Alain exceeds her³.

8) The OblO cannot be the controller/the target of reflexivization (as opposed, for instance, to the DirO and the IndirO).

9) The OblO cannot be the controller of the Secondary Actor in the causative FAIRE-construction.

10) The OblO cannot be the controller of an actantial-attributive complement.

11) The OblO cannot be the target of non-specific government (the clitics EN and Y do not have case, and the **obl-obj** SSyntRel does not impose a specific preposition).

12) The OblO cannot be the controller of the agreement of the MV.

13) The OblO follows the governing verb.

14) The OblO can be the target of pure left dislocation: *Sur ce point*, *j'insiste* lit. ^{(On this point, I insist⁾; \hat{A} mourir si jeune, elle ne s'attendait pas lit. ^{(To die so young, she did not expect⁾; *De travailler le matin*, *Alain n'enrage pas* lit. ^{(To work in the morning, Alain is not angered by⁾.}}}

NB: Some specific OblOs cannot be dislocated for purely semantico-communicative reasons: **En sanglots, Helen éclate* 'In sobs, Helen bursts out'; **De quelques mètres, Alain s'est approché* 'A few meters, Alain has approached' (cf., however, *De 3 mètres, Alain s'est approché de Helen; Danielle, c'était 5 mètres* '3 meters, Alain has approached Helen; for Danielle, it was 5 meters').

Formal types of OblO

1. A noun (in the broad sense) introduced by a preposition:

- a. A noun (including stressed pronouns): *insister sur le départ* 'insist on the departure', *en vouloir* [à N = IndirO] *de ces paroles* 'hold a grudge against N for having said this', s'approcher de quelques mètres 'approach a few meters', *penser à Helen* 'think of Helen', *éclater en sanglots* 'burst out in sobs', *rémonter à 1937* 'go back to 1937', *soumettre* [N] *à un test* 'submit [N] to a test', *échanger* [N] *avec Jean contre un vélo* 'exchange [N] with Jean against a bike', etc.
- b. A phrase equivalent to a noun (e.g., a headless relative or a quantitative phrase): *Je me moque de qui viendra* ^{(I} don't care who is coming⁾.

2. The clitic EN or Y: J'en raffole 'I am very keen thereon', J'y pense 'I think thereof'.

3. An infinitive introduced by a preposition: *Alain a consenti à travailler* ⁽Alain has agreed to work⁾, *Alain doute de pouvoir venir* ⁽Alain doubts that he can come⁾, *Alain se passe de dormir* ⁽Alain can do without sleeping⁾.

4. A subordinate QUE-clause: *douter que* CLAUSE 'doubt that ...', *prévenir* [N] *que* CLAUSE 'warn that ...'.

5. A prepositionless noun (with the verbs PARLER, CAUSER, ACHETER, VENDRE, PAYER, VOTER and a few others): *parler politique* $\langle b\acute{e}b\acute{e}s \rangle$ 'talk politics $\langle children \rangle$ ', *causer affaires* 'talk business', *vendre ses concombres 3 euros* 'sell one's cucumbers 3 euros', *payer cette robe 300 euros* 'pay this dress 300 euros', *voter Mitterrand* 'vote for Mitterrand'.

NB: The **obl-obj** SSyntRel is repeatable (for instance, *Leo a changé des marks en pesetas* [= OblO] *avec Marga* [= OblO] ^{(Leo has exchanged marks for pesetas with Marga⁽⁾).}

Justification

1. The constructions subsumed under the obl-obj SSyntRel satisfy all Criteria 1-3.

— Criterion 1: no Ds of the **obl-obj** SSyntRel that can appear with the same G contrast semantically while differing only by some syntactic means. Thus, consider the case of *dessiner au pinceau* 'draw with a brush' vs. *dessiner sur le pinceau* 'draw [something] on a brush'; here, *au pinceau* and *sur le pinceau* both are OblOs corresponding to different DSyntAs of DESSINER and they do contrast semantically. But these expressions do not contradict Criterion 1, because their difference is not in purely syntactic means: they are distinguished by prepositions, which are different lexemes.

— Criterion 2: all Ds are either implemented by a prepositional phrase, or are substitutable by a prepositional phrase, as in *parler politique* 'talk politics' ~ *parler de la politique* 'talk of the politics';²⁰ voter Mitterrand ~ voter pour Mitterrand 'vote for M.'; douter qu'il part 'doubt that he leaves' ~ douter de son départ lit. 'doubt of his departure'.

— Criterion 3: the **obl-obj** SSyntRel is repeatable (the possible number of OblOs is specified by the governing verb), cf. *Il m'invite au restaurant* [= OblO] *pour manger des moules* [= OblO] ^{(He invites me to a restaurant to eat mussels); *Il m'aide à constru-ire* [= OblO] *la maison avec son argent* [= OblO] ^{(He helps me build the house with his money); *Alain loue sa voiture pour 3 mois* [= OblO] *à 20 euros* [= OblO] *par jour* ^{(Alain rents his car for 3 months 20 euros a day).}}}

2. Traditional French grammar, as we have said, does not distinguish the IndirO and the ObIO. We, however, think that they should be separated: only the IndirO, but not the ObIO, is replaceable with the personal dative clitic **lui** 'to him/her'/**leur** 'to them', controls reflexivization and can itself be reflexivized. Cf. such indicative examples as the following ones (Blanche-Benveniste 1975: 39-40):

(18) a. Alain obéit à Helen [= IndirO] 'Alain obeys Helen'. ~ Alain lui obéit (*Alain obéit à elle).
b. Alain renonce à Helen [= OblO] 'Alain renounces Helen'. ~ Alain renonce à elle (*Alain lui renonce).

The verbs TÉLÉPHONER 'telephone' and APPARTENIR 'belong' take the phrase \hat{a} N as the IndirO, the verbs RÊVER 'dream [of N]' and RÉFLÉCHIR 'think over [N]'—as an

²⁰ The expressions *parler politique* and *parler de la politique* are not synonymous (*parler politique* is phraseologized and implies an exchange of political opinions), just as *parler affaires* 'talk business' is not synonymous with *parler des affaires* 'talk about business'. This, however, is irrelevant in the present context: the only important thing for us is the fact that the verb PARLER is used in both these expressions in the same sense ('talk').
OblO. The verb SERVIR 'serve' admits two \dot{a} N-phrases: one as an IndirO, the other as an OblO, cf. *Cela sert* \dot{a} *Alain* [= IndirO] *au nettoyage* [= OblO] *des toilettes* 'This serves Alain for the cleaning of the toilet³; the first is cliticizable by **lui**, the second does not cliticize (moreover, the OblO of SERVIR can be expressed by an infinitive: *Cela sert* \dot{a} *Alain* \dot{a} *nettoyer les toilettes* 'This serves Alain to clean the toilet³). See also above, No. 4, p. 189.

3. Neither can the Agent be subsumed under the obl-obj SSyntRel:

• The preposition that introduces an OblO is specified by the Government Pattern of the verb, while the choice of the basic preposition of the Agent—PAR—does not depend on the verb; it is imposed by the SSyntRel itself (this is a case of non-specific lexical government). As was already indicated—No. 6, Item 2, p. 193, this statement is true with the exception of a few verbs that require DE.

• In contrast to the OblO, left dislocation is problematic for the Agent:

(19) a. *Pour Alain* [= OblO], une telle lettre a été écrite la semaine dernière
 <sup>(For Alain, such a letter has been written last week⁾.
</sup>

vs.

[?]Par Alain [= Agent], une telle lettre a été écrite la semaine dernière ⁽By Alain, such a letter has been written last week⁾.

This is of course a consequence of the Agent being a result of communicative demotion, while left dislocation is a means to express (among others) Focalized Topicalization, i.e., communicative promotion. The Agent's demoted status is also seen in that the Agent discourages proleptization as well:

b. Alain, cette lettre a été écrite pour lui [= OblO], lit. ⁽Alain, this letter has been written for him⁾.
vs. ^{??}Alain, cette lettre a été écrite par lui [= Agent], lit. ⁽Alain, this letter has been written by him⁾.

• Criterion 3 does not allow us to collapse the Agent with the OblO: the **obl-obj** SSyntRel is repeatable, but the **agentive** SSyntRel is not (a Governor can have two, three, etc. OblOs—as a function of its Government Pattern, but there can be no more than one Agent per Governor).

Comment

In the expressions of the type *parler politique*, *causer affaires* and *voter Mitterrand* the D does not behave like a typical ObIO: it does not allow relativization, clefting or pure left dislocation. However, this is an extremely phraseologized construction (possible with just a few verbs), so that we need not require from these expressions to be fully similar to the prototype.

3.2 SSyntRels whose prototypical dependent is an infinitive

INFINITIVAL-OBJECTIVAL SSynt-Relations (8 - 10)

8. Direct-infinitival-objectival SSyntRel: G-dir-inf-obj→D

The prototypical D is a prepositionless infinitive.

Properties

The dependent member of the **dir-inf-obj** SSyntRel is the Direct-Infinitival Object [= DirInfO]—a prepositionless infinitive depending on the verbs *pouvoir* (can), *devoir* (must), *savoir* [*nager*] (know [how to swim]), *faillir* [*tomber*] (almost [fall]), *daigner* (deign), ...

The dir-inf-obj SSyntRel has two defining properties:

1) The DirInfO can be the target of cliticization by the "neuter" LE, but not with any governing verb: Je peux/dois partir 'I can/must leave' ~ Je le peux/le dois 'I can/must it'. Cf., however: Je sais nager 'I know how to swim' ~ *Je le sais; J'ai failli tomber 'I almost fell' ~ *Je l'ai failli; Il n'a pas daigné lire ce texte 'He did not deign to read this text' ~ *Il ne l'a pas daigné.

Although with SAVOIR, FAILLIR and DAIGNER the cliticization of the infinitive is impossible, we have decided to keep this infinitive in the DirInfO group, since it cannot be covered by any of the other Infinitival SSyntRels or by the **dir-obj** SSyntRel—in conformity with Criterion 2, and it seems counterintuitive to have for the governed infinitive of these three verbs a separate SSyntRel. On the other hand, the impossibility of cliticization of the governed infinitive is an individual property of the governing verb and has to be specified as a piece of lexicographic information in its Government Pattern.

NB: The impossible cliticization of the infinitive governed by the three above verbs does not contradict the first defining property of the **dir-inf-obj** SSyntRel (a contradiction would be a possible cliticization of a different type, e.g., by Y).²¹

2) In most cases, the DirInfO can be the target of left dislocation: **Partir**, je peux lit. 'To leave, I can'; coll. **Nager le crawl**, je sais bien 'To do the crawl, I know well'; **Lire ce texte**, il n'a pas même daigné lit. 'To read this text, he has not even deigned'. With some verbs, however, left dislocation is infelicitous or impossible: '**Partir**, je dois lit. 'To leave, I must'; ***Tomber**, il a failli à cause de sa nonchalance lit. 'To fall, he almost did because of his carelessness'.

The dir-inf-obj SSyntRel has the following descriptive properties:

3) The DirInfO corresponds to DSyntA II of the governing verb.

²¹ Our decision to subsume SAVOIR + V_{inf} under the **dir-inf-obj** SSyntRel is buttressed by the fact that the expression of the type [?]Je **lui** ferai savoir nager lit. ^{(I} will make to-him know how to swim⁾, although not ideal, is much better than *Je **le** ferai savoir nager. In this respect (= imposing the IndirO realization on the Secondary Actor of the FAIRE-construction), the V_{inf} with SAVOIR is thus closer to the nominal DirO.

4) The DirInfO is not obligatorily present in every clause (it is used only with a very small group of verbs).

5) The DirInfO does not obligatorily depend on the MV.

6) The DirInfO cannot be implicated in promotion or demotion.

7) The DirInfO cannot be the target of clefting: *C'est partir que je peux 'It is to leave that I can'.

8) The DirInfO cannot be the controller of the agreement of the MV.

9) The DirInfO follows the governing verb.

The dir-inf-obj SSyntRel is non-repeatable.

Formal types of the DirInfO

1. A prepositionless infinitive: pouvoir partir 'be able to leave'.

2. The "neuter" LE clitic: *Il le peut* 'He can it'.

Justification

This SSyntRel cannot be collapsed neither with the **dir-obj** SSyntRel, nor with the **copred-inf-obj** SSyntRel:

1) The DirInfO cannot be considered as a particular case of DirO, since it is not substitutable by the prototypical DirO—a noun (No. 3, Comment 5, p. 184).

2) The DirInfO cannot be considered as a particular case of CopredInfO (see below, No. 10, p. 201), since they feature different defining properties.

9. Oblique-infinitival-objectival SSyntRel: G-obl-inf-obj→D

The prototypical D is an infinitive introduced by a preposition.

Properties

The dependent member of the **obl-inf-obj** SSyntRel is the Oblique-Infinitival Object [= OblInfO]—an infinitive introduced by a preposition: *commencer* à V_{inf} (begin), *continuer* à/de V_{inf} (continue), se dépêcher de V_{inf} (hurry), réussir à V_{inf} (manage), finir par V_{inf} (finish), ... Exceptionally, the OblInfO can be a prepositionless infinitive: it happens only with the verbs PENSER (intend) and CROIRE (believe).²²

²² However, the construction PENSER- $\mathbf{r} \rightarrow que$ + CLAUSE is treated differently: we see here a different lexeme PENSER 'think', with which the QUE-clause is considered a DirO, since one can say **Qu'il faut se préparer** était pensé par tout le monde 'That it is necessary to prepare oneself was thought by everybody' and **Ce qu'Alain a pensé**, tout le monde le pense 'What Alain has thought, everybody thinks it'. Even if this verb does not accept the prototypical DirO—a regular noun, it takes nouns featuring a pronominal character: Alain pense la même chose 'Alain thinks the same thing', without mentioning the demonstrative CELA/ CE [QUE] or the negative RIEN 'nothing'. The same reasoning applies to CROIRE- $\mathbf{r} \rightarrow que$ + CLAUSE.

The **obl-inf-obj** SSyntRel does not have defining properties: within its group, it is characterized rather negatively, by opposition to other infinitival objects. (Let it be recalled that the same situation obtains with the **obl-obj** SSyntRel, No. 7, p. 194.)

The obl-inf-obj SSyntRel has the following descriptive properties:

- 1) The OblInfO corresponds to DSyntA II of the governing verb.
- 2) The OblInfO is not obligatorily present in every clause.
- 3) The OblInfO does not obligatorily depend on the MV.
- 4) The OblInfO cannot be promoted/demoted.
- 5) The OblInfO cannot be the target of cliticization.

Exception

With a few individual verbs, the cliticization by Y is possible: Alain pense préparer un cassoulet ^{(Alain} intends to prepare a cassoulet⁾. ~ Alain y pense; Alain a réussi à me convaincre ^{(Alain} has succeeded to convince me⁾. ~ Alain y a réussi.

6) The OblInfO cannot be the target of clefting: **C'est à travailler que je commence* lit. ⁽It is to work that I begin⁾.

7) The OblInfO cannot be the controller of the agreement of the MV.

8) The OblInfO follows the governing verb.

9) In most cases, the OblInfO cannot be the target of pure left dislocation (note that under left dislocation, an infinitive loses its preposition À, so that in the following examples this À is omitted): **Préparer le repas, Alain se dépêche* lit. 'To prepare the meal, Alain hurries'; ??*Partir pour la France, Alain pense depuis trois mois* lit. 'To leave for France, Alain has been intending for three months'; ?*Préparer le repas, Alain commencera dès demain* lit. 'To prepare the meal, Alain will begin tomorrow'. Cf., however, *Faire un bon repas, Alain réussit toujours* lit. 'To make a good meal, Alain always manages'.

Formal types of the OblInfO

- a. An infinitive with a preposition: *Alain a fini par accepter* lit. ^{(Alain has finished by to accept²).}
- b. A prepositionless infinitive: *Nous croyons avoir le droit de rester* 'We believe to have the right to stay'; *Je pense partir* 'I intend to leave'.

The obl-inf-obj SSyntRel is non-repeatable.

Justification

1. Strictly speaking, Criterion 2 does not allow us to subsume the constructions **a** and **b** under the same SSyntRel. However, since the problem is related just to two verbs— CROIRE and PENSER in the given senses, we decided to force the matters a bit, violating our own principles. The construction CROIRE/PENSER + V_{inf} cannot be subsumed under the dir-inf-obj SSyntRel, because this would contradict the defining properties of the latter: the infinitive in the dir-inf-obj SSyntRel is not cliticizable by Y and admits left dislocation, while with CROIRE/PENSER it is replaceable with Y, but cannot be dislocated. At the same time, it satisfies the properties of the obl-inf-obj SSyntRel.

2. The obl-inf-obj SSyntRel cannot be collapsed with none of the following four SSyntRels: the dir-obj SSyntRel, the obl-obj SSyntRel, the dir-inf-obj SSyntRel, and the copred-inf-obj SSyntRel:

1) The OblInfO cannot be considered as a particular case of the DirO, since it is not substitutable by the prototypical DirO—a prepositionless noun (Criterion 2), even in the case of a prepositionless infinitive: *Tout le monde pense partir* 'Everybody intends to leave' ~ **Tout le monde pense le départ* 'Everybody intends the departure'. In addition, the OblInfO does not satisfy the defining properties of the DirO (**Tout le monde le* [= *partir*] *pense* 'Everybody thinks it'; **Partir est pensé part tout le monde* 'To leave is thought by everybody', *Cela le* (**lui*) *fera penser partir* 'This will make intend to leave').

2) The OblInfO cannot be considered as a particular case of OblO, because it is not substitutable by a prepositional noun phrase: *Alain se hâte de partir* ⁽Alain hurries to leave⁾ ~ **Alain se hâte de/du départ* ⁽Alain hurries of the departure⁾. Moreover, the OblInfO and the OblO differ with respect to clefting: **C*'est *de partir* [= an OblInfO] *qu'Alain se hâte* ^{(II}t is to leave that Alain hurries⁾ ~ *C*'est *à partir* [= an OblO] *qu'Alain a consenti* ^{(II}t is to leave that Alain has agreed⁾; cf. *C'est au départ qu'Alain a consenti* ^{(II}t is to the departure that Alain has agreed⁾.

3) The OblInfO is different from the DirInfO with respect to cliticization and left dislocation: only the latter, but not the former, can be cliticized and admits pure left dislocation.

4) The OblInfO cannot be considered as a particular case of CopredInfO, since it contradicts the defining property of the latter (see immediately below).

The introduction of the last two infinitival objects—DirInfO and OblInfO—runs counter to the French grammatical tradition (cf. also Candito 1999). The corresponding infinitives are generally collapsed with nominal objects, i.e., with the DirO and the OblO. We, however, find this practice unfortunate: our infinitival objects combine with the verbs whose corresponding DSyntA (= II) expresses essentially a semantic fact (an event, an action, a state, etc.) rather than a semantic name; therefore, their prototypical D must be verbal rather than nominal. The governing verbs in question have very special ("semi-grammatical") meanings: modal, phasic, aspectual, etc. It is not for nothing that these verbs participate in paraphrases in which they become adverbial modifiers of the governed infinitive: *Il continue à écrire* ⁽He continues to write⁾ \approx *Il écrit toujours* ⁽He is still writing⁾; *Il se hâte de sortir* ⁽He hurries to go out⁾ \approx *Il sort en hâte* ⁽He goes out in a hurry⁾; *Il hésite à répondre* ⁽He hesitates to answer⁾ \approx *Il répond de façon hésitante* ⁽He answers in a hesitant way⁾.

10. Copredicative-infinitival-objectival SSyntRel: G-copred-inf-obj→D

The prototypical D is a prepositionless infinitive.

Properties

The dependent member of the copred-inf-obj SSyntRel is the Copredicative-Infinitival Object [= CopredInfO]—a prepositionless infinitive which depends on a verb of perception or else on FAIRE 'make', LAISSER 'let' or ENVOYER 'send'. This is the infinitive in the Accusativus Cum Infinitivo construction: Alain voit Helen traverser la rue 'Alain sees Helen cross the street'; Alain fait traverser la rue à Helen lit. 'Alain makes cross the street to Helen'; Alain laisse Helen traverser la rue 'Alain lets Helen cross the street'; Alain envoie Helen chercher du vin 'Alain sends Helen to bring some wine'; [?]Helen a été vue traverser la rue 'Helen has been seen to cross the street'. The traditional French-and, more generally, European-grammar does not have an established term for the dependent member of this SSyntRel. Le Goffic (1993: 275-276) considers it as the second direct object and proposes to call it "prédicat de l'objet;" this seems rather infelicitous, since the first direct object is not obligatorily present: Il entend chanter de vieilles chansons 'He hears sing old songs'. Yet the infinitive in question does have a semantic link to the DirO-it is, so to speak, predicated about it. In this respect, this infinitive is similar to copredicative attributes such as I drink my coffee hot, to pound the metal flat, He prefers Helen slim, etc. It is for this reason that we decided to call the infinitive in the constructions of the Accusativus cum Infinitivo type copredicative-infinitival object.

The CopredInfO has the following defining property:

1) The CopredInfO corresponds to DSyntA III of the governing verb, whose DSyntA II is realized as a DirO (which may be absent); this DSyntA II is, at the same time, coreferential with DSyntA I of DSyntA III.²³

Other relevant properties of the CopredInfO are:

2) The CopredInfO is not obligatorily present in every clause (it is used with a very small group of verbs).

3) The CopredInfO does not obligatorily depend on the MV.

4) The CopredInfO cannot be implicated in promotion or demotion.

5) The CopredInfO cannot be the target of cliticization.

6) The CopredInfO cannot be the target of clefting: **C'est traverser le fleuve à la nage que j'ai vu Alain* ^{(It} is to cross the river swimming that I have seen Alain).

7) The CopredInfO cannot be the controller of the agreement of the MV.

8) The CopredInfO follows the governing verb: *Il voit Helen danser/Il voit danser Helen* ⁽He sees Helen dance⁾.

²³ The latter fact explains why some researchers consider the DirO in this construction to be the Synt-Subject of the infinitive (e.g., Isaac 1986: 26-27). However, in French, it does not have any defining properties of the Subject (except for corresponding to DSyntA I), while it can be promoted to Subject and is replaceable by an accusative clitic, as all DirOs.

9) The CopredInfO cannot be the target of pure left dislocation: **Traverser le fleuve* à *la nage*, *j'ai vu Alain* lit. ⁽Cross the river swimming, I have seen Alain⁾ (in the correct sentence *Traverser le fleuve à la nage*, *j'ai vu Alain le faire* lit. ⁽To cross the river swimming, I have seen Alain to do it⁾, the infinitival phrase is a Prolepse, not a dislocated CopredInfO); **Se lever très tôt*, *j'ai fait Alain* lit. ⁽Get up early, I have made Alain⁾.

NB: The copred-inf-obj SSyntRel has an additional important property: its G is a *clitic-attracting* verb. This means that this verb admits the raising of the clitic from (a dependent of) its CopredInfO: *Il entend parler de leur voyage* lit. ^{(He} hears speak of their trip⁾ ~ *Il en entend parler* lit. ^{(He} hears speak thereof⁾; *J'enverrai prendre le livre* ^{(I'II} send to take the book⁾ ~ obsol. *Je l'enverrai prendre* ^{(I'III} send to take it⁾; *Cet air, je le laisse chanter* lit. ^{(This tune, I let sing it⁾.}

The copred-inf-obj SSyntRel is non-repeatable.

Justification

Criteria 2 and 3 forbid us to collapse this SSyntRel either with the **dir-obj** SSyntRel or the **obl-obj** SSyntRel:

• The CopredInfO is substitutable neither by a prototypical DirO (= a noun) nor by a prototypical OblO (= a prepositional phrase).

• Collapsing the **copred-inf-obj** SSyntRel with the **dir-obj** SSyntRel or the **obl-obj** SSyntRel would violate the value of their repeatability (it would allow for just two DirOs or two CopredInfOs). At the same time, their relevant properties are different.

Comments

1. The causative construction FAIRE-copred-inf-obj \rightarrow V_{inf} differs from other constructions covered by the copred-inf-obj SSyntRel in that it is characterized by a particular word order: it is so monolithic in this respect that only the reflexive SE of V_{inf}, a clitic governed by this V_{inf} (under very specific conditions: for instance, Cela la faisait me téléphoner tous les matins 'This made her call me every morning') or some adverbs and parentheticals can linearly separate FAIRE and the governed infinitive: *Alain fait Helen danser 'Alain makes Helen dance' vs. Alain laisse (voit) Helen danser 'Alain lets (sees) Helen dance⁾ (Tasmowski-de Ryck 1984). Furthermore, DSyntA II of the causative construction (= the Secondary Actor) is expressed either by the DirO or by the IndirO of FAIRE—as a function of the presence of a DirO with the Vinf; however, other Accusativus-Cum-Infinitivo constructions also show this feature, even if less frequently: for instance, Il a laissé faire le travail à ses assistants lit. 'He has let do the job to his assistants). All other linguistic properties which have been selected in this paper as relevant to the establishment of SSyntRels in French are shared by the causative FAIRE-construction and all the other Accusativus-Cum-Infinitivo constructions. We have therefore opted to subsume the causative FAIRE-construction under the copred-inf-obj SSyntRel and formulate additionally its individual properties (to be included in the lexicographic entry for the causative FAIRE)—rather than to create a special SSyntRel just for this construction.

2. The constructions with the semi-auxiliaries SE FAIRE and SE VOIR (*Il s'est fait tuer* 'He's got killed'; *Il s'est vu refuser l'accès* lit. 'He has seen himself to refuse the access' = 'He was refused the access') are also described via the **copred-inf-obj** SSyntRel: in spite of their semantic particularities, they have the same relevant syntactic properties.

As the preceding discussion has shown, the infinitive (prepositionless or introduced by a preposition) which depends on a verb can fulfill one of the following five SSyntroles: a Direct Object, an Oblique Object, a Direct-Infinitival Object, an Oblique-Infinitival Object, and a Copredicative-Infinitival Object. Let us give examples of corresponding verbs:

V-dir-obj \rightarrow V_{inf} (the infinitive is substitutable by the prototypical D of the dir-obj SSynt-Rel, i.e., a prepositionless noun):

```
apprendre à [parler chinois] 'learn to [speak Chinese]'

attendre de 'wait to' (cf. Être admis est attendu de plusieurs malades depuis

longtemps lit. 'To be admitted is awaited by several patients for a long time' and Alain

attend d'être admis 'Alain awaits to be admitted' ~ Alain l'attend)

chercher à 'try to' (cf. Alain cherche à convaincre ses opposants 'Alain seeks to con-

vince his opponents' ~ Il le cherche. The verb CHERCHER 'try to achieve ....' (as

opposed to CHERCHER 'look for') has no passive form)

craindre de 'be afraid of'

decider de 'decide to'

demander à [être admis] 'ask to [be admitted]'

exiger de 'require to'

jurer de [dire la vérité] 'swear to [tell the truth]'

promettre de 'propose to'

regretter de 'regret to', vouloir 'want', ...
```

 $V-obl-obj \rightarrow V_{inf}$ (the infinitive is substitutable by the prototypical D of the obl-obj SSynt-Rel, i.e., a noun introduced by a preposition):

consentir à 'agree to' *se décider* à 'bring oneself to' *douter de* [*pouvoir venir*] 'doubt to [be able to come]' *inviter* à 'invite to' *obliger* à 'oblige to' *parler de* [*partir*] 'talk of [leaving]' *soupçonner de* 'suspect of', ...

The verb RISQUER ^(risk) represents an interesting case: it governs two infinitives, one as a DirO and the other as an OblO, cf. *Tu risques* **d'avoir** [= DirO] *une contravention* **à** *te garer ici* [= OblO] ^{(You risk a ticket if you park here⁾.}

V-dir-inf-obj \rightarrow V_{inf} (the infinitive is the prototypical D of this SSyntRel, and cliticization via "neuter" LE is possible): *pouvoir* (can), *devoir* (must), ...

V-obl-inf-obj \rightarrow V_{inf} (the infinitive is the prototypical D, but cliticization via "neuter" LE is impossible):

commencer à/par 'begin to/by' *continuer de* 'continue to' *se dépêcher de* 'hurry to' essayer de 'attempt to' (this ESSAYER means 'attempt'; in essayer la nouvelle robe 'try the new dress' we have a different lexeme ESSAYER 'test') se hâter de 'hasten to' hésiter à 'hesitate to' persister à 'persist to' réussir à 'manage to' tâcher de 'attempt to' tarder à 'be long V-ing' terminer par 'finish by', ...

V-copred-inf-obj \rightarrow V_{inf} (the infinitive is the prototypical D; the Accusativus cum Infinitivo construction): *voir* (see), *entendre* (hear), *faire* (make), *laisser* (let), ...

3.3 SSyntRels whose prototypical dependent is an adjective

ATTRIBUTIVE-COMPLETIVE SSynt-Relations (11 - 13)

11. Copular-attributive-completive SSyntRel: G-cop-attr-compl→D

Properties

Its dependent member is the Copular-Attributive Complement [= CopAttrCo]. i.e., a clause element depending on a copular verb and corresponding to its DSyntA II: *Il est malade* ⁽He is sick⁾, *Elle reste directrice* lit. ⁽She remains she-director⁾, *Cet ensemble devient complet* ⁽This set becomes complete⁾, *La maison fait très grande* lit. ⁽The house does very big⁾ = ⁽... gives the impression of being very big⁾.

The copular verbs form a small closed set: ÊTRE 'be', DEVENIR 'become', RESTER (*malade*) 'remain (sick)', DEMEURER (*son ami*) 'remain (his friend)', SEMBLER 'seem', PARAÎTRE 'seem', FAIRE (*petit*) 'give the impression of (being small)'. A copula includes in its meaning the semantic component 'be' and has two DSyntAs which correspond to two SemAs of 'be'; thus, a copula is inherently intransitive. (This is a provisory characterization; the exact definition of copular verbs requires a special study.) On the contrary, a verb such as TROUVER in *Alain trouve Helen belle* 'Alain finds Helen beautiful' or RENDRE in *La robe rend Helen belle* 'The dress makes Helen beautiful' is by no means a copula, since although it also includes the semantic component 'be', it has three DSyntAs and is transitive. As a result, the phrase *trouver/rendre→belle* 'find/ make beautiful' is described by a different SSyntRel (= No. 12, actantial-attributive-completive, p. 208).

The CopAttrCo has the following two defining properties:

1) The CopAttrCo must be the target of subject agreement: it agrees obligatorily with the clause element that corresponds to DSyntA I of the governing copula: Il m'a ordonné d'être prête 'He ordered me to be ready [FEM]' (DSyntA I of ÊTRE is MOI 'I' [a woman]).

The agreement in question is either morphological (*Alain est intelligent* 'Alain is intelligent [MASC]' ~ *Aline est intelligente* 'Aline is intelligent [FEM]') or lexical-seman-

tic (*Alain a été directeur* 'Alan has been he-director' ~ *Aline a été directrice* 'Aline has been she-director').

2) The CopAttrCo can be the target of cliticization by the "neuter" LE: Je ne semble pas intelligente, je le $\langle *la \rangle$ suis lit. ^{(I} don't seem intelligent, I am it⁾; Elle n'est pas encore intelligente, elle le $\langle *la \rangle$ devient lit. ^{(She is not yet intelligent, she is becoming it⁾; Je ne suis pas encore une cantatrice, je le $\langle *la \rangle$ deviens lit. ^{(I} am not yet a she-singer, I am becoming it⁾; Elle n'est pas encore la directrice, mais elle le $\langle *la \rangle$ deviendra lit. ^{(She is not yet the she-director, but she will become it⁾.}}

Other properties of the CopAttrCo:

3) The CopAttrCo corresponds to DSyntA II of the governing copula.

4) The CopAttrCo is not obligatorily present in every clause.

5) The CopAttrCo does not obligatorily depend on the MV.

6) The CopAttrCo cannot be promoted/demoted.

7) If the CopAttrCo is a noun, it can be the target of relativization: *Voilà l'homme que tu deviendras si tu n'es pas sage* 'Here is the man that you'll become if you are not well-behaved'. (Cf. **Je déteste être malade, qu'Alain est toujours* 'I hate to be sick, which Alain always is', where the CopAttrCo is an adjective.)

8) If the CopAttrCo is a noun, it can be the target of clefting, although the construction is deemed non-elegant (and the judgments of the speakers diverge): C'est un grand linguiste qu'il deviendra 'It is a great linguist that he will become'.

9) In a special case—with the Subject CE 'this' and the verb ÊTRE in the identifying sense—the CopAttrCo is the controller of the agreement of the MV: *Ce sont mes amis* lit. 'This are my friends'.

10) The CopAttrCo linearly follows the governing verb.

Exceptions

The D precedes if 1) it is TEL 'such', cf. *Tel était son ordre* 'Such was his order'; or if 2) it is focalized, cf. *Plus dure sera la chute* lit. 'Harder will be the fall'.

11) The CopAttrCo cannot be the target of pure left dislocation: ^{?*}*Intelligent, il est peut-être* lit. ^{(Intelligent, he is maybe⁾; [?]*Intelligent, il n'est pas encore, mais il le devien-dra* lit. ^{(Intelligent, he isn't as yet, but he will become it⁾; [?]*Intelligent, il deviendra peut-être* lit. ^{(Intelligent, he will become maybe⁾; **Intelligent, il semble à tout le monde, sauf moi* lit. ^{(Intelligent, he seems to everybody, except me⁾.}}}}

Formal types of the CopAttrCo

We give here the maximal set of possibilities: the formal types of the CopAttrCo that depends on ÊTRE; with other copulas the CopAttrCo may lack some of these formal types (the respective possibilities are specified in the Government Pattern of the copula).

1. An adjective: *Il est gentil* 'He is nice'.

2. A prepositionless noun: *Il est médecin* 'He is a doctor'; *C'est une linguiste* [FEM] lit. 'This is a she-linguist'.

3. A headless relative: *Elle sera qui elle veut être* 'She will be who she wants to be'.

4. The "neuter" LE or a personal clitic: *Il l'est* lit. ^(He is it).

5. A prepositional phrase: *Il est sans argent* 'He is without money'.

6. A noun introduced by the conjunction COMME: *Il est comme les autres* 'He is like others'.

7. A numeral: Ils étaient cinq 'They were five'.

8. An adverb: *Il est debout* (*bien, mieux*) 'He is standing up (well, better)'.

9. An infinitive:

- a. Without preposition: *Refuser serait l'offenser* 'To refuse would be to insult him'; *Cette expérience semble condamner votre hypothèse* 'This experiment seems to condemn your hypothesis'.
- b. Introduced by a preposition: *Notre but est de décrire la conjugaison* 'Our goal is to describe the conjugation'.

10. A QUE-clause: Son désir est qu'on le laisse tranquille lit. 'His desire is that they leave him alone'.

The cop-attr-compl SSyntRel is non-repeatable.

Comment

The verb ÊTRE admits several CopAttrCos—if they are implemented by adverbial or prepositional phrases; thus, in (20) we have three such phrases:

(20) a. *Alain était à Paris (là-bas) sans argent dans un état désespéré* ⁽Alain was in Paris (there) without money in a desperate state⁾.

However, this does not contradict our statement about the non-repeatability of the **cop-attr-compl** SSyntRel: it is not true that each of the three prepositional phrases depends directly on the copula ÊTRE, which would produce parallel, i.e., repeatable, dependencies of the same type. Were this the case, we would face a theoretically impossible situation: one DSynt-valence (= DSyntA II of ÊTRE) being implemented more than one time. It is only the leftmost of the prepositional phrases that depends on ÊTRE via the **cop-attr-compl** SSyntRel; each of the others depends on its predecessor via the **quasi-coordinative** SSyntRel:

b. ÊTRE−cop-attr-compl→À PARIS−quasi-coord→SANS ARGENT−quasi-coord→DANS UN ÉTAT ...

207

Similarly to the **coordinative** SSyntRel, the **quasi-coordinative** SSyntRel links the clause elements that fulfill the same semantic role, but without being really conjoined.

12. Actantial-attributive-completive SSyntRel: G-act-attr-compl→D

Properties

Its dependent member is the Actantial-Attributive Complement [= ActAttrCo], which bears semantically either a) on the Subject, or b) on the Direct Object:

a) *Il s'appelle Alain* lit. 'He calls himself Alain', *Elle s'appelle Aline* lit. 'She calls herself Aline'; *Cet ensemble est dit maximal* lit. 'This set is said maximal'; *Aline est considérée intelligente* lit. 'Aline is considered intelligent [FEM]'; *Élu directeur, Alain est parti en Espagne* lit. 'Elected he-director, Alain has left for Spain'.

b) On l'a appelé Alain lit. 'They called him Alain', On l'a appelée Aline lit. 'They called her Aline'; On considère Aline intelligente lit. 'They consider Aline intelligent [fem]'; On a élu Alain directeur lit. 'They elected Alain he-director'; On l'a faite trop grosse lit. 'They have made it too big [FEM]'; On l'a surnommé Le Barbu lit. 'They nick-named him 'The Bearded'.

The two defining properties of the ActAttrCo are:

1) The ActAttrCo must be the target of subject/object agreement. More precisely, it agrees obligatorily:

• either with the Subject of the governing V, if this V has no DirO (in particular, if it is in the passive);

• or with the nominal SSynt-governor of the governing V, if this V is in the form of past participle (= reduction of a clause with the passive: *J'ai lu un roman, [qui a été] considéré par tout le monde fort intéressant* ^{(I} have read the novel, [which has been] considered very interesting by everybody⁾;

• or with the DirO of V, if it is present.

The agreement in question is either morphological (*intelligent* ~ *intelligente*) or lexico-semantic (*Alain a été nommé directeur* lit. 'Alain has been nominated he-director' ~ *Aline a été nommée directrice* lit. 'Aline has been nominated she-director'; On a nommé *Alain directeur* lit. 'They have nominated Alain he-director' ~ On a nommé Aline directrice lit. 'They have nominated Aline she-director').

2) The ActAttrCo cannot—in contrast to the CopAttrCo—be the target of cliticization: *Il est considéré intelligent* 'He is considered intelligent' ~ **Il l'est considéré* 'He is considered it' vs. *Il est intelligent* 'He is ill' ~ *Il l'est* lit. 'He is it'.

Other properties of the ActAttrCo:

3) The ActAttrCo corresponds to DSyntA II or III of the governing verb.

Verbs that take the ActAttrCo as their DSyntA II: s'appeler 'be called', s'annoncer 'announce itself', s'avérer 'turn out to', se percevoir 'be perceived', se présenter (comme ...) 'appear (as ...)', être dit 'be said', tomber (malade, amoureux) 'fall (ill, in love)', ...

209

- Verbs that take the ActAttrCo as their DSyntA III: considérer 'consider', élire 'elect', nommer 'nominate', ... (Tout le monde considère Alain génial 'Everybody considers Alain a genius' and Alain est considéré génial par tout le monde 'Alain is considered a genius by everybody').
 - 4) The ActAttrCo is not obligatorily present in every clause.
 - 5) The ActAttrCo does not obligatorily depend on the MV.
 - 6) The ActAttrCo cannot be promoted/demoted.

7) The ActAttrCo cannot be the target of relativization: *Le directeur de l'usine, qu'Alain a été nommé $\langle qu$ 'on a nommé Alain a beaucoup de moyens 'The plant director, which Alain has been nominated (which they have nominated Alain), has many means'.

8) Even if the ActAttrCo is a noun, it cannot be the target of clefting: **C'est un* grand linguiste qu'Alain est considéré lit. 'This is a great linguist that Alain is considered'; **C'est un grand linguiste qu'on considère Alain* lit. 'This is a great linguist that they consider Alain'.

NB: If the ActAttrCo is an adjective, it admits of "concessive detachment:" *Tout intelligent qu'Alain est, il ne le comprendra pas* ⁽However intelligent Alain is, he will not understand this⁾.

9) The ActAttrCo cannot be the controller the agreement of the MV.

10) The ActAttrCo linearly follows the governing verb.

Exception

The ActAttrCo precedes the MV in the concessive detachment construction, see above.

11) The ActAttrCo cannot be the target of pure left dislocation: **Intelligent* (*Un grand linguiste*), *Alain est considéré* (*on considère Alain*) 'Intelligent (a great linguist), Alain is considered (they consider Alain)³.

Formal types of the ActAttrCo

1. An adjective or a participle: *Helen est considérée intelligente* 'Helen is considered intelligent'; *Helen, considérée intelligente par ses collègues, peut obtenir ce qu'elle veut* 'Helen, considered intelligent by her colleagues, can obtain what she wants'; *Alain considère Helen intelligente* 'Alain considers Helen intelligent'; *Alain regarde Helen traversant la rue* 'Alain watches Helen crossing the street'; *Je voudrais voir cette affaire clarifiée* 'I would like to see this business clarified'.

2. A prepositionless noun: *Il s'appelle Alain* lit. ^{(He} calls himself Alain⁾, *Alain est élu directeur* lit. ^{(Alain is elected he-director⁾; *Alain nomme son fils Igor* ^{(Alain calls his son Igor⁾; *On l'a bombardé président* ^{(He} was suddenly thrust into the position of president⁾.}}

3. A noun or an adjective introduced by the conjunction COMME 'as' or the preposition DE: *Alain est proposé (On propose Alain) comme directeur* 'Alain is proposed (They propose Alain) as director'; *Cette théorie est considérée (On considère cette théorie)*

comme fort intéressante 'This theory is considered (They consider this theory) as very interesting'; *Alain a traité Helen de menteuse* 'Alain has called Helen a liar'; *Alain a qualifié ce journal de tendancieux* 'Alain has labeled this newspaper as tendentious'.

4. A noun introduced by the preposition EN TANT QUE, A TITRE DE OR POUR: Alain est proposé (On propose Alain) en tant que (à titre de) directeur (Alain is proposed (They proposed Alain) as director); On tient Alain (Alain est tenu) pour un génie lit. They hold Alain (Alain is held) for a genius).

5. A prepositional phrase: *Elle croit Alain* à **Paris** (sans argent) lit. 'She believes Alain in Paris (without money)'; *Je l'ai laissé sans un sous* 'I have left him without a cent'.

6. An adverb: Helen croyait Alain dehors lit. (Helen believed Alain outside).

7. A relative clause: *Alain les voit qui traversent la rue* lit. ^{(Alain sees them who are crossing the street⁾; *Ils étaient vus qui traversaient la rue* lit. ^{(They were seen who were crossing the street⁾.}}

The types 5 and 6 appear only with the verb CROIRE 'believe', the type 7—with perception verbs.

The act-attr-compl SSyntRel is non-repeatable.

Justification

1. For the **act-attr-compl** SSyntRel Criterion 2 is satisfied only partially, that is, not for all governing verbs: some Ds indicated above are not substitutable by the prototypical Act-AttrCo—an adjective. The ActAttrCo with the verbs APPELER ^(call), PROPOSER ^(propose), ÉLIRE ^(elect) and NOMMER ^(call/nominate) can only be a noun (*s'appeler Alain* ^{(be} called Alain); *être proposé comme directeur* ^{(be} proposed as director); thus, we admit an ActAttrCo which is not substitutable by the prototype. Our decision is justified by the high degree of similarity of the syntactic behavior of this non-referential noun and all "normal" ActAttrCos and our unwillingness to postulate a separate SSynt-Rel just for these verbs otherwise.

On the other hand, we see an alternative solution: to posit a different **attr-compl** SSyntRel, because one of the **attr-compl** SSyntRels presupposes morphological agreement of the ActAttrCo (when the latter is an adjective), while the other does not.

2. We do not see other SSyntRels with which the act-attr-compl SSyntRel could be collapsed. Thus, the act-attr-compl SSyntRel is opposed to the subject-copredicative and the object-copredicative SSyntRels²⁴ (Criterion 1):

(21) a.

subj-copred ______
Petit, Alain était considéré intelligent
lit. ⁽Little [MASC], Alain was considered intelligent [MASC]⁾

 $^{^{24}}$ The subject-copredicative SSyntRel and the object-copredicative SSyntRel (see below) are not actantial and therefore not considered in this paper.

b.

v obj-copred _____
 Petite, Alain la considérait intelligente
 lit. ^cLittle [FEM], Alain considered her intelligent [FEM]^o

Comments

1. In some approaches, the distinction between the clause elements covered by the **copular** and the **act-attr-compl** SSyntRels is drawn in a different way. Thus, for English, Quirk *et al.* (1991: 728-729) distinguish Subject Complement, which includes complements of copulas and of the verbs of the type *be considered* or *be called* (because they bear semantically on the Subject), and Object Complement—for the cases of the type *find* [N] *beautiful* (because this complement bears on the Object). The same treatment is traditionally proposed for French: for instance, Le Goffic 1993 (passim) et Baylon & Fabre 1995: 198-199. Our decision to oppose Copular Complements to Actantial Complements (without distinction between Subject and Object Complements) is based on the following two considerations:

1) CopAttrCos and ActAttrCos differ in their relevant properties:

• CopAttrCos, but not ActAttrCos, admit of cliticization: *Il l'est* lit. 'He is it' vs. **Il l'est considéré* lit. 'He is considered it'.

• CopAttrCos admit of pure left dislocation, while ActAttrCos do not: *Intelligent*, *il* est lit. 'Intelligent, he is' vs. **Intelligent*, *il* est considéré lit. 'Intelligent, he is considered'/**Intelligent*, on le considère lit. 'Intelligent, they consider him'.

• Unlike ActAttrCos, a CopAttrCo can control the agreement of the MV (with ÊTRE 'be'): *Ce sont mes amis* lit. 'This are my friends'.

2) Typologically, copular constructions have in syntax the pride of place—they are (near-)universal and have special properties in many languages; therefore, we prefer to keep them separate.

2. The distinction between Subject and Object Complements, as we have seen, is fully determined by the presence of a DirO (with the governing verb) and thus can be dispensed with. This means that we can do with one **act-attr-compl** SSyntRel.

3. An interesting case of ActAttrCo: an adjective bearing semantically on an infinitive which is introduced by the preposition DE (this infinitive is the DirO of the governing verb), for instance: *Alain trouve plaisant de faire un cassoulet* ^{(Alain finds} [it] pleasant to prepare a cassoulet².

4. Similar to the **cop-attr-compl** SSyntRel, the **act-attr-compl** SSyntRel can also have seemingly multiple Ds expressed by prepositional phrases or adverbs:

(22) Helen croyait Alain à Paris sans argent dans un état désespéré (Helen believed Alain in Paris without money in a desperate state). These consecutive phrases are treated the same way as Ds of this type with the copula ÊTRE (see above, No. 11, Comment, p. 207): only the first one is taken to be an ActAt-trCo, while the others depend on their preceding neighbor via the **quasi-coord** SSyntRel:

CROIRE−[*Alain*]**−cop-attr-compl**→À PARIS**−quasi-coord**→SANS ARGENT**−quasi-coord**→DANS UN ÉTAT DÉSESPÉRÉ

13. Predicative-attributive-completive SSyntRel: G-pred-attr-compl→D

The prototypical—and the only possible D—is an adjective (masculine, singular).

Properties

Its dependent member is the Predicative-Attributive Complement [= PredAttrCo], which is possible with a number of verbs of the type SENTIR 'smell' [intrans], COUPER 'cut', COÛTER 'cost', CHANTER 'sing', RAPPORTER 'yield, bring in': *Ça coûte cher* [= II] lit. 'This costs expensive'; *Les draps sentent bon* [= III] *la menthe* lit. 'The bedsheets smell good the mint'; *couper gros/menu* [= III] *la viande* lit. 'cut meat big/small'; *Alain chante cet air juste* [= III] lit. 'Alain sings this tune correct'; *Son entreprise rapporte gros* [= II] lit. 'His enterprise brings big' = '... gives him a good return'; *Il porte haut* [= III] *la tête* 'He carries the head high' (see the list of such verbs in Le Goffic 1993: 367-368).

The PredAttrCo has one defining property:

1) The PredAttrCo is the target of non-specific government, namely, it is an adjective frozen in the masculine singular form: *Les roses sentaient bon* ^(The roses smelled good).

Other properties:

2) The PredAttrCo corresponds to DSyntA II, III or IV (X vend Y à Z cher [= IV] lit. (X sells Y to Z expensive) of the governing verb.

3) The PredAttrCo is not obligatorily present in every clause.

4) The PredAttrCo does not obligatorily depend on the MV.

5) The PredAttrCo cannot be promoted/demoted.

6) The PredAttrCo cannot be the target of cliticization: **Cette robe le* [= *cher*] *coûte* ⁽This dress costs it [= expensive]⁾; **Cette caisse le* [= *mauvais*] *sent* ⁽This box smells it [= bad]⁾.

7) The PredAttrCo cannot be the target of the agreement.

8) The PredAttrCo cannot be the controller of the agreement of the MV.

9) The PredAttrCo follows the governing verb; it precedes all other objects, including the DirO. 10) The PredAttrCo cannot be the target of pure left dislocation: **Bon, le pain sent* 'Good, the bread smells'.

The pred-attr-compl SSyntRel is non-repeatable.

3.4 SSyntRels whose prototypical dependent is an adverbial expression

THE CIRCUMSTANTIAL and COMPARATIVE SSynt-Relations (14 - 15)

The two SSyntRels below cover all Circumstantials and Comparative Expressions, including all free modifiers of verbs. However, in what follows we characterize them strictly within the limits of our data—that is, only those Circumstantials and Comparative Expressions which express DSyntAs of the verb.

14. Circumstantial SSyntRel: G-circum→D

The prototypical D is an adverb.

Properties

Its dependent member is a Circumstantial [= Circum]: an adverbial expression that specifies the place, the time, the duration of an action, the manner in which the action is carried out, etc. But what we are interested in here is a particular case of Circum: a Circum that expresses a DSvntA of the verb (cf. Adverbials in the role of *compléments* essentiels in Le Goffic 1993: 355). Some of these Circums—the Circums of place, of time and of manner-appear rather restrictedly, i.e., with a handful of verbs, such as HABITER à Paris 'live in Paris', VIVRE rue du Dahomey 'live in Dahomey Street', SE TROUVER *là-bas* 'find oneself there', ATTENDRE [N] *demain* 'expect [N] tomorrow', SE COMPORTER bien (de façon amicale, avec une générosité incrovable, royalement) (behave well (in a friendly manner, with an incredible generosity, royally), TRAITER [N] très bien (royalement) (treat [N] very well (royally), ÉVALUER [N] très haut (evaluate [N] very highly', CARACTÉRISER [N] positivement 'characterize [N] positively', or RECEVOIR [N] amicalement 'receive [N] friendly', etc. Some other Circums-for instance, the Circums of measure—are more widespread (as an expression of a DSyntA: DSyntA II, most of the time): MANGER beaucoup 'eat much/a lot', LIRE plus qu'elle (read more than she), BOIRE *trop* (drink too much), etc.²⁵

The circum SSyntRel has two defining properties:

1) If the Circum is a locative phrase, it can be the target of cliticization by EN or Y.

2) If the Circum is a noun phrase, it can be the target of relativization: *la générosité avec laquelle Alain nous a reçus* 'the generosity with which Alain has received us'; *la ville où Alain habite* 'the city where Alain lives'.

The descriptive properties of the Circum considered here are:

²⁵ The Circum of Measure does not express DSyntA II with all transitive verbs; cf.:
(i) a. *Ils ont beaucoup mangé (construit)* <sup>(They have eaten (built) a lot².
</sup>

vs.

b. **Ils ont beaucoup préparé* $\langle coupé \rangle$ ^{(They have prepared $\langle cut \rangle$ a lot⁾.}

This fact shows the DSynt-actantial role of the Circum of Measure in the sentences in (i-a).

3) The Circum corresponds to DSyntA II or III of the governing verb.

4) The Circum is not obligatorily present in every clause.

5) The Circum does not obligatorily depend on the MV.

6) The Circum cannot be promoted/demoted.

7) The Circum can be the target of clefting: *C'est très amicalement qu'Alain nous a reçus* ⁽It is very friendly that Alain has received us².

8) The Circum (we speak only of Circums that express DSyntAs!) follows the governing verb.

9) The Circum cannot be the target of pure left dislocation: **Très amicalement, il* s'est comporté envers son frère, pas envers moi 'Very friendly, he has behaved with respect to his brother, not me³; ?*À Paris/Là-bas, il habite depuis longtemps 'In Paris/ There, he has been living for a long time³.

Formal types of the Circum

1. An adverb (*ici* 'here', *là* 'there', *bien* 'well', *amicalement* 'friendly').

2. A prepositional phrase (*au bord de la mer* 'at the sea-side', *sous le pont* 'under the bridge', *avec amitié* 'with friendship', *sans cérémonie* 'without ceremony', *de façon* Adj 'in an Adj way', ...).

- 3. The clitic Y or EN.
- 4. A prepositionless noun phrase:
 - a. a proper name of a place in a city ([*Alain habite*] *Place de la Concorde* lit. ⁽[Alain lives] Place de la Concorde³);
 - b. a noun denoting a day or a date ([*Noël tombe*] *un dimanche* ⁽[Christmas falls] on a Sunday³).

Generally speaking, i.e., taking into account all Circums rather than only those that express the DSyntAs of the Governor, the **circum** SSyntRel is repeatable:

(23)

Alain s'est très bien←circum−*comporté*−circum→*hier à l'école* ⁽Alain has behaved very well yesterday at school⁾.

Justification

1. The circum SSyntRel must be distinguished from the sentential-circumstantial SSyntRel in conformity with Criterion 1 (semantic contrast). Cf. (24):

(24)

Avec beaucoup de méchanceté, il se comporte-circum \rightarrow de façon inattendue 'With much wickedness, he behaves in an unexpected manner³.

— sent-circum –

VS.

sent-circum _______ *De façon inattendue, il se comporte*−circum→*avec beaucoup de méchanceté* ⁽In an unexpected manner, he behaves with much wickedness⁾.

2. The **circum** SSyntRel with a D implemented by a prepositional phrase must also be distinguished from the **obl-obj** SSyntRel. Although our relevant properties do not distinguish them, a linguistically valid distinction exists. As a rule, the preposition that introduces an OblO is uniquely determined by the governing verb: in its Government Pattern a specific preposition is indicated; the corresponding prepositional phrase cannot be attached to every verb. Thus, in *louer* [*une voiture*] *pour trois mois* ⁽rent [a car] for three months⁾, the boldfaced phrase is an OblO (although, semantically, it could be considered a Circumstantial). With a Circum, the choice of the specific preposition is not imposed by the governing verb: cf. *aller à/vers/devant/derrière/dans/hors de l'école* ⁽go to/towards/in front of/behind/in/outside of the school⁾; the corresponding prepositional phrases in *aller de Montréal à Singapour par le Pacifique* ⁽go from Montreal to Singapore via the Pacific⁾ are considered to be Circums. (Of course, this does not prevent these phrases to be DSyntSs of ALLER.)

More difficult is establishing the syntactic difference between *au concert* ⁽to a concert⁾ [= OblO] and *au restaurant* ⁽to a restaurant⁾ [= Circum] with the verb INVITER ⁽invite⁾. In both cases, the preposition is not determined by the verb alone (*inviter au concert/au restaurant/en France/sur la terrasse/chez Alain* ⁽invite to a concert/to a restaurant/to France/on the terrasse/at Alain's⁾, and the prepositional phrase easily combines with any verb (e.g., *mourir au concert (au restaurant)* ⁽die at a concert/in a restaurant⁾). However, a telling distinction can be found even here: *le restaurant où Alain m'a invitée* ⁽the restaurant where Alain has invited me⁾ vs. ^{??}*le concert où Alain m'a invitée* ⁽the concert where Alain has invited me⁾ [correct expression: *le concert auquel Alain m'a invitée*], which shows that a locative Circum and an OblO admit of different types of relativization. (Unfortunately, this test does not work for all verbs. Thus, *emmener au concert* ⁽take to a concert⁾ is described by the **obl-obj** SSyntRel, similarly to *inviter au concert*, but the relativization with *où* is here quite natural: *le concert où Alain m'a emmenée*.)

A Circum expressed by a prepositionless noun of the type *tomber un dimanche* ⁽fall on a Sunday⁾ cannot be considered an OblO, because it is not replaceable with a prepositional phrase (the prototypical D of the **obl-obj** SSyntRel), but is replaceable with an adverb (the prototypical D of the **circum** SSyntRel): *Pâques tombe tard/tôt cette année* ⁽Easter falls late/early this year⁾.

3. With some transitive verbs, the Circum of Measure expresses DSyntA II and therefore does not combine with an obvious DirO:

(25) **Il a beaucoup mangé la soupe* lit. ⁽He has much eaten the soup⁾.²⁶

However, even in such cases, the Circum of Measure cannot be subsumed under the **dir-obj** SSyntRel since it violates the defining properties of the latter:

• It does not readily passivize: **Trop a été mangé* 'Too much has been eaten', **Plus que normal a été bu* 'More than normal has been drunk'.

• It does not impose the IndirO role on the Secondary Actor of the causative FAIRE-construction: On $le \langle *lui \rangle$ fait manger beaucoup trop 'They make him eat too much'.

The quantitative adverbs of the type BEAUCOUP, PEU, PLUS [que], TROP, etc. are not considered traditionally as Circums of Measure; the only example found in Grevisse 1993: 476—allonger une robe de dix centimètres 'lengthen a dress by 10 cm³—presents in the point of fact an OblO (because in this phrase, the preposition DE is uniquely determined by the verb). Le Goffic (1993: 234) notes the semantic relatedness between quantitative adverbs and DirOs with some verbs, without, however, making his proposal specific enough.

Comment

We by no means imply that one **circum** SSyntRel is sufficient for the description of the French syntax. It is quite possible that several particular types of Circums should be distinguished based on purely SSynt-considerations. We, however, allow ourselves to make abstraction from this problem, which is marginal within the present context.

15. Comparative SSyntRel: G-compar→D

The prototypical D is a nominal phrase introduced by the comparative conjunction COMME (as).

Properties

Its dependent member is a Comparate [= Compar]: a nominal phrase introduced by the conjunction COMME 'as' [COMME \rightarrow N] and expressing the comparison either with the Subject or with the DirO of the governing verb V. Let it be reminded that here we consider only a particular case of the Compar, namely, a Compar that expresses a DSyntA of the verb; French has only a few verbs of the corresponding type, such as SE COM-PORTER 'behave' or TRAITER 'treat': *Alain s'est comporté comme un héros* 'Alain has behaved as a hero'; *Alain a traité Helen comme une reine* 'Alain has treated Helen as a queen'.

The compar SSyntRel has two defining properties:

- 1) The Compar cannot be the target of cliticization.
- 2) The Compar cannot be the target of relativization.

The descriptive properties of the Compar considered here are:

²⁶ The expression of the type *Il a beaucoup mangé de soupe* lit. ^{(He} has much eaten of soup⁾ represents the "split" of the phrase *beaucoup de soupe* (much soup⁾ (which is a DirO) rather than the combination of two codependents BEAUCOUP and SOUPE (cf. **Il a mangé de soupe*).

3) The Compar corresponds to DSyntA II or III of the governing verb.

4) The Compar is not obligatorily present in every clause.

5) The Compar does not obligatorily depend on the MV.

6) The Compar cannot be promoted/demoted.

7) The Compar can be the target of clefting: C'est comme un ami qu'Alain nous a reçus lit. 'It is as a he-friend that Alain has received us'; C'est comme une amie qu'Alain a reçu Helen lit. 'It is as a she-friend that Alain has received Helen'.

8) The Compar follows the governing verb.

9) The Compar cannot be the target of left dislocation: *Comme un héros, Alain s'est comporté ⁽As a hero, Alain has behaved⁾. (Comme un héros Alain s'est comporté pendant le voyage au Népal ⁽As a hero Alain has behaved during the trip to Nepal⁾ is possible with a descending contour on the boldfaced phrase—as pointed out to us by C. Blanche-Benveniste; in this case, the phrase in question is a fronted Circumstantial.)

The comparative SSyntRel is non-repeatable.

Justification

1. Distinguishing the **compar** SSyntRels from the **circum** SSyntRel is motivated by the following three considerations:

• They do not have the same prototype: the prototypical Circum is an adverb, while the prototypical Compar is a COMME-phrase (which is in fact the only possible expression of D). Therefore, Criterion 2 does not allow the unification.

• They have diverging relevant properties: the Compar does not admit cliticization, relativization nor left dislocation, which are possible (even if partially) for the Circum.

• Typologically, comparative constructions present a host of specific properties. They are related to coordination and, like the latter, are implicated in complex transformations irrelevant for simple Circums. Cf., for instance, *J'aime Alain comme Helen* ^(I) love Alain as Helen⁾, which means either ^(I) love Alain as I love Helen⁾ (*Helen* is parallel to *Alain*) or ^(I) love Alain as Helen loves him⁾ (*Helen* is parallel to *je*).

2. Consider the significative opposition in (26):

(26) a. *Alain traite Helen comme un roi* ⁽Alain treats Helen as a king⁾. vs.
b. *Alain traite Helen comme une reine* ⁽Alain treats Helen as a queen⁾/

Alain traite ses amis comme des rois 'Alain treats his friends as kings'.

One can have the impression that (26) constitutes a clear case of application of Criterion 1 (= semantic contrast). However, in point of fact, the contrast here is not related to different syntactic means of expression: the derivational difference in gender (ROI 'king' ~ REINE 'queen') is lexical-semantic and the inflectional difference in number

(roi 'king' ~ rois 'kings') is semantic. Both differences are taken care of at the deeper level of representation (this is, so to speak, semantic agreement). Therefore, they should not have any impact on SSyntRels. As a result, we have no need to postulate two different SSyntRels (for instance, subj-compar vs. obj-compar).

3.5 SSyntRels whose prototypical dependent is Direct Speech

THE QUOTATIVE-OBJECTIVAL SSyntRel (16)

16. Quotative-objectival SSyntRel: G-quot-obj→D

Properties

Its dependent member is Direct Speech, or the Quotative Object [= QuotO]: « *J'aime le poulet à l'estragon* », annonça Alain⁽'I love tarragon chicken,' announced Alain⁾; « *Tu es déjà là ? », s'étonna Helen* lit. ^{('}You are already here?,' became astonished Helen⁾.

The QuotO has just one defining property:

1) The QuotO has special prosody/punctuation.

Other (descriptive) properties of the QuotO:

2) The QuotO corresponds to DSyntA II of the governing verb.

3) The QuotO is not obligatorily present in every clause.

4) The QuotO does not obligatorily depend on the MV.

5) The QuotO can be the target of promotion by passivization, with possible subsequent Impersonalization: *Il a été annoncé : « Gardez vos places »* lit. ^{(It} has been announced: 'Remain in your seats'); *C'est alors que fut annoncé : « Gardez vos places »* lit. ^{(It} is then that was announced: 'Remain in your seats').

6) The QuotO imposes the IndirO realization of the Secondary Actor in the causative FAIRE-construction, but—unlike the DirO—only with nominal phrases; with clitics, both realizations remain possible:

(27) a. Alain fait crier à Helen : « Nous sommes foutus ! » lit. ⁽Alain makes to Helen shout: 'We are finished!')~
^{??} Alain fait crier Helen : « Nous sommes foutus ! » lit. ⁽Alain makes Helen shout: 'We are finished!')
vs.
b. Alain lui/la fait crier : « Nous sommes foutus ! » lit. ⁽Alain makes to-her/her shout: 'We are finished!')

7) The QuotO can both follow and precede the MV (= a speech verb); in case it precedes the MV, it entails the inversion of the MV and the Subject: *« Gardez vos places », a dit Alain* lit. ⁽'Remain in your seats,' has said Alain⁾. The quotative-objectival SSyntRel is non-repeatable.

Justification

The **quot-obj** SSyntRel cannot be subsumed under the **dir-obj** SSyntRel in conformity with Criterion 1 (semantic contrast):

(28) Alain me dit trois mots [= DirO] 'Alain says to me three words'.
vs.
Alain me dit : « Trois mots » [= QuotO] 'Alain says to me: 'Three words''.

Accordingly, a special quotative SSyntRel was proposed in Iordanskaja & Mel'čuk 1981.

On the DSynt-level, Direct Speech is taken to be a particular type of DSyntA II of the governing speech verb: namely DSyntA II_{dir-sp} . On the SSynt-level, the Direct Speech and the DirO are considered as two different expressions: 1) either an utterance (*« Trois mots » ""Three words")*, or else 2) a description of an utterance (*trois mots * three words*) or of the content of an utterance (*qu'il veut manger un morceau that he wants to have a bite*). The DSyntRel II_{dir-sp} in the DSyntS indicates the presence of a literal reproduction of an utterance (no changes in it, in particular no paraphrasing, are possible). Thus, such a sentence as [*Alain*] *dit : « Je veux manger un morceau »* [Alain] says: 'I want to have a bite' receives the following DSyntS:



Je veux manger un morceau

However, as we have just seen, the SSynt-behavior of Direct Speech is very similar to that of the DirO: two of DirO's defining properties (passivization and the control of the Secondary Actor in the FAIRE-construction) are shared by the QuotO, although in a weakened form. In addition, Indirect Speech, which is semantically and syntactically close enough to Direct Speech, is considered a DirO: *Alain dit*-dir-obj- $\gamma qu'il$ voulait manger un morceau ⁽Alain says that he wanted to have a bite³. It is thus not astonishing that grammarians hesitate in regard to syntactic representation of Direct Speech. Cross-linguistically, Direct Speech manifests as well this duality. Thus, in Georgian, the Subject of the Speech Verb + Direct Speech construction is in the ergative, what is expected of the Subject of a transitive verb—so that Direct Speech is treated in Georgian as a normal DirO: (29a). But in Chukchee, which also features an ergative construction, the Subject of an intransitive verb; therefore, we see that in Chukchee Direct Speech is by no means a DirO: (29b).

(29)	a. Georgian							
	«Sadili mz	zada+a!)	» — deda + n	n tkv	а			
	dinner rea	idy is	mother E	RG said				
	⁽ The dinne	r is read	y,' said Motl	her ⁾ .				
	b. Chukchee		-					
	«ŋottqenən	a ənŋe	ejmevke!» -	— ikv?i	pipiqəl y ək	æj+ Ø	mel'ota	ı+γtə
	that.one	not	come.close	said	mouse	NOM	hare	DAT
	('Don't con	ne close	to this one!,'	' said the	e mouse to	the hare ⁾ .		

Interesting data about particularities of the syntactic behavior of Direct Speech in different languages can be found in Munro 1982.

4 The summary: Synoptic tables of French valence-controlled SSyntRels

To make the results of our description more surveyable, we will now offer four tables (for the first four classes of SSyntRels) which present, in a compact form, a characterization of the considered SSyntRels/the corresponding clause elements in terms of their relevant properties.

The value of a property is given according to the prototypical D' of the SSyntRel in question; wherever it seems important to specify a different value for a non-prototypical D we do so, using a slash and indicating (in brackets) the part of speech of D. [?] means that the corresponding value is not stable—either speakers are not unanimous with respect to it, or it varies as a function of different governors. If a property is inapplicable to the prototype of **r** it is indicated by a zero.

4.1 Subject, Objects, Agent: the prototypical dependent is a noun

The clause elements represented in the table below are ordered—from left to right according to the number of positive values (of the properties) they feature: Subj has 13 positive values, DirO has 11, IndirO has 6, Pseudo-DirO 6, Agent 5, and OblO 4. The IndirO is put before the Pseudo-DirO, which has the same number of positive values, because the IndirO is linguistically closer to the preceding elements according to a more important property: reflexivization. (The Quasi-Subject is an exception: since it is the result of a transformation, it is not considered in the subsequent discussion; it is positioned in the rightmost column in order not to obscure the picture.) Interestingly, such ordering leads to a hierarchy: if a lower element features the positive value of a property, then either any higher element also does or the property is not applicable to it, but not vice versa. (Exception: the Pseudo-DirO precludes pure left dislocation while the higher IndirO does not.) And this hierarchy corresponds to the hierarchy of nominal clause elements mentioned on p. 156: the Keenan-Comrie Hierarchy (1977).

The Keenan-Comrie hierarchy was established using one parameter only: accessibility of a clause element for (a particular type of) relativization. However, it was shown that the same hierarchy obtains according to several other parameters: thus, an element higher in the hierarchy determines the syntactic role of the Causee (= the Secondary Actor under causativization; Comrie 1975) while lower elements do not; the accessibility of clause elements for promotion via grammatical voices also corresponds to this hierarchy (if a lower element can be promoted, then all the higher elements necessarily can; Gary & Keenan 1977); the same hierarchy manifests itself in the ability of clause elements to control reflexive and reciprocal anaphora (Pollard & Sag 1992: 266) and to admit extraction (Abeillé 1997a: 23); etc. Now, our results confirm once again the validity of this hierarchy—based, however, not just on one parameter, but rather on a full set of parameters relevant in this framework.

Clause elements Properties of SSyntRels	Subj	DirO	IndirO	Pseudo- DirO	Agent	OblO	Quasi- Subj
1. Corresponding DSyntA	I	Π	II/III	П	I/II	any but I	I
2. Obligatory pres- ence	+	_	_	_	_	_	_
3. Dependence on the MV only	+	_	_	_	_	_	+
4. Target of pro- motion/demotion	demotion	promotion	_	-	_	_	_
5. Target of clitici- zation	Cl _{pers}	Cl _{pers} /EN/ LE _{neu} [V _{inf}]	Cl _{pers}	Cl _{pers}	EN	EN, Y	EN [N]
6. Target of rela- tivization	+	+	+	+	+	+	-
7. Target of cleft- ing	+	+	+	+	+	+	+
8. Controller/Tar- get of reflexiviza- tion	controller	controller/ target	controller/ target	_	_	_	controller
9. Controller of the SSynt-role of the Secondary Actor	0	+	_	_	_	_	0
10. Controller of actantial-attribu- tive complements	+	+	_	_	_	_	[?]—
11. Target of non- specific govern- ment	nomina- tive	accusative	dative	accusative	PAR	-	_
12. Controller of agreement of MV	finite verb	past participle	-	_	_	_	_
13. Precedes MV	+	_	_	—	_	_	_
14. Impossibility of pure left dislo- cation	+	+	_	+	[?]-	_	+

Thick cell borders indicate the defining properties of the respective clause element.

Here is the hierarchy such as it results from our study:

```
Subject > DirO > IndirO > Pseudo-DirO > Agent > OblO
```

Note that it includes two more clause elements, which were not considered by Keenan and Comrie: the Pseudo-DirO (a kind of degenerate DirO) and the Agent. Two lower elements in the Keenan-Comrie hierarchy are omitted, since they do not belong to the verb active valence and therefore are outside our scope. To show how this hierarchy can play a role in the identification of nominal actantial clause elements, we present at the end of the paper the corresponding identification tree: see p. 228.

4.2 Infinitival Objects: the prototypical dependent is an infinitive

Clause elements			
Properties of SSyntRels	DirInfO	OblInfO	CopredInfO
1. Deep-Syntactic Actant	II	II	III
2. Obligatory presence	_	_	_
3. Dependence on the MV only	_	_	_
4. Target of promotion/demotion	_	-	_
5. Target of cliticization	LE _{neutre}	-	—
6. Target of clefting	-	-	_
7. Control of agreement of the MV	_	_	_
8. Precedes the MV	_	_	_
9. Impossibility of left dislocation	—	[?] +	+

As one sees from this table, the three Infinitival Objects do not differ very much from each other. However, their distinctions, although not numerous, are rather important and seem to justify our differentiating them.

Clause elements Properties of SSyntRels	CopAttrCo	ActAttrCo	PredAttrCo
1. Deep-Syntactic Actant	п	11/111	II/III/IV
2. Obligatory presence	_	_	_
3. Dependence on the MV only	_	_	_
4. Target of promotion/demotion	_	_	_
5. Target of cliticization	LE _{neut}	_	_
6. Target of relativization	0/+[N]	0/-[N]	_
7. Target of clefting	0/[?]+[N]	0/-[N]	0
8. Target of non-specific government	accusative [N]	0	masc, sg
9. Target of agreement	subject agreement	subject/object agreement	-
10. Control of agreement of the MV	0/+[N, with CE as Subj]	0/[N]	0
11. Precedes the MV	_	_	_

4.3 Attributive Complements: the prototypical dependent is an adjective

4.4 *Valence-controlled Circumstantials/Comparates: the prototypical dependent is an adverb/a* COMME-phrase

Clause elements Properties of SSyntRels	Circum	Compar
1. Deep-Syntactic Actant	II/III	II/III
2. Obligatory presence	_	—
3. Dependence on the MV only	_	_
4. Target of promotion/demotion	_	_
5. Target of cliticization	Y, EN	_
6. Target of relativization	0/+[N]	_
7. Target of clefting	+	+
8. Precedes the MV	_	_
9. Impossibility of left dislocation	+	+

Similarly to what has been just said about InfOs, our Circums and Compars are not very different with respect to the established list of relevant properties. This is due to the fact that we consider only a particular case—these clause elements as expression of DSyntAs of the governing verb. If we take into account the full range of possibilities open to the comparative construction in French, the differences would be more visible.

5 Conclusion

Put in a nutshell, the content of this paper is as follows: While working on the inventory of valence-controlled SSyntRels governed by a verb, we have tried to take into account the complex interaction between a SSyntRel, formal types, i.e., syntactic classes, of its dependents²⁷ and individual lexemic features of its governor. In other words, we tackle the problem of an optimal distribution of linguistic information between these three types of linguistic entities. Our guiding principle has been to avoid both extreme semanticism and extreme formalism. We do not want to base the system of SSyntRels on their semantic functions nor on the syntactic distribution of their Ds; what we are looking for is a subtle equilibrium between both these aspects—such that it is in conformity with the lexicographic information contained in the lexical entries of the Gs. Cf. Ju. Apresjan's insistence on the necessity of a perfect agreement between the grammar and the lexicon, "which must be tuned to each other" (Apresjan 1986: 57).

Our central tool in this endeavor is the prototypical Dependent of a given SSynt-Rel—an idea that is itself by no means new (it goes back to Jespersen and Tesnière and is shared by many others), but that has been applied to our material in a rigorous and systematic way. Non-prototypical dependents of a SSyntRel are determined based on their similarity with the prototype; we suggest a more precise interpretation of the notion of similarity, p. 160.

And last, but not least, as has been already said, we made a special effort to ensure the typological plausibility of the proposed system of SSyntRels.

Appendices

List of lexemes and constructions considered in the paper

Causative FAIRE-construction Clitic as interrogative marker FAIRE (meteorol.) with a governed noun FAIRE, SE, with a governed infinitive FAILLIR with a governed infinitive FALLOIR with a governed noun and infinitive IL Y A with a governed noun Measure Verb with a governed noun	: p. 177 : p. 171 : p. 182 : p. 203 : p. 198 : p. 182 : p. 182 : p. 182 : p. 181	Modal Verb with a governed infinitive <i>parler politique</i> PAYER with a governed numeral phrase Phasal Verb with a governed infinitive SAVOIR with a governed infinitive VOIR, SE, with a governed infinitive VOULOIR, with a governed infinitive	: p. 198 : p. 181 : p. 182 : p. 199 : p. 198 : p. 203 : p. 185
--	--	--	--

²⁷ Technically speaking, to account for the special nature of different Ds of a SSyntRel, we do the following: For each relevant property of any SSyntRels, it is explicitly indicated to which syntactic classes this property is applicable. For instance, the property of specific linear placement (= precedes/follows the MV) is not applicable to clitics, whose linear position is not determined by their subordinating SSyntRel.

List of major clause elements with their possible SSynt-roles

Noun without preposition

An N can be: 1. Subject: No. 1, p. 172.

- 2. Quasi-Subject [Il pleuvait des coups]: No. 2, p. 175.
- 3. Direct Object: No. 3, p. 180.
- 4. Pseudo-Direct Object [with some special verbs: *Il pesait 60 kg*; *Ça sent le hareng*]: No. 5, p. 191.
- 5. Oblique Object [*parler politique*, *payer cette robe 300 euros*; with some special verbs: *parler, causer, acheter, vendre, payer, voter*...)]: No. 7, p. 196.
- 6. CopAttrCo [with copulas: Il est médecin]: No. 11, p. 297.
- 7. ActAttrCo [with some special verbs: *Il s'appelle Alain; Pierre est élu directeur*]: No. 12, p. 209.
- 8. Circum [with some special verbs: *Il habite Place de la Concorde*; *Noël tombe un dimanche*]: No. 14, p. 214.

Noun with preposition

A PREP + N can be:	1. Direct Object [the preposition DE, after a negated verb: On
	n'a pas trouve de livre]: No. 3, p. 180.
	2. Indirect Object [the preposition À: <i>Il a donné son livre à</i>
	<i>Marie</i>]: No. 4, p. 188.
	3. Agent [the prepositions PAR and DE, with a passive form:
	Elle a été reçue par son ami ; Aimée de tout le monde , elle]:
	No. 6, p. 194.
	4. Oblique Object [insister sur l'expulsion, échanger avec
	Jean contre un vélo]: No. 7, p. 195.
	5. CopAttrCo [with copulas: <i>Il est sans argent</i>]: No. 11,
	p. 207.
	6. Circum [On l'a trouvé au bord de la mer; se trouver à
	<i>Montréal</i>]: No. 14, p. 214.
	7. ActAttrCo [with prepositions EN TANT QUE, À TITRE DE,
	POUR and some others: Pierre est proposé à titre de directeur;
	Elle croyait Pierre à Paris (sans un sous); Il a traité Marie de
	<i>menteuse</i>]: No. 12, p. 210.

Clitic

A CL _{NOM} can be:	1. Subject: No. 1, p. 172.
A CL _{ACC} can be:	 Direct Object: No. 3, p. 180. Pseudo-Direct Object [<i>Cette robe la</i> [= cette somme] coûte]: No. 3, p. 191.
A CL _{DAT} can be:	1. Indirect Object [Je lui donne ce livre]: No. 4, p. 188.
Impers. IL can be:	1. Subject [<i>II pleut</i>]: No. 1, p. 172.

EN can be:	 Quasi-Subject [<i>Il en</i> [= des camions] arrive]: No. 2, p. 175. Direct Object [<i>Il en trouve partout</i>]: No. 3, p. 180. Agent [<i>Il en était adoré</i>]: No. 6, p. 194. Oblique Object [<i>J'en raffole</i>]: No. 7, p. 195. Circum [<i>Il vient d'en sortir</i>]: No. 14, p. 214.
Y can be:	 Oblique Object [<i>J'y pense</i>]: No. 7, p. 195. Circum [<i>Il y habite depuis longtemps</i>]: No. 14, p. 214.
LE _{neutre} can be:	1. DirInfO [<i>Il le peut</i>]: No. 8, p. 199. 2. CopAttrCo [<i>Il l'est</i>]: No. 11, p. 207.

Infinitive without preposition

A V_{INF} can be:	1. Subject [Courir fatigue Pierre]: No. 1, p. 173.
	2. Direct Object [Il préfère partir]: No. 3, p. 180.
	3. DirInfO [<i>Il peut chanter</i>]: No. 8, p. 199.
	4. OblInfO [Je pense partir]: No. 9, p. 200.
	5. CopredInfO [Il la voit traverser la rue; Il fait (laisse) M
	danser]: No. 10, p. 201.
	6. CopAttrCo [Refuser serait l'offenser]: No. 11, p. 207.

Infinitive with a preposition

A PREP + V_{INF} can be:	1. Subject [De voir	• ça m'a bouleversé	2]: No. 1, p. 173.
----------------------------	---------------------	---------------------	--------------------

2. Quasi-Subject [Il fatigue Pierre de courir]: No. 2, p. 175.

Marie

- 3. Direct Object [*Il apprend à parler chinois*; *Il propose de partir*]: No. 3, p. 180.
- 4. Oblique Object [*Il me conseille* à travailler; *Il se passe* de dormir]: No. 7, p. 195.
- 5. OblInfO [*Il a fini par accepter*; *commencer à parler*; *cesser de parler*]: No. 9, p. 200.
- 6. CopAttrCo [Notre but est de décrire ...]: No. 11, p. 207.

Infinitive governing an interrogative pronoun

A WH + V_{INF} can be: 1. Direct Object [*Il sait à qui parler*]: No. 3, p. 180.

Adjective

An ADJ can be:	1. CopAttrCo [with copulas: <i>Il est malade</i> ; <i>Ils sont cinq</i>]:
	No. 11, p. 207.
	2. ActAttrCo [with some special verbs: Marie est considérée
	intelligente]: No. 12, p. 209.

3. PredAttrCo [with some special verbs: *sentir bon*]: No. 13, p. 212.

Adverb

An ADV can be:	 CopAttrCo [<i>Il est debout</i>]: No. 11, p. 207. Circum [<i>attendre qqn demain</i>, <i>se trouver ici</i>, <i>recevoir amicalement</i>]: No. 14, p. 214.
Conjunction COMME	
COMME + N can be:	 CopAttrCo [with copulas: <i>Il est comme les autres</i>]: No. 11, p. 207. ActAttrCo [with some verbs: <i>Pierre est proposé comme directeur</i>]: No. 12, p. 209. Compar [<i>se comporter comme un héros</i>; <i>traiter Marie comme une reine</i>]: No. 15, p. 216.
COMME + ADJ can be:	 ActAttrCo [<i>Cette théorie est considérée comme fort intéressante</i>]: No. 12, p. 209.
Subordinate clause	
QUE + PROP can be:	 Subject [<i>Que Marie soit venue étonne Pierre</i>]: No. 1, p. 173. Quasi-Subject [<i>Il semble que Marie est venue</i>]: No. 2, p. 175. Direct Object [<i>Il sait que Marie est venue</i>]: No. 3, p. 180. Oblique Object [<i>Pierre doute que Marie vienne</i>]: No. 5, p. 195. CopAttrCo [with copulas and some special verbs: <i>Son désir est qu'on le laisse tranquille</i>]: No. 7, p. 207.
WH + PROP can be:	 Subject [<i>Pourquoi Pierre a dit cela reste un mystère</i>]: No. 1, p. 173. Direct Object [<i>Il sait pourquoi Pierre a dit cela</i>]: No. 3, p. 180.
SI _{interr} + PROP can be:	1. Direct Object [<i>Il veut savoir si nous partons demain</i>]: No. 3, p. 180.

Direct Speech

«PROP» can be:	1. Subject [« On va le faire demain » est son énoncé
	<i>préféré</i>]: No. 1, p. 173.
	2. Quasi-Subject [Il a été annoncé « Restez tranquilles ! »]:
	No. 2, p. 176.
	3. QuotO [« Je pars », annonça Pierre]: No. 16, p. 218.

Main nominal actantial clause elements: identification tree



Acknowledgments

The research reported in this paper was financed by the following organizations: Fonds CRSH (grant No. 410-99-1076), Fonds FCAR (grant No. 99-ER-2741), and Fonds CRSNG/NSERC (grant No. OGPO200713).

The first draft has been read, criticized and commented upon by S. Kahane; its final shape owes a lot to his efforts. A. Polguère, with infinite patience and invariably good humor, with-stood our relentless interrogations concerning French data; long stimulating discussions with him substantially helped us on the starting—and the most difficult—stages of our research. The subsequent versions of the paper underwent the scrutiny of C. Blanche-Benveniste, S. Kahane, J. Milićević, A. Polguère, and L. Wanner; a major part of it was gone through by D. Beck and Y.-Ch. Morin. We did our best to take their remarks into account, and we are pleased to express here our heart-felt grat-

itude to all of them. All mistakes and drawbacks that remained are of course our responsibility.

Endnotes

^[1] (p. 167) The construction Mes enfants, je leur permets tout is represented in a different way:

> _____ prolept _____ Mes enfants, je leur—indir-obj-permet tout. lit. ⁽My children, I permit them everything)

Here, we see a special SSyntRel used for a Fronted Dislocated Topic, which can be called, in syntactic terms, a prolepsis.

^[2] (p. 174) S. Kahane has indicated to us an example which can seem problematic in this respect:

(i) *Il devrait arriver trois personnes* lit. ^(It should arrive three people).

If the Quasi-Subject in (i) is cliticized, the clitic attaches to arriver rather than to devrait:

(ii) Il devrait **en** arriver trois (*Il **en** devrait arriver trois)

'It should arrive three thereof'.

Does (ii) argue in favor of subordinating trois personnes to arriver and not to the MV devrait? We do not think so, because clitics show quite a specific syntactic behavior. In particular, they do not necessarily depend syntactically on the same element as their source; therefore, clitics are not relevant when determining the dependency of a non-clitic sentence element. As far as the clitic EN is concerned, many other similar cases are known when EN does not syntactically depend on the same element as its source. N. Ruwet (1972: 50-51) gives a whole series of corresponding examples:

(iii) *L'auteur de ce livre va devenir célèbre*

'The author of this book will become famous'. ~

L'auteur va **en** devenir célèbre (*L'auteur **en** va devenir célèbre).

La solution de ce problème doit être simple 'The solution of this problem must be simple'. \sim *La solution doit en être simple (*La solution en doit être simple).*

As we see, EN is put into a position where it has the infinitive as its host-much like the situation with EN what replaces the Quasi-Subject in (ii). (This is explainable by the fact that the verbs ALLER, DEVOIR, etc. are not clitic-attracting, see p. 203, NB.)

^[3] (p. 178) Our short discussion of the causative FAIRE-construction requires three additional remarks.

• As indicated in Morin 1980, the rule formulated here is, in point of fact, more complex. Thus, in some cases the Secondary Actor of a causative FAIRE-construction can appear as an IndirO of FAIRE in the absence of a DirO of the lexical verb: Cela leur fait penser à leurs enfants lit. 'This makes to them think of their children'. In some other cases, the Secondary Actor appears as a DirO even in the presence of a DirO of the lexical verb: Cela la fait se poser de nombreuses questions lit. 'This

229

makes her ask of-herself numerous questions³. (Such "deviations" are possible only with clitics: **Cela fait penser à leurs enfants à tous les parents qui* ... lit. 'This makes think of their children to all parents who ...³ and **Cela fait se poser de nombreuses questions cette pauvre femme qui* ... lit. 'This makes ask of-herself numerous questions this poor woman who ...³.)

• Similar facts concerning different SSynt-realizations of the Secondary Actor are observed in some other constructions with embedded infinitives, for instance, with LAISSER 'let' and VOIR 'see'—as a function of word order:

- (i) Je laisse Alain lire le livre 'I let Alain read the book'. ~
 Je laisse lire le livre à Alain lit. 'I let read the book to Alain'.
- (ii) Je vois Alain lire le livre 'I see Alain read the book'. ~
 - Je vois lire le livre à Alain lit. 'I see read the book to Alain'.

• The property of imposing the SSynt-role of the IndirO on the Secondary Actor of the causative FAIRE-construction is shared by the Quotative Object, i.e., Direct Speech (Kayne 1977: 203): *Alain fait dire à Helen/Alain lui fait dire : « J'ai tort. »* lit. 'Alain makes to Helen/to-her say: 'I am wrong''. At the same time, unlike the DirO, the QuotO systematically allows the Secondary Actor of the causative construction to be realized as a DirO as well: *Alain la fait dire : « J'ai tort. »* 'Alain makes her say: 'I am wrong''—but again, only if this DirO is a clitic: **Alain fait dire Helen : « J'ai tort. »* 'Alain makes Helen say: 'I am wrong''.

References

Abeillé, Anne 1997a. "Fonction ou position objet ?". Au gré des langues 11: 8-29.

Abeillé, Anne 1997b. "Fonction ou position objet ? (II et fin)". Au gré des langues 12: 8-33.

Apresjan, Jurij 1986. "Integral'noe opisanie jazyka i tolkovyj slovar' [An Integrated Description of Natural Language and the Lexicon]". *Voprosy jazykoznanija* 2: 57-70.

Apresjan, Jurij, Iomdin, Leonid & Percov, Nikolaj 1978. "Ob"ekty i sredstva modeli poverxnostnogo sintaksisa russkogo jazyka [Objects and Means of A Model of the Surface Syntax of Russian]". *International Review of Slavic Linguistics* 3 (3): 249-312.

Apresjan, Jurij, Boguslavskij, Igor', Iomdin, Leonid, Lazurskij, Aleksandr, Percov, Nikolaj, Sannikov, Vladimir, & Cinman, Leonid 1984-1985. *Lingvističeskoe obespečenie sistemy francuzsko-russkogo avtomatičeskogo perevoda ÈTAP-1* [Linguistic Software for a System of French-to-Russian Automatic Translation: ETAP-1]. Preprinty Problemnoj gruppy po èksperimental'noj i prikladnoj lingvistike, vyp. 154-155, 159-160, 166-167, 174. Moskva: Institut Russkogo jazyka AN SSSR.

Apresjan, Jurij, Boguslavskij, Igor', Iomdin, Leonid, Lazurskij, Aleksandr, Percov, Nikolaj, Sannikov, Vladimir, & Cinman, Leonid 1989. *Lingvističeskoe obespečenie sistemy ÈTAP-2* [Linguistic Software of the ETAP-2 System]. Moskva: Nauka.

Baylon, Christian & Fabre, Paul 1995. *Grammaire systématique de la langue française*. Paris: Nathan.

Blanche-Benveniste, Claire 1975. Recherche en vue d'une théorie de la grammaire française. Essai d'application à la syntaxe des pronoms. Paris: Champion.

Boons, Jean-Paul, Guillet, Alain & Leclère, Christian 1976. La structure des phrases simples en français. Constructions intransitives. Genève: Droz.

Borg, Albert & Comrie, Bernard 1984. "Object Diffuseness in Maltese". In Plank (ed.) 1984: 109-126.

Bresnan, Joan 2001. Lexical Functional Grammar. London: Basil Blackwell.

Candito, Marie-Hélène 1999. Organisation modulaire et paramétrable des grammaires électroniques lexicalisées. Application au français et à l'italien. [PhD thesis] Paris: Université Paris 7.

Comrie, Bernard 1975. "Causatives and Universal Grammar". *Transactions of the Philological Society* 1974: 1-32.

Dalrymple, Mary 2001. Lexical Functional Grammar. New York etc.: Academic Press.

Dryer, Matthew 1986. "Primary Object, Secondary Objects, and Antidative". *Language* 62 (4): 808-845.

Eynde, Karel van den & Mertens, Piet 2003. "La valence : l'approche pronominale et son application au lexique verbal". *Journal of French Language Studies* 13 (1): 63-104.

Gaatone, David 1993. "Les locutions verbales et les deux passifs du français". In Gross (ed.) 1993: 37-52.

Gaatone, David 1998. "L'objet direct comme notion formelle dans la formulation des règles syntaxiques". *Travaux de linguistique* 35: 13-20.1.

Gary, Judit & Keenan, Edward 1977. "On Collapsing Grammatical Relations in Universal Grammar". In *Grammatical Relations* [= *Syntax and Semantics* 8], P. Cole & J. Sadock (eds.), 83-120. New York *etc.*: Academic Press.

Gladkij, Aleksej 1983. "Toward a Formal Definition of Grammatical Case and Gender of Nouns". In A. Gladkij & I. Mel'čuk, *Elements of Mathematical Linguistics*, 188-218. Berlin/New York: Mouton.

Grevisse, Maurice 1993. Le bon usage. Grammaire française [13th edition]. Paris/ Louvain-la-Neuve: Duculot.

Gross, Gaston (ed.) 1993. Sur le passif. Paris: Larousse [= Langages 109].

Gross, Maurice 1968. *Grammaire transformationnelle du français. Syntaxe du verbe.* Paris: Larousse. [Reprinted: 1986, Paris: Cantilène.]

Gross, Maurice 1975. Méthodes en syntaxe. Paris: Hermann.

Guillet, Alain & Leclère, Christian 1992. La structure des phrases simples en français. Constructions transitives locatives. Genève/Paris: Droz.

Hudson, Richard 1992. "So-called 'Double Objects' and Grammatical Relations". *Language* 68 (2): 251-276.

Iordanskaja, Lidija & Mel'čuk, Igor 1981. "On a Class of Russian Verbs that Can Introduce Direct Speech: Lexical Polysemy or Semantic Syntax?". In *Slavic Verb*, P. Jakobsen & H. Krag (eds.), 51-66. Copenhague: Rosenkilde and Bagger. [Reprinted in: Mel'čuk, Igor 1995. *The Russian Language in the Meaning-Text Perspective*, Vienna: Wiener Slawistischer Almanach/Moskva: Jazyki russkoj kul'tury, 215-231.] Iordanskaja, Lidija & Mel'čuk, Igor 2000. "The Notion of Surface-Syntactic Relation Revisited (Valence-Controlled Surface-Syntactic Relations in French)". In *Slovo v tekste i v slovare. Sbornik statej k semidesjatiletiju Ju.D. Apresjana*, L. Iomdin & L. Krysin (eds.), 391-433. Moskva: Jazyki russkoj kul'tury.

Isaac, Luc 1986. *Syntaxe dépendentielle du français* [unpublished report]. Utrecht: BSO Research.

Kahane, Sylvain 1997. Bubble Trees and Syntactic Representations. In *Proceedings MOL'5*, T. Becker & H.-U. Krieger (eds.), Saarbrücken: DFKI, 70-76.

Kahane, Sylvain 2000. Extractions dans une grammaire de dépendance lexicalisée à bulles. In *Les grammaires de dépendances* [= *t.a.l.*, 41:1], S. Kahane (ed.), 211-243.

Kahane, Sylvain & Mel'čuk, Igor 1999. Synthèse des phrases à extraction en français contemporain. *t.a.l.*, 40:2, 25-85.

Kayne, Richard 1977. Syntaxe du français. Le cycle transformationnel. Paris: Seuil.

Keenan, Edward L. 1976. "Towards a Universal Definition of «Subject»". In *Subject and Topic*, Ch. Li (ed.), 303-333. New York: Academic Press.

Keenan, Edward L. & Comrie, Bernard 1977. "Noun Phrase Accessibility and Universal Grammar". *Linguistic Inquiry* 8 (1): 63-98.

Keenan, Edward L. & Comrie, Bernard 1979. "Noun Phrase Accessibility Revisited". *Language* 55 (3): 649-664. [See the reprints of Keenan's publications in: Keenan, E.L. 1987, *Universal Grammar: 15 Essays*, London *etc.*: Croom Helm.]

Kunze, Jürgen 1975. Abhängigkeitsgrammatik. Berlin: Akademie-Verlag.

Lazard, Gilbert 1994a. "L'actant H: sujet ou objet ?" Bulletin de la Société de linguistique de Paris 89 (1): 1-28.

Lazard, Gilbert 1994b. L'actance. Paris: Presses Universitaires de France.

Le Goffic, Pierre 1993. Grammaire de la phrase française. Paris: Hachette.

Leclère, Christian 1979. "Syntactic Datives and Ethic Datives". *Statistical Methods in Linguistics* 1 (2): 122-147.

Leclère, Christian 1993. "Classe de constructions directes sans passif". In Gross 1993: 7-31.

Maxwell, Dan & Schubert, Klaus (eds.) 1989. *Metataxis in Practice. Dependency Syntax for Multilingual Machine Translation*. Dordrecht/Providence, RI: Foris.

Mel'čuk, Igor' 1962. "Ob algoritme sintaksičeskogo analiza jazykovyx tekstov. Obščie principy i nekotorye itogi [An Algorithm for Syntactic Analysis of Texts (General Principles and Some Results]". *Mašinnyj perevod i prikladnaja lingvistika* 7: 45-87.

Mel'čuk, Igor' 1963. "Avtomatičeskij analiz tekstov (na materiale russkogo jazyka) [Automatic Text Analysis (for Russian)]". *Slavjanskoe jazykoznanie*, 477-599. Moskva: Izd-vo AN SSSR.

Mel'čuk, Igor' 1964. Avtomatičeskij sintaksičeskij analiz. Tom I: Obščie principy. Vnutrisegmentnyj sintaksičeskij analiz [Automatic Syntactic Analysis. Vol. I: General Principles. Intrasegmental Syntactic Analysis]. Novosibirsk: Izd-vo Sibirskogo Otd. AN SSSR.
Mel'čuk, Igor' 1974. *Opyt teorii lingvističeskix modelej tipa Smysl \u2278 Tekst* [Outline of a Theory of 'Meaning-Text' Linguistic Models]. Moskva: Nauka. [Reprinted: 1999, Moskva: Škola «Jazyki russkoj kul'tury».]

Mel'čuk, Igor 1979. Studies in Dependency Syntax. Ann Arbor, MI: Karoma.

Mel'čuk, Igor 1988. *Dependency Syntax: Theory and Practice*. Albany, N.Y.: State University of New York Press.

Mel'čuk, Igor 1993. "Agreement, Government, Congruence". *Lingvisticæ Investiga-tiones* 17 (2): 307-373. [See also in Mel'čuk 2006: 31*ff*.]

Mel'čuk, Igor 1997. Vers une linguistique Sens-Texte. Paris: Collège de France.

Mel'čuk, Igor 2004. "Actants in Semantics and Syntax II: Actants in Syntax". *Linguistics* 42 (2): 247-291.

Mel'čuk, Igor 2006. *Aspects of the Theory of Morphology*. Berlin—New York: Mouton de Gruyter.

Mel'čuk, Igor & Pertsov, Nikolaj 1987. Surface Syntax of English. A Formal Model within the Meaning-Text Framework. Amsterdam/Philadelphia: John Benjamins.

Mel'čuk, Igor' & Savvina, Elena 1974. "O formal'noj modeli sintaksisa aljutorskogo jazyka". In *Preprinty Problemnoj gruppy po èksperimental'noj i prikladnoj lingvistike* 55: 15-32. [Published in English: 1978. "Toward a Formal Model of Alutor Surface Syntax: Predicative and Completive Constructions". *Linguistics*, Special Issue: 5-39; see also in Mel'čuk 1988: 267-392.]

Morin, Yves-Charles 1980. "Les bases syntaxiques des règles de projection sémantique : l'interprétation des constructions en *faire*". *Lingvisticæ Investigationes* 4 (1): 203-212.

Morin, Yves-Charles 1985. "On the Two French Subjectless Verbs VOICI and VOILÀ". *Language* 61 (4): 777-820.

Munro, Pamela 1982. "On the Transitivity of 'Say' Verbs". In *Syntax and Semantics* 15, P. Hopper & S. Thompson (eds.), 301-318. New York *etc.*: Academic Press.

Plank, Frans (ed.) 1984. *Objects. Towards a Theory of Grammatical Relations*. London *etc.*: Academic Press.

Pollard, Carl & Sag, Ivan 1992. "Anaphors in English and the Scope of Binding Theory". *Linguistic Inquiry* 23 (2): 261-303.

Quirk, Randolph, Greenbaum, Sydney, Leech, Geoffrey, & Svartvik, Jan 1991. A Comprehensive Grammar of the English Language. London/New York: Longman.

Ruwet, Nicolas 1972. Théorie syntaxique et syntaxe du français. Paris: Seuil.

Schubert, Klaus 1989. "Dependency Syntax for Parsing and Generation". In Maxwell & Schubert (eds.) 1989: 7-15.

Tasmowski-de Ryck, Liliane 1984. "?*Lui faire téléphoner quelqu'un d'autre; une stratégie ?". Lingvisticæ Investigationes 8 (2): 403-427.

Van Valin, Robert & LaPolla, Randy 1997. *Syntax. Structure, Meaning and Function.* Cambridge: Cambridge University Press.

Zwicky, Arnold 1977. On Clitics. Bloomington, IN: Indiana University Linguistics Club.

Linear Placement of Serbian Clitics in a Syntactic Dependency Framework

Jasmina Milićević

1 Introduction

The paper describes linear placement of Serbian second-position clitics. The theoretical framework adopted is the Meaning-Text theory (Žolkovskij & Mel'čuk 1967, Mel'čuk 1974 and 1986, Steele (ed.) 1990, Kahane 2003), in particular the Meaning-Text dependency syntax approach (Mel'čuk & Pertsov 1987, Mel'čuk 1988, 2003 and this volume). Clitic placement is considered from the viewpoint of linguistic synthesis (i.e., in the direction from meaning to text).

A set of rules is proposed which allows for synthesizing Serbian sentences containing clitics, starting from their syntactic representations: 1) cliticization rules and 2) clitic linearization rules, the latter subdivided into clitic cluster building rules and rules for the positioning of the clitic cluster in the morphological representation of the sentence.

The paper is structured as follows. Section 2 presents clitics in general. Section 3 introduces basic facts about Serbian second-position clitics and their linear placement. Section 4 proposes a formalization of these facts by means of corresponding rules within a Meaning-Text linguistic model. Section 5—the Conclusion—offers an overview of second-position clitic placement in Serbian and briefly discusses the role of syntactic and prosodic factors in this operation. Clitic placement is viewed as a syntactic operation, crucially, albeit indirectly, driven by syntactic dependencies; the role of prosodic factors is recognized inasmuch as the prosodic features of sentence elements (\approx constituents) influence their capacity to host the clitics.

2 What are clitics?

A clitic is a prosodically deficient wordform, i.e., a wordform which does not inherently carry stress/tone and has to depend prosodically on a full-fledged [= stressed/ intoned] wordform or a phrase in the sentence, called the *host* of this clitic.

A typical example of clitics are clitic forms of personal pronouns in Romance languages; the French sentence in (1) features a sequence of three such clitics (in boldface) leaning on their host, the auxiliary AVOIR ^(have) (underlined). (1) Fr. *Je te l'<u>ai</u> déjà dit* lit. ⁽I to-you_{sg} this have already said⁾ = ⁽I have already told you this⁾.

In this paper I consider only the clitics which have special properties with respect to other wordforms of a language, that is, clitic pronouns, verbs and particles (clitic prepositions and conjunctions are left out). Special nature of these clitics is manifested in both their syntax and morphonology.

Syntactically, or more precisely, from the viewpoint of their linear placement, clitics exhibit the following particularities: clustering, rigid ordering within the sentence—even in the so-called free word-order languages, and sensitivity to prosody (these features will be discussed and illustrated below in connection with Serbian clitics).

Morphonologically, clitics form a prosodic unit with the host and may phonologically interact with it (allomorphic variation, sandhis, deletion, etc.), i.e., they display the behavior typical of affixes. As a result, it is not always easy to determine whether we are dealing with a clitic [= a wordform] or with an affix [= a part of a wordform]. Consider, for instance, the following Serbian sentences featuring the 1sg future tense marker \acute{cu} ^{(I]} will⁾:

- (2) Serbian
 - a. Radiću ^{([I]} work-INF FUT.1.SG⁾ [$\Leftarrow Radi+ti \acute{c}u$] Cf. Croatian *Raditi ću* (*Radit' ću*) [idem].
 - b. *Reći ću* ⁽[I] tell-INF FUT.1.SG².

In (2a), **ću** phonologically interacts with its host—it provokes the truncation of the infinitive ending **-ti**—and looks very much like an affix; in (2b), however, where it is hosted by an infinitive that ends in **-ći**, no such interaction happens and **ću** behaves rather as a (clitic) wordform. (For arguments in favor of treating Serbian future markers as clitics, rather than as suffixes, see Milićević 1999: 249-251 and 2005.)

Because of their bizarre properties,¹ resulting from this complex interaction of syntax and morphology/morphonology, clitics represent an interesting area of research.

Clitics can be classified along several axes (two influential classifications of clitics were proposed in Zwicky 1977 and Klavans 1995):

According to the type of host: adverbal clitic, as in (1), vs. second-position clitics, as in (2) and (3).

(3) Serb. <u>Ona</u> ga je, nažalost, napustila lit. ^{(She him is_(Aux), regrettably, having-left) = ^{(Unfortunately, she has left him).}}

According to the type of clisis (= direction of attachment): enclitics (postclitics), which follow the host, as in (2)-(4), vs. proclitics (preclitics), which precede the host, as in (1) and (5), vs. endoclitics (meso- or intraclitics), which are inserted into the host; cf.

¹ Cf. Spencer's (1991: 358) expression *schizophrenia on the part of the clitics* and the title of a paper on clitics by H.-H. Hock: *What is a nice word like you doing in a place like this?*

(6), where the pronominal indirect-object clitic **lhe** leans on the radical of the verbal wordform and precedes the inflectional suffix **-ei** ⁽¹sg⁾;

- (4) Sp. *Quiero* <u>ver</u>*te* lit. ⁽[I] want to-see you_{sg}⁾ = ⁽I want to see you⁾.
- (5) Sp. *Te* <u>quiero</u> ver lit. $(You_{sg} [I] want to-see) = (I want to see you).$
- (6) Peninsular Port. <u>Screver</u>*lheei* lit. ⁽write+FUT+to-him/her+1.SG⁾ = ⁽I will write to him/her⁾.

According to the domain of placement: clausal clitics, positioned within their clause, cf. (1)-(6), *vs.* phrasal clitics, positioned within their phrase, cf. (7).

(7) Lat. *Senatus populusque romanus* lit. ^{(Senate people-and Roman⁾ = ^{(The Senate and the people of Rome⁾.}}

According to the existence of a corresponding full [= stressed] wordform: clitic lexemes, containing only clitic forms, as in (7), *vs.* clitic wordforms, which co-exist within a lexeme with full wordforms. Clitic wordforms are further divided into inflectional clitics, cf. (1)-(6),² and morphonological clitics, cf. (8), featuring the reduced forms of the English future auxiliary WILL and the personal pronoun THEY.

(8) <u>I'll tell</u>'em.

The basic differences between adverbal clitics (as in Romance languages) and second-position clitics (as in Slavic) can be summarized as follows.

Syntactic class of clitics

Adverbal clitics are pronouns and particles (e.g., the negative NE and the reflexive SE in French, the interrogative particle TU in Quebec French), while second-position [= 2P] clitics are more heterogeneous—they include auxiliaries, pronouns and particles.

Host of clitics

Adverbal clitics are necessarily hosted by a verb, while 2P clitics attach to any appropriate constituent of their clause—typically, the one in the clause-initial position.

Syntactic governor of clitics

The host of an adverbal clitic (in the Deep-Morphological Structure of the clause) always corresponds to the governor of this clitic (in the Surface-Syntactic Structure).³ In contrast, the host of a 2P clitic and its syntactic governor do not necessarily correspond:

² For a lexeme which contains both clitic and full forms, the opposition $(_{CLITIC}) \sim (_{FULL})$ is inflectional (since it opposes lexes of the same lexeme). This opposition corresponds to the category of tonicity (Mil-ićević 1999: 238). Thus, a clitic and the corresponding full form of a personal pronoun in Serbian have the same grammemes of number, gender and case, but different grammemes of tonicity.

³ On the other hand, the surface-syntactic governor of a clitic does not necessarily correspond to the governor of the source of this clitic in the Deep-Syntactic Structure of the clause. Thus, in the Deep-Syntactic Structure of (1) above, the sources of the clitics are governed by the lexical verb DIRE ^(tell) (the auxiliary does not exist at this level of representation).

in fact, if they do, this is purely coincidental. For instance, the sequence of clitics **ga je** in (3) is hosted by the personal pronoun $ON_{(3p)fem, sg}$ (she), which, in the syntactic structure of (3), does not govern either of the clitics. In point of fact, this pronoun is governed itself by one of the clitics—the present-tense auxiliary **je** (is), which is the top node of the corresponding Surface-Syntactic tree, that is, the head of its clause. (On the status of clitic auxiliaries as syntactic heads, see Milićević, to appear.)

Type of clisis

Adverbal clitics can be either proclitics, enclitics or endoclitics, depending on syntactic, communicative, stylistic, etc., conditions; cf. (4) and (5) above, where the pronoun **te** is a proclitic, respectively an enclitic. As for 2P clitics, they are always enclitics (but see the Conclusion, p. 271, for some conflicting evidence from Serbian).

The literature on clitics is huge, so I will mention just a few basic references. On clitics in general, see Zwicky (1977), Spencer (1991: 351-393), Klavans (1995), Halpern (1995, 1999), Dixon & Aikhenvald (2002) and Anderson (2005). Romance pronominal clitics are discussed, for instance, in Borer, ed. (1986); Iordanskaja (1982) offers a Meaning-Text based description of the linear placement of French pronominal clitics. For 2P clitics, in particular Serbian/Croatian clitics, see Halpern & Zwicky, eds., (1996), Kaiser, ed., (1997), Ćavar (1999), Franks & Holloway King (2000) and Progovac (2005: 125-166). A classic work on Serbian/Croatian clitics is Browne (1974) and clitic placement in literary Serbian/Croatian is described in Popović (1997: 283-364). The only dependency-based description of Serbian/Croatian clitics I am aware of is Čamdžić & Hudson (2002), within the Word Grammar framework.

3 Basic facts about Serbian clitics

In this section, I will discuss the forms of the clitics, the constitution of the clitic cluster and the options available for the linear positioning of the cluster within a clause.⁴

Before I proceed, a word of caution is in order. Grammaticality judgments of sentences containing clitics differ significantly among native speakers. This is due, among other things, to interferences between different variants (Croatian, Bosnian, Serbian) and styles (literary, journalistic, colloquial), featuring rather subtle differences in clitic placement, as well as to individual differences. I will try to stick to what can be called standard Serbian, but, some fellow speakers will no doubt disagree with some of my judgments (just as I found a number of examples of clitic placement in linguistic literature or in texts available on the WWW less than acceptable).

3.1 Forms of clitics

In Serbian, 2P clitics include:

⁴ Some of the examples of Serbian sentences containing clitics cited in the paper are my own, others are taken from Serbian texts (novels, newspapers, WWW) or papers on Serbian clitics (in particular Browne 1974, Radanović-Kocić 1996 and Popović 1997). Sources of examples will not be specifically indicated in the text.

- 1) clitic forms of auxiliary verbs and copula/locative verbs: HTETI_(Aux) 'will', BITI_(Aux) 'be' and BITI_(Copula/Locative) 'be';
- 2) clitic forms of personal pronouns (in the dative, genitive and accusative);
- 3) reflexive adjunct SE [= REFL], which is a voice marker and is also used to mark lexical reflexives (= inherently reflexive verbs);⁵
- 4) interrogative particle LI1 [= INTERR] and emphatic particle LI2 [= EMPH].

All clitics except the reflexive SE and the emphatic particle L12 have corresponding full forms; here are fragments of paradigms of a personal pronoun and of an auxiliary verb:

 $JA_{(Pron.Pers 1.sg)} (I) =$..., **mi** clitic, dat ~ **mèni** full, dat, **me** clitic, acc ~ **mène** full, acc, ...⁶ $BITI_{(Aux)} (be) =$

sam clitic, pres, 1, sg ~ jèsam full, pres, 1, sg, ..., je clitic, pres, 3, sg ~ jèste full, pres, 3, sg, ...

The auxiliaries in the negative do not have clitic forms, that is, they are always full (= stressed): n(sam (I] am not), ..., n(je (he/she/it] is not), etc.

The operation whereby a clitic form of a lexeme is chosen in the process of speaking is called *cliticization*. In Serbian, cliticization is licensed by communicative and/or syntactic factors; thus, communicatively unmarked pronouns as direct dependents of a verb and auxiliaries must appear in a clitic form, while a pronoun functioning as a prepositional object is necessarily full; see p. 260, where a sample cliticization rule is provided.

There are instances of clitic incompatibility which prevent the cliticization; these will be illustrated in the next subsection.

3.2 *Constitution of the clitic cluster*

In order to be linearly positioned, all clitics of the same clause are gathered in a clitic cluster. There is one clitic cluster per clause, except in some cases involving the Main Verb [= MV] with an infinitive complement that has its own clitic dependents; then there can be two clusters in the clause (see immediately below).⁷ The clitic cluster cannot be interrupted by a non-clitic element.

The structure of the clitic cluster is specified by the following template:

LI	Aux $<$ Cop, Loc $> \neq$ je 'is'	PRONDAT	PRON _{ACC/GEN}	PRON _{GEN/ACC}	REFL	Aux $<$ Cop, Loc $> =$ je 'is'
1	2	3	4	5	6	7

⁵ The reflexive adjunct behaves in the same way with respect to linear placement regardless of whether it marks voices (objectal reflexive, pronominal passive or subjectal suppressive—according to the typology of voices proposed in Mel'čuk 1993 and 2006: 181-263) or inherently reflexive verbs.

⁶ Serbian has four tonal accents: short-falling [`], short-rising [`], long-falling [^], and long-rising [']; accent symbols do not appear in current texts.

⁷ In the Meaning-Text approach, an infinitive never heads a clause, since a clause is taken to be an expression that necessarily contains a finite verb; thus, in the present framework, we can speak only of infinitival phrases.

The forms of the auxiliary and the copula/locative verbs BITI ^(be) (fully homophonous) are distributed in the following way: all the forms go into the position 2, except the 3sg form, je ^(is), which must be placed in the position 7.

The forms of personal pronouns in the dative include pronouns functioning as indirect object of verbs (e.g., DATI N_{dat} (give to N) and oblique objects of adjectives (e.g., VERAN N_{dat} (faithful to N), which are inherent actants of verbal/adjectival lexemes (i.e., part of their Government Pattern), as well as the so-called Free Datives, Possessor's Datives and Ethical Datives, which are not inherent actants of verbs. The latter, illustrated in (9a-c), will not be considered in this paper. Note only that the Ethical Dative clitics (but not the Free Dative and the Possessor's Dative ones) can co-occur with the indirect-object datives, as shown in (9d). Thus, strictly speaking, the Ethical Dative clitic should be assigned a separate position in the clitic-cluster building template.

- (9) a. Jezik mu_[Free Dative] je bio sredstvo za ...
 lit. 'Language to-him is_(Copula) having-been a means for ...' = 'To him (In his view) language was a means for ...'
 - b. *Kako ti*_[Possessor's Dative] *je otac?* lit. 'How to-you_{sg} is_(Copula) father?' = 'How is your father doing?'
 - c. *Kako si* **mi**_[Ethical Dative]? lit. 'How are you_{sg} to-me?' = 'How are your doing, sweety?'
 - d. *Nemoj* **mi**_[Ethical Dative] **joj**_[Indirect-Object Dative] *ga oduzeti!* lit. ⁽Do.not to-me to-her it take.away!⁾ = ⁽Do not take it away from her on me!⁾

A 3p fem accusative pronominal clitic \mathbf{je} (her) has a variant \mathbf{ju} , which has to be used before the homophonous pres 3sg auxiliary clitic \mathbf{je} (is)_(Aux) or the homophonous 3p fem gen pronominal clitic \mathbf{je} (of-her); cf.:

- (10) a. $Ko ju_{(Pron)} je_{(Aux)} \langle *je_{(Pron)} je_{(Aux)} \rangle$ video? lit. (Who her is having-seen?) = (Who saw her?)
 - b. Ko $ju_{cl, acc} je_{cl, gen} \langle * je_{cl, acc} je_{cl, gen} \rangle lišava?$ lit. 'Who her of-her [e.g. of freedom_(N, fem)] is depriving?' = 'Who is depriving her of it?'

A 3p dative pronoun and a 1/2p accusative pronoun cannot co-occur if they are both in the clitic form—that is, clitic sequences meaning 'me to him', 'you to them'; etc., are banned⁸; cf.:

- (11) a. Ko *[mu_{(3p)cl,dat} me (nas)_{(1p)cl, acc}] je predstavio? lit. 'Who to-him me (us) is having-introduced?' = 'Who has introduced me (us) to him?' VS.
 b. Ko [mu_{(3p)cl,dat} ga (ih)_{(3p)cl, acc}] je predstavio? lit. 'Who to-him him (them) is having-introduced?' =
 - ^{(Who has introduced him \langle them \rangle to him?⁾}

⁸ In fact, this is only a rough formulation of the constraint in question: some such sequences are not outright ungrammatical but I cannot enter into this problem here. For more on the incompatibility of dative and accusative clitics, see Milicević (2007: 108-109); for this phenomenon in Romance languages, see Miller & Monachesi (2003: 87*ff*).

The genitive and the accusative clitics (whose masculine and neuter forms are homophonous) have highly restricted and problematic co-occurrence: they co-occur only when used with a handful of verbs (e.g. OSLOBODITI ^{(liberate/rescue/exonerate N of/from N⁾, LIŠITI ^{(deprive N of N⁾}, UDOSTOJITI ^{(honor N with N⁾}) and some of their combinations are unacceptable (homophonous sequences ***ga+ga**, ***ih+ih**, ***je+je**, as well as sequences ***je**+CL_{acc/gen}). Furthermore, even the acceptable combinations are avoided, being rather difficult to process. Indeed, speakers tend to disagree about the mutual order of these clitics. I believe that this is so because, with the exception of two cases below, they can actually be used either in the "acc+gen" or in the order "gen+acc".}

The order "acc+gen" is obligatory in case of 3sg fem clitics: $\mathbf{ju}_{acc}+\mathbf{je}_{gen} vs$. * $\mathbf{je}_{gen}+\mathbf{je}_{acc}$; cf. again (10b), as well as the unacceptability of Ko * $\mathbf{je}_{cl, gen} \mathbf{ju}_{cl, acc}$ lišava?, having the same meaning as (10b).

The order "gen+acc" is obligatory if the genitive clitic functions as a complement of a quantity word of type MNOGO 'many N_{gen} ', NEKOLIKO 'some N_{gen} ', KOLIKO 'how.many N_{gen} ', VEĆINA 'majority N_{gen} ', etc., which is itself the Subject of a verb having another complement in the accusative. (The genitive clitic, due to the semantics of its governor, is necessarily in the plural.) Thus, cliticization of the nouns in *Koliko ljudi*_{(N)gen} *je posetilo muzej*_{(N)acc}? lit. 'How.many of-people is_(Aux) having-visited museum?' = 'How many people visited the museum?' results in *Koliko ih*_{gen} ga_{acc} *je posetilo*? lit. 'How.many of-them him is_(Aux) having-visited?' = 'How many of them visited it?', the inverse order of the clitics being ungrammatical: *Koliko* * ga_{acc} *ih*_{gen} *je posetilo*? lit. 'How.many him of-them is_(Aux) having-seen?'

In all other cases, the order of these clitics is free (e.g., either $\mathbf{ga}_{acc}+\mathbf{ih}_{gen}$ or $\mathbf{ih}_{gen}+\mathbf{ga}_{acc}$); however, the order "acc+gen" is somewhat easier to understand and may be considered as the default case.

- (12) a. ACC+GEN
 Lišili su ga [e.g., Petra_{(N, masc)sg, acc}] ih [e.g., gradjanskih prava_{(N, neut)pl, gen}]
 lit. ⁽[They] deprived him [e.g., Peter] of-them [e.g., civil rights]³.
 - b. GEN+ACC *Lišili su ih* [e.g., gradjanskih prava_{(N, neut)pl, gen}] *ga* [e.g., Petra_{(N, masc)sg, acc}] [idem].⁹

Situations of incompatible clitic co-occurrence are taken care of by special filter rules, which are beyond the scope of this paper (the interested reader can consult Milićević 2007).

The following clitic combinations are excluded by general syntactic rules of Serbian (the numbers refer to the positions in the clitic cluster building template above, p. 239;

⁹ The free order of acc/gen clitic and the homophony of their masc/neut forms give rise to ambiguity. Thus, sequences of clitics in (12) can be also read the other way around, i.e., (12a) can be interpreted as ⁽[They] deprived of-it [e.g., the right to work] them [e.g., Peter and Marija]⁾, and (12b) as ⁽They deprived them [e.g., Peter and Marija] of-it [e.g., the right to work]⁾.

positions 4 and 5 are taken as default positions for the accusative, respectively genitive, clitics):

- *2+7, since there can be only one clitic auxiliary or copula/locative verb per clause;
- *4+6, since Serbian has no reflexive verbs taking an accusative complement.¹⁰

Combination 3+5 is restricted. There are no lexemes taking simultaneously a dative and a genitive complement, but the sequence "dat+acc" can appear if a verb which takes a dative indirect object has as its Subject a quantity word governing a genitive clitic (cf. above, the discussion of the relative order of the accusative and genitive clitics) or if it has a quantified Noun Phrase as its Direct Object. Here are the corresponding examples: *Većina* **mu**_{dat} **ih**_{gen} *se protivi* 'Majority of-them to-it REFL opposes_(Vrefl)² = 'A majority of them [e.g., of voters] is opposed to it [e.g., to this proposal]³ and *Daj* **joj**_{dat} **ih**_{gen} *pet* 'Give to-her of-them five³ = 'Give her [e.g., (to) Maria] five of them [e.g., of those books]³. (Otherwise, this combination is possible if the dative clitic is a Possessor's Dative or an Ethical Dative, i.e., a non-inherent verbal complement; cf.: *Nema* **mi**_{dat} **ga**_{gen} 'There.is.not to-me of-him³ = 'He is absent on me³, where **mi** is an Ethical Dative.)

Due to the above restrictions, the maximum number of clitics in a cluster is four for an interrogative sentence, three for an affirmative sentence and two for a negative sentence (minus LI clitics and the auxiliaries, which, as previously mentioned, are always stressed in the negative); for instance:

(13) a. *Kad* **li** sam mu ga poslao? 1 2 3 4 lit. ^{(When EMPH am_(Aux) to-him it having-sent?⁾ = ^{(When on earth did I send it to him?⁾}}

b. <i>Juče</i>	sam	тu	ga poslao
	2	3	4
lit. ⁽ Yeste	erday ar	n _(Aux)	to-him it having-sent ^{i} = ⁽ I sent it to him yesterday ⁾ .

c. $Nisam_{(Aux)}$ **mu ga** poslao 3 4 lit. ⁽Am-not to-him it having-sent⁾ = ⁽I did not send it to him⁾.

If we take into account non-inherent dative complements, there can be up to five clitics in a cluster, as in the following one, including a Possessor's Dative clitic: *Kako li* su mu ga se [1+2+3+5+6] dokopali? 'How EMPH $\operatorname{are}_{(Aux)}$ to-him it REFL [they] havinggot_(Vrefl)?' = 'How on earth did they manage to lay their hands on it (something belonging to him)?'. However, such extended clusters are rare.

If the MV of the clause has an infinitive complement, the infinitive's own clitic dependents may in some cases stay within the infinitival phrase, where they form a separate clitic cluster, instead of being raised to the general clitic cluster of the clause; this is illustrated in (14a). In (14b), we see an instance of clitic *pseudo-climbing*: the

¹⁰ Croatian has a voice, marked with SE, that does not exist in Serbian—the subjectless suppressive (Mel'čuk 2006: 203-204); verbs in this voice do take an accusative object. Thus, the combination 4+6 is possible in Croatian; cf.: *Redovito ih se kontrolira* regularly he-CL.MASC.PL.ACC REFL control-IND.PRES.3.SG lit. ^{(Regularly [it]} controls itself them⁾ = ^(They are regularly controlled).

accusative pronominal clitic **te** (you_{acc}) , which depends syntactically on the infinitive *videti* (see), becomes part of the general clitic cluster of the clause.¹¹

- (14) a. Lepo $je_{(Copula)}$ [videti $te_{(Pron)}$ ovde]_{Inf.Phrase} lit. 'Nice is to-see you_{sg} here' = 'It is nice is to see you here'.
 - b. Lepo $te_{(Pron)} je_{(Copula)} [videti ovde]_{Inf.Phrase}$ lit. 'Nice you_{sg} is to-see here' = 'It is nice is to see you here'.

Clitic pseudo-climbing can be obligatory, optional or impossible, depending on the syntactic role of the infinitive within the clause. Roughly speaking, the situation is as follows. If the infinitive is an Object of the MV, its clitics must join those of the MV in the general clitic cluster of the clause; that is, pseudo-climbing is obligatory; cf. (15). However, if the infinitive is the Subject, pseudo-climbing is allowed/disallowed depending on the position on the infinitive within the clause: if the Subject infinitive is linearly the first constituent of the clause, pseudo-climbing is impossible and the clitics of the MV cannot join those of the infinitive, either; that is, two clitic clusters must be constructed; cf. (16). If the Subject infinitive is not clause-initial, pseudo-climbing is optional; this was already shown in (14).

- (15) a. Oni su mi hteli [reći da ...]_{Inf.Phrase} lit. 'They are_(Aux) to-me having-wanted to-tell that ...' = 'They wanted to tell me that ...'
 VS.
 b.*Oni su hteli [reći mi da ...]_{Inf.Phrase} lit. 'They are_(Aux) having-wanted to-tell to-me that ...' = 'They wanted to tell me that ...'
 (16) a. [Videti te_(Pron) ovde]_{Inf.Phrase} lepo je_(Copula) lit. 'To-see you_{sg} here nice is² = 'It is nice is to see you here'. VS.
 - b. *[*Videti ovde*]_{Inf.Phrase} *lepo te*_(Pron) *je*_(Copula) lit. 'To-see here nice you_{sg} is² = 'It is nice is to see you here².
 - c. *[*Videti* $te_{(Pron)} je_{(Copula)} ovde]_{Inf.Phrase} lepo$ lit. 'To-see you_{sg} is here nice' = 'It is nice is to see you here'.

Clitic pseudo-climbing will not be considered in this paper. However, evidence from clitic pseudo-climbing will be used in the Conclusion to substantiate the claim of the predominant role of syntactic factors in Serbian clitics linear placement. (Remarks on clitics pseudo-climbing in Serbian/Croatian, referred to as *climbing*, can be found in Browne 1974: 122-123 and Progovac 2005: 146-147.)

¹¹ In the Meaning-Text framework, climbing proper is raising to a higher governor in the syntactic tree (during the transition from the Deep- to the Surface-Syntactic Structure); e.g., Fr. *J'ai vu Sylvain y aller* lit. ^{(I} have seen Sylvain there go² = ^{(I} have seen Sylvain go there² [the adverbial clitic y (there² depends on *aller* (go², which in its turn depends on *voir* (see², which depends on *avoir* (have²] \Rightarrow *J'y ai vu aller Sylvain* ^{(I} there have seen go Sylvain² = ^{(I} have seen Sylvain go there² [y depends on *avoir*]. This, however, does not happen in Serbian: in (14b) and (15a), the pronoun does not change its syntactic governor, only its position in the morphological string. That is why we cannot speak of the real clitic climbing but of something reminiscent of it—hence the qualifier *pseudo*-.

3.3 Linear placement of the clitic cluster within a clause

As mentioned above, Serbian 2P clitics are linearly positioned within a clause with respect to a constituent 12 —a wordform or a phrase—with the appropriate hosting capability, called a *hosting constituent*, or *host*.

3.3.1 Basic types of hosting constituents

There are two basic types of hosting constituents: 1) an absolute host, i.e., the actual host in all clauses (in which it appears), irrespective of its position within the clause, and 2) a potential host, i.e., the actual host in some clauses, depending on its properties and the structure of the clause. These two types of host are illustrated, respectively, in (17) and (18):

(17)	a. [<i>Da</i>] _{[abs}	host] [sam][cl.cluster]	samo	znao!
	that _(Conj)	a	m _(Aux)	only	having-known
	⁽ If only I	knew!)	、 ,		
	b. [<i>Samo</i>]	$[da]_{[abs host}$	[] [sam] _{[c] cluster}]	znao!	
	only	that _(Coni)	am _(Aux)	having	-known
	[idem]	(conj)	(Turi)	C	

- c. *Ali samo* [*da*]_[abs.host] [*sam*]_[cl.cluster] *znao!* but only that_(Conj) am_(Aux) having-known 'But if I only knew!'
- (18) a. [*Tu*]_[host] [*su se*]_[cl.cluster] *iznenada javila dva slučaja oboljenja* there are REFL suddenly appear two cases of-illness ⁽Two cases of illness have suddenly appeared there⁾.
 - b. [Dva slučaja oboljenja] [tu]_[host] [su se]_[cl.cluster] iznenada javila [idem].
 - c. [Dva slučaja oboljenja]_[host] [su se]_[cl.cluster] tu iznenada javila [idem].
 - d. [Javila]_[host] [su se]_[cl.cluster] tu iznenada dva slučaja oboljenja [idem].

In (17), the subordinate conjunction DA is the actual host in all clauses: no matter which position DA occupies within the clause, the clitic cluster must follow it; otherwise, the result is ungrammatical: $*Samo [sam]_{[cl.cluster]} [da]_{[abs.host]} znao!$, $*Ali [sam]_{[cl.cluster]} [da]_{[abs.host]} samo znao!$, $*Znao [sam]_{[cl.cluster]} [da]_{[abs.host]} samo!$ (In clauses without the absolute host the constituents SAMO, ALI and ZNATI can host the

¹² I will be relying on an intuitive understanding of the notion of constituent until Section 4.3.2, p. 261, where it will be characterized. For the time being, note that constituents in the Meaning-Text approach, although they "physically" most often correspond to those used in the phrase-structure approach to syntax, are conceptually quite different: Meaning-Text constituents are not part of the syntactic representation of the sentence; rather, they are created and exploited by linearization rules in the process of mapping unordered syntactic trees onto fully ordered morphological strings.

clitics.) In (18), however, the deictic adverb TU is the actual host only in (18a-b); in (18a), it must host the cluster, in (18b), it may, but not need, host it—cf. (18c-d)—and in (18d) it cannot host it: $*[Javila] tu_{[host]} [su \ se]_{[cl.cluster]} iznenada dva slučaja oboljenja.$

Accordingly, two types of clitic placement are distinguished: fixed (= grammaticized) placement and variable placement (cf. Popović 1997: 289). In addition, a special placement option, available both in fixed and variable placement, must be considered the insertion of the clitic cluster into a hosting constituent.

3.3.2 Fixed placement

The cluster is placed after the absolute hosting constituent.

Absolute hosts are of the following types: 1) subordinate conjunctions, such as DA ^(that), NEKA ^(that), PRE NEGO ŠTO ^(before), POŠTO ^(after), AKO ^(if), etc.;¹³ 2) WH-words and constituents of the form "WH-word+X" (cf. *clause introducers* of Browne 1975). These two types of host are illustrated, respectively, in (19) and (20).

- (19) a. [*Ne želim* [[*da*]_[abs.host] *ti se žalim*]_{Completive}]_{Matrix} lit. 'Not [I] want that_(Conj) to-you_{sg} REFL complain² = 'I do not want to complain to you².
 - b. Ali [tek (odmah) pošto]_[abs.host] se odselio u Kanadu, ... lit. 'But as.soon.as (immediately) after_(Conj) REFL [he] having-moved_(Vrefl) to Canada, ...' = 'But as soon as he moved to Canada, ...'
 - c. [*Neka*]_[abs.host] *ti je* sa srećom! lit. 'That_(Conj) to-you_{sg} is_(Copula) with luck!' = 'May you have luck!'
 - d. *Sve* [*neka*]_[abs.host]*je prokleto*! lit. 'Everything that_(Conj) is_(Copula) damned!' = 'Damn everything!'

In (19a), we see the conjunction DA introducing a completive clause, and in (19b), the conjunction POŠTO introducing a temporal subordinate clause. In (19c-d), the conjunction NEKA is used in a special type of subordinate clause that might be called "optative;" the use of the conjunction DA in another special type of subordinate clause, expressing a counterfactual, is illustrated in (17) above.

- (20) a. [*Kad*]_[abs.host] *je kod kuće, ...* lit. ^{(When [he/she] is_(Copula) at home, ...⁾ = ^{(When he/she is at home, ...⁾}}
 - b. *Kod kuće* [*kad*]_[abs.host] *je*, ... [idem]
 - c. *Ne znam* [*zašto* (*kako, gde*)]_[abs.host] *su nestali* lit. ⁽Not [I] know why (how, where) [they] are_(Aux) having-disappeared⁾ = ⁽I do not know why (how, where) they disappeared⁾.

¹³ Some coordinate conjunctions are absolute hosts as well; cf. Popović (1997: 297ff).

- d. *Devojka* [*čijih očiju*]_[abs.host] *se seća* ... lit. ⁽Girl whose eyes REFL remembers_(Vrefl) ...⁾ = ⁽The girl whose eyes [he/she] remembers ...⁾
- e. [*Čiji drug*]_[abs.host] *će doći sutra?* lit. ^{(Whose friend FUT.3.SG to-come tomorrow?) = ^(Whose friend will come tomorrow?)}

Thus, in the fixed placement, the syntactic class of a constituent is the only factor relevant for determining its hosting capability.¹⁴

3.3.3 Variable placement

The clitic cluster is placed after a (potential) hosting constituent—so as not to violate the skipping conditions (see below).

The hosting constituent may be of any syntactic class, but must satisfy some specific requirements, having to do with 1) its syntactic role/composition, 2) its communicative markedness (e.g., focalization, contrastiveness), and 3) its prosodic heaviness [= the number of stresses it carries].

Some hosting constituents are illustrated in (21):

(21) a. [*Moj drug Marko*]_[host] *će doći sutra* lit. ⁽My friend Marko FUT.3.SG to-come tomorrow⁾ = ⁽My friend Marko will come tomorrow⁾.

or

Moj drug Marko [*doći*]_[host] *će sutra* [idem].

b. [*Sutra ujutru*]_[host] *ćemo ići u Beograd* lit. ⁽Tomorrow morning FUT.1.PL to-go to Belgrade⁾ = ⁽Tomorrow morning we will go to Belgrade⁾.

or

Sutra ujutru [ići]_[host] ćemo u Beograd [idem].

c. [*Sapiru*]_[host] *je jezik sredstvo* ... lit. ⁽For-Sapir language is_(Copula) a means ...⁾ = ⁽For Sapir language is a means ...⁾

or

Sapiru [jezik]_[host] je sredstvo ... [idem]

d. [*NAGRADU*]_[host] *je dobio Marko* lit. ⁽Award-ACC.SG having-got is_(Aux) Marko-NOM.SG⁾ = ⁽It is an award that Marko got⁾.

¹⁴ When there are two (or more) absolute hosts in the clause, the situation is as follows. If both hosts are WH-words, the cluster is positioned after the first; cf.: $[Ko]_{abs.host-1} je [koga]_{abs.host-2} tužio?$ ^{(Who is}_(Aux) whom having-sued?) = ^(Who sued whom?); otherwise, it is positioned after the second, cf.: *Devojka* $[koja]_{abs.host-2} [kao da]_{abs.host-2} je bila ... ^{(Girl} who as if is_(Aux) having-been ...) = ^(The girl who seemed to be ...)$

Three necessary properties of a (potential) hosting constituent are as follows; all of them must be satisfied at the same time:

- 1) It is NOT a constituent set off from the rest of the clause by an obligatory pause/ pauses. Such pauses can be induced either by the constituent's syntactic role (e.g., a parenthetical, a non-restrictive apposition) or its communicative role (e.g. a circumstantial expressing a communicative specifier¹⁵). Pause-inducing constituents can be characterized as +detached; this characterization, which represents a prosodic generalization over syntactic/communicative properties of constituents, can then be used to determine the hosting capability thereof.
- 2) It is NOT a completive clause functioning as the Syntactic Subject of the MV.
- 3) It is NOT a $V_{non-fin}$ if hosting a cluster containing the interrogative particle L11.

The constituents which do not have these properties are incapable of hosting the clitic cluster, i.e., they are non-hosting constituents; cf.:

- (22) a. *[Moj drug, Marko,]_{NP with Non-Restr.Appos = +detached} će doći sutra lit. ⁽My friend, Marko, FUT.3.SG to-come tomorrow⁾ = ⁽My friend, Marko, will come tomorrow⁾.
 - b. *[*Poslednjih dana*,]_{Comm.Specifier = +detached} *smo* svedoci neobične pojave ... lit. ⁽In-recent days, [we] are_(Copula) witnesses of-unusual phenomenon ...⁾ = ⁽Lately, we are observing an unusual phenomenon ...⁾
 - c. *[*Da dodje*]_{Completive.Clause, Subject} *se nije usudio* lit. ^{(That}_(Conj)comes REFL not.is having-dared⁾ = ^{(To come he didn't dare⁾.}
 - d. *[*Video*]_{Verb, non-fin} *li si ga*_{Cluster containing LI ^(whether)? lit. ⁽See-ACT.PART.SG.MASC whether be_(Aux) him?⁾ = ⁽Have you_{sg} seen him?⁾}

¹⁵ A communicative specifier is an element of the Communicative Structure of the sentence (see below, p. 256), namely the one outside the communicative core (Theme and Rheme); roughly speaking, it specifies a circumstance of the event(s) constituting the communicative core.

Compare (22a), where an NP containing a non-restrictive apposition (= a + detached constituent) is unfit to host the clitic cluster, with (21a), where an NP containing a restrictive apposition (= a - detached constituent) functions as host.

To express the meaning intended in (22d), it is necessary to reconstruct the sentence, by using either the full form of the interrogative particle (= DA LI) or the full form of the auxiliary; cf., respectively, $Da \ li_{\text{[host]}} si \ ga \ video$? and $J \approx li_{\text{[host]}} li \ ga \ video$?

Skipping

There can be more than one hosting constituent in the clause. In such a case, some of them can be ignored when looking for the landing site for the clitic cluster; that is, they may be *skipped* (term introduced in Halpern 1995). A hosting constituent which can be skipped is called a *skippable host*. For instance, the initial constituent in each of the sentences in (21) is a skippable host.

A skippable host has one of the following properties:

- 1) It is a prosodically heavy constituent, i.e., a constituent containing more than one stressed wordform, such as [*Moj drug Marko*] in (21a) and [*Sutra ujutru*] in (21b); a constituent of this type is assigned the feature +heavy.
- 2) It is a communicatively marked constituent, i.e., a constituent expressing focalization of the Theme or the Rheme, contrast or emphasis, to which the feature +contrastive is assigned, and it is not (the lexical part of) the MV. Two such constituents are the Free Dative [*Sapiru*] in (21c), functioning as the focalized Theme of the sentence (which is expressed by fronting), and the direct object NP [*NAGRADU*] in (21d), appearing in the role of the focalized Rheme (which is expressed by a particular prosodic contour, symbolized here by capital letters). The MV, even if marked as +contrastive, cannot be skipped; cf.: **DOBIO* [*nagradu*][host] *je Marko* lit. ⁽Having-got award-ACC.SG is_(Aux) Marko-NOM.SG² = ⁽Marko did get the award²).¹⁶
- 3) It is a lexically marked item, such as the conjunctions ALI 'but' in (21e) and JER ('for' = 'because') in (21f).

¹⁶ The condition 2, which states that a contrastive constituent can be skipped as long as it does not function as the MV of the clause is actually not quite correct, for at least the following two reasons. First, some +contrastive subjects seem more difficult to skip than others; thus, **SAPIR* (**ON*) *jezik je smatrao sredstvom* ... 'SAPIR (HE) language is_(Aux) having-considered as a means ...² = 'It was Sapir (he) who considered language as a means ...² is unacceptable, but *NAGRADA dodeljena je Marku* 'AWARD having-been-given is_(Aux) to-Marko² = 'It is an award that was given to Marko² is not. The same observation holds for circumstantials; cf. the ungrammatical **SUTRA oni će doći* 'TOMORROW they will come² *vs*. the acceptable *SUTRA knjigu mi donesi* 'TOMORROW book-ACC.sg to-me bring-IMPER.2.sg² = 'Bring me the book TOMORROW². An intriguing situation, since we are dealing here with the same circumstantial. What this could mean is that global word order within the clause or properties of other constituents may be a factor in determining conditions for the skipping of contrastive constituents. Second, examples of skipped **-heavy** and **-contrastive** constituents can be found (through a Google search); cf. [?]Nekada to *je bilo normalno* 'Once that is_(Copula) having-been normal² = 'At one time, this was considered normal² is (note that *nekada* the is NOT followed by a pause here!). For the time being, I am unable to make this condition more precise.

Skipping can be repeatable; this is shown in (23), where the first three constituents are skippable hosts (by virtue of being +heavy) and have actually been skipped; the cluster thus appears after the fourth constituent of the clause.

(23) [U tim okolnostima] [ideje Čomskog] [dalju razradu] [doživele] su u delu njegovih sledbenika ...
lit. ⁽In those circumstances ideas of-Chomsky further elaboration having-experienced are_(Aux) in work of-his followers ...⁾ = ⁽Under those circumstances, Chomsky's ideas were further elaborated in the work of his followers ...⁾

Skipping is always optional, with preferences; see immediately below.

A hosting constituent which cannot be skipped is called *non-skippable host*. A non-skippable host has the following properties:

It is **-heavy** AND is either **-contrastive** or the lexical part of the MV (but see note 16, p. 248).

The initial constituents in (24) are non-skippable hosts.

- (24) a. *[*Marko*]_{-heavy, -contrastive} *doći će sutra* lit. 'Marko to-come FUT.3.SG tomorrow' = 'Marko will come tomorrow'.
 - b. *[*Možda*]_{-heavy, -contrastive} *doći će sutra* lit. 'Maybe to-come FUT.3.SG/FUT 3.PL tomorrow' = 'Maybe he/they will come tomorrow'.
 - c. *[$DO\acute{CI}$]-heavy, +contrastive, Lex.part of the MV Marko \acute{ce} (, kad ti kažem) lit. 'To-come Marko FUT.3.SG (, when you_{sg} [I] tell)² = 'Marko WILL come (I assure you)².

Preferences for skipping

As indicated above, skippable hosts are either heavy constituents or contrastive ones. Let me take them in turn.

Heaviness of a constituent is a gradient: the more stressed wordforms a constituent contains, the heavier it is. Thus, *moj prijatelj* 'my friend' is less heavy than *takvo interesovanje lingvistike za matematiku* 'such interest of the linguistics for the mathematics', which is, in its turn, less heavy than *svima onima koji žele da se podrobnije upoznaju sa najnovijim dostignućima generativne gramatike* 'to all those who want to get more closely acquainted with the latest achievements of the generative grammar'.

The heavier a constituent is, the more preferable it is to skip it. Thus, for the first constituent above, the skipping preference is rather low, for the second one it is relatively high, while for the third one this preference is very high, verging of obligatoriness.

Here is a more developed illustration of skipping preferences for heavy constituents:

(25) a. [Svim studentima] se preporučuje da ...

lit. 'To-all students REFL recommends that_(Conj) ...' = 'It is recommended to all students to ...'

EQUALLY GOOD

[*Svim studentima*] *preporučuje* **se** *da* ... [idem]

b. [Svim studentima lingvistike] se preporučuje da ... lit. 'To-all students of-linguistics REFL recommends that_(Conj) ...' = 'It is recommended to all students of linguistics to ...'

SOMEWHAT BETTER

[*Svim studentima lingvistike*] *preporučuje* **se** *da* ... [idem]

latest achievements of the generative grammar are advised to ...

c. [Svim studentima lingvistike koji nisu čitali Čomskog] se preporučuje da ... lit. ⁽To.all students of-linguistics REFL who haven't read Chomsky recommends that_(Conj) ...⁾ = ⁽It is recommended to all students of linguistics who haven't read Chomsky to ...⁾

BETTER

[Svim studentima lingvistike koji nisu čitali Čomskog] preporučuje **se** da ... [idem]

d. [?][Svima onima koji žele da se podrobnije upoznaju sa najnovijim dostignućima generativne gramatike] se preporučuje da ...
lit. ⁽To-all those who want that REFL more-thoroughly get-acquainted with latest achievements of-generative grammar.⁾ = ⁽All those who want to get a better idea about the

DEFINITELY BETTER

[Svima onima koji žele da se podrobnije upoznaju sa najnovijim dostignućima generativne gramatike] preporučuje **se** da ... [idem]

When it comes to preferences for the skipping of contrastive constituents, the situation is less clear, since with these constituents several interacting factors that may impede or favor the skipping are at play; these factors, which include context, style and global word order, require a more careful study.

Clause-final position

With the exception of non-raised infinitive clitics (i.e., the clitics which have not undergone pseudo-climbing out of the infinitival phrase; cf. p. 243), which can be clausefinal, Serbian clitics tend to stay away from the right edge of the clause (cf. Browne 1974: 120, Radanović-Kocić 1996: 438, Popović 1997: 333). Clitics can be placed clause-finally under specific communicative and/or stylistic conditions, into which I cannot delve here;¹⁷ sentences in (26) illustrate the clause-final placement of the copula/the auxiliary BITI ^(be):

- (26) a. *Razlozi ovoj pojavi višestruki su* lit. 'Causes to-this phenomenon multiple are_(Copula)' = 'This phenomenon has multiple causes'.
 - b. *Svi pokušani načini obmanuli su* lit. ^{(All} attempted ways having-failed $are_{(Aux)}^{(Aux)} = {}^{(Everything that has been tried has failed)}$.

Otherwise, they are in the clause-final position only if 1) this cannot be avoided and 2) the preceding non-host is of a particular type.

- (27) a. *[*Moj drug Marko*][skippable host] *posetio me je* $\langle odsutan je \rangle$ lit. ⁽My friend Marko having-visited me is_(Aux) $\langle absent is_{(Copula)} \rangle$ ⁾ = ⁽My friend Marko visited me $\langle is absent \rangle$ ⁾.
 - b. $[Na \check{z}alost,]_{[non-host]}$ posetio **me je** $\langle odsutan je \rangle$ lit. ^{(Regrettably, having-visited me is_(Aux) $\langle absent is_{(Copula)} \rangle$) = ^{(Unfortunately, he visited me $\langle is absent \rangle$).}}
 - c. [?][*Moj drug, Marko*,]_[non-host] posetio **me je** $\langle odsutan je \rangle$ lit. ⁽My friend, Marko, having-visited me is_(Aux) $\langle absent is_{(Copula)} \rangle$ ⁾ = ⁽My friend, Marko, visited me $\langle is absent \rangle$ ⁾.

Sentence (27a) illustrates the illicit skipping of an otherwise skippable host, leaving the cluster clause-final; cf. correct sentences [*Moj drug Marko*]_[skippable host] posetio **me** *je juče* $\langle odsutan je od juče \rangle$ ^{(My friend Marko visited me yesterday \langle has been absent since yesterday \rangle), where the skipping of the same constituent is permitted since it does not have the effect of placing the clitics at the very end of the clause. In order to avoid the undesirable result illustrated in (27a), two strategies are available: do not skip or insert (see below).}

In (27b) and (27c) the clause-initial constituent is a non-host—because it induces a pause after itself—and the only available position for the cluster is clause-final. However, the effect of placing the cluster clause-finally in different in these two cases. The non-host in (27b) is a sentential adverb and the sentence is perfect. In (27c), the nonhost is an NP containing a restrictive apposition and the result is less felicitous: a better way to express the meaning of sentence (27c), which is highly marked (it signals the surprise on the part of the speaker) is *Moj drug, Marko, — on me je posetio* (*on je odsutan*) ^{(My friend, Marko, he visited me (he is absent); that is, a reformulation is preferable.}

¹⁷ It seems that the copula clitic is more easily accepted in the clause-final position than the auxiliary clitic. Also, there may be a difference in behavior of the copula in case a state *vs*. a property is predicated (e.g., *be absent* vs. *be smart*; cf. the ESTAR ~ SER opposition in Spanish): in the former case, the clause-final position seems more acceptable.

3.3.4 Special placement: insertion into a host

The cluster placed after a special type of host, called *insertable host*, can be inserted into this host—after the host's first stressed word; cf.:

(28) [*Moj me je drug Marko*]_[insertable host] *posetio juče* lit. ⁽My me is_(Aux) friend Marko visited yesterday⁾ = ⁽My friend Marko visited me yesterday⁾.

Remarks

- Insertion of the clitic cluster into a constituent is distinct from disruption of a constituent by another constituent, under the impact of the communicative structure. Compare *Koliko vam je godina?* 'How.many to-you_{pl} is_(Copula) years?' = 'How old are you?' (insertion of two clitics into the constituent *Koliko godina*) with *Koliko vi imate godina?* 'How.many you_{pl} have years?' = 'How old are you?' (disruption of the same constituent by two full fledged-wordforms, each forming a constituent in its own right).
- 2. In older language or in literary style, examples of insertion not after the first stressed wordform of the constituent can be found. (According to examples in Ćavar 1999: 158, this is still possible in modern Croatian.)

There are two cases of insertion: optional and obligatory. Optional insertion concerns all clusters and all hosting constituents—absolute and potential—of the required syntactic composition, while the obligatory insertion concerns only the clusters containing the copula BITI ^(be) and the hosting constituents in a specific syntactic role, namely, those implementing the predicative-attributive construction.

Optional insertion

The insertable host is a +heavy constituent; it can be a subordinate or a coordinate phrase.

1) Insertion into a subordinate phrase

The insertable host is of the form "Premodifier+X(+Postmodifier)"; this can be either an absolute host, as in (29a-b), or a potential host, as in (29c-g).

- (29) a. [*Kojim je putem*] *došao*? lit. ^{(By-which is_(Aux) way [he] having-come?⁾ = ^{(By which way has he come?⁾}}
 - b. [Koje su boje] gradski autobusi u Beogradu?
 lit. 'Of-which is(Copula) color city buses in Belgrade?' = 'What color are the city buses in Belgrade?'
 - c. [*Suviše je malo*] *živeo* lit. ⁽Too is_(Aux) shortly having-lived⁾ = ⁽His life was too short⁾.
 - d. [*Mnogo su veći uspeh*] *na tom polju imali blumfildovci* lit. ⁽Much are(Aux) bigger success in that field having-had Bloomfieldians⁾ = ⁽The proponents of Bloomfield's approach were much more successful in that field⁾.

- e. [*U tom se smislu*] *može govoriti o* lit. ⁽In that REFL sense can to-speak of ...⁾ = ⁽In that sense one can speak of ...⁾
- f. [*Svi* su studenti elektrotehinike] učestvovali u protestu lit. ⁽All are(Aux) students of-electrical-engineering having-participated in protest⁾ = ⁽All students of electrical engineering participated in the protest⁾.
- g. [*Te su ideje Comskog*] *dalju razradu doživele u* ... lit. 'Those are(Aux) ideas of-Chomsky further elaboration having-experienced in ...' = 'Those ideas of Chomsky were further elaborated in ...'

vs.

- (30) a. *[Studenti su elektrotehnike] učestvovali u protestu
 lit. 'Students are(Aux) of-electrical-engineering having-participated in protest² = 'students of electrical engineering participated in the protest².
 - b. *[*Ideje su Čomskog*] *dalju razradu doživele u ...*lit. 'Ideas are(Aux) of-Chomsky further elaboration having-experienced in ...' = 'Those ideas of Chomsky were further elaborated in ...'

2) Insertion into a coordinate phrase

The insertable host is a constituent of the form " $X+Z_{(Coord,Conj)}+Y$ ".

- (31) a. [*Tudja je i neprijatna*] *ta kuća* lit. ⁽Alien is_(Copula) and unpleasant that house⁾ = ⁽That house is alien and unpleasant⁾.
 - b. [*Tiho su ali živo*] *razgovarali* lit. 'In-low-voice [they] are_(Aux) but lively having-talked' = 'They were talking in a low voice but lively'.

Obligatory insertion

- (32) a. [*Studenti su elektrotehnike*] *od 2003* lit. ⁽Students are_(Copula) of-electrical-engineering since 2003⁾ = ⁽They have been students of electrical engineering since 2003⁾.
 - b. *[*Studenti elektrotehnike*] *su od* 2003 [idem].

Compare the ungrammatical (32b) with a correct [*Studenti elektrotehnike*]_{Subject} $su_{(Copula)}$ od 2003 na čelu protesta 'Students of-electrical-engineering are since 2003 athead of protest² = 'Students of electrical engineering have been spearheading the protest since 2003', where the NP *Studenti elektrotehnike* functions as the Subject of the clause and can be followed by the clitic.

Some speakers (myself included) would prefer to avoid the ungrammatical result in (32b) not by resorting to insertion, as in (32a), but rather by realizing the pronominal subject ONI ^(they), dropped in both sentences of (32), and use it as the host of the copula (in spite of the fact that pronominal subjects are normally dropped under neutral

communicative conditions): Oni $su_{(Copula)}$ [studenti elektrotehnike]_{Predicative Attribute} od 2003. This tendency to avoid insertion indicates that this placement option is loosing ground in Serbian.¹⁸ However, it is still fairly frequently used in some styles, especially in the language of the press.

Preferences for insertion

Insertion preferences will be given only for subordinate phrases. The relevant factors are syntactic composition of the insertable host, its prosodic heaviness and communicative markedness. The number of clitics in the cluster does not seem to be relevant.

1) Syntactic Composition

Insertion into a constituent of the form "Premodifier+X", in the order of decreasing preference: Adv+Adj/Adv+Adv, Adj+N, N+N.

Insertion into a constituent of the form "Premodifier+X+Postmodifier" is disfavored, especially if the postmodifier is a noun complement in the genitive. Thus, [*Te su ideje* $\check{C}omskog$] ... lit. 'Those are_(Aux) ideas of-Chomsky ...' is worse than [*Te ideje* $\check{C}om$ skog] *su* ... lit. 'Those ideas of-Chomsky are_(Aux) ...', etc.

2) Prosodic Heaviness

Insertion into a very heavy constituent is avoided (here we have the same preferences as for skipping of a very heavy constituent). However, the insertion into such a constituent can be forced, as it were, if this is the only way to avoid putting the cluster clause-finally; cf. (33). Thus, the non-final position constraint is stronger than the present one.

(33) Protiv **smo** dijaloga velikih na štetu trećih zemalja

lit. ⁽Against [we] are_(Copula) of-dialog of-big [ones] on detriment of third countries⁾ = ⁽We are against the dialog of big powers at the expense of third countries⁾.

3) Communicative Markedness

Insertion into a constituent whose first stressed wordform carries a communicative load is preferred. Two such cases will be indicated.

• The constituent's first word is interrogative (a subset of obligatory hosts)

For instance, [$\check{C}iji$ **je** drug] došao? lit. ^{(Whose is_(Aux) friend having-come?) = ^(Whose friend came?) is better than [$\check{C}iji$ drug] **je** došao? lit. ^{(Whose friend is_(Aux) having-come?) In a WH-question, the interrogative pronoun expresses the focal part of the Rheme, even though it is not prosodically or otherwise marked (in this connection, see Browne 1976 and Halpern 1995: 86).}}

¹⁸ For instance, sentences such as $[Lav je_{(Copula)} Tolstoj]_{NP}$ veliki ruski pisac lit. ^{(Leo is Tolstoy [a]} great Russian writer⁾, where the cluster is inserted into a complex proper name, are now obsolete, but were not uncommon in older language.

• The constituent's first word is contrastive

For instance, [*MOJ je drug*] *došao* lit. ^{(MY is}_(Aux) friend having-come⁾ = ^{(It is MY friend who came⁾, where the modifier *MOJ* expresses the contrastive focus of the sentence, is better than [*MOJ drug*] *je došao* lit. ^{(MY friend is_(Aux) having-come⁾. However, preferences are equal for [*Moj je DRUG*] *došao* lit. ^{(My is}_(Aux) FRIEND having-come⁾ = ^{(It is my FRIEND who came⁾ and [*Moj DRUG*] *je došao* lit. ^{(My FRIEND is_(Aux) having-come⁾, since in both sentences it is the second wordform of the insertable constituent (rather than the first) that expresses the contrastive focus.}}}}

To sum up, a description of the linear placement of Serbian second-position clitics has to account for the following thee operations: 1) the cliticization, 2) the construction of the clitic cluster, and 3) the linear placement of the cluster within the clause, the latter operation involving, in its turn, a) the identification of hosting constituent(s) and b) the choice of the actual host, with possible insertion into the host. Before going on to formally describe these operations in the Meaning-Text framework, the following remark is in order.

As it should be clear from the preceding discussion, the term *second position* [= 2P] used to describe the placement of Serbian clitics is just a conventional label. In point of fact, 2P is the default position for the cluster in clauses with no absolute hosts, in the sense that the cluster can always be placed after the first hosting constituent. But the first hosting constituent is not necessarily the first constituent of the clause; cf., for instance, examples (22a-c), where the first hosting constituent is the second one in the clause. And even if it is the first constituent, it sometimes can be skipped; cf., for instance, examples (21).

4 Serbian second-position clitics in the Meaning-Text framework

Meaning-Text Theory [= MTT] considers language to be a set of rules establishing a correspondence between any given meaning and all synonymous texts which implement it (and vice versa). MTT proposes to describe this correspondence by means of functional models of languages, called *Meaning-Text Models*.

A Meaning-Text Model (of a language) is a synthesis-oriented, semantics-driven, dependency-based stratificational model. It presupposes seven levels of representation of utterances—semantic, deep/surface syntactic, deep/surface morphological, deep/surface phonological—and consists of six sets of rules [= modules] that establish correspondences between representations of adjacent levels.

Clitic placement rules are part of the Surface-Syntactic module, which maps a Surface-Syntactic Representation (of a sentence) to the corresponding Deep-Morphological Representation(s).

MILIĆEVIĆ

4.1 The input for the clitic placement operation: SSyntR

The Surface-Syntactic Representation [= SSyntR] of a sentence is a set of four structures: the Surface-Syntactic Structure [= SSyntS], the Surface-Syntactic Communicative Structure [= SSynt-CommS], the Surface-Syntactic Anaphoric Structure [= SSynt-AnaphS] and the Surface-Syntactic Prosodic Structure [= SSynt-ProsS].

The SSyntS, the basic structure of the SSyntR, is a linearly unordered dependency tree. The nodes of this tree are labeled with actual lexemes of the sentence subscripted with meaning-bearing inflectional values (in Serbian, grammemes of nominal number and of verbal voice, mood and tense); its branches are labeled with names of surface-syntactic dependency relations [= SSyntRel], which are language-specific. Surface-syntactic relations in Russian and English, are described, respectively, in Mel'čuk (1974: 237-260) and Mel'čuk & Pertsov (1987: 212-440); see also Mel'čuk in this volume, Subsection 4.8, p. 52. An inventory of surface-syntactic actantial relations in French can be found in Iordanskaja and Mel'čuk (2000) and in their paper in the present volume.

Serbian clitics appear in the following SSynt-roles:¹⁹

- $V_{(Aux)/(Copula/Locative)}$ is the top node of the SSyntS of a clause.
- Pronouns depend on a lexical verb via the following SSyntRels: direct-objectival (accusative/genitive pronouns), indirect-objectival (dative pronouns) and oblique-objectival (genitive pronouns); pronouns governed by an adjective are linked to it via the oblique-objectival SSyntRel and those governed by a quantifying lexeme of type MNOGO 'a-lot', KOLIKO 'how-much/how-many', etc. (only genitive pronouns) are linked to it via the completive SSyntRel.
- The reflexive adjunct SE depends on a lexical verb via the auxiliary-reflexive SSyntRel.
- The interrogative particle L11 depends on the top node (a lexical verb or an auxiliary), while the emphatic particle L12 depends on the lexeme which it "emphasizes"—both via the adverbial SSyntRel.

SSynt-CommS consist of markers of mutually exclusive values of communicative oppositions (Mel'čuk 2001), such as Thematization (Theme ~ Rheme ~ Specifier), Giveness (Given ~ New), Focalization (Focalized ~ Non-Focalized), Perspective (Foregrounded ~ Backgrounded ~ Neutral), etc., characterizing subtrees of the SSyntS. (In other frameworks, the communicative organization of sentences is known as Functional Sentence Perspective, Information Structure or Information Packaging; see, for instance, Sgall *et al.* 1986, Lambrecht 1994 and Chafe 1994.) The SSynt-CommS plays a crucial role in the linearization and prosodization of the dependency tree, in particular in languages with the so-called free word order, such as Serbian (for more on this, see below).

The SSynt-AnaphS specifies co-referential links between appropriate nodes of the SSyntS.

The SSynt-ProsS consists of a set of markers of meaning-bearing prosodies: declarative, interrogative, exclamative; ironic, pathetic, etc.

¹⁹ We should speak rather of the sources of the clitics, since the cliticization has yet to be performed.

Let there be the following sentences, each featuring the same cluster, consisting of two clitics—su ^(be)_(Aux)-IND.PRES.3.PL se REFL:

- (34) a. *Poslednjih dana*, [*javila*]_[host] *su se u njihovim kućama dva slučaja oboljenja* lit. ⁽In-last days, having-appeared are_(Aux) REFL in their houses two cases of illness⁾ = ⁽Recently, in their households there have been two cases of illness⁾.
 - b. *Poslednjih dana, u njihovim kućama [javila*]_[host] su se dva slučaja oboljenja [idem].
 - c. *Poslednjih dana*, [*u njihovim kućama*]_[host] *su se javila dva slučaja oboljenja* [idem].
 - d. *Poslednjih dana*, [*u njihovim* **su** se kućama]_[host] javila dva slučaja oboljenja [idem].

Sentences in (34) are propositionally synonymous (i.e., they express the same Semantic Structure). They also feature the same communicative orientation: they are all-rhematic sentences (i.e., not having the Theme) reporting on a particular state of affairs, namely the existence of an illness at a specific place and time.²⁰

The SSyntR of sentences in (34) is as follows:





The identical underlying communicative orientation of the four sentences in (34) is reflected in the (almost) identical linear order of their full-fledged (= non-clitic) elements. In all sentences, the SSynt-Subject NP *dva slučaja oboljenja* is close-final due to the fact that is expresses the Rhematic Focus (in a Serbian declarative sentence, the sen-

 $^{^{20}}$ The Rheme ~ Theme division of a sentence can be tested by finding an underlying question for it, i.e., a question to which this sentence can be an appropriate answer. Each of the sentences in (34) can be an appropriate answer to the underlying question *What is going on?*, used to elicit all-rhematic sentences.

tence element expressing the most informative part of the rheme is realized closefinally). The circumstantial *poslednjih dana* is realized in the clause-initial position due to the communicative feature **Foregrounded** characterizing the corresponding SSynt-subtree (foregrounding of a sentence element signals its psychological prominence for the Speaker). The MV *javila* and the circumstantial *u njihovim kućama* are not additionally marked for communicative features, so that their linear position is allowed to vary.²¹

As for the order of the clitic elements (with respect to the non-clitic ones), it is not free—it does not depend on the communicative choices of the Speaker but is contingent upon the previously selected order of full-fledged sentence elements.

4.2 *The output of the clitic placement operation: DMorphR*

The Deep-Morphological Representation [= DMorphR] of a sentence is a set of two structures: the Deep-Morphological Structure [= DMorphS] and the Deep-Morphological Prosodic Structure [= DMorph-ProsS].

The DMorphS, the basic structure of the DMorphR, is a fully ordered string of morphological representations of all wordforms of this sentence, i.e., the names of the corresponding lexemes subscripted with all relevant inflectional values (= both the semantic grammemes, carried over from the surface-syntactic level, and the syntactic grammemes—those of agreement/government and tonicity). For the sentences with clitics, the DMorphS contains the clitic cluster placed in the appropriate position.

DMorph-ProsS consists of a set of markers of all pauses and prosodies (= both the semantically and the syntactically induced ones). These are not, however, the final sentence prosodies: in the transition towards the phonological structure, some phonologically conditioned rephrasing may happen.

From the SSyntR above, the rules of the surface-syntactic module will produce the four DMorphRs below. (The clitic cluster is boxed; the syntactic grammemes are in boldface. The only element of DMorph-ProsS represented are the pauses: "||" stands for a longer pause and "|" for a shorter one.)

DMorphR of (34a)

POSLEDNJI_{masc, pl, gen} DAN_{pl, gen} || JAVITI_{part, act, neut, pl}BITI_{ind, pres, 3, pl, cl} SE U NJIHOV_{fem, pl, loc} KUĆA_{pl, loc} | DVA_{masc, nom} SLUČAJ_{sg, gen} OBOLJENJE_{sg, gen}

²¹ The role of communicative factors in linearization of SSynt-Structures can be illustrated by comparing sentences in (34) with a propositionally synonymous sentence featuring a different communicative organization, for instance, (i) *Ta dva slučaja oboljenja javila su se posledjih dana u njihovim kućama* 'Those two cases of illness have appeared lately in their households', in which the Subject, *ta dva slučaja oboljenja* (two cases of illness), expresses the semantic Theme, the rest of the sentence expressing the semantic Rheme. This communicative difference between sentences in (34) and sentence (i) is manifested by their different word orders. Formally speaking, sentence (i) cannot be synthesized from the same underlying representation as sentences in (34)—even if we disregard the minor propositional difference between them (presence *vs.* absence of the demonstrative).

DMorphR of (34b)

POSLEDNJI_{masc, pl, gen} DAN_{pl, gen} || U NJIHOV_{fem, pl, loc} KUĆA_{pl, loc} |JAVITI_{part, act, neut, pl} BITI_{ind, pres, 3}, pl, cl SE|DVA_{masc, nom} SLUČAJ_{sg, gen} OBOLJENJE_{sg, gen}

DMorphR of (34c)

POSLEDNJI_{masc, pl, gen} DAN_{pl, gen} || U NJIHOV_{fem, pl, loc} KUĆA_{pl, loc} BITI_{ind, pres, 3, pl, cl} SE JAVITI_{part, act, neut, pl} | DVA_{masc, nom} SLUČAJ_{sg, gen} OBOLJENJE_{sg, gen}

DMorphR of (34d)

POSLEDNJI_{masc}, pl, gen DAN_{pl}, gen || U NJIHOV_{fem}, pl, loc BITI_{ind}, pres, 3, pl, cl SE KUĆA_{pl}, loc JAVITI_{part}, act, neut, pl | DVA_{masc}, nom SLUČAJ_{sg}, gen OBOLJENJE_{sg}, gen

Linear placement of 2P clitics poses an interesting problem for a dependency framework. Clitics positioning is determined not only by syntactic dependencies and the communicative structure as it is the case for full-fledged words. Normally, a dependent member of a SSynt-relation is placed with respect to its governor (and perhaps its codependents), taking into account the relevant communicative information. Not the clitics. Here, some additional factors—in particular, the prosodic properties of sentence elements (\approx constituents)—are at play. As we have seen, such prosodic properties of a constituent as heaviness (= the number of stressed wordforms it contains) and its capacity to induce a pause/pauses are relevant to the identification of clitic-hosting constituents. This fact indicates the existence of an involved relationship between dependency structure and prosody in the clitic placement. The study of this relationship occupies an important place in the literature on 2P clitics; indeed, the role of prosodic factors in clitic placement has been given increasingly more prominence, so much so that looking for a prosodic account has become the hottest point. This question will be briefly addressed in the Conclusion, where I will argue, on the basis of evidence adduced in the paper, that in Serbian 2P clitics linear placement syntactic factors play a decisive role.

4.3 Rules required to carry out the SSyntR-to-DMorphR mapping

The three major types of rules of the SSynt-module are: 1) morphologization rules, further divided into government rules (e.g. case assignment), agreement rules and cliticization rules; 2) linearization rules; 3) prosodization rules.

All these rules, except cliticization rules, are needed to synthesize not only sentences containing clitics, but all sentences. Thus, case-assignment rules assign case to all nominals, i.e., nouns and pronouns (both clitic and full), etc.

I will discuss here only cliticization and linearization rules, leaving prosodization rules aside (but see the Conclusion).

The rules are either of the form $X_{[level n]} \Leftrightarrow Y_{[level n+1]} | C$, where X and Y are fragments of representations of adjacent levels and C the conditions under which the correspondence holds, or else they have the form of logical implication: if X, then Y.

4.3.1 Cliticization rules

These rules assign the grammeme 'CLITIC' to all lexemes in the SSyntS which have to be cliticized. Which lexemes will be cliticized is determined by the rules themselves. In most cases, cliticization in Serbian is obligatory; however, it can be prevented by communicative or syntactic factors.

There are cliticization rules for auxiliary verbs, personal pronouns, and the interrogative particle L11. (Recall that the emphatic particle L12 and the reflexive adjunct SE are clitic lexemes, having the feature "clitic" in their syntactics.²²) A sample cliticization rule, which takes care of the cliticization of personal pronouns, follows. (The shadowing indicates the context of the application of the rule.)



r = direct-objectival, indirect-objectival **or** oblique-objectival

Illustration of the conditions

1) If a pronoun in the SSyntS is communicatively marked (as, say, emphatic or contrastive), it will be assigned the grammeme (FULL) (by a different rule) and will eventually surface in the full form; cf.:

(35) Video sam njega (*ga) (a ne njegovog brata)
lit. 'Having-seen am_(Aux) he-FULL.MASC.SG.ACC (*he-CL.MASC.SG.ACC) (and not his brother)'
⁽[I] saw HIM (and not his brother)² = ⁽It was him that I saw (, not his brother)².

2) Pronouns used in coordination are always full, cf.:

(36) Video sam njega $\langle *ga \rangle$ i nju $\langle *je \rangle$ lit. ⁽Having-seen am_(Aux) he-FULL.MASC.SG.ACC $\langle *he-CL.MASC.SG.ACC \rangle$ and she-FULL.FEM.SG.ACC $\langle *she-CL.FEM.SG.ACC \rangle^{2} = ([I])$ saw both him and her².

Remarks

- 1. If \mathbf{r} = subjectival or prepositional, L₂ is not cliticized; in other words, a pronoun appearing in the role of the SSynt-subject is always full, and so is a pronoun functioning as a prepositional object; this latter case is illustrated in (37):
 - (37) To sam čuo od njega $\langle *ga \rangle$

⁽That $am_{(Aux)}$ having-heard from_(Prep) he-FULL.MASC.SG.GEN (*he-CL.MASC.SG.GEN)⁾ = ⁽I heard that from him⁾.

²² Syntactics is a component of a linguistic sign (along with the signified and the signifier); it contains the information about the combinatorial properties of the sign (e.g., part of speech, declension/conjugation group, Government Pattern, collocations).

2. Cliticization of L_2 can lead to situations of impossible/dubious clitic co-occurrence (involving specific combinations of dat/acc and gen/acc clitics, cf. p. 240); in such cases, the present rule has to operate in conjunction with filter rules, mentioned above, p. 241, which ban or allow the cliticization, as appropriate.

Similar conditions—namely, neutral communicative marking—apply to the cliticization of the auxiliaries. In the case of our sample sentences (34a-d), p. 257, the auxiliary will be assigned the grammeme 'CLITIC', since it is communicatively unmarked.

4.3.2 Linearization rules

Linearization rules determine the actual linear order of words in the sentence, based on syntactic, communicative and prosodic information in the SSyntR.

Elements of a Serbian sentence (\approx constituents) can be linearly arranged in several different ways, depending on complex conditions—above all communicative, but also syntactic and lexical ones. Moreover, as we have seen, the clitic cluster is positioned after an appropriate constituent; this means that all constituents have to be built and linearized before we can proceed with the linearization of the clitics. For this reason, the linearization in Serbian (and other languages with 2P clitics) is, I believe, best viewed as a two-stage operation: linearization of non-clitic elements and linearization of clitic elements. But, before describing these operations, I have to say a few words on the notion of constituent in the Meaning-Text approach.

As already mentioned, MTT constituents are conceptually different from constituents used in the phrase-structure approach to syntax—so much so that, in order to avoid confusion, it would probably be better not to call them constituents at all.²³ However, given the fact that the term *constituent* has become so familiar in studies of 2P clitic placement, I find it difficult to dispense with.

The phrase-structure representation—at least in its classical form—tries to represent together syntactic hierarchy (i.e., syntactic relations) and word order (\approx constituency) of sentence elements; see, for instance, Jackendoff 1977. The MTT approach separates these two aspects of sentence organization very sharply. A dependency-based syntactic representation, such as the one I have been using in this paper, has as its basic structure a linearly non-ordered tree, which reflects only the hierarchy of sentence elements; word order (viz. constituents), as a means of expressing this hierarchy, cannot be part of the syntactic structure. Basic information on dependency representation can be found in Mel'čuk 1988: 12-42 and in this volume, p. 1*ff*; for a comparison between the phrase-structure and dependency formalisms see, for instance, Hudson 1980 and Rambow & Joshi 1997.

Under the present view, constituents are entities used for the task of computing word-order and prosody of the sentence from its SSyntS; they are sentence-building blocks, created and exploited by linearization rules, and their sole purpose is to serve as "intermediaries" between the syntactic tree and the morphological string. Being the output of linearization rules, constituents can be visualized only at the morphological level

²³ This is the viewpoint of I. Mel'čuk, who consistently uses the term *group* instead of *constituent*; cf., for instance, the discussion of word-order rules for Russian in Mel'čuk 1974: 268*ff*.

of representation—in particular in an intermediate representation, called the partial DMorphR (see below).

I am not in a position to offer a rigorous definition of the notion of MTT constituent and must limit myself to a rather crude characterization thereof. For a recent attempt at theorizing this notion within the MTT approach, albeit from an "unorthodox" MTT perspective, see Gerdes & Kahane 2007.

An MTT constituent is a string of (fully ordered) wordforms which corresponds to a SSynt-subtree, i.e., is its continuous projection, and behaves as a whole from the view-point of linearization and prosodization (= the string is moved around and prosodized as a whole).

In other words, a constituent is a string of wordforms which underlies a single prosodic group [\approx a phrase]; prosodic groups are computed based on such constituents.

One important case, not covered by the above characterization involves conjunctions in Serbian. Since in this language conjunctions are capable of hosting the clitic cluster (cf. examples in Section 3.3, p. 244), a SSynt-subtree headed by a conjunction does not project into a single constituent; rather, the conjunction (perhaps with an adverbial dependent) forms a constituent in its own right.

Following Mel'čuk (1974: 268-300), two basic subtypes of constituent will be distinguished: an *initial constituent*, which is a string made up of wordforms linked by local SSynt-relations, and a *final constituent*, a string made up of initial constituents whose heads are linked by semi-local SSynt-relations, such that it is a dependent of the MV (i.e., the absolute head of the clause).²⁴

For instance, the initial and the final constituents for the sentence *Sutra dolazi Marko, moj prijatelj iz Beograda* lit. ^{(Tomorrow} comes Marko, my friend from Belgrade⁾ = ^{(Marko, a friend of mine from Belgrade, is coming tomorrow⁾ are as follows: Initial constituents:}

[SUTRA]_{Advp}, [DOLAZI]_{VP}, [MARKO]_{NP}, [MOJ PRIJATELJ]_{NP}, [IZ BEOGRADA]_{PP}

Final constituents:

[SUTRA]_{Advp}, [DOLAZI]_{VP}, [[MARKO] [MOJ PRIJATELJ] [IZ BEOGRADA]]_{NP}

An initial and a final constituent can coincide (just as, say, a sentence can coincide with a clause); this is the case with the AdvP and the VP above.

A final constituent corresponds to what is traditionally called a *sentence element* (subject, predicate, object, circumstantial, etc.). It is the final constituents that are relevant for the placement of clitics.

 $^{^{24}}$ A SSynt-relation between two lexemes L_1 and L_2 is local if the mutual order of L_1 and L_2 can be determined without reference to other lexemes making up the sentence. L_1 and L_2 are linked by a semilocal SSynt-relation if determining their mutual order has to take into account their relations with other lexemes, namely with their co-dependents. There exists a third type of SSynt-relations—global SSynt-relations, which hold between the heads of final constituents and are used, together with the communicative/ prosodic information, to determine the order of final constituents within a clause.

Let us now go back to the operation of linearization.

Linearization of non-clitic elements

Linearization of non-clitic elements of the clause is taken care of by two sets of rules: constituent building rules and rules for linear arrangement of constituents; cf. word-order rules for Russian in Mel'čuk 1967 and 1974: 268-299.

1) Constituent building rules

These rules map subtrees of a SSyntS onto corresponding constituents, and determine linear order of wordforms within each constituent. Here we have local and semi-local linearization rules, which work recursively, building first the initial and then the final constituents of the clause. Dependencies have a crucial role in this process; the impact of the communicative structure is much less significant.

Here are the final constituents for sentences (34):

- 1. [POSLEDNJImasc, pl, gen DANpl, gen]_{NP} +heavy
- 2. [U NJIHOV fem, pl, loc KUĆApl, loc]_{PP} +heavy
- 3. [[DVAmasc, nom SLUČAJsg, nom]_{NP} [OBOLJENJEsg, gen]_{NP}]_{NP} +heavy, +detached
- 4. [JAVITIpart, act, pl, neut]_{VP}

Each final constituent is supplied with features which reflect its syntactic, communicative and prosodic properties relevant to clitic placement, i.e., to determining its hosting capability. (More generally, these features are needed to compute the prosodic structure of the sentence, i.e., they are assigned to constituents in all sentences, not only in those containing clitics.) These are the three features—**±heavy**, **±detached** and **±contrastive**—that have been used during the discussion in Section 3.3.3; for their more detailed characterization see immediately below.

Constituents 1, 2 and 3 above are assigned the feature +heavy (since each contains more than one stressed wordform), and constituent 3 is additionally assigned the feature +detached (since it corresponds to a SSynt-subtree communicatively marked as Foregrounded; cf. the SSyntR of our sample sentences, p. 257). Constituent 4, for which the values of all three features are negative, bears no explicit marking.

2) Rules for linear arrangement of constituents

These are global linearization rules, which determine the order of final constituents within the clause (and the order of clauses within the sentence); here, the communicative factors play a crucial role. Thus, in Serbian, fronting of a constituent is a way of expressing a specific communicative role, namely the Focalization of the Rheme; a constituent placed clause-finally expresses the focal part of the Rheme; etc.

For communicative factors determining the linear ordering of sentence elements in the case of our sample sentences, see p. 257.

The result of this stage of linearization is a Partial DMorphR of the sentence—a sequence of constituents (supplied with relevant features), plus a set of clitics that have yet to be ordered and linearly positioned in their turn.²⁵

Here is the Partial DMorphS of (34b-d):

[POSLEDNJImasc, pl, gen DANpl, gen]_{+heavy}, _{+detached} [U NJIHOVfem, pl, loc KUĆApl, loc]_{+heavy} [JAVITIpart, act, pl, neut] [DVAmasc, nom SLUČAJsg, gen OBOLJENJEsg, gen]_{+heavy}

+

{BITIind, pres, 3, pl, cl SE}

Linearization of clitics

Here, again, two sets of rules are needed: rules which build the clitic cluster and rules which position the cluster into the Partial DMorphR of the clause.

1) Clitic cluster building rules

These rules build the clitic cluster according to the clitic cluster building template, p. 239; more precisely, they assign the appropriate position in the cluster to all the lexemes in the SSyntS which meet the conditions for cliticization (as determined by cliticization rules; cf. one such rule above, p. 260) or are lexically marked as clitics.

The two clitic cluster building rules involved in the construction of the cluster we see in sentences (34) follow.

- 1. (a) BITI_(V) clitic, not [pres, 3, sg] \rightarrow 2; (b) BITI_(V) clitic, pres, 3, sg \rightarrow 7
- 2. SE_(Refl.Adjunct, clitic) $\rightarrow 6$

The rule 1 contains two subrules: the subrule (1a) assigns position 2 in the cluster to all clitic forms of $BITI_{(Aux)/(Copula/Locative)}$ (be), except the 3sg.pres clitic [= **je** (is)], and the subrule (1b) assigns position 7 to the 3sg.pres clitic form of $BITI_{(Aux)/(Copula/Locative)}$ (be). The Rule 2 assigns position 6 to the Reflexive Clitic Adjunct SE.

In our case, since the form of the auxiliary is not pres, 3, sg, the order of the two clitics is 2+6.

2) Clitic cluster linear positioning rules

The rules which position the clitic cluster within the clause are of three types: hostidentifying rules; clitic cluster placement rules proper, including insertion rules; and preference rules for skipping and insertion. In addition, there is a rule which bans the clitics from being placed in the clause-final position, when this is required.

²⁵ The Partial DMorphR is an intermediate representation, obtained as the output of a subset of SSyntrules before the transition towards the DMorphR is completed. Unlike other representations used by a Meaning-Text model, the Partial DMorphR is not a representation of a real sentence. I cannot provide here a justification for introducing it, apart from its obvious usefulness for the description of 2P clitics placement. For a discussion of representation levels presupposed by a Meaning-Text model, see, for instance, Mel'čuk 2000: 28-41.

2a) Host-identifying rules

These rules identify, among the constituents in the Partial DMorphR of a clause, those which are fit to host the clitic cluster. The rules will not be stated formally.²⁶

As previously mentioned, there are two major types of hosting constituents: absolute hosts (involved in fixed placement) and potential hosts (involved in variable placement). Absolute hosts are simply given by a list, while potential hosts are identified using features \pm heavy, \pm detached and \pm contrastive, representing (partial) generalizations over prosodic, syntactic and communicative properties of constituents.

The feature +heavy is assigned to a constituent containing more than one stressed wordforms (cf. Zec & Inkelas 1990: 373*ff*), which identifies it as skippable host; cf., for instance, examples (21a-b), p. 246, and (25), p. 250.

The feature **+detached** is assigned to a constituent inducing a pause (after itself or both before and after), which identifies it as a non-host. Such a constituent either fulfils a specific syntactic role (an address, a parenthetical, a constituent containing a non-restrictive apposition, a detached modifier, a converbal phrase, etc.) or has been made communicatively prominent by the Speaker (e.g., a circumstantial expressing a communicative specifier, a sentential adverb, autonomous circumstantial, etc.); thus, the prosodic break(s) induced by the constituent appear(s) as a marker of its syntactic/communicative role. Cf. examples (22), p. 247, and the following ones:

- (38) a. [Umoran,]_{Detached.modifier = +detached = non-host} zaspao je odmah lit. 'Tired, having-fallen.asleep is_(Aux) immediately' = 'Tired, he fell asleep immediately'.
 - b. [*Vraćajući se kući*,]_{Converbal.phrase = +detached = non-host} *sreo je Marka* lit. 'Returning home, having-met is_(Aux) Marko² = 'On his way home, he met Marko².
 - c. [*Marija*,]_{Address = +**detached** = non-host volim te lit. 'Marija, [I] love you_{sg}' = 'Marija, I love you'.}
 - d. [*Posle tri dana*,]_{Autonomous.circumstantial = +detached = non-host odustali su lit. 'After three days, having-given.up are_(Aux)' = 'After three days, they gave up'.}
 - e. [*Nažalost*,]_{Sentential.adverb = +detached = non-host} odustali su lit. 'Regrettably, having-given.up are_(Aux)' = 'Unfortunately, they gave up'.

Compare (38d), where the constituent *Posle tri dana* has the role of an autonomous circumstantial (i.e., is +detached and thus a non-host), with the sentence *Posle tri dana su odustali*, where the same constituent functions as a non-autonomous circumstantial (i.e., is -detached and thus a possible host).

The feature +contrastive is assigned to a constituent expressing a positive value of one of such communicative oppositions as Focalization (of the Theme/the Rheme) or

²⁶ Here is an example of a formal statement of such rules: C+heavy \rightarrow skippable host; C+heavy of the form "Premodifier+X+(Postmodifier)" \rightarrow insertable host; etc.

Emphasis. A constituent of this type carries a particular prosodic contour and is, in most cases, a skippable host; cf. examples (21), p. 246, and the following:

(39) [*Ćamila*]_{Focalized Theme = +contrastive = skippable host *svi su znali* lit. '*ćamil-sg.acc* all are_(Aux) having-known' = 'As for *ćamil*, everybody knew him'.}

A constituent which is neither of these, i.e., for which the values of all three features are negative, is, again in most cases, a non-skippable host; cf. (24a-b), p. 249, and the following:

(40) *[*Ćamil*]_{non skippable host *bio je brat sultana* ... lit. '*Ćamil having-been is*_(Copula) brother of sultan ...' = '*Ćamil was the brother of the sultan ...*'}

In some cases, these features are not sufficient to determine the hosting capability of a constituent: the syntactic role or the syntactic class of a constituent can override them.

First, although a contrastive constituent is normally a skippable host, a contrastive verb is a non-skippable host; cf. (24c) and the following example:

(41) *[*IZNAJMILI*] +contrastive, Verb = non skippable host *kuću su (a ne kupili)* lit. 'Having-rented house are_(Aux) (and not having-bought)' = 'They RENTED a house (rather than bought one)'.

Second, even though constituents bearing the unmarked values of the three features above are normally non-skippable hosts, there are some exceptions. This is the case, for instance, of coordinate conjunctions JER ^(for = because) and ALI ^(but), which are skippable hosts; cf. examples (21e-f), p. 247. Items such as these have to be specified by a list or else be lexically marked for their hosting capability.

A subset of clitic-hosting constituents—those marked as +heavy—have to be additionally marked for insertability. This is done as a function of their syntactic composition (possible for constituents of the form "Premodifier+X(+Postmodifier)" or "X+Z_(Coord.Conj)+Y") or their syntactic role (obligatory for constituents in the role of the Predicative Attribute); such constituents are assigned features "Insert" and "Insert!", respectively ("!" means ⁽obligatory⁾).

(42) a. *[*Ideje su Čomskog*] +heavy, not "Premodifier+X" = not Insert *bile* ... lit. 'Ideas are _(Aux) of-Chomsky having-been ...' = 'Chomsky's ideas were ...'

[*Te* su ideje Čomskog] +heavy, "Premodifier+X" = Insert bile ... lit. ⁽These are _(Aux) ideas of-Chomsky having-been ...⁾ = ⁽These ideas of Chomsky were ...⁾

b. *[Sasvim jasan] +heavy, Pred.Attr = Insert! mi je njegov stav

lit. 'Completely clear to-me is_(Aux) his position.' = 'His position is completely clear to me'. vs.

[Njegov stav] +heavy, not Pred.Attr = not Insert! *mi je sasvim jasan*

lit. 'His position to-me is_(Aux) completely clear.' = 'His position is completely clear to me'.

VS.

Examples in (42) illustrate only potential hosts which are insertable. Absolute hosts can be insertable, too; see p. 252*ff* for examples and discussion.

The result of the application of host-identifying rules to the Partial DMorphR of sentences (34b-d) is as follows:

[POSLEDNJI_{masc}, pl, gen DAN_{pl}, gen]**non-host** [U NJIHOV_{fem}, pl, loc KUĆA_{pl}, loc]**host**, **skip**, **insert** [JAVITI_{part}, act, pl, neut]**host**, **non-skip** [DVA_{masc}, nom SLUČAJ_{sg}, gen OBOLJENJE_{sg}, gen]**host**, **skip**, **insert**

Since in this Partial DMorphR there are no absolute hosts, variable placement applies. The initial constituent, bearing the feature +detached, is identified as a non-host. The second one, a +heavy constituent of the form "Premodifier+X", is a skippable and insertable host. The third one, a -heavy constituent and the MV of the clause, is a non-skippable host. The fourth constituent is, just like the second one, a +heavy constituent of the form "Premodifier+X", so it is a skippable and insertable host. This gives us a total of three logically possible landing sites for the clitic cluster, one of which—after [DVA-masc, nom SLUČAJ_{sg}, gen OBOLJENJE_{sg}, gen]—must be discarded by the next set of rules.

2b) Clitic cluster linear placement rules

Basic Rules

- 1. If, in a clause, there is a constituent C identified as the absolute host, C_{HOST!}, then the clitic cluster is placed after it.
- 2. If no $C_{HOST!}$ is present,

then the clitic cluster is placed a) after the first C of its clause identified as a nonskippable host, $C_{HOST, NON-SKIP}$, or b) after any C identified as C_{HOST} and preceding the first non-skippable host $C_{HOST, NON-SKIP}$.

Additional conditions for clusters containing the interrogative clitic particle LI1.

- 1. If the clitic cluster contains the interrogative particle L11, **then** it cannot be hosted by a non-finite verb.
- 2. If the clitic cluster contains the LI1 and its host is not a verb, then the style of the utterance is literary or dated.

For an illustration of the first condition, see (22d), p. 247, and (43a). The second condition is illustrated in (43b).

- (43) a. *[Doći] li je želeo?
 lit. 'To-come whether is_(Aux) having-wanted?' = 'Did he want to come?'
 - b. literary/dated [*Marka*] *li* čekaš? lit. ⁽Marko-SG.ACC [you_{so}] whether wait?⁾ = ⁽Are you waiting for Marko?⁾

In the Partial SMorphR above, there are two possible landing sites for the clitic cluster: either after the second constituent (a skippable host) or after the third one (a nonskippable host). The first possible landing site is used in the sentence (34b) and the second one in the sentences (34c)-(34d), with, in the case of the sentence (34d), the subsequent insertion.

No-clause-final position rule

If, in a clause, there exists a non-final position available for the clitic cluster, then the cluster should not be placed in the clause-final position.

The above rule is approximate: as we have seen (p. 251), some sentences with the cluster placed clause-finally are fine, but such cases need to be studied more carefully before the rule can be made more precise.

This rule does not apply in the case of sentences (34b-d); see examples (27), p. 251, for an illustration of cases where it is relevant.

2c) Insertion rules

The insertion of the clitic cluster into a hosting constituent means that the cluster is put after the first stressed wordform of this constituent.

1. If a cluster is positioned after a constituent C_{INSERT!} and contains a clitic form of the copula BITI ^(be),

then it must be inserted into this constituent.

2. If a cluster is positioned after a constituent C_{INSERT} , then it can be inserted into it.

See Section 3.3.4, p. 252ff, for examples of insertion.

In our case, there is one insertable host: *U njihovim kućama*, [poslednjih su se dana] javila... (If the first constituent were not detached, it would be insertable: [*U njihovim su se kućama*] poslednjih dana javila...); the insertion is optional, since the cluster does not contain the clitic form of the copula and the hosting constituent is not marked as "Insert!" (i.e., does not have the syntactic role of the predicative attribute).

3) Preference rules for the placement of the clitic cluster

These rules select the actual landing site for the clitic cluster, if two or more possible landing sites have been identified in the Partial DMorphR of the clause. More specific-ally, they take care of the following two operations: 1) skipping of an optional host, and 2) optional insertion of the cluster into an insertable host.

Preferences for optional skipping

Only a rule for skipping heavy constituents will be given. (Contrastive constituents can be skipped, too, as shown in (21c-d), p. 246, and (40), p. 266, but I still haven't worked out the skipping conditions that apply to them.)
Preference rule for skipping of heavy constituents

If the skippable host $C_{HOST.SKIP}$ is relatively heavy (= contains more than three stressed wordforms),

then the skipping of $C_{HOST,SKIP}$ is preferred.

See Section 3.3.3, p. 250 for examples of preferred skipping.

In our case, there is only one skippable host that actually can be skipped: the second constituent of the clause. Since this is not a relatively heavy constituent, skipping preference is not strong; thus, sentence (34b), with skipping, and sentence (34d), without skipping, are equally good (p. 257).

Optional skipping of a host must not interfere with the No-clause-final-position rule (cf. 2b, p. 268). Thus, if the constituent preceding the clause-final one is a skippable host, it nevertheless should not be skipped; see example (27a), p. 251.

Preferences for optional insertion

Only a preference rule for optional insertion of the clitic cluster into a subordinate phrase will be given (insertion into a coordinate phrase being ignored).

Preference rule for optional insertion

- If the constituent C_{INSERT} after which the clitic cluster is positioned a) is of the form "Interrogative Premodifier+X" or b) carries communicative load, then prefer the insertion.
- 2. If the constituent C_{INSERT} after which the clitic cluster is positioned a) is of the form "Interrogative Premodifier+X+Postmodifier" or b) is relatively heavy, then avoid the insertion.

See p. 254 for examples and discussion.

With our sample sentences, there is no strong preference for insertion or against it; thus, sentence (34c), without insertion, and sentence (34d), with insertion, are equally good.

5 Conclusion

The overall picture of Serbian 2P clitics placement emerging from the present description is as follows. Unlike full-fledged words, Serbian 2P clitics are not linearly positioned with respect to their syntactic governors. That is, the host of a clitic in the morphological structure of a clause need not (and often does not) correspond to the governor of this clitic in the syntactic structure. Understandably enough, since all the clitics of the same clause are linearly positioned together, as a cluster, while in the syntactic structure they have different governors or are governors themselves (a finite auxiliary verb is considered to be the syntactic head of its clause). This, however, does not mean that syntactic dependencies are not relevant for Serbian 2P clitic placement. Even

though they are not explicitly mentioned in the clitic cluster building and positioning rules, syntactic dependencies, together with the communicative oppositions, play a crucial role in determining constituents and their properties necessary for clitic placement. In fact, syntactic dependencies and communicative oppositions underlie all the properties of constituents to which clitic placement is sensitive, including also their prosodic properties; thus, a prosodic break induced by a constituent or its prosodic contour is a marker of this constituent's specific syntactic/communicative role, and the prosodic heaviness of a constituent a consequence of its syntactic composition. The present account of Serbian 2P clitics placement is, then, in terms of the previously mentioned syntax vs. prosody debate raging in recent literature on the topic,²⁷ unambiguously a syntactic one.

This is not to say that certain generalizations in terms of prosody are not possible. In point of fact, many different syntactic/communicative underlying factors can be reduced to the same prosodic expression; this is what I tried to model by the features \pm heavy, \pm detached and \pm contrastive, used to determine the hosting capability of constituents. However, these features are not sufficient for the task: first, they play no role in the fixed clitic placement, where only syntactic considerations are relevant (namely, the syntactic class of a constituent), and, second, they can be overridden by syntactic factors in the variable placement. Here are some cases in which, in the variable clitic placement, syntax "tramps" prosody.

- 1) Skipping of a contrastive constituent can be precluded by its syntactic role; such is the case of a contrastive constituent functioning as the Main Verb of the clause; cf. example (41), p. 266.
- 2) Insertion of the clitic cluster into a constituent is allowed/disallowed or obligatory as a function of its syntactic role and its syntactic composition; cf., respectively, examples (29) and (30), p. 252*ff*.
- 3) There is a syntactic constraint on hosts of clusters containing the interrogative particle L11: the host of such clusters cannot be a non-finite verb; cf. examples (22d) and (43a), p. 247 and p. 267.
- 4) Clitic pseudo-climbing (cf. p. 242) is allowed/disallowed by syntactic factors: the syntactic role of the infinitival phrase.
- 5) Some items have to be lexically marked for their hosting capability, irrespective of their characterization in terms of the above features; cf. examples (21e-f), p. 247.

Furthermore, there are instances of violation of prosodic constraints in clitic placement that indicate that in this operation syntactic factors have precedence over prosodic

²⁷ A paradigmatic example of a syntactic account of Serbian/Croatian clitic placement is Progovac 1996, where clitic landing sites are defined in terms of a structural position (Complementizer) in the syntagmatic tree to which clitics are moved (from their base-generated positions) by syntactic rules. Halpern 1995 is basically a syntactic account as well, but makes use of some prosodic factors: 2P placement (= after the first phrase) is treated as a syntactic operation, while 2W placement (= after the first word, corresponds to my insertion) is described as a result of "prosodic inversion" between the clitics and their host. A typical prosodic account can be found in Radanović-Kocić 1988 and 1996, where the landing sites for the clitics are defined in terms of prosodic units of the sentence: clitics are placed after the first phonological word or phrase) of their intonational phrase. Bošković 1995 and Hock 1996 contain similar proposals. For a review of syntactically- *vs.* prosodically-minded literature on clitic placement in Serbian/Croatian, see Ćavar 1999: 117-185.

ones. That is, in some contexts, Serbian clitics, which by all accounts should be ENclitics, actually behave as PROclitics. Consider, for example, the following sentence:

(44) $Ideje \check{C}omskog | su ga inspirisale$ lit. (Ideas of-Chomsky are_(Aux) him having-inspired) = (Chomsky's ideas inspired him).

The clitics are positioned after the first constituent, and this is the only possible landing site for them. (Although heavy, the initial constituent cannot be skipped, since this would put the clitics in an illegitimate clause-final position; this constituent is not insertable, either, as it does not have the required syntactic composition.) Thus, the clitics must be placed after the first constituent and should form a prosodic unit with it. However, if a small pause is automatically made after this constituent, the clitics, which are syntactically enclitics, prosodically procliticize to the following constituent.

Also, the language of the press has a tendency to put the clitics after a constituent in the role of a non-restrictive apposition or a parenthetical, even though such a constituent is obligatorily followed by a pause; thus, sentences such as the following one, stigmatized by the normative grammar (cf., for instance, Klajn 2000), are common occurrence in newspapers:

(45) Jedan od glavnih optuženih, pored J. Obradovića, || je smenjeni predsednik kragujevačke opštine lit. ^{(One of principal persons-charged, besides J. Obradović, is sacked mayor of-Kragujevac municipality⁾ = ^{(The sacked mayor of the municipality of Kragujevac is one of the principal persons charged, along with J. Obradović⁾.}}

Here, the clitic copula je ^(is) is actually pronounced as a proclitic to the following phrase. (According to the norm, only *jeste*, i.e., the full form of the copula, is admitted in this position.)²⁸

As a final example, consider the difference in the placement of the clitic cluster in the following sentences:

(46) a. *Poslednjih dana su se* || , *u njihovim kućama*, || *javila dva slučaja oboljenja* [-detached] [+detached]

lit. ^{(In-last days $are_{(Aux)}$ REFL, in their houses, having-appeared two cases of illness⁾ = ^{(Recently, in their households there have been two cases of illness⁾.}}

²⁸ This fact was noted in Popović 1997: 313: "In speech the so-called enclitics appear after internal pauses (pronounced as proclitics or with a separate accent)" [translation mine—JM]. Cf. also the following Czech example (Toman 1986: 125): Czech, coll. *Ten doktor, co mu duvěřuješ, se neholi* ^{(The} doctor that_(Conj) he-DAT [you_{sg}] trust REFL not-shaves³ = ^{(The} doctor, whom you trust, does not shave³. According to Toman, "[...] the phenomenon just described is accounted for naturally once we regard Czech clitics as prosodically neutral and as having the direction of clisis determined locally within particular prosodic context" (ibid.: 127). It should be noted, however, that the parallel with Czech does not hold quite well, since in standard Czech, but not in standard Serbian, clitics can occupy the clause-initial position, where their proclitic status is obvious. (In Serbian, examples of clitics in the clause-initial position are found only in colloquial speech; cf. coll. *Si normalan?* ^{(Are}(Copula)</sup> [you_{sg}] normal?³; coll. *Sam ti rek'o!* ^{(Am}(Aux)</sup> to-you_{sg} having-told!³ = ^(Haven't I told you so!³)

b. **Poslednjih dana su se*, || *u njihovim kućama javila dva slučaja oboljenja* [+detached] [-detached] lit. ⁽In-last days are_(Aux) REFL, in their houses having-appeared two cases of illness⁾ = ⁽Recently, in their households there have been two cases of illness⁾.

Since in (46a), the pause after the circumstantial *poslednjih dana* is not induced by this constituent, but rather by the following parenthetical, *u njihovim kućama*, the circumstantial can host the clitic cluster. In (46b), however, the pause after *poslednjih dana* is induced by the circumstantial itself, so it cannot host the clitics. This means that, in order to properly position the clitics in the morphological string, it is not sufficient to know just where the pauses are—it is necessary to make reference to the specific constituents which induce them (which is done in the present description by means of the three features specifying the relevant properties of constituents).

The evidence adduced above suggests that the crucial role in Serbian 2P clitics placement is played by syntactic dependencies and communicative oppositions—they are expressed, among other things, by prosodies, but cannot be completely reduced to prosodies.

Acknowledgments

I would like to thank the following people for their insightful comments on preliminary versions of the paper: Wayles Browne, Lidija Iordanskaja, Sylvain Kahane, Igor Mel'čuk and Alain Polguère. I am especially indebted to Wayles Browne and Igor Mel'čuk.

References

Anderson, Stephen 2005. Aspects of the Theory of Clitics. Oxford: Oxford University Press.

Borer, Hagit (ed.) 1986. *The Syntax of Pronominal Clitics* [Syntax and Semantics 19]. Orlando/San Diego, etc.: Academic Press.

Bošković, Željko 1995. "Participle Movement and Second-Position Cliticization in Serbo-Croatian". *Lingua* 96: 245-266.

Browne, Wayles 1974. "On the Problem of Enclitic Placement in Serbo-Croatian". In *Slavic Transformational Syntax, Michigan Slavic Materials* 10, R. Brecht & C. Chvany (eds.), 36-52. Ann Arbor, MI: The University of Michigan Press.

Browne, Wayles 1975. "Serbo-Croatian Enclitics for English-Speaking Learners". *Kontrastivna analiza engleskog i hrvatskog ili srpskog jezika*, 105-134. Zagreb: Institut za lingvistiku Filozofskog fakulteta.

Browne, Wayles 1976. "Two WH-fronting Rules in Serbo-Croatian". Južnoslovenski filolog 32: 195-204.

Chafe, Wallace 1994. Discourse, Consciousness and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing. Chicago/London: The University of Chicago Press.

Ćavar, Damir 1999. Aspects of the Syntax-Phonology Interface. Cliticization and Related Phenomena in Croatian [Ph.D. thesis]. University of Potsdam.

Čamdžić, Aneta & Hudson, Richard 2002. "Serbo-Croat-Bosnian Clitics and Word Grammar". *UCL Working Papers in Linguistics* 14: 321-353.

Dixon, Robert & Aikhenvald, Alexandra (eds.) 2002. *Word. A Cross-linguistic Typology*. Cambridge etc.: Cambridge University Press.

Franks, Steven & Holloway King, Tracy 2000. *A Handbook of Slavic Clitics*. New York/Oxford: Oxford University Press.

Gerdes, Kim & Kahane, Sylvain. 2007. "Phrasing it Differently". In *Selected Lexical* and Grammatical Issues in the Meaning-Text Theory. In Honor of Igor Mel'čuk, L. Wanner (ed.), 297-335. Amsterdam/Philadelphia: Benjamins.

Halpern, Aaron 1995. On the Placement and Morphology of Clitics. Stanford, CA: CSLI.

Halpern, Aaron 1999. "Clitics". In *The Handbook of Morphology*, A. Spencer & A. Zwicky (eds.), 101-123. Oxford: Blackwell.

Halpern, Aaron & Zwicky, Arnold (eds.) 1996. *Approaching Second: Second Position Clitics and Related Phenomena*. Stanford, CA: CSLI.

Hock, Hans-Heinrich 1996. "Who's on first? Toward a Prosodic Account of P2 Clitics". In Halpern & Zwicky (eds.) 1996: 199-270.

Hudson, Richard 1980. "Constituency and Dependency". Linguistics 18 (3/4): 179-198.

Iordanskaja, Lidija 1982. "Le placement linéaire des clitiques pronominaux en français contemporain". *Lingvisticae Investigationes* 6 (1): 145-188.

Iordanskaja, Lidija & Mel'čuk, Igor 2000. "The Notion of Surface-Syntactic Relation Revisited (Valence-Controlled Surface-Syntactic Relations in French)". In *Slovo v tekste i v slovare. Sbornik statej k semidesjatiletiju Ju.D. Apresjana.* L.L. Iomdin & L.P. Krysin (eds.), 391-433. Moskva: Jazyki russkoj kul'tury.

Jackendoff, Ray 1977. X-bar Syntax: A Study of Phrase Structure. Cambridge, MA: MIT Press.

Kahane, Sylvain 2003. "The Meaning-Text Theory". In Dependency and Valency. An International Handbook of Contemporary Research, vol. 1, Agel, V., Eichinger L.M.,

Eroms, H-W., Hellwig, P., Heringer, H.J. & Lobin, H. (eds.), 546-570. Berlin/New York: De Gruyter.

Kaiser, Lizianne (ed.) 1997. Yale A-Morphous Linguistics Essays. Studies in the Morphosyntax of Clitics. New Haven, CT: Yale University.

Klajn, Ivan 2000. "Muke sa enklitikom" [Troubles with the Enclitic]. *Stranputice smisla*, 132-133. Beograd: NIN.

Klavans, Judith 1995. On Clitics and Cliticization: The Interaction of Morphology, Phonology and Syntax. New York/London: Garden Publishing Inc.

Lambrecht, Knud 1994. Information Structure and Sentence Form. Topic, Focus and the Mental Representation of Discourse Referents. Cambridge: Cambridge University Press.

Mel'čuk, Igor 1967. "Ordre des mots en synthèse automatique des textes russes". *T.A. Informations* 8 (2): 65-84.

Mel'čuk, Igor 1974. *Opyt teorii lingvističeskix modelej Smysl* \Leftrightarrow *Tekst* [Outline of a Theory of 'Meaning-Text' Linguistic Models]. Moskva: Nauka. [Reprinted in 1999, Moskva: Jazyki russkoj kul'tury]

Mel'čuk, Igor 1986. "Semantic Bases of Linguistic Description (Meaning-Text Linguistic Theory)". *The 12th LACUS Forum 1985*: 41-87.

Mel'čuk, Igor 1988. *Dependency Syntax: Theory and Practice*. Albany, N.Y.: State University of New York Press.

Mel'čuk, Igor 1993. "The Inflectional Category of Voice: Towards a More Rigorous Definition". In *Causatives and Transitivity*, B. Comrie & M. Polinsky (eds.), 1-46. Amsterdam/Philadelphia: Benjamins.

Mel'čuk, Igor 2000. *Cours de morphologie générale*, vol. 5. Montréal/Paris: Les Presses de l'Université de Montréal/C.N.R.S. Éditions.

Mel'čuk, Igor 2001. *Communicative Organization in Natural Language*. Amsterdam/ Philadelphia: Benjamins.

Mel'čuk, Igor 2003. "Levels of Dependency in Linguistic Description: Concepts and Problems". In *Dependency and Valency. An International Handbook of Contemporary Research, vol. 1*, Agel, V., Eichinger L.M., Eroms, H-W., Hellwig, P., Heringer, H.J. & Lobin, H. (eds.), 188-229. Berlin/New York: De Gruyter.

Mel'čuk, Igor 2006. *Aspects of the Theory of Morphology*. Berlin/New York: Mouton de Gruyter.

Mel'čuk, Igor & Pertsov, Nikolaj 1987. Surface Syntax of English. A Formal Model within the Meaning-Text Framework. Amsterdam/Philadelphia: Benjamins.

Miller, Phillip & Monachesi, Paola 2003. "Les pronoms clitiques dans les langues romanes". In *Les langues romanes, problèmes de la phrase simple*, Godard, D. (ed.), 67-123. Paris: CNRS Éditions.

Milićević, Jasmina 1999. "Pronominal and Verbal Clitics in Serbian: A Morphological Description". *Wiener Slawistischer Almanach* 43: 231-256.

Milićević, Jasmina 2005. "Clitics or Affixes? On the Morphological Status of the Future Tense Markers in Serbian'. In *Morphology and its Demarcations. Selected Papers from the 11th Morphology Meeting, Vienna, February 2004*, W. U. Dressler, D. Kastovsky, O. E. Pfeiffer & F. Rainer (eds.), 39-52. Amsterdam/Philadelphia: Benjamins.

Milićević, Jasmina 2007. "Co-occurrence of Serbian Second-Position Clitics. Syntactic and Morphonological Constraints". In *Selected Papers from the 5th International Morphology Meeting Décembrettes 05, December 2-4, 2005, Toulouse*, 99-120. München: Lincom Europa.

Milićević, Jasmina (to appear). Serbian Auxiliary Verbs—Syntactic Heads or Dependents? In *Proceedings of the The 31st Annual Conference of the Atlantic Provinces Linguistics Association (APLA'31), November 2-3, 2007, Fredericton.*

Popović, Ljubomir 1997. *Red reči u rečenici* [Word Order in the Sentence]. Beograd: Društvo za srpski jezik i književnost Srbije.

Progovac, Ljiljana 1996. "Clitics in Serbian/Croatian: Comp as the Second Position". In Halpern & Zwicky (eds.) 1996: 411-428.

Progovac, Ljiljana 2005. *A Syntax of Serbian. Clausal Architecture*. Bloomington, IN: Slavica Publishers.

Radanović-Kocić, Vesna 1988. *The Grammar of Serbo-Croatian Clitics: A Synchronic and Diachronic Perspective* [Ph.D. Thesis]. University of Illinois at Urbana-Champaign.

Radanović-Kocić, Vesna 1996. "The Placement of Serbo-Croatian Clitics: A Prosodic Approach". In Halpern & Zwicky (eds.) 1996: 429-446.

Rambow, Owen & Joshi, Aravind 1997 "A Formal Look at Dependency Grammars and Phrase-Structure Grammars, with Special Consideration of Word-Order Phenomena". In *Recent Trends in Meaning-Text Theory*, L. Wanner (ed.), 167-190. Amsterdam/ Philadelphia: Benjamins.

Sgall, Peter, Hajičová, Eva & Panevová, Jarmila 1986. *The Meaning of the Sentence in its Semantic and Pragmatic Aspects*. Dordrecht: Reidel.

Spencer, Andrew 1991. Morphological Theory: An Introduction to Word Structure in Generative Grammar. Oxford: Blackwell.

Steele, John (ed.) 1990. *Meaning-Text Theory: Linguistics, Lexicography and Implications*. Ottawa, etc.: University of Ottawa Press. Toman, Jindřich 1986. "Cliticization from NPs in Czech and Comparable Phenomena in French and Italian". In Borer (ed.) 1986: 123-145.

Zec, Draga & Inkelas, Sharon 1990. *The Phonology-Syntax Connection*. Chicago: The University of Chicago Press.

Zwicky, Arnold 1977. On Clitics. Bloomington, IN: Indiana University Linguistics Club.

Žolkovskij, Aleksandr & Mel'čuk, Igor 1967. "O semantičeskom sinteze [On Semantic Synthesis]". *Problemy kibernetiki* 19: 177-238. [French translation: "Sur la synthèse sémantique" (1970). *TA Informations* 2: 1-85.]

General Index: Notions, Names, Languages

2P [= second-position] 237

A

Abkhaz [language] 64 absolute head (of a syntactic structure) 44 absolute host (of a clitic) 244 accusativus-cum-infinitivo construction 176, 203 actant xviii, xx, 6, 13, 18, 43, 49, 112, 116, 118, 120, 124 ~s vs. circumstantials 50 deep-syntactic ~ 49, 153, 157 semantic ~ 10, 21, 49 syntactic ~ 49, 82, 113 actantial-attributive complement 208 agent 192 agreement 16 \sim class 16 Akhvakh [language] 17 Albanian [language] 38 Alutor [language] 61 analytical \sim exponent of a grammeme 77, 78 ~ form 91 ~ verbal form 79 Arab grammarians (on dependencies) 22 Arabic [language] 64 argument (of a predicate) 3

B

basic diathesis 84 Bloomfield, L. 51 branch (of a syntactic structure) 6 bubble (in a syntactic structure) 92, 94 Bushoong [language] 21

С

causative FAIRE-construction 161, 177, 181, 183, 185, 188, 203, 219, 229, 230 centrifugal (language) 84 centripetal (language) 84 Chechen-Ingush [language] 89 Chinese [language] 31 Chukchee [language] 43, 44, 219, 220 circumstantial 213, 217 ~ of measure 213, 215

actants vs. ~s 50 clause 3 ~ element 153, 154 clause-final position 250, 268 clefting pseudo-~ 158 climbing 243 clitic pseudo-~ 242 clitic 78, 229, 235 ~ climbing 243 ~ cluster 239, 264, 267, 268, 271 ~ cluster template 239 ~ co-occurrence 241 ~ pseudo-climbing 242 \sim raising 62 ~, "neuter" Fr. LE 198, 199, 206, 207 ~, direct reflexive Fr. sE 185, 190 ~, DirO 178 ~, enclitic vs. proclitic 271 ~, Fr. EN 180, 186, 214, 229 ~, Fr. y 214 ~, indirect reflexive Fr. SE 190 ~, its syntactic governor 189, 237 ~-attracting verb 203, 229 ~s, Serbian, their syntactic roles 256 fixed ~ placement 245 host of a ~ 189, 229, 235 linearization of ~s 264 possessor's dative ~ 242 resumptive $\sim 38, 166$ second-position ~ 235 variable ~ placement 245, 246 cliticization 239 ~ rules 260 collocation 84, 91, 92 communicative specifier 247 comparate 216 complement copular-attributive ~ 205 predicative-attributive ~ 212 complementation 49 compound numeral 32 congruence 20 conjoined nominal phrase 80 conjunct 52 constituency 89 constituent 244, 248, 261 ~ building rules 263

~ linear arrangement rules 263 hosting ~ 244 immediate ~ analysis 117 construction (syntactic) 151 accusativus-cum-infinitivo ~ 176, 203 causative FAIRE-~ 177, 181, 183, 185, 188, 203, 219, 229, 230 elliptical ~ 82 endocentric ~ 52 exocentric ~ 52 coocurrence control 42 coordinability of two sentence elements 98 coordinate expression 71 coordination 50, 92 copredicative-infinitival object 202 copular verb 63, 205 copular-attributive complement 205 coreferentiality 9 criterion (for SSyntRels) repeatability ~ 166 semantic contrast ~ 162 syntactic substitutability ~ 163

D

D [= dependent (syntactic)] 151 dative ethical ~ 240 free ~ 240 possessor's ~ 240 dativus ethicus [= ethical dative] 189 deep vs. surface distinction xvii, 5 defining properties (of a SSyntRel) 159 dependency 5 ~ grammar 111, 112 ~ grammar vs. ~ representation 22 ~ syntactic structure xiii ~ vs. constituency 89 "double ~" 67 "insufficient ~" 73 "mutual ~" 72 "no ~" 72 morphological ~ 8, 12 semantic $\sim 8, 10$ syntactic ~ 6, 8, 21, 40 syntactic ~, properties of 41 dependent (syntactic) xiv, 9, 112, 151 prototypical ~ (of a given SSyntRel) 159, 164, 169 descriptive properties (of a SSyntRel) 160 determiner 77 DG [= dependency grammar] 111 diathesis basic ~ 84 direct object 4, 176, 186 direct speech 219 direct-infinitival object 198 DirO [= Direct Object] 4 dislocation

left ~ 158 DMorph-ProsS [= Deep-Morphological Prosodic Structure] 7 DMorphS [= Deep-Morphological Structure] 7 double accusative 97 double dependency 67 DSyntRel [= Deep-Syntactic Relation] 6 DSyntS [= Deep-Syntactic Structure] 6 Dutch [language] 30

E

ellipsis 88 elliptical construction 82 endocentric construction 52 English [language] 5, 11, 20, 35, 38, 39, 42, 51, 52, 70, 77, 78, 82, 86, 89 ethical dative 240 exocentric construction 52 expression coordinate ~ 71 expressive means (of natural language) 24 extraction 91, 111 wh~ 133, 136

F

feature syntactic ~ 18 fixed clitic placement 245 floating quantifier 59 free dative 240 French [language] 20, 28, 38, 48, 62, 64, 78, 86, 96, 98, 99, 100, 128, 129, 141, 167

G

G [= governor (syntactic)] 151 Georgian [language] 18, 219, 220 German [language] 30, 48, 97 government 18 non-specific ~ 157 governor (syntactic) xiv, 9, 112, 151 ~ of a clitic 189, 237 grammar dependency ~ 111, 112 phrase structure ~ 111 grammeme 4 grouping (in a syntactic structure) 94, 100

H

Hawaiian [language] 77 head (syntactic) xiv, 9, 112 ~ Feature Principle 117 absolute ~ (of a syntactic structure) 44 headless relative 70, 137 Hindi [language] 19 host see *hosting constituent* 244 host (of a clitic) 189, 229, 235 ~-identifying rules 265 absolute ~ 244 potential ~ 244 skippable ~ 248 hosting constituent 244, 246 HPSG [= head-driven phrase structure grammar] 111 Hungarian [language] 19

I

idiom 91 immediate constituent analysis 117 impersonal ~ verb 173 impersonalization 170, 174, 183 incorporability 43 indirect object 4, 186 ~ vs. oblique object 196 IndirO [= Indirect Object] 4 inflectional category 4, 16 insertion (of the clitic cluster) 245, 252, 268 insufficient dependency 73 intensifier 84 island constraint 127 izafa suffix 64

J

Jakobson, R. 21 Japanese [language] 59, 62, 85

K

Kayardild [language] 17 Keenan, E. 155 Keenan's principle 66 Keenan-Comrie hierarchy (of clause elements) 156, 220 Kinyarwanda [language] 96 Korean [language] 40, 97 Kunze property 35, 163 quasi-~ 36, 164

L

Latin [language] 63, 76, 87, 97, 99 ~ preposition CIRCA 76 left dislocation 158 lexical function 84, 92 Lezgian [language] 45 light verb [see *support verb*] 91 linearization ~ of non-clitic elements 263 linearization rules 261 locality principle 130 Lushootseed [language] 47, 61, 99

Μ

Maasai [language] 60 Macchiavelli, N. 44 measure noun 60, 63, 74 minimal pair test 162 modification 49 modifier 112, 118, 120 Morph-**D** [= morphological dependency] 8 morphological ~ contact point 29 ~ dependency 8, 12 mutual dependency 72 mutual substitutability 166

Ν

Nias [language] 73 no dependency 72 non-configurationality 88 noun measure ~ 60, 63, 74 null transferer 128 numeral ~ phrase 74 compound ~ 32 Nunggubuyu [language] 89

0

object copredicative-infinitival ~ 202 direct ~ 4, 176, 186 direct-infinitival ~ 198 indirect ~ 4, 186 indirect ~ vs. oblique ~ 196 oblique ~ 194 oblique-infinitival ~ 199 pseudo-direct ~ 190 quotative ~ 218 oblique object 194 indirect object vs. ~ 196 oblique-infinitival object 199 Old Georgian [language] 17 omissibility (of a sentence element) 42

P

part of speech ~ (according to L. Tesnière) 122 passivization 183 Persian [language] 39, 64 phasic verb 63 phrase 3 conjoined nominal ~ 80 minimal ~ xiv numeral ~ 74 status of ~ 111, 142 phrase structure grammar 111 possessor's dative clitic 240, 242 potential host (of a clitic) 244 predicative-attributive complement 212 principle "Always-the-same-SSyntS" ~ 70 "As-Little-DSynt⇒SSynt-Restructuring-as-Possible" ~ 70 Head Feature ~ 117 locality ~ 130 projectivity 85 pronoun 260 relative ~ 67 prototypical dependent 164, 169 pseudo-clefting 158 pseudo-direct object 190 PSG [= phrase structure grammar] 111

Q

quasi-Kunze property 36 *qui-que* alternation 128 quotative object 218

R

raising 71 clitic ~ 62 realization verb 84 reflexive adjunct 239 relation deep-syntactic 6 surface-syntactic ~ 6, 114, 127, 141, 151 surface-syntactic ~s, degree of obliqueness 121 surface-syntactic ~s, relevant properties of 156 surface-syntactic ~s, their inheritance 142 surface-syntactic ~s, their inventory 155 relative headless ~ 70, 137 relative pronoun 67 relativizer 125 repeatability (of syntactic roles) 37, 167 repeatability criterion 166 representation sentence ~ 4 restricted lexical cooccurrence 84 resumptive clitic 38 Rheme ~ Theme division 257 Russian [language] 19, 30, 46, 48, 51, 59, 63, 74, 76, 89, 98 ~ preposition OKOLO 76 coordination of pronouns in ~ 81

S

Sanskrit [language] 97 satellite 10 secondary actor 177, 181, 185 second-position clitic 235 semanteme 5 semantic

~ dependency 10 \sim name 3 ~ predicate 3 ~ Structure 5 semantic contrast criterion 162 Sem-**D** [= semantic dependency] 8 semi-auxiliary ~, Fr. SE FAIRE 203 ~, Fr. SE VOIR 203 SemS [= Semantic Structure] 5 sentence 3 \sim representation 4 Serbian [language] 87, 97, 235 skipping 248, 249 slashed element 126 Southern Tiwa [language] 43 Spanish [language] 38, 62, 167 SSyntRel [= Surface-Syntactic Relation] 6 SSyntS [= Surface-Syntactic Structure] 6 structure central [= bearing] ~ (of a linguistic representation) 5 Deep-Morphological ~ 7 Deep-Morphological Prosodic ~ 7 Deep-Syntactic ~ 6 Semantic ~ 5 Surface-Syntactic ~ 6 subject 170 quasi-~ 170, 174 subjectless suppressive (voice) 242 substitution test 165 support verb [see light verb] 84 syntactic ~ construction 151 ~ dependency 21, 40 ~ feature 4, 18, 95 ~ governor 9 ~ group 92 \sim head 9 ~ relation xiii, xiv ~ structure xiii ~ substitutability 35 syntactic substitutability criterion 163 syntactics 4, 260 Synt-**D** [= syntactic dependency] 8

Т

Tabassaran [language] 60 Tagalog [language] 77 template (of a clitic cluster) 239 Tesnière, L. 100, 113, 122 test (for -emic units) cooccurrence ~ 168 minimal pair ~ 162 substitution ~ 165 *that*-trace effect 128 Tiwa [language] Southern ~ 43 top node xiv, 151 transferer 124, 130 null ~ 128 Turkish [language] 47, 99

U

unit -emic ~ 152, 162 utterance 2

V

```
valence
active syntactic ~ 151
passive syntactic ~ 4, 27, 95, 151
value (of a syntactic feature) 4
variable clitic placement 245, 246
verb
copular ~ 63
impersonal ~ 173
```

```
light ~ [see support verb] 91
phasic ~ 63
realization ~ 84
support ~ [see light verb] 84
verbless sentence 46
voice 84
subjectless suppressive ~ 242
```

W

Welsh [language] 85 wh-extraction 133, 136 word combination operation 120 word order 84 wordform 3

Z

zero copula 99 zero verb form 46