

Institut für Sprachwissenschaft

Universität zu Köln

Topics, Presuppositions, and Theticity:
An Empirical Study of Verb-Subject Clauses
in Albanian, Greek, and Serbo-Croat

Inauguraldissertation

zur

Erlangung des Doktorgrades

vorgelegt der

Philosophischen Fakultät der Universität zu Köln

von

Dejan Matic

April 2003

Erstgutachter: Prof. Dr. Hans-Jürgen Sasse

Zweitgutachter: Prof. Dr. Fritz Serzisko

Contest

1. Introduction	7
2. State of the art	10
2.1. Lexical properties	11
2.2. Given and new	14
2.2.1. Grades of newness	15
2.2.2. All-new utterances	19
2.3. Discourse functions	23
2.4. Semantic integration	27
2.5. Theticity research and related approaches	29
2.5.1. Ancient history: Brentano and Marty	29
2.5.2. Conceptualist approaches to theticity	33
2.5.3. Semantics and theticity	36
2.5.4. Discourse-pragmatic reinterpretation of theticity	38
2.6. Topic and focus: presupposition/assertion-based approach	43
2.6.1. New topics	44
2.6.2. Sentence focus	46
2.6.3. Spatio-temporal argument as topic	52
2.7. Skeptics	57
Appendix: VS order in generative grammar	60
3. Aims and announcements	65
3.1. Forms and functions	65
3.2. VsX-construction	68
3.3. Inversion	70
3.4. vS-construction	71
4. Information structure: presupposition, assertion, topic and focus	76
4.1. The incremental model of communication	76
4.2. Presupposition and assertion	78
4.2.1. Basic issues	78
4.2.2. Two notes on presupposition	87
4.2.3. Two kinds of presuppositions	88
4.2.4. Sources of presuppositions	90
4.2.5. Definitions	94
4.3. Topic	95
4.3.1. Basic issues	95
4.3.2. Discourse properties of topics	102
4.3.3. Some formal properties of topics	106
4.3.4. Definitions	109
4.4. Focus	110
4.4.1. Basic issues	110
4.4.2. Types of foci	115

4.4.3. Some formal properties of focus	121
4.4.4. Definition	126
4.5. Terminological issues	126
4.5.1. Four levels	126
4.5.2. Focus domains and indirect topics	129
4.6. Why is information structure formally marked?	132
5. Preliminaries	136
5.1. Preliminary notes on Albanian, Modern Greek and Serbo-Croat	136
5.2. Heaviness and cataphoric relevance	141
5.3. Some technical details	143
5.4. Corpus	145
5.5. Preliminary statistics	148
5.5.1. General statistics and cross-linguistic variation	148
5.5.2. Intralinguistic variation	151
5.5.3. Clause types and VS order	152
5.5.4. Summary	156
6. VsX construction	157
6.1. Statistics	158
6.2. Ratified topics, discourse frames, and the Balkan syntax	160
6.3. Formal properties of VsX construction and information structure	163
6.3.1. Length and intonation	163
6.3.2. Elements of the construction	165
6.3.2.1. Verb and preverbal elements	165
6.3.2.2. Subject and other postverbal elements	166
6.3.2.2.1. Serbo-Croat	167
6.3.2.2.2. Albanian and Modern Greek	168
6.3.2.3. Complex predicate forms	171
6.3.2.4. Formal properties and information structure: A summary	173
6.4. Semantic and informational features of subjects and verbs	174
6.4.1. Subject	174
6.4.2. Verb	179
6.5. Discourse functions of VsX construction	182
6.5.1. Polarity focus	183
6.5.1.1. Polarity focus in illocutionary contexts	184
6.5.1.2. Polarity focus as a text-cohesion device	191
6.5.2. Focus domain: broad and narrow foci	193
6.5.3. Discourse functions of VsX construction: Summary and conclusion	203
6.6. Non-structural factors influencing the use of VsX	205
6.6.1. Albanian and Serbo-Croat	206
6.6.2. Modern Greek	209
6.7. VsX: Conclusion	210

7. Inversion: Introduction	216
8. Focal Inversion	218
8.1. Quotation Inversion	218
8.1.1. Syntax and prosody	220
8.1.1.1. Basic features	220
8.1.1.2. Elements of the construction	221
8.1.1.2.1. General tendencies	221
8.1.1.2.2. Independent elements	223
8.1.1.2.3. Dependent elements	224
8.1.1.2.4. Lexical solidarities	227
8.1.1.2.5. Elements of the construction: Summary	230
8.1.1.3. Complex predicate forms and complex sentences	232
8.1.1.4. Summary	235
8.1.2. Information structure and discourse function	236
8.1.3. Semantics of quotation inversion and the syntactic status of the quote	241
8.1.4. Alternatives to quotation inversion	246
8.1.5. Discourse-pragmatic differences between Alb., MG, and SC	249
8.1.6. Conclusion	250
8.2. Wh-inversion	251
8.2.1. The pragmatic role of question words	252
8.2.2. Syntax, prosody, and information structure	256
8.2.2.1. General rules	256
8.2.2.2. All-ratified constituent questions	258
8.2.2.3. Constituent questions with non-ratified elements	262
8.2.2.4. Complex predicate forms and complex sentences	266
8.2.2.5. Ratified and non-ratified, short and long: A summary	269
8.2.3. Alternatives to wh-inversion	272
8.2.4. Discourse-pragmatic differences between Alb., MG, and SC	274
8.2.5. Conclusion	275
8.3. Fronted focus inversion (ff-inversion)	276
8.3.1. Syntax, prosody, and information structure	277
8.3.2. Meaning of fronted narrow focus expressions	280
8.3.2.1. Semantic and informational features of fronted narrow foci	281
8.3.2.2. Discourse functions of fronted narrow foci	287
8.3.2.3. Summary	298
8.3.3. Alternatives to ff-inversion	299
8.3.4. Conclusion	301
9. Subordinate inversion	303
9.1. Relative inversion	303
9.1.1. Balkan relative clauses	304
9.1.2. Statistics	306
9.1.3. Adjacency constraint and restrictive vs. appositive relative clauses	307
9.1.4. Information structure, restrictive relative clauses, and relative inversion	311
9.1.5. Syntax, prosody, and information structure of relative inversion	313
9.1.5.1. General rules	313
9.1.5.2. All-ratified restrictive relative clauses	314
9.1.5.3. Restrictive relative clauses with a non-ratified element	317
9.1.5.4. Summary	318

9.1.6. Semantic and informational properties of subjects and verbs	319
9.1.7. Alternatives to relative inversion	321
9.1.8. Non-structural differences between Alb., MG, and SC	322
9.1.9. Conclusion	323
9.2. Embedded inversion	323
9.2.1. Balkan embedded clauses	324
9.2.2. Embedded clauses and presuppositions	326
9.2.3. Statistics	330
9.2.4. Adjacency constraint	331
9.2.5. Syntax, prosody, and information structure of embedded inversion	335
9.2.5.1. Clauses with the feature [-adjacent]	336
9.2.5.2. Clauses with the feature [+adjacent]	339
9.2.5.3. Summary	340
9.2.6. Semantic and informational properties of subjects and verbs	341
9.2.7. Alternatives to embedded inversion	342
9.2.8. Conclusion	345
9.3. Generalization patterns	345
9.3.1. Formal and functional explanations of frequency phenomena	346
9.3.2. Generalization patterns	347
9.3.2.1. VS structure	347
9.3.2.1.1. Superfluous inversion	347
9.3.2.1.2. Excursus: Some paths of generalization	351
9.3.2.1.3. An interrupted process	353
9.3.2.2. SV structure in SC	355
10. Inversion: Conclusion and summary	357
10.1. Discourse pragmatics	357
10.1.1. Presuppositional contexts	357
10.1.2. Ratification	359
10.2. Formal restrictions	362
10.3. Syntactic pattern, deviations, and configurationality	364
11. vS Construction	366
11.1. Statistics	368
11.2. Formal properties and information structure	372
11.2.1. Basic features: Intonation, length, complex predicates and complex clauses	372
11.2.2. Topical elements	375
11.2.2.1. Clitics	375
11.2.2.2. Preverbal topic expressions	378
11.2.2.2.1. Objects as preverbal topic expressions	378
11.2.2.2.2. Free datives	379
11.2.2.2.3. Setting adverbials	379
11.2.2.2.4. Predicatives and manner adverbs	380
11.2.2.3. Topic expressions between the verb and the subject	382
11.2.2.3.1. Objects and free datives	382
11.2.2.3.2. Setting adverbials	383
11.2.2.4. Topic expressions after the subject	384
11.2.2.5. Clauses without clausal topic expressions	385
11.2.2.6. Topical elements: A summary	387

11.2.3. Further elements of the focus domain _____	388
11.2.3.1. Modal adverbials _____	389
11.2.3.2. Preverbal elements of the focus domain _____	389
11.2.3.3. Elements of the focus domain between the verb and the subject _____	390
11.2.3.4. Elements of the focus domain after the subject _____	392
11.2.3.5. Further elements of the focus domain: A summary _____	394
11.2.4. Formal properties and information structure: Conclusion _____	395
11.3. Semantic and informational properties of the elements of vS _____	396
11.3.1. Semantic and informational properties of topic expressions _____	396
11.3.2. Semantic and informational properties of the subject _____	401
11.3.3. Semantic and informational properties of the verb _____	405
11.3.3.1. Basic features _____	405
11.3.3.2. Meanings in context _____	410
11.3.3.2.1. Lexical reinterpretation _____	410
11.3.3.2.2. Lexical solidarities _____	416
11.3.3.3. Semantic and informational properties of the verb: A summary _____	420
11.4. Semantic/informational properties and focus construal _____	421
11.4.1. Prototypical verbs and false generalizations _____	421
11.4.2. Lexical and informational features triggering narrow focus construal _____	424
11.4.3. Lexical solidarities and focus construal _____	430
11.5. Discourse functions of vS clauses _____	432
11.5.1. vS clauses with narrow focus construal _____	432
11.5.1.1. Fronted focused subjects and vS _____	432
11.5.1.2. The narrower domain of narrow foci in vS _____	433
11.5.2. vS clauses with broad focus construal _____	437
11.5.2.1. vS clauses with locational topics _____	437
11.5.2.1.1. Descriptive function _____	438
11.5.2.1.2. Introductory function _____	443
11.5.2.2. vS clauses with situational topics _____	447
11.5.2.2.1. Annuntiative function _____	448
11.5.2.2.2. Interruptive function _____	451
11.5.2.2.3. Explanatory function _____	455
11.6. Three essential properties of vS clauses: An attempt of interpretation _____	459
11.6.1. Untypical topics _____	459
11.6.1.1. Locational topics _____	459
11.6.1.2. Situational topics _____	463
Excursus: Am I re-inventing the spatio-temporal argument? _____	466
11.6.1.3. Other untypical topics _____	467
11.6.2. Untypical subjects _____	469
11.7. Theticity and vS _____	471
Sources _____	475
References _____	477

1. Introduction

- (1-1) *Fovus ton Turkokiprion provalun i "Fainansial Taims"*
fears of-the Turkish-Cypriots show the Financial Times
'›Financial Times‹ speaks of the fears of the Turks in Cyprus.' (Elefterotipia, p.6)
- (1-2) *Nga lugina frynte një erë e lehtë.*
from valley-the blew a wind the cool
'A cool WIND was blowing from the valley.' (Bishqemi, p. 6)
- (1-3) *Iznad šume diže se crven oblak koji naveštava jesenje vetrove.*
above forest rises itself red cloud which announces autumnal winds
'Above the forest rises a red CLOUD announcing autumn winds.' (Kiš, p. 12)

The object of this study are sentences with verb–subject (VS) order in three languages spoken on the Balkan peninsula: Modern Greek (1-1), Albanian (1-2) and Serbo-Croat (1-3). Its purpose is what (interesting) linguistics is all about: to establish a connection between a particular linguistic form – VS order – and a particular meaning¹.

There are basically two ways of mapping forms to meanings. One can start from what is visible, the form, and by investigating the contexts in which it occurs eventually find out what it conveys. Apart from the general difficulty of finding and defining meanings, in using this procedure one has to be aware of the following two possible problems: (a) What at first sight seems to be a unitary form may on closer examination turn out to be a cluster of different constructions which are only superficially similar (the homonymy problem), and (b) Not all forms necessarily carry particular meanings; certain forms may turn out to be semantically/pragmatically redundant formal devices in the language under consideration (the zero-meaning problem). Keeping these provisos in mind, one can perform a fine piece of linguistic analysis using the form-to-meaning approach. The other possibility is to delimit a certain field of 'meaning' in advance and then look for the ways in which it is expressed. This type of approach may be useful in the initial stages of the research, since intuitions of meaning can bring to light formal differences which would otherwise remain unnoticed, or simply give a general direction in which one is to move. However, a complete investigation

¹ 'Meaning' should not be understood in its narrow semantic sense, but rather as a hyperonym of the truth-conditionally defined 'meaning' and pragmatically defined 'function'; the purpose of this vagueness will become clear in the course of the study.

conducted on the basis of these intuitions without reverting to linguistic form is permanently in danger of forcing one's own conceptions of what the meaning is onto the language. Furthermore, the meaning-to-form approach implies that there are meanings which exist independently of the language form, which is, to say the least, an unwarranted belief.

I obviously prefer to use the form-to-meaning procedure: coupled with a number of intuitions on the meaning of VS order, it seems to be the best means to achieve a sound empirically based account of the construction(s) in question. However, the state of the art in the field of VS-research is such that it does not allow for free diving in the virginally pure waters of empiry, for the simple reason that these waters are not virginally pure. Namely, there is an established meaning commonly associated with VS order, first conceived of in the philosophical tradition, i.e. independently of linguistic form, and only later paired with various constructions in different languages, among others with VS-sentences. This meaning is nowadays usually labeled **theticity**, a name given to it by one of the first researchers in the field, the Prague philosopher Anton Marty.

One would expect that the existence of a preestablished meaning should make my task almost too simple: the form and the meaning are already there, the only thing I have to do is to show that they form a linguistic sign in the Saussurean sense of the term and adorn the whole thing with a couple of nice details concerning discourse properties of the construction, so as to give my presentation an irresistible empirical flavor. Unfortunately, this is not the case, since theticity is everything but a clear-cut notion. The intuition that some natural language utterances are somehow 'special', i.e. 'thetic', is formalized and explained in a dozen different, usually contradictory, ways, and there are at least as many attempts to show that these utterances are really nothing special, so that theticity should, on this view, be considered only an empty label. Not only the intension, but also the extension of the notion of theticity (if there is such a thing at all) is rather vague: even if the core instances of thetic utterances are relatively indisputable, the assignment or non-assignment of less central examples to the category is often simply a matter of taste.

Confronted with this chaotic state of affairs (the details of which are presented in Section 2), one can either ignore the whole theticity issue and behave as if it were not there at all, concentrating only on VS order and its distribution, or try to assume a dual perspective and show the connectedness of a form which is perhaps not a unitary category at all and a meaning which is almost impossible to grasp. Basically, I have chosen the former option, but not in its orthodox form, i.e. with certain compromises: the basic method I used was to

determine meanings and constructions on the basis of the occurrences of VS forms in context, but I did not neglect the existent literature on theticity either.

Namely, on one hand, many VS sentences are in no way compatible with the current definitions of this category; on the other, for some VS sentences, there is indeed a form-meaning correspondence which in one way or the other resembles different descriptions given for it. Hence I shall try to find out if there exists the *differentia specifica* of that meaning and see if it is necessary to postulate a specific category ‘thetic’, or if the meaning in question is reducible to some superordinate principle. Then I shall compare my results with the current theories of theticity.

Thus, although my approach is basically the form-to-meaning one, the complicated and interesting issue of theticity is accounted for with a combined form-to-meaning/meaning-to-form method. Note that the perspective I assume is only partly dual. What I want to know is what VS order means, and what, if anything, the theticity is. As these two questions partly overlap, I concentrate on both of them in the part where they overlap; elsewhere, I deal only with the first question.

This partial duality will be most prominent in the following section, intended to give a short overview of the state of the art in both the theticity and the VS research. Section 3 contains a short assessment of the data presented in Section 2 and, against this background, a more precise account of the problems and aims of my own work. A chapter on the issues of topic and focus, terminology and methodology are followed by the main part of the study, the empirical investigation of VS order in Balkan languages.

2. State of the art

As indicated above, the presentation of the previous research I am offering may at the first sight look like a somewhat heterogeneous collection of opinions on VS order on one hand, and on such diverse phenomena like the German expletive *es*, the Japanese particle *ga* and accented subjects in English on the other. This colorful diversity is, of course, triggered by the dual perspective assumed in the present study: although there is indeed a strong association between theticity and VS order in many languages, this is not universally the case, so that some other structures on the basis of which the cross-linguistic status of the notion of theticity has been postulated have to be included as well.

Those who accept the existence of theticity tend to understand it as a semantic, pragmatic or psychological primitive, i.e. as an unanalyzable basic notion, which in some languages happens to be expressed by VS order. On the other hand, many approaches to VS order operate with notions other than theticity, trying to reduce the whole issue to some superordinate, more general principle. In these approaches, VS order is considered to be ‘nothing special’, i.e. it is believed to be a consequence of certain regularities which also apply to other constructions, whereby these regularities can be formal, semantic, or pragmatic, depending on the personal preferences of the analyzing linguist. A careful sifting of the relevant literature reveals that these are, as a matter of fact, two basic attitudes towards the VS phenomenon: it is either a consequence of a primitive semantic or pragmatic feature or somehow reducible to a more general rule or a set of rules. I shall call the first attitude the **thetic attitude** (although some linguists adhering to this view do not use this term) and the latter the **systemic attitude**.

In presenting the rich literature on the topic, I shall combine the chronological and the thematic principle, in order to illustrate both the historical development and the principal schools of interpretation of VS order and theticity. My state of the art report is meant to be detailed, but it does not pretend to be exhaustive, a fact easily understandable in view of the quantity of books and papers wholly or partially devoted to the problem¹.

¹ My classification of the approaches to the VS order and theticity draws heavily on the state of the art reports by Sasse (1987, 1996), Lambrecht (1987) and Haberland (1994).

2.1. Lexical properties

Perhaps the first thing one notices while exploring VS order in any of the familiar Indo-European languages is the relatively restricted number of predicates regularly occurring in VS sentences. This seems to have struck the first researchers in the field as the criterial property of these sentences: in the works of Wackernagel (1892), Delbrück (1900), Berneker (1900), Kieckers (1911), Frisk (1932), Behagel (1932), and others, one repeatedly reads that the reason for subject inversion is the ‘weakness’ of certain verbs, i.e. their very general meaning. Thus Frisk (1932: 36) says: ‘Diese [Verben] enthalten allgemein bekannte Begriffe, die der Verfasser eben darum als Ausgangspunkt nimmt, um dann das unbekannte Subjekt als das Ziel des Satzes folgen zu lassen’. A typical instance of the ‘allgemein bekannte Begriffe’ are verbs of existence, which are always adduced as the prototypical examples for VS order, as in the following Ancient Greek sentence (Frisk 1932: 62):

(2-1) *Esti tēs Krētēs Oaxsos polis, en tēi egeneto Etearkhos.*

is of-the Crete Oaxus city in which was-born Etearchus

‘There is in Crete a town named Oaxus, where Etearchus was born.’ (Hdt. 4.154.1)

The idea of the semantic properties of predicates as the triggers of VS order became the standard explanatory procedure in much of the work done on the Romance inversion, (e.g. Blinkenberg 1928, Le Bidois 1952, Hatcher 1956, Wandruszka 1982, papers in Fuchs 1997, etc.) and, to a lesser extent, in the literature on VS word order in Slavonic (Adamec 1966, Kovtunova 1976, Robblee 1994), with lexical semantics figuring either as the only trigger of the VS phenomenon or at least as a more or less essential part of a multicausal explanation. The principal progress consisted in the fact that the descriptions of the predicate classes believed to trigger inversion gradually became more detailed. The classical quotation from Hatcher (1956: 7) nicely illustrates this: ‘the verbs ... tell us only or mainly that the subject exists or is present; is absent, begins, continues, is produced, occurs, appears, arrives’. Apart from enumerating the most frequent types, however, Hatcher made an important generalization, according to which most of VS sentences are existential statements in a wider sense, since they express that a situation contains or lacks something, whereby this something is coded as the subject (see Sasse 1987:532, 1996:6 for an assessment). This line of reasoning reached its peak with Allerton and Cruttenden’s 1979 paper, in which the explanation in terms of existential semantics was extended to the English construction with accented subjects and deaccented predicates (for the sake of brevity, I shall use Lambrecht’s (1994) term **prosodic inversion** for this construction henceforth). The principal result of their analysis is that the

construction in question is triggered by the low informational value of the predicate with respect to the subject, the only predication conveyed being that of different modes of existence, coming into existence or of misfortune. The most important development to be observed between Hatcher and A&C is contained in the phrase ‘with respect to the subject’, as it allows for the predicates whose primary meaning is not that of existence to be included in the group of existentials. Consider (2-2):

(2-2) *The SUN’s shining.*

The KETTLE’s boiling. (A&C 1979: 52)

These sentences are treated as existential on a par with such prototypical existential statements as *There is a FLY in my soup* (A&C 1979: 51), because verbs like ‘shine’ and ‘boil’ denote the prototypical mode of existence of the sun and of the kettle, or, as A&C put it, ‘the emptiness of a verb depends on the identity of the subject.’ (1979: 51). The old concepts of semantic predictability (a term used by Adamec 1966 as a designation for a very common pattern of the semantic relationship between the verb and the subject in the Russian VS construction) and lexical solidarity (Coseriu 1967), which gave rise to yet another reinterpretation of VS order and prosodic inversion, that of semantic integration (see 2.4), are here reinterpreted as a covert predication of existence.

The lexical approach to VS order and related phenomena is not confined to the existential interpretation: at least two other approaches, developed in the Eighties and Nineties, can with some restrictions be classified as lexical. The first one is based on the semantic role of the subject and for the greater part overlaps with the extensive research done on unaccusativity since Perlmutter’s seminal paper (1978). Roughly, this approach is based on the observation that subjects in VS construction and prosodic inversion are frequently nonagentive, with the semantic roles ranging from Patient or Theme to Location. Most of the work based on this intuition was conducted in the generative framework and will be reviewed in greater detail in the appendix to Section 2.7. Suffice it here to say that the principal objects of this line of research are English and French locative inversion (*In the garden was a big dog*) and VS constructions in Italian and Spanish. It is mostly concerned with the theory-internal problems regarding the locus of base-generation and case-assignment: oversimplifying considerably, VS order is treated a direct consequence of the assumed base-generation of the subjects of the lexical class of unaccusatives VP-internally and of the assumed capability of unaccusative verbs to assign case directly. In an attempt to get to grips with apparent counterexamples for the unaccusative hypothesis, Drubig (1992) proposed a more sophisticated version of this account which combines the stage/individual level distinction (see below) with

unaccusativity. The only two content-oriented approaches to the question of subject nonagentivity in VS construction I am aware of, those of Lambrecht (1995, 2000) and Kennedy (1999), combine this semantic observation with the pragmatic notions of topic and focus, i.e. treat it as one of the consequences of a particular pragmatic or cognitive construal of propositions. Thus Lambrecht claims that the typical cases of VS order and prosodic inversion involve unaccusative verbs (2000: 617); apparent exceptions to this, like the sentence *JOHN called* (with its Italian VS pendent *Ha telefonato GIOVANNI*) are explained as instances of semantic reinterpretation: ‘due to the presentational function of the construction in which it occurs, the subject ... is not conceptualized as an agent but as an entity whose presence in the discourse is manifested via the activity in which it is involved’ (Lambrecht 2000: 673, n.15).

The second recently developed approach based on lexical properties of verbs takes the distinction between stage level and individual level predicates (Carlson 1977) as its starting point. Predicates that denote temporally stable properties (like *be altruistic*) are called individual level predicates, those that denote more transitory properties (*be available*) stage level predicates. Carlson conceived of this distinction in purely semantic terms; Kratzer (1995) and Diesing (1992) added a little syntax: stage level predicates, but not individual level predicates, have an additional (covert) argument denoting the spatio-temporal location of the eventuality described. This assumption is then used to explain a number of asymmetries in the syntactic behavior of the two predicate types. The first to apply this argumentation to the problem of the utterances called *thetic* was Ladusaw (1994; see also Drubig 1992), who exemplified his ideas with Japanese *ga* sentences and English locative inversion and the expletive *there*. His interpretation is basically semantic (or rather at the interface between semantics and ontology), and will therefore, together with the work of his followers, be dealt with in the section devoted to this interpretative type (2.5.3). Another direction of research on VS order and related phenomena which is based on the stage/individual contrast is more pragmatic in nature: roughly, the covert spatio-temporal argument of stage level predicates is assumed to play the topic role, allowing for the subject-verb complex to be interpreted as a unified comment on this topic, which is then assumed to be a sufficient account for certain exceptional formal properties of the sentences containing stage level predicates, VS order being one of them. Since individual level predicates do not have a spatio-temporal argument, they do not license VS order (Byrne 1997, Jäger 1997, 2001, Erteschik-Shir 1997, Pinto 1997, Etchegoyhen and Tsoulas 1998, Tortora 1999, Longobardi 2000; more on this approach in Section 2.6.3). Once again, the fact that it is not only predicates unequivocally denoting

transitional states of affairs that are used in VS sentences and related constructions (Rosengren 1997) is solved by lexical reinterpretation: in appropriate contexts, every individual level predicate can be interpreted as belonging to the stage level group. Thus Kratzer (1995: 155) claims that *Henry was French* can be treated as stage-level due to the past tense: one can imagine that Henry used to be French, but is now an American citizen.

The most striking feature of these three lexical interpretations of VS order and related phenomena, all of them unequivocally systemic (as opposed to thetic) in their basic attitude, is that they are all, at some level of analysis, forced to resort to some kind of lexical reinterpretation: nonexistential verbs become existential, verbs with agentive subjects are understood nonagentively, individual level predicates are coerced into stage level predicates. Though the existence of the principle of reinterpretation is in itself indisputable, and though this kind of analysis is, for some of the examples adduced, intuitively appealing and doubtless correct, the bulk of recalcitrant (non-existential, agentive, individual level) predicates occurring in VS construction are only with great difficulties to be explained as somehow constructionally mapped onto the prescribed type of meaning. Thus, I do not see how the boiling of a kettle is to be understood existentially, or how the agentivity of Giovanni who phoned me up is diminished by the verb-subject inversion. The description of the phenomenological field in terms of lexical semantics is rather to be treated as a description of *tendencies*, not of criterial properties of VS order: postposed subjects do tend to occur with the verbs liable to existential and stage level interpretations, and they do tend to be nonagentive, but none of these is the necessary condition for the use of the construction, let alone the decisive cause of its existence.

2.2. Given and new

The criterion of givenness vs. newness was first adduced as a method of explaining word order variation by Henry Weil (1844): what is given, or old, i.e. already known to the participants in communication, comes first, while new elements, which are somehow unpredictable in the given context and therefore represent the ‘purpose of speech’ (*but du discourse*) tend to follow. During more than one and a half centuries that followed, the notions of given and new were given different names and received different specifications and redefinitions. They were coupled with certain discourse functions and there were even attempts to free them from their impressionistic nature by establishing statistical criteria, etc.,

but the simple idea of different grades of familiarity which are to be made responsible for the inversion remained intact. Until the early Fifties, however, the given/new distinction was prominent only in one rather isolated linguistic island – in Prague, where Vilém Mathesius incorporated it in his theory of Actual Sentence Bipartition (*aktuální členění větné*, later known as Functional Sentence Perspective)¹. The Praguean version of the given/new distinction had a vast influence – first in the Czech and Slovakian linguistics, then in Slavonic studies in general, and finally, with the publication of Mathesius' works in English (Vachek 1964, Mathesius 1975), in the greater part of the linguistic community. Mathesius combined the textual criterion of givenness/newness with the logical notion of predication, as expressed in the distinction between grammatical and psychological/logical subject. Those sentence elements which serve as psychological subjects are generally given, by physical and textual context or encyclopedic knowledge, and are called *theme*, whereas the psychological predicate, *rheme*, is of necessity new, i.e. somehow unpredictable and unknown. In the later development of the Prague School, the criterion of givenness/newness gradually ousted the criterion of predication; the terminology changed, so that what is given became *contextually bound* (e.g. Hajičová-Sgall 1982); the theme-rheme dichotomy gave place, at least in the research conducted or influenced by Jan Firbas from the late sixties on, to the idea of the gradual increase of *communicative dynamism* (CD), roughly identifiable with newness, so that the elements more to the left have the smallest CD, whereas in the rightmost ones the grade of CD tends to be maximal (Firbas 1992).

2.2.1. Grades of newness

The most obvious way to fit the verb-subject order into the idea of the gradual rise of newness from left to right is to assume that the verb is old or given, and is thus the theme, whereas the

¹ In fact, the old-new distinction was not applied as an explanatory mechanism for word-order variation only by the Pragueans, but also in classical philology, probably due to the fact that Weil's book was primarily dedicated to Ancient Greek and Latin and thus directly in classicists' focus of interest. Since this line of research found little echo outside of its narrow philological field and did not contribute anything essentially new to the debate, I mention it only cursorily. The frequent occurrence of VS sentences in classical languages is generally considered to be a result of the givenness of the verb, caused either by the contextual 'oldness' or by the lack of content (*Drucklosigkeit* in Frisk 1932, see 2.1), with the subject carrying the new information, or being otherwise emphasized (Loepfe 1940, Dover 1960, Morocho Gayo 1985).

subject, being somehow new, plays the role of the rheme. If the sentence contains a locative adverbial, it forms the theme together with the verb.

(2-3) [(adverbial +) verb]_{given/theme} [subject]_{new/rheme}

VS order in Slavonic was the first to be subject to the interpretation represented in (2-3): with minor variations, this type of analysis of VS order is endorsed by Trávniček (1939, 1951), Daneš (1957) and Firbas (1964, 1992) for Czech, by Mistrík (1966) for Slovakian, by Čistjakova (1954), Braun (1962) and, in a somewhat updated form, by Kompeer (1992) for Russian, Szober (1933 [=1979]) and Penčev (1980) for Bulgarian, etc. Thus, in the sentences like:

(2-4) Czech (modified from Firbas 1992: 59)

Tam stála hlava rodiny, sám starý Jolyon.

there stood head of-family self old Jolyon

‘There stood the head of the family, old Jolyon himself.’

(2-5) Russian (Kompeer 1992: 218)

Nastupaet vesna, taet sneg.

comes spring melts snow

‘The SPRING is coming, the SNOW is melting.’

(2-6) Bulgarian (Szober 1933: 279)

Minuvat dnite.

pass days-the

‘The DAYS are passing by.’

the verbs (+/- adverbials) ‘there stood’, ‘is coming’, ‘is melting’ and ‘pass by’ are believed to be given by the context or by the general knowledge, and therefore thematic, and the subjects ‘the head of the family’, ‘the spring’, ‘the snow’, ‘the days’ are felt to be new in the discourse, and therefore rhematic. This interpretation was, with some exceptions, generally abandoned in Russian and Slavonic linguistics after the appearance of Adamec 1966 (see below). Nevertheless, old verbs and new subjects found their way out of the Slavonic studies relatively early, with Bolinger’s interpretation of inverted subjects in Spanish (Bolinger 1954a, 1954b): according to Bolinger, the verb is more predictable and less informative than the subject, and is therefore placed first.

(2-7) Why did you throw your letter away?

Se rasgó el papel.

itself tore the paper

‘The PAPER has torn.’ (Bolinger 1954b: 47)

The inversion in (2-7) is explained by the fact that ‘the reference to the paper as a cause is unexpected’, wherefore the subject is more new than the verb. More important than this scalar interpretation of VS construction is Bolinger’s observation (1954a) that VS order in Spanish has approximately the same function as the construction with accented subjects and deaccented verbs, i.e. prosodic inversion, in English (*JOHNSON died*): this was the first step in establishing a cross-linguistic category expressed by different linguistic means, a tendency which was to reach its culmination with Sasse (1987) and Lambrecht and Polinsky (1997). At first, however, it was the given/new distinction that attracted most attention. In Romance linguistics, where a long tradition of attributing subject inversion to certain lexical properties of the verbs had existed (2.1), Bolinger’s work introduced an additional dimension. Thus Hatcher, although basically a lexicalist, adduces in her famous study of Spanish existentials (1956) the presence of the verb in the underlying question (= givenness) as one of the factors triggering VS order. Similar combinations of the lexical and given-new criteria (old verb, new subject) are to be found in Le Bidois (1952: 350sq.), Combettes (1992) and Fournier (1997) for French, Grupo di Padova (1974) and Wandruszka (1982: 3sq) for Italian, Ocampo (1990, 1995) for Spanish¹. Here are some more Romance examples:

(2-8) French (Fournier 1997: 122)

En 1939 commence la reconstruction du moulin.
 in 1939 began the reconstruction of-the mill
 ‘In 1939 began the reconstruction of the mill.’

(2-9) Italian (Grupo di Padova 1974: 153)

Suona il campanello!
 rings the bell
 ‘The BELL is ringing.’

(2-10) Spanish (Ocampo 1995: 425)

Fueron los chicos.
 went the kids
 ‘The KIDS left.’

In English linguistics, probably due to the influence of the integration analysis proposed by Chafe (1974), Bolinger’s explanation of prosodic inversion in terms of the given/new distinction did not find much resonance: the only interpretation along these lines I am aware

¹ The most original explanation for the thematic status of the verb is offered by Grupo di Padova: all sentences of this type are used with the presupposition that something happens; since it is the verb that denotes ‘happening’, it is of necessity more given than the subject, and therefore thematic.

of is Huckin (1977: 170-3), who claims that in *the SUN is shining*, uttered out of the blue, the verb *is shining* is given, the subject *the sun* new. Two other English constructions, however – sentences with the expletive *there* and the so-called locative inversion – have been frequently accounted for in this way, notably by Firbas (1964, 1966, 1992), Bolinger (1971), (1977), Hartvigson-Jakobsen (1974), Penhallurick (1984), etc. Thus, in the following sentences from Bolinger (1977: 93), the location (‘across the street’) and the fact that something exists in that location (‘be’) are considered to be given, i.e. thematic:

(2-11) *Across the street there is a grocery.*

(2-12) *Across the street is a grocery.*

Apart from the analysis in terms of givenness of the verb and newness of the subject, the given/new distinction gave impulse to yet another interpretation, first proposed by Trávníček (1961). This approach is more firmly based on linearity: if there is a preverbal element, it is the theme, if not, the theme is the verb. Consider the following Czech sentences (Trávníček takes them from Mathesius 1939):

(2-13) *V jedné zemi panoval král.*

in one country ruled king

‘In one country there ruled a king.’

(2-14) *Byl jeden král.*

was one king

‘There was a king.’

In (2-13), the theme is the adverbial expression *v jedné zemi*, and the verb-subject combination *panoval král* is the rheme, whereas in (2-14) the verb *byl* is the theme, the subject *jeden král* being the only rheme. A comparable proposal is to be found in Halliday (1967: 212). The resulting scheme is more template-like than (2-3):

(2-15) [**1st element**]_{theme} [**rest**]_{rheme}; (1st element = Adv or V, rest = V+S or only S)

The latest development of the given/new interpretation of the VS order runs roughly along the lines depicted above. In a number of publications (Birner 1994, 1996; Birner-Mahootian 1996; Birner-Ward 1992, 1993, 1996, 1998; Ward 1990, 1998), Betty Birner and Gregory Ward give a corpus-based account of expletive constructions, locative inversion and simple subject inversion in English, Italian and Farsi. Elaborating on the fine-grained differentiation of the notions ‘given’ and ‘new’ developed by Prince (1981, 1992), Birner and Ward show that different VS structures, both in English and crosslinguistically, are sensitive to different types of absolute and relative familiarity (givenness), whereby the preposed element (a locative expression or the verb) has to be either absolutely given (i.e. present in the preceding

text) or only more given than the inverted subject. In some constructions, it is the type of givenness that is decisive: thus, in *there*-sentences, the inverted subject has to be discourse-new, but it need not be hearer-new, etc.

Descriptions of the VS phenomenon mentioned in this section clearly belong to the group of interpretations I subsumed under the name of systemic attitude: VS order and the related constructions are held to be derivable from a general principle of linear arrangement according to which the elements which are known, or given, appear first and are followed by the elements that are new. If the verb is given, or more given than the subject, the VS order appears, merely as a consequence of the superordinate first-given-then-new principle.

2.2.2. All-new utterances

The approach I shall call ‘all-new approach’ also operates with the given/new distinction, but differs from the previous one in being categorial instead of scalar. Neither the verb nor the subject are given, or theme, but they together represent a unified piece of new information, rheme, instead. If a sentence contains a preposed element, an adverb or a prepositional phrase, then it is only this preposed element that is given, i.e. theme. Schematically:

(2-16) **([adverbial]_{theme/given}) [verb+subject]_{rheme/new}**

The basic difference between (2-16) on one hand and (2-3) and (2-15) on the other is the categorial status of VS order: whereas in (2-3) and (2-15) the relative order of verb and subject is triggered by a higher-order principle of givenness/newness or by its relationship to linearity, the interpretation (2-16) envisages VS construction as a primitive with a specific meaning of its own: it marks the denotatum of the verb-subject combination as a new information.

The first description of VS sentences given along these lines is to my knowledge that of Mathesius (1939), who analyzes (2-13) and (2-14) (*v jedné zemi panoval král* and *byl jedan král*) as instances of verb-subject rhemes, allowing only for the adverbial *v jedné zemi* to have the thematic, i.e. given status. Mathesius’ thesis was accepted by Adamec (1966) in his influential study on Russian word order, which established it as the predominant interpretation of VS order not only in Russian linguistics (e.g. Kovtunova 1965, 1976, Sirotinina 1974, Krylova-Khavronina 1986), but also in the studies of some other Slavonic languages, notably Bulgarian (Georgieva 1974, Dyer 1992) and Serbo-Croat (Popović 1997). The broader popularization of the idea of all-new sentences, however, came from different

quarters. In one of the early attempts to incorporate pragmatic notions into the generative paradigm, Kuno (1972) coined the term ‘neutral description’ to designate Japanese sentences with the particle *ga* and the corresponding English sentences with prosodic inversion. His neutral descriptions intensionally strongly resemble Mathesius’ sentence rhemes. Thus the sentence

(2-17) *Oya, John ga kita*

oh John *ga* came

‘Oh, JOHN came!’ (Kuno 1972: 273)

is qualified as a neutral description because ‘the entire sentence conveys new information’. This type of characterization became popular in the Seventies and Eighties. Apart from the above mentioned works on word order in Slavonic, it is adduced as the explanation of English prosodic inversion by Schmerling (1976), Gussenhoven (1983a), Selkirk (1984), Rochemont and Culicover (1990), etc.; Italian VS sentences are treated as ‘all-rheme’ or ‘all new’ by Alisova (1972: 136), Lonzi (1974: 198), and, if I understand her correctly, Wehr (2000: 273) (in contrast to Wehr 1984, see below). The all-new definition occurs sporadically even today (cp. e.g. Grenoble’s (1998: 159-175) description of Russian VS sentences as ‘all-comment sentences’).

In contrast to the proponents of the verb-theme approach, who as a rule treat VS order as a single construction with one meaning, most of the people working with the notion of all-new utterances operate with more diversified structures. Thus, it is almost commonplace in the publications on VS order in Slavonic written after the appearance Adamec (1966) that this construction has at least two meanings – all-rheme or all-new on one hand, and verb-theme plus subject-rheme on the other, so that it covers both the interpretation given in (2-16) and that in (2-3). Accordingly, the Russian sentence

(2-18) *Gudit veter.*

whistles wind

‘The WIND is whistling.’ (Adamec 1966: 53)

can be understood either as an all-new (if pronounced out of the blue) or as a subject-new utterance (‘it is the wind that whistles, [not me]’). Exactly the same type of constructional polysemy appears in Kuno’s (1972) description of Japanese *ga*-sentences, which are said to convey both neutral descriptions (all-new utterances) and subject-rhemes (‘exhaustive listing’ in his terminology).¹ VS construction and its counterparts are thus ambiguous between the

¹ In some studies, chiefly in those dealing with Slavonic languages (e.g. , Adamec 1966, Krylova-Khavronina 1986, Jacenik-Dryer 1992, Popović 1997, Grenoble 1998), some further subtypes of the VS construction are

verb-subject-rheme and subject-rheme interpretations. One may argue that this ambiguity (or, for that matter, vagueness) renders the categorial status of all-new utterances doubtful and points to some kind of reducibility to a higher-order principle. Unfortunately, this issue has never been seriously discussed in the all-new literature. My impression is that there are two (implicit) ‘schools’, or rather tendencies. On one hand, Kuno (1972), for instance, seems to be firmly convinced that all-new utterances represent a specific pragmatic construal and that their formal coincidence with new-subject utterances is due to pure chance; he thus falls under my label ‘thetic attitude’. On the other hand, Adamec (1966) claims that whether a VS sentence will be interpreted with a verb-subject rheme (2-16) or with a subject rheme (2-3) depends on a number of factors, like verb semantics, semantic role of the subject, its textual boundness, etc. The all-new interpretation is thus understood as a consequence of the fulfillment of certain semantic and pragmatic conditions, so that Adamec’s theory is, if only tentatively, to be numbered among the systemic approaches to the VS phenomenon.

2.2.3. Given and new: An assessment

In its orthodox, binary form, the given-new approach to the information structure phenomena belongs to the past. The notions of ‘given’ and ‘new’ and, accordingly, the notions of theme and rheme, have never, or only seldom, been precisely defined. Are elements of propositions given in themselves, i.e. literally present in the given context or in the minds of the interlocutors, resembling thus Chafe’s activated elements (Chafe 1987, 1992), or are they given only in relation to the rest of the proposition? (*Mutatis mutandis*, the same question may be asked about the new elements.) In other words, if I say *PETER arrived*, and analyze it as [Peter]_{new/rheme} [arrived]_{old/theme}, does it mean that my interlocutor and I have just been thinking of someone arriving and did not think of Peter at all, or is it the case that ‘arriving’ is given with respect to the proposition, whereas the newness of ‘Peter’ is only a function of the unpredictability of this referent with respect to ‘arriving’, although we might have just been thinking of good old Peter? The greater part of the given-new literature I sifted through does not bother about notional clarifications but simply accumulates attributes like ‘given’, ‘known’, ‘predictable’, ‘contextually present’, ‘less important’, etc. vs. ‘new’, ‘unexpected’, ‘unpredictable’, ‘more important’, etc. The analyses one gets are thus usually extremely

identified, notably the type with given, thematic subjects and a colloquial or ‘folkish’ flavor, used predominantly in narrative contexts. The relevance of the multiple meanings of VS order will be discussed in more detail in Sections 2.3 and 3.

vague. In a few cases where the terminology did receive some explanation, it is more often the former alternative that is chosen, doubtless because of its empirical flavor: if givenness simply means that a referent is activated in the mind of the interlocutors, then it can be objectively determined by looking for the presence of the referent in the previous discourse.

Now, the problem with this highly mechanistic version of information structure is that language does not seem to function that way: recently activated, i.e. explicitly given referents have the terrible habit of appearing every now and then in the positions where only completely new elements should stand, and vice versa; speakers seem to be awfully careless about the proper treatment of things they have just mentioned, pronouns occur in all-new contexts, etc.¹ Since the theory in itself does not work, it is no wonder that the explanations of the VS phenomenon given within this framework do not work either. To give just one example: it is perfectly possible to say *the TRAIN arrived*, or, in Modern Greek, *irthe to TRENO*, in Serbo-Croat *došao je VOZ*, in Russian *prišël POEZD* (all three VS), etc., to a person one has just been talking to about the time the train (a particular train) is going to arrive. If the activation status of the subject or of the verb-subject combination as a whole were responsible for VS order and prosodic inversion, these sentences would be infelicitous, which is not the case. In sum: objectively understood context and absolute notions of givenness and newness fail to account for the way pragmatics influences sentence form. Accordingly, the given/new approach to the VS phenomenon, both in its verb-theme and all-new forms, falls short of explaining the object of its research.

However, the intuitions hidden behind this approach are essentially correct: the verbs and the subjects of VS sentences often do seem to form a kind of unified information chunk which somehow conveys new information, the subjects are frequently indefinite and somehow new to the discourse, there is indeed a feeling of some sort of systematic ambiguity between the interpretations with a ‘new’ verb-subject complex and with a ‘new’ subject only, etc. The problem is that newness and givenness simply do not represent the appropriate theoretical background to account for these intuitions.

¹ For a more thorough and therefore more devastating critique of the given-new dichotomy, see Pasch 1981, Reinhard 1982, Keijsper 1985 and Sasse 1987; see also Section 4.

2.3. Discourse functions

In addition to their basic characterizations of VS order as a lexical or information structure phenomenon, most of the authors cited in 2.1 and 2.2 mention the typical discourse functions of the construction: it is generally assumed to introduce new discourse referents into discourse, which is why it so often used as a text- or paragraph-opening device, or it somehow conveys a state of affairs as a whole, and is thus frequent in out-of-the-blue utterances and journal article titles. In a number of works devoted to VS order and related constructions, these discourse functions are taken not as consequences of a more general lexical or information structure principle, but as primitives which directly determine the sentence form.

The first decisively discourse-oriented explanation of VS order I am aware of is Hetzron's 'presentative movement', developed in a series of papers dating from the early Seventies (1971, 1975). A sentence element can be furnished with the feature [+presentative] when the speaker intends to give it a 'status of prominence in the short-range memory' (Hetzron 1975: 347), because it is to play an important role in the subsequent discourse, or because it provides a basis for the following contrast, or for some other reason. The presentative element is formally marked by a movement towards the end of the sentence. This presumably universal principle is then used as a means of explaining such seemingly different phenomena as English pseudo-clefts, Hungarian expletive *ott*-constructions, resumptive pronouns in Amharic, immediately preverbal subjects in Turkish and Japanese – and locative inversion in English and French, *there*-sentences in English, locative-subject and verb-subject orders in Hungarian, Modern Hebrew, Russian, Finnish, etc. Here are some examples (Hetzron 1975: 351):

(2-19) Finnish

Talossa on mies.

house-in is man

(2-20) Hungarian

A házban van egy ember.

the house-in is a man

(2-21) Arabic

fi:l-bayti rağulun.

in-the-house man-a

'There is a man in the house.'

In all these VS sentences, the referent ‘a man’, being presentative, is placed last. So the intended cataphoric relevance of the subject is a direct trigger for VS order, without intermediate stages of information structure, givenness and newness, or similar: only the iconic motivation counts. A comparable analysis is applied to Sinhalese and Tamil postverbal subjects by Herring and Paolillo (1995), and, within the LFG framework, on locative inversion in Chichewa and English by Bresnan (1994).

The broad typological account of VS order put forth in the influential studies by Givón (1977, 1983b) and Hopper (1975, 1979, 1986, 1987) ascribes this construction a completely different function. Givón’s analysis of data from Biblical Hebrew (1977) leads him to the conclusion that VS order expresses continuous chains of events, whereas SV plays a disruptive role in discourse. The same result, rephrased in terms of topic continuity, is reached in his study of word order phenomena in Ute (1983b). The following Hebrew example (from Myhill 1992a: 268) illustrates this:

- (2-22) *Va-yoʔmer elohīm la-ʔshah: »Mah-zot ʔasita?« Va-toʔmer ha-ʔshah:*
 and-said God to-the woman what-this you-did? and-said the-woman
»Hana[ʔash hishi ʔani.«
 the-serpent deceived-me
 ‘And God said to the woman: »What have you done?«, and the woman said: »The serpent deceived me.«...’ (Genesis 3:13)

The turns in the conversation are temporally sequenced and continuous, and are therefore VS; the action of the snake deceiving Eve is outside the actual sequel of events, disruptive and discontinuous, which is why it is SV.

Hopper’s account of languages with SV/VS variation (mostly based on data from Old Germanic and Malay) operates with the notion of transitivity as defined by Hopper and Thompson (1980): VS sentences carry the main story line and mark the foreground, meaning that they are more transitive, event-oriented, furnished with greater topic continuity, etc.; SV sentences give the background information, are less transitive, predication-oriented, express topic discontinuity, etc. This variation is said to be parallel in function to such diverse phenomena as voice in Tagalog and aspect variation in Slavic. Compare the following passage from the *Anglo-Saxon Chronicle*, representing the entry given for the year 755 (Hopper 1979: 222):

- (2-23) *Her Cynewulf benam Sigebryht his rices.*
 here Cynewulf took Sigebryht his kingdom.
 ‘In this year, Cynewulf took Sigebryht his kingdom.’

(A description of the events preceding this usurpation: Cynewulf had sent Sigebryht's brother Cyneheard to exile; he came back to the country.)

Ond þa geascode hē þone cyning lytle werode ...

and then found he the king with-small company ...

‘And then he (sc. Cyneheard) found the king with a small escort.’

The SV syntax is confined to the introduction, which presents a prior history of Cynewulf, Sigebryht and Cyneheard and thus falls out of the sequential narrative order. The VS syntax begins where the actual events are narrated, i.e. where the continuity of events has to be marked.

The simple notional apparatus and the objective-looking methodology of these generalizations made the analysis in terms of background vs. foreground or of continuity vs. discontinuity popular in language descriptions: Hopper's foreground-background approach is used by Dahlgren (1998) and Brustad (2000) to describe the SV/VS alternation in colloquial Arabic, by Luraghi (1995) to explain the VS construction in a number of old Indo-European languages (Hittite, Ancient Greek, Latin, Old Germanic), etc. Givón's continuity approach has been exceedingly popular. It has been instrumental in the descriptions of the SV/VS alternation in various languages, like Chamorro (Cooreman 1983, 1987, 1992), Colville (Darnell 1995), Nez Perce (Rude 1992), Spanish (Bentivoglio 1983), Tagalog (Fox 1985), etc.

Givón's and Hopper's results are diametrically opposite to the usual characterizations of VS sentences, independent of the feature made responsible for the inversion: whereas VS order is usually considered to introduce new elements into discourse, give background descriptions, etc., here it is treated as a narrative device *par excellence*. This difference in interpretation is only partly triggered by the idiosyncratic terminology these authors use, the main reason lying in the fact that Givón and Hopper, partly due to the choice of languages on which they base their description (Semitic, Ute, Malay, Old Germanic), describe a completely different type of VS construction than most of the people writing on Romance, Slavic, Finno-Ugric, etc. inversion or Modern Germanic expletive constructions.

This conclusion was first reached by John Myhill, who in a series of important studies (1984, 1985, 1986, 1992a, 1992b) tried to establish a relationship between the word order type and the meaning of SV/VS alternation and thus free this alternation from its undifferentiated universalistic-iconic background. The basic idea is simple: if a language has more than 60 % VS sentences, than this word order serves to signal sequentiality, marking topic continuity and uninterrupted chains of events; if the percentage of VS sentences is lower, they occur mostly in different types of existential, descriptive or interruptive contexts.

In other words, VS languages use VS order primarily for narration, SV languages primarily for presentation or description. Since the notions ‘VS language’ and ‘SV language’ are not absolute but rather gradient (a language can be VS or SV to a greater or smaller extent), it is often the case that both the narrative and the descriptive/existential type are found in a language, which is then only a logical consequence of its mixed overall grammatical structure. Jacennik and Dryer’s study of the Polish VSX construction (1992) and Longacre’s (1995, 1999) discussion of discourse functions of word order in a number of VSO languages follow approximately the same line. Though to a certain extent schematic, conceptually unclear and insufficiently radical, Myhill’s approach to the polysemy of VS structure and its relation to the word order typology represents a major breakthrough not only in the research of VS order, but in word order typology in general, as it is the first large-scale attempt to establish a relationship between formal properties of languages and their discourse structures. Its major flaw is, in my view, its understanding of discourse functions as unanalyzable primitives, inherited from the main-stream functional paradigm.

The idea of discourse functions as direct triggers of VS construction has its roots in the understanding of grammar as a perpetually emerging, unfinished system which is in constant adjustment to the cognitive (iconic motivation) and communicative (discourse motivation) needs of its speakers. It is implied or explicitly claimed that there is no such thing as conventionalized grammar, but at best some sort of compromise between the existent state and the competing motivations. In this way, discourse needs of the speakers can exercise an immediate, unmediated influence on the sentence form. That Hetzron on one hand, and Givón, Hopper and their followers on the other, attribute completely different functions to similar word order phenomena, is, as indicated, a consequence of the simple fact that they describe different construction types in different language types; the underlying understanding of grammar is basically the same.

The denial of the systematic nature of language was characteristic of the early years of functionalism, when it was necessary to establish a clear alternative to the systemomania of different formalist approaches (e.g. Garcia 1979, Givón 1979a); meanwhile, it has been abandoned even by some of its most prominent proponents (Givón 1979c). More particularly, a number of empirically based studies, starting with Myhill (1984), have clearly shown that some apparently similar constructions, supposed to be the universal response of the languages to the challenges of discourse, vary in their application from one language to another, and are thus rather conventionalized, language-specific expressions of pragmatic values, dependent on the overall grammatical structure of particular languages (see especially Sasse 1995a, 1996,

Birner and Ward 1998, Ward 1998 and Prince 1998). As indicated above, much of this empirical research, has been hampered by the idea of the unanalyzable nature of discourse functions (more on this in Section 6.6.). Nevertheless, the strong emphasis given to discourse considerations by Hetzron, Givón, Hopper and others is basically correct: combined with a detailed account of formal restrictions holding in a language, it is a firm basis not only for linguistic description, but also for cross-linguistic comparison, as Myhill has shown.

Finally, all interpretations of VS order reviewed in this section are systemic: Hetzron's, Givón's and Hopper's, because VS order is directly derived from higher-order iconic principles, Myhill's, because it assumes that the grammatical type determines the use and meaning of a particular word order pattern.

2.4. Semantic integration

In some of the lexicalist approaches reviewed in 2.1, a close semantic tie between the verb and the subject is taken to be one of the relevant factors triggering VS order and related phenomena. The semantic closeness between the two sentence elements is interpreted existentially, i.e. basically nonexistential verbs are understood as existential with respect to certain subjects.

This is, however, not the only possible treatment of the lexical solidarity that often exists between the verb and the subject. Namely, in assuming a lexical reinterpretation of the verb one actually postulates a change in the semantic value of certain elements of the proposition. However, it is also possible to keep the semantic values intact, and to assume that the same semantic content is processed in a different manner. The first to propose this idea was Chafe (1974): in prosodic inversion, for instance in the sentence *The BUTTER melted*, the subject-verb combination is said to form a conceptual unity for the speaker. In contrast, sentences with two intonation peaks (*The BUTTER MELTED*), or, for that matter, those with 'normal' intonation (*The butter MELTED*), are felt to convey the same information in separate information chunks.

Anna Fuchs (1976, 1980, 1987) coined the term for this intuition: integration. Integration is not confined to subject-verb complexes: it is a cross-categorial principle applying to all types of verb-noun and noun-adjective combinations, as well as to coordinate structures. In order for integration to take place, both elements of the complex have to be somehow 'informationally relevant' or 'new'; the speaker then has a choice whether s/he will present the information in separate chunks or integrate its parts into one, 'globally new unit' (Fuchs

1980: 449). What is interesting in Fuchs's approach to the subject-verb type of integration is that she embeds this phenomenon into broader discourse considerations. The discourse function of integrated (i.e. prosodically inverted) sentences is to give information about a deictically centered point of relevance, which is achieved by creating a relevance relationship between this point and the *whole situation* denoted by the subject-verb complex. This is the reason why the subject and the verb have to be semantically close: as Fuchs (1980: 457) puts it, 'the predicate ... denotes an action, state, property, etc. (typically one out of a limited number) that the subject is frequently mentioned in connection with, in certain context of use'. If the subject and the verb are not semantically close in the sense described above, the internal relationship between them would be felt to be relevant, so that the interlocutors would not be able to establish a holistic relationship between an external relevance point and the subject-verb complex as a whole. Let me illustrate this with an example (modified from Fuchs 1980: 458):

(2-24) Are you coming for a drink?

(a) *Nee du, HELENA ist krank, ich will lieber gleich heim.*

(b) *Nee du, HELENA ist SCHWIMMEN, ich muß sie gleich abholen.*

In (a), the relationship between the subject and the verb is, according to Fuchs's definition, that of semantic closeness, since the illness of a child is frequently mentioned in the context of turning down an invitation. For this reason, the subject-verb complex is conceptualized as a 'globally new' information conveyed about an external point of relevance, namely the speaker's refusal to come for a drink. In (b), the subject and the verb are not closely semantically tied (swimming is not a frequent reason for turning down an invitation), so that the hearer has to interpret the relationship holding between them as a relevant piece of information. The sentence has therefore a double intonation peak. Since one utterance can convey only one relevance relationship, the external point of relevance (the speaker's rejection of the invitation) is not directly predicated about. Semantic integration is thus conceptualized only as a prerequisite for a certain type of information transmission.

The notion of integration was taken over by Jacobs (1993), who renamed it 'informational nonautonomy' in a recent paper (1999) and furnished it with a more formal definition. The idea is that integration is a matter of compositionality: the difference between the informationally autonomous and the integrated readings of a phrase (NP, PP, sentence, etc.) lies in the number of steps taken in the calculation of the meaning of the phrase from the meanings of its elements. In the informationally autonomous reading, it is first the meanings of the elements that are recognized, and only thereafter the meaning of the phrase; in the

integrated reading, the meaning of the phrase is somehow computed without previously recognizing the meanings of the elements. This difference in information processing has a significant number is reflected in syntax and intonational phonology and is subject to a number of formal restrictions.

Although the integrational interpretation has been applied only to prosodic inversion in English and German, the intuitions underlying it also hold for the related constructions, including VS structure. The feeling that the situation denoted by the verb-subject complex is somehow unanalyzed is present in many other approaches as well (the all-new approach and manythetic interpretations); what is specific about the semantic integration is that the holistic nature of the situation is located in semantics. The exact way the supposed integration takes place is, despite Jacob's efforts, not entirely clear, but Fuchs's strongly pragmaticized explanation of the semantic tie between the verb and the subject as a sort of uncontroversiality in the given context seems to me to be one of the most interesting proposals regarding the relationship between the two principal elements of VS construction. The parallel drawn between the verb-subject complex and 'integrative' processes in other syntactic contexts exerted some influence on the subsequent literature (e.g. Lambrecht and Polinsky 1997). Of course, this parallel unequivocally qualifies the integrational interpretation as one of the systemic approaches to the VS phenomenon.

2.5. Theticity research and related approaches

2.5.1. Ancient history: Brentano and Marty

The Aristotelian division of human judgement in subject (*hupokeimenon*) and predicate (*katēgoroumenon*), generally accepted as an axiom for more than two thousand years, suddenly became less self-evident somewhere around 1850, probably due to the growing awareness of linguistic facts¹. The problem can be formulated as follows: how is it that in a number of sentences the grammatical structure consisting of a 'grammatical' subject and a verb does not correspond to the underlying 'logical' or 'psychological' structure consisting of a 'logical' or 'psychological' subject and a predicate? To take an example from that period, how is one to explain that in the sentence *il arrive deux étrangers* 'the subject is *il arrive* while *deux étrangers* is the predicate' (Meyer-Lübke 1899, quoted in Seuren 1999: 355)?

¹ For details see Seuren 1998:120-133, 1999.

Two basic positions are possible: either one disposes of the subject-predicate dichotomy altogether or one restricts its applicability only to *some* sentences (or judgments, depending on whether one is a linguist or a logician). The former option was to become victorious in logic and much of formal semantics, due to the enormous influence exercised by Frege and Russell. The latter, developed by Franz Brentano and especially by his pupil Anton Marty, was more or less forgotten until Kuroda revitalized it (1972), though it did leave some indirect traces in much of what was written on topic and focus, theme and rheme, etc.

The Brentano-Marty theory¹ deals primarily with judgment forms, whereby ‘judgment’ is only partly coterminous with what modern logic deals with, namely propositions. Apart from abstract relationships that hold between terms and predicates (which is what propositions are), it also comprises the cognitive acts included in forming a proposition (‘the genesis of the propositional thought, as distinct from its substance’, Seuren 1999: 360). The basic idea is that there are two types of judgments, only one of which reflects the Aristotelian subject-predicate scheme. The following oft-quoted passage from Kuroda (1972: 154) is a good summary of Brentano and Marty’s argumentation:

This theory assumes, unlike either traditional or modern logic, that there are two different fundamental types of judgments, the categorical and thethetic. Of these, only the former conforms to the traditional paradigm of subject-predicate, while the latter represents simply the recognition or rejection of material of judgment. Moreover, the categorical judgment is assumed to consist of two separate acts, one the act of recognition of that which is to be made the subject, and the other, the act of affirming or denying what is expressed by the predicate about the subject. With this analysis in mind, thethetic and the categorical judgments are also called the simple and the double judgments.

The ontological basis – two cognitive acts, presentation (*Vorstellung*) and judgment (*Urteil*), and two modes of judgment, affirmation and denial – is the same for both categorical andthetic judgments. The differences between them concern the way these cognitive primitives function and interact.

Both the categorical and thethetic judgments are based on two modes of judgment, affirmation and denial. The difference lies in the fact that inthetic judgments, a presentation (*Vorstellung*) is affirmed or denied (*anerkannt* or *verworfen*) in itself, whereas in categorical judgments, a predicate (presumably also to be understood as a presentation, *Vorstellung*, but see Ladusaw 2000: 238) is affirmed or denied **of** the presentation of a subject (*zuerkannt* or

¹ The account of the theory I give is of necessity short and incomplete. For more detail see Marty (1918), Ulrich (1985), and Ladusaw (2000 [=1994]); for Anton Marty’s language philosophy in general see papers in Mulligan (1990) and Raynaud (1992).

aberkannt), i.e. relatively. The corollary to this is that, although the same kind of thing, a presentation, is affirmed or denied in both judgment types, it is only in the categorical judgments that the predication relationship exists, as it is only here that the judgment is made relatively to an object or eventuality. So we can say, to use some of Marty's examples, that the thetic judgment underlying the sentence *Es regnet* does not contain a predication but simply represents an affirmation of the presentation of the rain falling, whereas *Diese Blume ist gelb* conveys a categorical judgment in which the presentation (?) of being yellow is affirmed **of** the presentation of a particular flower, so that the act of predication is present, due to the relational nature of the judgment.

Every categorical judgment consists of two judgments: the first recognizes or rejects the existence of the presentation of the subject notion by satisfying this presentation with an object or eventuality of the outer world, the second is the predication of a property relative to this object or eventuality. A thetic judgment simply recognizes or rejects the existence of a presentation by satisfying it with an object or eventuality.

Now, there is an obvious parallelism between the thetic judgment and the first step in the categorical judgment: they both recognize the existence of a presentation by matching it to an object or eventuality. There are some difficult points, though, concerning this parallelism. First the ontological problem. Brentano and Marty do not mention it, but are only at a step's distance from the question of subjects with zero reference: what happens if the first part of the categorical judgment is not a recognition of the subject, but its rejection?

This problem, in the form of the Russellian bald king of France, was to determine the theory of reference throughout the twentieth century. The second problem concerns the complex nature of the judgment material in thetic judgments. The existence of the subject of a categorical judgment is independently asserted ('judged', i.e. recognized): the presentation of 'this flower' in *diese Blume ist gelb* is satisfied by an object. The presentation which is the material of a thetic judgment may contain further presentations, as in *Da schläft eine Katze*, where the judgment material is the eventuality 'a cat sleeps' which contains a further presentation 'cat'. Now, how is this presentation ('cat') satisfied by an object if the affirmation of its existence is not an object of a separate cognitive act, i.e. a separate judgment? The answer Brentano offers to this is what is today called unselective existential closure (Ladusaw 2000: 237): the existence of the subordinate presentation included in the

presentation constituting the material of thetic judgment is asserted indirectly, *via* affirmation of the presentation of which it is a part¹.

The final point worth mentioning in the context of the Brentano-Marty theory is the strict distinction between the judgment form and the language form: language is only an imperfect reflection of thought. The thetic/categorical distinction is therefore not directly derivable from the sentence form. Consequently, Marty's examples of thetic judgments comprise such divergent sentence types as existentials (*Gott ist, Es gibt gelbe Blumen*), impersonal sentences (*Es regnet, Es dünkt mir*) and sentences with universally quantified subjects (*Alle Blumen sind schön*) – the latter because the underlying judgment in such sentences is supposed to be a negative thetic judgment ('there are no flowers which are not beautiful').

Brentano and Marty located the thetic/categorical distinction on the level of thought and conceptualized it as an extralinguistic phenomenon. Nevertheless, the theory addresses, in an explicit or implicit form, some of the issues which still represent major problems of the linguistic theory. Since most of the approaches to theticity that ensued are based on different solutions given to these issues, or on the emphasis given to one of them, this seems to be the right place to enumerate them.

(i) **Predication.** What is predication? An automatic semantic relationship between a term and a predicate on the level of propositional thought, or a pragmatic relationship established by the speaker between a discourse referent and another referent or denotatum on the level of utterance? (Brentano and Marty use the term in the former sense.) Or are we dealing with two distinct phenomena here? And, depending on the answer to this question, on what level should the notions of subject and predicate be defined and how are they to be labeled? (For my proposal, see 4.3.1.)

(ii) **Integration.** How are complex 'presentations', i.e. concepts consisting of more than one simple individual term, built and understood? How do we affirm or deny the 'judgment material' of thetic judgments and, for that matter, the 'predicate' of categorical judgments?

(iii) **Reference.** How do we establish the reference of terms, or, the other way around, how do we commit ourselves to the existence of terms? Is it really a separate cognitive act in the case of 'subjects' of categorical judgments and merely an indirect commitment in the case of parts

¹ The question of semantic compositionality of the allegedly unary propositional material with which Brentano and Marty were confronted is, although both the terminology and the intellectual tradition are different, strongly reminiscent of the problem the integrational approach (2.4) has in explaining the nature of semantic integration of verbs with subjects.

of ‘judgment material’ in thetic judgments, as Marty claims? Or is there some universal mechanism underlying both cases?

(iv) **Linguistic form.** If thetic judgments do not display a subject-predicate distinction, why are sentences which represent them more often than not formed with this distinction? Is there any relationship between the thetic/categorical distinction and the linguistic form? (Marty’s answer to this question is negative.)

2.5.2. Conceptualist approaches to theticity

The first descent of theticity from the transparent air of speculation into the rough world of linguistic empiry (apart from a short note in Mathesius 1929; see Sasse 1987: 512) – I am referring to Kuroda’s paper from 1972 – was marked by two tendencies: to keep the philosophical foundation of the notion intact on one hand, and to confirm its actual existence by natural language data on the other. These two tendencies dominate the otherwise rather diversified line of theticity research which I label conceptualist.

Theticity is understood as a phenomenon on the level of thought. Now, it is not very popular to do linguistic research based on non-linguistic notions, and I suppose the people mentioned in this section would very much like to give a purely semantic account of the thetic/categorical distinction. The problem is that it is very difficult to describe the difference between, say, *Peter CAME* (‘categorical’) and *PETER came* (‘thetic’) in terms of truth-conditions. The solution offered by Kuroda and others is a slight displacement of theticity from semantics to a higher level: the thetic/categorical distinction is basically a matter of conceptualizing states and events, meaning that it is a manner of thinking, but somehow very close to semantics. The linguistic relevance of theticity is another important feature of the conceptualist approach: though in principle allowing for underspecification of the thetic/categorical distinction in some contexts, the proponents of this approach use, in contrast to Brentano and Marty, the language form as the most important criterion for telling apart the underlying thetic and categorical judgments, a fact that has some important consequences with respect to the extension and intension of the notion of theticity.

The conceptualist framework was established by Sige-Yuki Kuroda in a paper dealing with the Japanese particles *wa* and *ga* (Kuroda 1972; see also Kuroda 1976, 1987, 1990): the thetic/categorical distinction is said to be unequivocally marked in Japanese, with the particle

wa attached to an argument denoting the categorical, and the particle *ga* to the one denoting thethetic judgment, as in the following examples:

(2-25) *Inu ga hasitte iru.*

dog GA running is

(2-26) *Inu wa hasitte iru.*

dog WA running is

‘A/the dog is running’

According to Kuroda, (2-25) is a simple recognition of the perceived state of affairs ‘a dog is running’, loosely paraphrasable with ‘there is a dog running’. In contrast, (2-26) implies two judgments: first, the recognition of the fact that there is a certain dog, and second, the assessment of the predicate ‘is running’ with respect to the referent ‘dog’ established by the first judgment. Thus, *ga* in (2-25) marks thetheticity, whereas *wa* in (2-26) is a signal for the categorical judgment. If there indeed exist immediately observable linguistic forms corresponding to the rather abstract thetic/categorical distinction, then we can find out more about this distinction by exploring the properties of these linguistic forms. So Kuroda used his intuition that *ga/wa* = thetic/categorical in order to modify the Brentano-Marty theory. In his first theticity paper (1972), it is the extension of the notion of thetheticity that is changed; in the most recent one (1990), its intension is also touched upon.

First the extension. (a) Universal statements (‘all flowers are beautiful’) are not thetic. Their subjects have to be marked by *wa*, which demands a theory of reference different from that adopted by Brentano and Marty. Thus Kuroda states that it is not real objects that we set up in our minds as bases (‘subjects’) for categorical judgments, but rather ideas of entities, or ‘intentional meanings’; (b) The particle *ga* is used not only in existential and ‘impersonal’ sentences, but with a much wider range of verbs (see (2-25)). Consequently, thetheticity is not restricted to a couple of well-defined situation types but is rather a general way of conceptualizing all types of states of affairs; (c) The choice between *ga* and *wa* is given only in what Kuroda calls specific sentences (roughly, those pertaining to particular occurrences of events); in what he somewhat confusingly labels generic sentences (habitual, constant, generic *stricto sensu*, etc., states of affairs), only *wa* is allowed. Thus, thetheticity is restricted to conceptualizing particular occurrences of states of affairs.

The reinterpretation of thetheticity Kuroda (1990) undertakes is based on what I in 2.5.1 called the reference issue: How do speakers commit themselves to the existence of terms in thetic and categorical judgments? If I understand his argumentation correctly, Kuroda claims that terms in thetic judgments exist in speakers’ minds only within the judgement material:

‘the dog’ in (2-25) exists in the speaker’s mind only as an entity defined by the event of running and has no separate existence as ‘substance’ beyond the confines of this event. In contrast, ‘the dog’ in (2-26) is subject to multiple confirmation of its independent existence: first by a judgment recognizing its existence obliquely, through an event, and then by the recognition that the term introduced by this judgment is a separate, independent ‘substance’. The final stage in the categorical judgment is, of course, predication of a property to the term that has been established as ‘subject’ by the two previous steps. Thus, ‘the dog’ in thethetic judgment/sentence only has a fuzzy existence, via being perceived within the event of running; ‘the dog’ in the categorical judgment/sentence exists independently both of the event and of our perception of the event. These two types of existential commitments are related to two cognitive domains, perception and memory/experience, with the ‘thetic’ existence derived from the former, the ‘categorical’ from the latter. The ‘thetic’ dog exists for us only as a part of the perceived event; the ‘categorical’ dog exists for us because we keep it stored in our memory, i.e. knowledge.

Unfortunately, the simple equation $ga = \text{thetic}$ does not seem to hold, at least not absolutely: many *ga*-phrases are dangerously subject-, or, if you prefer, topic-like, and some *wa*-phrases are impossible to interpret as subjects/topics (see e.g. Clancy and Downing 1987, Ueno 1987, Portner and Yabushita 1998). Nevertheless, Kuroda had followers. Shibatani (1991), in a paper with a strong psychologizing slant, describes thethetic/categorical distinction, which he simply equates with the *wa/ga* opposition in Japanese, in terms of the opposition between perception and experience (*ergo*, we are again dealing with two types of judgments):

Topicless (=thetic, D. M.) sentences present witnessed events and states as straightforwardly as they are perceived without analyzing their parts and without reflecting the speaker’s experiential judgment regarding the relationship between the analyzed parts. Topic (=categorical) sentences, on the other hand, separate out topics, which are experientially judged in terms of their relationship to the rest of the sentence (Shibatani 1991: 101).

Thus, theticity is a phenomenon of cognition (the modern equivalent of ‘thought’) reflected directly in the language by means of the lack of ‘topic’ (= *wa* phrase). A ‘generic’ statement cannot bethetic because it cannot be based on perception, but only on experience.

Apart from the fact that it made the logical notion of theticity somewhat more concrete by binding it to a particular form, Kuroda’s (and Shibatani’s) conceptualist interpretation had the merit of being the first to raise two important issues:

(i) **Perception vs. experience.** Sentences that have been characterized as thetic seem to be somehow based on perception, and to exclude experience, or knowledge. Kuroda explains this perceptual effect by the purely perceptual nature of thetic judgment (I think what I see), but there are also other explanations for this.

(ii) **Genericity.** ‘Generic’ – ‘non-episodic’ would perhaps be a better term – statements seem to be only marginally or not at all possible in the constructions that are characterized as thetic. For Kuroda and Shibatani, this is a logical corollary of the fact that generic statements cannot be based on perception, but only on experience, the latter being the basis of categorical judgments only. The relatedness of this issue with the interpretation of theticity in terms of the distinction between stage level and individual level predicates (2.1, 2.5.3, 2.6.3) is obvious.

2.5.3. Semantics and theticity

The decisive step towards the pure semantic conception of theticity, made possible by the development of dynamic semantics in the 1980s, was taken by William Ladusaw in his influential 1994 paper (= Ladusaw 2000; see also Drubig 1992). He replaced Kuroda’s vaguely philosophical terminology with the instrumentarium of situational semantics and rooted the thetic/categorical distinction in the current semantic discussion, using it to derive Milsark’s generalization on strong and weak quantifiers (individual level predicates must have strong, i.e. definite, specific or generic subjects) and incorporating Kratzer’s idea of the presence of a spatio-temporal argument in the valence of stage-level predicates into the description of theticity.

Ladusaw operates with the following inventory: objects and eventualities, descriptions of objects and eventualities, and properties. Thetic judgments are based on descriptions of eventualities which are affirmed or denied. Since the basis of the thetic judgement is a description of eventuality, the predicate in the thetic proposition must have an event argument which binds it spatio-temporally. As indicated above, Kratzer (1989=1995) postulated the existence of this argument for stage level predicates, which are thus licensed for thetic judgments. Individual level predicates do not qualify for thetic judgments for two reasons: (a) They denote permanent properties and do not have an event argument, so that they cannot serve as descriptions of eventualities; and (b) They denote properties, which are ontologically different from descriptions, and it is only descriptions that may serve as bases for thetic judgments. Categorical judgements have a saturated description of an object as their bases.

Properties, prototypically individual level predicates, are affirmed or denied relative to this basis. Properties can be derived from descriptions of eventualities, which explains the capability of stage level predicates to be used in categorical judgments. Only strong subjects (i.e. those with a definite or specific reference) are saturated descriptions of objects, which is why Milsark's generalization holds for individual level predicates (and for stage level predicates used as properties). Note that the explanation is basically lexical: theticity is based on stage level predicates, which are semantically parametrized event descriptions; categorical judgments are based on individual level predicates, which are properties.

Despite my brutal simplification of Ladusaw's argumentation, it is clear why this theory was attractive for some formal linguists: first, the intuitively appealing thetic/categorical distinction is translated into familiar semantic terms; second, a number of important problems regarding reference and verb semantics are solved in a simple and elegant way; finally, some popular syntactic devices (Kratzer's event argument, Diesing's VP-internal subject generation with some predicates) are kept intact and embedded into a broader semantic perspective. Laduslaw's theory of theticity has been frequently used as a ready-made solution for various syntactic and semantic problems: Byrne (1997) and Etchegoyhen and Tsoulas (1998) use it as a starting point for an explanation of definiteness effects in Spanish and French VS sentences, Basilico (1999) expands it in order to account for some peculiarities of the placement of unaccented pronouns in English locative inversion, while Ogawa (2001) gives an account of nominal semantics based on the stage/individual level distinction. The greatest impression, however, was made by the very fact that the thetic/categorical distinction was for the first time phrased out in formal terms. For the consequences of this see 2.6.3.

One of the weak points of the theory is, as in Kuroda's case, its empirical adequacy. Rosengren (1997) has shown that many naturally occurring sentences, although formally 'thetic', are construed with individual level predicates. Furthermore, the very basis of the theory, the distinction between stage and individual level predicates, does not seem to be so certain today as it was at the beginning of the Nineties: a number of studies have proven the formal and semantic differences between the two groups to be reducible to other, more fundamental features of predicates, like stativity (see McNally 1994 and Jäger 2001 and the references quoted therein). Finally, Milsark's generalization can be derived in a simpler way, by taking resort to the fundamental pragmatic principles of assertivity (Jäger 2001, see Section 3.2).

2.5.4. Discourse-pragmatic reinterpretation of theticity

Although it chronologically precedes Ladusaw's semantic attempt, I am dealing with the pragmatic reinterpretation of theticity only after Ladusaw, principally for the reasons of cataphoric coherence: the idea of the pragmatic roots of the thetic/categorical distinction had the greatest impact on the subsequent research, so that the pragmatic interpretation is today enshrined as *the* interpretation of theticity, as will become patent in Section 2.6.

The thetic/categorical distinction was first used to explain the VS/SV alternation by Vattuone (1975). However, since his main concern is to find a suitable definition of a 'neutral' sentence, which he identifies with thetic judgment, his theoretical contribution to the problem is negligible. The first important step toward the redefinition of theticity aimed at increasing its descriptive applicability was a series of articles by Sasse (1982, 1984, 1987) and a book on VS order in Rumanian by Ulrich (1985). Since Sasse's broad typological account is both more theoretically elaborated and more influential, I shall concentrate on it in what follows.

Although functional parallels between formally distinct constructions in different languages had been noticed before (Bolinger 1954a, see 2.2.1), Sasse (1987) is the first to explicitly establish a universal category of theticity. He makes a strong claim to the effect that the thetic/categorical distinction is 'universally reflected in sentence structure in a way as basic to the syntax of human languages as, say, the distinction between declarative, interrogative, and imperative sentences' (p. 518). That is, the thetic/categorical distinction is obligatorily encoded in every language, albeit in various ways, depending on the overall grammatical structure of a language. The following examples illustrate the major types of expressing theticity he identifies.

(2-27) **VS order**; Albanian (also Romance, Slavonic, Greek, Hungarian, Mandarin, etc.)

Leh qeni.

barks dog-the

'The DOG is barking.' (Sasse 1987: 537)

(2-28) **prosodic inversion**; German (also English)

Die KATZE hat miaut.

'The CAT has meowed.' (Sasse 1987: 527)

- (2-29) **split structures**; Mandarin (also Egyptian Arabic, Welsh, French, Boni, etc.)
Yǒy rén gěi nǐ dǎ-diànhuà.
 EXIST person to you hit-telephone
 ‘Somebody phoned you.’ (Sasse 1987: 542)
- (2-30) **‘thetic’ particles**; Hebrew (also Arabic)
Whinnēh malʾāḳē elohīm ʾolīm wyordīm bō.
 and-THET.PART angels of-god ascending and-descending on-it
 ‘And behold the angels of God ascending and descending on it.’ (Sasse 1987:544)
- (2-31) **subject incorporation**; Onondaga (also Oneida, Mohawk, Boni, etc.)
ʔo-nōhs-aték-ha.
 SUBJ-house-burn-STAT
 ‘The HOUSE is burning.’ (Sasse 1987: 550)
- (2-32) **nominalization**; Tongan (also Tagalog and other Austronesian languages)
Ko e ui 'ae tangata.
 EXIST the call ALIEN.POSS.-the man
 ‘The MAN is calling.’ (Sasse 1987: 552)

The extension of the notion of theticity thus became much wider than Brentano and Marty could have ever dreamed. In a way, Sasse created a research tradition of thetic sentences, in that he united various research traditions of apparently extremely variegated phenomena under the label of the thetic/categorical distinction.

What is the nature of this distinction? It is clearly discourse-pragmatic, partly intersecting with the given/new distinction, but – this is an important point after decades of given verbs and new subjects – only in an epiphenomenal way. The first approximation offered is the notion of communication perspective, conceived as ‘the general shape a speaker gives the state of affairs which he is about to convey in a sentence’ (p. 518).

Two different communication perspectives, the thetic and the categorical, are derived from a more fundamental distinction, that between nonpredicative and predicative assertions. Assertion is understood as a superordinate notion denoting every kind of information transmission via propositions. Predication is the assertional act whereby a property is ascribed to an entity, called a predication base; this is how categorical statements work. The nonpredicational assertion is a simple recognition of a state of affairs, without a reference

point, i.e. without a predication base; this is the definition of thetic statements¹. If an entity is involved in the thetically presented state of affairs, it is not picked out as the predication base but is presented as part of the event:

While an entity serving as a predication basis is always autonomous, that is, independent of and OUTSIDE the predicated event – this must be so since the event is presented as its property – an entity involved in a simple ‘recognition’ is INSIDE the event and may not be conceived of as an entity at all. (Sasse 1987: 555)

An important distinction is drawn between the semanto-pragmatic and overt grammatical predicativity: while the former is a result of the speech act of property-ascribing, the latter simply represents the result of a process whereby in many languages the categorical sentence type is grammaticalized as the default sentence type. In other words, in many languages, the subject-predicate dichotomy, originally stemming from the pragmatic articulation of categorical utterances into predication base and property, has become the obligatory sentence scheme regardless of the pragmatic articulation (‘subject-prominence’, Sasse 1982). Thus, because of formal constraints holding in many languages, thetic utterances often must display grammatical predicativity, although on level of content nothing of the sort is present in them. The question of linguistic form, raised as early as Brentano and Marty (2.5.1), is thus resolved in terms of a diachronically explainable mismatch between form and content.

The differentiation between pragmatic and grammatical predicativity is the starting point for a unified functional explanation of different types of formal markings of thetic sentences: they all share the same feature, that of diminishing the obligatorily present grammatical predicativity to a certain extent. Thus, VS order, which is in many cases accompanied by lack of agreement, or even by a non-nominative case of the subject, displays less prototypical features of grammatical predicativity (initial position of the subject, nominative, agreement) than SV order; in prosodic inversion, the assignment of the sentence stress has the same function; incorporation of the subject into the verbal complex is understood as a formal sign of its being ‘inside’ the reported event, i.e. of its not being the predication base, etc. To put it bluntly, speakers find a way to cheat on the predication-based structures of languages in order to express nonpredicative (thetic) assertions.

¹ Note that this is an answer to the question I raised in 2.5.1, in connection with Brentano and Marty’s theory: predication is understood as a (pragmatic) mode of assertion in which a relationship between an entity and a property is established and as such distinguished from another mode of assertion, in which an eventuality is presented as whole. My answer to this question and the corresponding terminological solution differ significantly from Sasse’s proposal (see 4.3.1).

Yet another distinction is drawn that is relevant for the formal properties of thetic sentences, the one between entity-central and event-central thetic statements. The former posit the existence of an entity, the latter the existence of an event. Languages vary as to the extent to which the one or the other type is preferred in expressing thetic statements. The incorporation strategy, in which nouns lose their referentiality, is at the event-central end of the scale, the split structure strategy, which inevitably draws attention to an entity, at the entity-central end.

The predicational/nonpredicational dichotomy is embedded in a wider discourse perspective: the use of the two assertion types is determined by needs of communication, more precisely, by the expectations the speaker creates or assumes on the part of the hearer. Consider (2-33) and (2-34):

(2-33) What's new? *HARRY's coming*.

(2-34) What's going on outside? *HARRY's SINGING*. (Sasse 1987: 521)

Though textual and situational contexts are practically the same, the answer in (2-33) is construed as a thetic utterance (prosodic inversion), that in (2-34) as a categorical one. Apart from being a nice case against the explanation of the thetic/categorical distinction in terms of contextual givenness, this minimal pair illustrates how different assumed expectations influence the choice between the thetic and the categorical perspective. The answer expected in (2-33) is of the form 'what is new is the event of Harry's coming', because it is not any coming but only Harry's that is relevant; both elements are of equal communicative value. In contrast, in asking the question in (2-34), one expresses one's interest in the noise heard from outside, not in the entity producing it (or at least the speaker understands it that way); the expected form is something like 'it is singing you hear, and (by the way) Harry is the performer'. The elements are of different communication value. When the same communicative value is assumed to be expected for all elements of a proposition, the unitary, nonpredicational thetic construction is chosen; when there is an informational separation of an entity and the rest of the proposition, it is the categorical statement that is uttered:

[T]he thetic type of statement is used whenever the speaker assumes that the hearer expects unitary information to be given about the whole situation in question, and the categorical type of expression is used whenever he assumes that the hearer expects information units about constituent parts of the situation to be built up successively. (Sasse 1987: 568)¹

¹ It is interesting to compare this explanation for the choice of thetic articulation (one of the few that have been proposed in the thetic research at all) to Fuchs's proposal (2.4). They both operate with the notion of integrated reading, but Sasse assumes that the integration is purely discourse-pragmatic, having to do with the expected

Many formal and lexical restrictions and tendencies of VS sentences and related structures described in the relevant literature – monoargumentality, ‘weakness’ of the verbs, ‘newness’ of the subjects, incompatibility with generic statements – are derived from this basic pragmatic feature (pp. 524-6 and *passim*).

Sasse’s universalist approach had a decisive impact on the future research of VS order and other phenomena he included in the thetic category. First, his reinterpretation of the logical, or cognitive, thetic/categorical distinction in discourse-pragmatic terms became, albeit with significant modifications, the generally accepted basis of theticity research. Second, he demonstrated the crosslinguistic formal relevance of the assumed distinction and gave a simple and logical functional explanation of this relevance, thus initiating a line of typological research interested in discovering formal means of marking theticity, both in particular languages (see e.g. Wu 1992 for Chinese, Polinsky 1990, 1993 for Chukchee, Kinyarwanda and other languages, Maslova 1997, 2003, for Yukagir, etc.) and as a basis of broader typological generalizations, the most important being Lambrecht and Polinsky’s principle of subject-object neutralization (Lambrecht and Polinsky 1997; see 2.6.3) and the latest version of É. Kiss’s discourse configurability theory (É. Kiss 1998a), which takes the formal differentiation of thetic and categorical sentences as the criterion of the topic-configurability of a language. Sasse gave a unified and intuitively appealing account of many formal and lexical tendencies and restrictions holding in a number of ‘marked’ constructions in various languages and offered a plausible explanation for their distribution in discourse. His solution is an elegant one, since it disposes of the unnecessary formal and cognitive apparatus often stipulated for the distinction in question.

Whether all these advantages of this approach rectify the postulation of a universal pragmatic dichotomy is another question, which I hope to answer in the empirical part of this study. There is only one rub apart from this, but a fundamental one. Namely, if thetic statements do not contain predication (understood as ascription of a property to an entity), how come that, on hearing a thetic utterance like that in (2-33), we know that the property expressed by the grammatical predicate holds for the subject which is not a predication base just as well as when we hear a categorical utterance like (2-34), in which the subject is the predication base and the grammatical predicate plays the role of the pragmatic predicate? In

communicative values of the sentence elements, whereas Fuchs considers it to be a consequence of a non-controversial relationship between an entity and an event in a given context. In Sasse’s approach, the integration is a phenomenon in its own right; Fuchs understands it as the necessary prerequisite for the reported event to be interpreted as relevant for an external relevance point.

other words, when you hear (2-33) and (2-34), you will know about Harry both that he is coming and that he is singing, although Sasse's theory predicts a different result. This problem (recognized as such by Sasse himself in his 1996 paper) had been addressed already some time before Sasse's paper by Allerton and Cruttenden ('In *The TAP is leaking* the tap is surely the topic', 1979: 53) and by Wehr (1984; see below), but neither of them offered a satisfactory explanation. Actually, the relationship between the property assignment and the assertional behavior (see 2.5.1) is a major problem of the information structure research in general, not only of thethetic/categorical distinction. I shall lay out my solution in 4.3.1 in some detail. To expound my results in advance, I think that Sasse's intuitions are right, but his notional apparatus needs some refinement in order to capture these intuitions in a non-contradictory way.

2.6. Topic and focus: presupposition/assertion-based approach

In the Eighties, the terms *topic*, *comment* and *focus* were still ambiguous between the given/new and 'predicational' readings. If you found the word 'topic' in a linguistic text, you were never sure whether the author meant the old information or something like Sasse's predication base; 'comment' was either a piece of new information or Sasse's pragmatic predicate. To make things worse, 'focus' was used either as a synonym for 'comment' or as a designation for a prosodically highlighted constituent. It is for these reasons that Sasse decided to use a somewhat idiosyncratic but less ambiguous terminology, 'predication base' instead of 'topic' and 'pragmatic predicate' instead of 'comment' or 'focus'. Due to the general decline of the given/new approach and to Lambrecht's redefinition of 'focus' as assertional part of proposition (Lambrecht 1986, 1987, 1994), the situation became somewhat clearer: 'topic' is nowadays generally used in the meaning of 'predication base', as a label for the element of the proposition *about* which the proposition is construed, whereas 'focus' (sometimes still called 'comment') denotes the part of the proposition conveying Sasse's assertion (both predication andthetic 'recognition').¹ Accordingly, the topic-focus-based interpretations of the VS order and related phenomena to be reviewed in this section, despite terminological differences, do not differ from Sasse's approach in their ontological basis.

¹ For more detail on topic and focus see Section 4, especially 4.3 and 4.4.

2.6.1. New topics

In her account of the relationship between discourse and syntax in Romance, Wehr (1984) put forward an explanation of VS order and cognate constructions based on a combination of the topic-comment articulation and the given/new distinction. According to Wehr, VS order in Romance is used only in all-new contexts, meaning that both the verb and the subject have to be new to the discourse (though other, optional, elements need not necessarily be new). However, she distinguishes between two types of VS order along the parameter of topic-comment articulation. The following examples (Wehr 1984: 24) illustrate the first type:

(2-35) French

Il était trois petits enfants.

it was three small children

‘There were once three small children.’

(2-36) Italian

Era arrivato da Buenos Aires il padre di Guido.

was arrived from B.A. the father of G.

‘Guido’s FATHER arrived from Buenos Aires.’

Sentences which contain an existential or a predicate of appearance on the scene are said to contain no topic, and serve to introduce those new elements which are to become relevant discourse referents in the subsequent discourse. Obviously, this interpretation does not differ essentially from the usual all-rheme or all-new accounts of VS (2.2.2). The second group, labeled ‘neutral description’ (after Kuno 1972), can contain any kind of predicate and is used when the new element is not going to be relevant in the subsequent discourse.

(2-37) Italian (Wehr 1984: 54)

What happened? *Chiamó una ragazza.*

cried a girl

‘A GIRL cried.’

(2-38) Spanish (Wehr 1984: 56)

Do me a favor and don’t mess up these papers: *qe se va a enfadar Mario.*

himself will get-angry M.

‘MARIO will get angry.’

The difference in relation to existential, ‘topicless’, all-new sentences is said to lie in the fact that the inverted subject in neutral descriptions is not only a new discourse referent, but also the sentence topic, the latter not being the case in existentials. The comment about this

sentence topic is conveyed by the preposed verb. ‘Neutral descriptions’ are thus interpreted as predicational structures (in Sasse’s terms) in which the information is presented in two steps: a new discourse referent is established, like in existentials, but, in contrast to existentials, this referent is at the same time commented upon, the comment about it being given within the same sentence. On the other hand, the difference between the VS neutral descriptions and the canonical SV sentences lies in the fact that in SV sentences the topic is already established, i.e. it need not be posited.¹ Wehr finds the formal support for her analysis of the VS sentences with nonexistential predicates in crosslinguistic comparison: in the contexts in which VS ‘neutral descriptions’ are used in Italian or Spanish, French uses a subclass of clefts called all-new clefts (i.e. those in which the clefted element cannot be interpreted as narrow focus), like the following:

(2-39) What’s the matter? *C’est maman qui me bat.*

it is mum who me hits

‘My MUM is beating me’ (Wehr 1984: 65)

Now, (2-39) indeed has a form which seems to reflect the interpretation Wehr gives for VS ‘neutral descriptions’: the existence of a discourse reference is asserted (‘it is my mother’) and a comment is given about this new topical referent (‘she is beating me’)². But this does not have to mean that sentences used in similar contexts in other languages must have the same underlying topic-comment (or topic-focus) articulation, unless one assumes that this articulation is automatically determined by the context, which is a typical feature of the given/new approach, explicitly rejected in its pure form by Wehr. Italian, Spanish, Latin, etc., VS constructions with nonexistential verbs do not bear any formal sign of double articulation, but are rather formally identical to the type with existential verbs, assumed to be topicless. Thus I see no compelling reason to posit different informational articulations of VS sentences in these languages along the existential/nonexistential line.

¹ Apart from all-new existentials and neutral descriptions, Wehr (1984: 22) recognizes two more distinct types of VS in Romance: verb-topicalization structures with given verbs (lat. *Eiectus sum. Eicitur et Proselenus.* – ‘I was thrown out. PROSELENUS was thrown out, too.’) and subject-focus sentences (ital. [in a café] *Pago io.* – ‘I’m paying’). In a way, these two minor types strongly resemble the given-verb-new-subject interpretation of the Prague School approach.

² This bifunctional construction is not confined to French; it occasionally occurs in other Romance languages (Wehr 1984: 67), in English (*there was a farmer had a dog*, Lambrecht 1988), and Chinese (LaPolla 1995:315). Lambrecht (1988, 1994:180) calls this structure ‘presentational amalgam’ or ‘biclausal presentational construction’ and ascribes its existence to the principle of separation of role and reference.

Wehr's interpretation (modified to the standard all-new approach in Wehr 2000, see 2.2.2) tries to capture the intuitive difference between 'empty' verbs (existence, appearance, etc.) and 'full' verbs like *die*, *leak* or *get angry*. Since the former do not seem to ascribe any significant property to the entity they are grammatically predicated of, the constituent denoting this entity is somehow felt to be non-topical. The latter convey distinctive, 'informative' properties, so that mental representations of their subjects are enriched by them and are therefore felt to be topics. This is an old problem in a new disguise (see 2.5.1 and 2.5.4): our knowledge seems to increase indiscriminately, both concerning topical and non-topical referents, although only the former should theoretically be the case. Wehr tried to solve this apory by resorting to semantic properties of predicates and crosslinguistic parallelisms. The result is interesting, but it falls short of solving the problem.

2.6.2. Sentence focus

VS sentences and related constructions seem to convey unitary pieces of information (see 2.4). Furthermore, this information is felt to be somehow highlighted. If we combine these two intuitions, we get the idea of focussing the whole utterance. The first to express it was to my knowledge Müller-Hauser, who characterized the French cleft construction exemplified in (2-39) as 'mise en relief de la phrase entière' (1943: 225); a similar definition for the same construction was given by Rothenberg (1971). McCanna (1973: 104) ascribes the Welsh 'abnormal sentences' the property of 'emphasizing the total statement', which is said to be 'characteristic of explanatory statements such as occur commonly in response sentences'. The notion of focus these authors use is rather ill-defined ('highlighting' and 'emphasizing' are not definitions, but metaphorical circumscriptions), so that the sentence-focus interpretation remains vague. It was only with the refinement of the theory of focus by Lambrecht (1986, 1994) that 'focus' became an operational notion in language descriptions, so that his interpretation of VS order and related constructions in terms of sentence focus (Lambrecht 1987, 1994, 1995, 2000, Lambrecht-Polinsky 1997) is the first one with clear contours.

Lambrecht recognizes only one type of information transmission through propositions, pragmatic assertion (roughly corresponding to Sasse's use of the term); predication is in his framework a semantic relationship holding between terms and predicates and as such not directly relevant to the issues of information packaging. The scope of assertion varies across utterances: depending on communicative intentions, an argument, a predicate complex, or the

whole proposition may represent the assertive material in an utterance. These different scopes of assertion are formally marked by focus structures, or focus constructions. Accordingly, three focus constructions, or simply three focus types, are distinguished in grammar: argument focus, predicate focus, and sentence focus. Note that Lambrecht's focus types are based on constituency, as understood in the American structuralist tradition: possible scopes of assertion are defined on the basis of possible syntactic constituents, NP vel sim., VP, and sentence. How is this related to prosodic inversion, VS sentences, 'all-new' clefts and similar phenomena?

There are three relevant levels in Lambrecht's analysis. First, he apparently (e.g. 2000: 620) postulates the existence of thetic propositions (corresponding to Marty's and Kuroda's thetic judgments), i.e. propositions with a non-binary structure. The second and most important level are grammatical structures labeled sentence-focus structures (VS order, prosodic inversion, etc.), defined as follows: 'Sentence construction[s] formally marked as expressing a pragmatically structured proposition in which both the subject and the predicate are in focus.' (Lambrecht 2000: 617). The third level is that of focus construal, i.e. of giving an underspecified focus structure a particular scope reading. Since only grammatically marked meanings are relevant for the linguistic theory, it is argued that it is only the second level, that of sentence-focus construction, that is of interest to a linguist (the third level being at best interesting for a discourse analyst). The first level, that of thetic proposition (which, especially in the 2000 paper, has a somewhat extralinguistic flavor) is explicitly ruled out from the field of linguistic analysis. Namely, thetic, non-binary propositions (or judgments), may but need not be mapped onto sentence-focus structures, depending on the idiosyncratic rules of grammar of particular languages. Thus the same propositional content can be coded as sentence focus in one and as predicate focus in another language:

(2-40) Russian

Idet DOŽD'.

goes rain

(2-41) English

It's RAINING. (Lambrecht 2000: 619)

The propositional content 'it is raining' is, according to Lambrecht, coded as a sentence-focus construction in Russian (VS order), but not so in English, where the predicate focus is used. What is of relevance is the formal marking of the assertional scope which includes both the subject and the predicate. In Russian, this assertional scope is marked by the VS order; in English, by prosodic inversion. Since (2-41) does not display the latter phenomenon but a

typical predicate-focus intonational structure (sentence stress on the predicate), i.e. since it is grammatically identical to *He is WALKING* and not to *PETER is coming*, it does not count as a sentence-focus utterance, although it presumably expresses athetic judgment.

Now, it is often the case that a sentence-focus structure like (2-40) can be read not only as expressing athetic proposition, with both subject and predicate under the scope of assertion, but also as an argument-focus construction, with only the subject conveying the assertional material. Thus, (2-40) can be used in a corrective context, when someone thought that it was snowing: *Idet DOŽD* then means (literally) ‘goes RAIN (not SNOW)’. This reveals the relevance of the third level, that of focus construal: one focus construction can allow for more than one focus construal. In the case of the Russian VS construction, one can say that it allows for the sentence-focus and the argument-focus construals, where the argument under focus is the subject.

What linguistics should deal with is: (a) What means do languages use to formally express the sentence-focus articulation? (b) Under which conditions does a sentence-focus structure express the sentence-focus construal as opposed to the argument-focus construal? And, related to (b), (c) when do speakers use sentence-focus structures with sentence-focus construal, and why?

In answering the first question, Lambrecht combines Sasse’s notion of diminished predicativity with a sort of markedness analysis (see especially Lambrecht-Polinsky 1997 and Lambrecht 2000): the underlying functional motivation for different encodings of the sentence-focus construction (subject included in the focus domain together with the verb) in different languages is to mark it as distinct from the unmarked predicate-focus construction, in which the subject is outside the scope of assertion and the focus domain is the VP. What counts is creating a clear formal difference between the former and the latter structure. Thus VS order is used as a sentence-focus construction only in SV languages, where SV order codes predicate-focus, prosodic inversion (Sv) in those languages, in which the predicate focus is marked by the sentence stress on the verb (sV), etc. The emphasis is laid on the conventional, non-iconic nature of the sentence-focus marking: depending on the overall grammatical structure and language-specific grammatical restrictions, languages use the one or the other construction with the only aim to mark it as distinct from the unmarked predicate-focus structure (which is, in turn, also coded according to the overall grammatical structure). The most striking example of this basically structuralist account is the claim (Lambrecht-Polinsky 1997) that, provided there is some word order flexibility in a language, when predicate focus is coded by the SV order, sentence focus is coded by the VS order (Russian,

Italian), and when predicate focus is coded by the VS order, sentence focus is marked by the SV order (Welsh, Arabic).

As far as question (b) is concerned (the relationship between focus structure and focus construal), Lambrecht notes (like Adamec 1966, Fuchs 1980 and Sasse 1987) that from the two readings, sentence and subject focus, the former is much more restricted, both lexically and textually. The conclusion he draws is that only certain predicates with certain semantic roles of the subject allow for the sentence-focus construal without difficulties, whereas the others have to be reinterpreted in an appropriate manner in order to be construed in this way (see 2.1). This is where question (c), concerning the discourse properties of sentence-focus constructions, enters the scene. The function of these constructions is primarily presentational (Lambrecht 1987, 1994), that is, they introduce new discourse referents. In contrast to the existential interpretation of VS construction (2.1), this does not imply that sentence-focus constructions necessarily assert the existence of discourse referents: the act of presentation can be perfectly well performed by presenting an entity as involved in a certain action. This explains the frequency of inherently presentative predicates in the sentence-focus construction as well as the preferably nonagentive nature of subjects: agentive subjects tend to be understood as separated from the event, whereas the nonagentive ones are somehow within it. Lexical reinterpretation, whereby the meaning of the predicate and the semantic roles of the arguments are adjusted to the meaning of the construction, accounts for the instances that deviate from the prototype defined by presentative predicates and nonagentive subjects: it is thus claimed not only that in *The SUN's shining* the verb *shine* represents a sort of event frame within which the sun is presented in the discourse (1987: 375), but also that in *JOHN called* John is introduced into discourse as someone who phoned, whereby he loses his agentivity because he is presented as a part of the calling event (1995: 167), not as the performer of the action.

The emphasis given to formal features in this approach, and particularly the claim that only those assumed pragmatic differences that are observable in grammar are linguistically relevant had (or should have had) important consequences in the theticity research, the most important being the exclusion of sentences with double intonation peak from the phenomenological field of theticity, i.e. of sentence focus (a step foreshadowed in Keijsper 1985). Namely, since sentences like Chafe's famous example *The BUTTER MELTED* may appear in all-new contexts ('What's the matter?', or similar), it has often been assumed that they are also instances of sentence focus (e.g. Gussenhoven 1983a, Selkirk 1984, Rochemont and Culicover 1990, Cinque 1993, Pinto 1997). Lambrecht notes that this intonation pattern is

much more often used in order to express the unmarked predicate focus structure, albeit with new, ‘non-ratified’ or contrastive subject-topic, for instance in a context like ‘What happened to the things you bought yesterday?’, where sentences with double peak may be used to convey a series of predicate-focus utterances with contrastive topics (*The BUTTER MELTED, the STRAWBERRIES are INEDIBLE, and the CHEESE is full of WORMS.*). The accent on the subject is defined not as a focus accent, but as an activation accent (Lambrecht 1994) or a topic-ratifying accent (Lambrecht and Michaelis 1998). Accordingly, the construction with two intonation peaks is classified as a predicate focus construction which sometimes may be construed as sentence focus (recall the difference between the focus construction/structure and the focus construal), and excluded from the discussion of theticity (Lambrecht 2000: 620-22). *Mutatis mutandis*, this generalization holds also for the languages in which VS order is assumed to mark sentence focus: SV sentences used in all-new contexts are simply predicate focus constructions with a sentence focus construal.

The basic difference between Sasse’s theticity theory and Lambrecht’s sentence focus lies in the level of linguistic structure in which the presumed thetic/categorical (or sentence/predicate focus) distinction is placed. As noted above, Lambrecht posits only one type of information transmission via propositions, which he labels assertion, whereas Sasse assumes that there is a fundamental difference between the predicational and the ‘thetic’ assertion. Sasse’s theticity is thus a pragmatic primitive which triggers the use of specific formal markings. Within Lambrecht’s system, though he does not object to the assumption that there is some kind of difference between ‘categorical’ (binary) and ‘thetic’ (non-binary) propositions, the relevant distinction is that of linguistic coding: those constructions which mark the subject as falling under the scope of assertion, i.e. as (a part of) focus are instances of sentence focus (‘thetic’), and those in which the subject is the topic about which the assertion is conveyed, with the VP as focus domain, are predicate focus constructions (‘categorical’). Oversimplifying a little, one may say that the essence of theticity in Lambrecht’s view is the inclusion of the element designed to be topic into the scope of assertion.

This view is not without problems. As Lambrecht himself notes, his sentence focus constructions can contain topics, i.e. elements outside the scope of assertion, provided they are not subjects. Thus the Italian sentence (2-42) clearly asserts something about the direct object ‘she’:

(2-42) *L' ha lasciata il MARITO.*

her has left the husband

'Her HUSBAND left her.' (Lambrecht 2000: 648)

One may wonder why this sentence should be included in the category of sentence focus constructions, when it is not the whole sentence but only the verb-subject complex that is in the scope of assertion. In other words, why use the term 'sentence focus' if it is only verb-subject focus that is meant? This is not merely a question of nomenclature but rather a more fundamental question of the categorial status of Lambrecht's sentence focus: if the only difference between the predicate and the sentence focus is that in the former it is the object (or some adjunct) and in the latter the subject that is, together with the verb, included into the focus domain, then it is not clear why we need two ontologically distinct categories. This is where Lambrecht's dependence on a specific understanding of syntax becomes apparent. Namely, if the scope of assertion, i.e. focus, has to be mapped onto constituents, then only what is considered to be a constituent can be focus. In one widely accepted understanding of constituency, verbs and objects form a constituent, the verb phrase, but verbs and subjects do not. Thus, if a verb and an object are in the scope of assertion, one is entitled to consider the VP to be focus; we are dealing with an instance of Lambrecht's predicate focus. On the contrary, if an assertion is represented by a verb and a subject, one cannot speak of the VP as focus domain, since subjects are believed not to belong to VP. One has to go one node up in order to find a constituent built up by the verb and the subject. This constituent is the sentence. This is why verb-subject focus domains *must* be considered sentence foci even when sentences contain unequivocally nonassertional material.

There are two ways to solve this problem. The first is to alter the understanding of constituency depicted above; the second is to abandon the idea that focus has to map on constituents. Both solutions are partly adopted by Kennedy (1999). Her analysis of the English prosodic inversion strongly resembles that of Lambrecht, with two important differences: (a) some subjects (those of unaccusative and stage level predicates) constitute VPs with their verbs, and (b) focus does not have to correspond to syntactic constituents. The relationship between focus and syntax is rather an indirect one, meaning that the subjects assumed to be generated VP-internally (unaccusatives, stage level predicates) are more easily interpreted as constituting a focus domain with the verb, because they form a constituent, but this does not prevent the subjects believed to be VP-external (unergatives, individual level predicates, transitive verbs) to build focus domains with their verbs under appropriate conditions. This, if I understand Kennedy's argumentation right, implies that the verb-subject

focus domain (assertional scope) is more probable and easier to construe if syntactic and assertional structures agree than in the case of mismatch between these two structures, this being only a matter of statistical preferences, not of categorical distinctions. This dissociation of constituency and assertional structure would rectify the step Kennedy does not undertake (at least not explicitly), namely the complete abandonment of the sentence focus category and its replacement with a simple broad focus interpretation of the verb-subject combination, parallel to the verb-object, verb-adjunct and similar complexes.

The sentence focus interpretation of VS order and related constructions influenced much of the work on intonational phonology (see chapter 5 in Ladd 1996 and Kennedy 1999 for an overview) and some descriptive studies, like LaPolla's analysis of VS order in Chinese (LaPolla 1995) and the account of the same word order pattern in Basque by Bellver and Michaelis (1999). The theory of theticity proposed by Rosengren (1997) is in essence a translation of Lambrecht's sentence focus interpretation into the language of the Minimalist program, albeit with a somewhat greater emphasis on the extralinguistic nature of the thetic/categorical distinction. The main problem of Lambrecht's theory is, as I indicated above, the syntactic nature of focus he postulates and the inconsistencies resulting from this assumption¹. If one removes this obstacle, however, his interpretation (basically a further development of the ideas expressed in Sasse 1987) represents a sound basis for the analysis of VS sentences and related constructions.

2.6.3. Spatio-temporal argument as topic

Although the enthusiasm about Carlson's stage/individual level distinction has somewhat dwindled in the last couple of years (see 2.5.3), Kratzer's postulation of a Davidsonian event (spatio-temporal) argument for stage level predicates (Kratzer 1995 [=1989]) still plays an

¹ Önnarfors's study of verb-first sentences in German (1997) is a good illustration of the dangers of the equation of syntactic and assertional structures. Having defined sentences like *Fragt der Konditor die Friseur...* as sentence-focus clauses on the basis of the fact that the position in front of the verb is not occupied, he goes on to analyze them as 'thetic', although their discourse functions and lexical properties strongly diverge from everything that is usually subsumed under this term (and strongly resemble the VS sentences which express narrative continuity in Hebrew or Old English as described by Givón and Hopper, see 2.3). 'Theticity' (or sentence focus) is thus reduced to a meaningless label for sentences with a certain syntactic property. This kind of misunderstanding has been avoided by people like Lambrecht, doubtless due to their sound linguistic intuitions, but is, I believe, inevitable if one seriously takes focus and constituency to be mutually dependent.

important role in the interpretation of ‘thetic’ sentences, threatening to outlive the distinction it was originally devised to explain.

It was noted as early as Adamec (1966) that VS sentences and related constructions tend to have a spatio-temporally bound reading, i.e. that most of them somehow automatically pertain to *hic et nunc*, in contrast to canonical SV sentences, an observation frequently repeated in the literature on Slavonic VS constructions since (Bonnot and Fougeron 1982, 1983, Keijsper 1985, Kompeer 1992). The case in point are sentences like (2-40), which can only be interpreted as descriptions of the current state of affairs (it is raining now), not as habituais or similar (it rains every day, etc.). Furthermore, a highly interesting interaction between aspect and VS order has been observed by Bonnot and Fougeron (1982) and Kompeer (1992). Here is a Russian example illustrating this (adapted from B&F 1982: 319):

(2-43) At that moment Pjotr woke up.

(a) *Zvonil TELEFON.*

rang: IMPF telephone

(b) *Zazvonil TELEFON.*

rang: PERF telephone

(c) *Telefon ZAZVONIL.*

telephone rang: PERF

First the contrast between the two VS sentences, (a) and (b). Sentence (a), with an imperfective verb, is interpreted to mean that the action of the telephone ringing had started before and was taking place simultaneously with the action of Pjotr’s waking up, the former thus being the cause of the latter; the appropriate English translation would be *The TELEPHONE was ringing* (with the implication ‘it had started to ring before Pjotr woke up’). On the other hand, (b), with a perfective verb, means that the telephone started to ring only after Pjotr had woken up: *The TELEPHONE rang* or, better, *The TELEPHONE started to ring*. In contrast to these VS constructions, the SV sentence (c), although with a perfective verb itself, can have both temporal interpretations, i.e. it can mean that the telephone had started to ring before Pjotr woke up, similar to (a), and that it started to ring after Pjotr had woken up, like (b). This difference between the aspectually induced temporal sensitivity of the VS sentences and the temporal underspecifiedness of the SV type is taken to indicate the same kind of spatio-temporal boundedness of VS order as the example (2-40) described above. Now, recall that a similar observation was made with reference to the Japanese *ga* sentences by Kuroda and Shibatani (2.5.2): they cannot have a ‘generic’ (non-episodic) reading and are said to be based

on perception. Thus we are again facing a similar semanto-pragmatic characterization of formally completely unrelated constructions.

This is where Kratzer's spatio-temporal argument enters the scene. It has often been claimed that only stage level predicates occur inthetic statements (see 2.1 and 2.5.3), which perfectly fits Slavonic and Japanese data discussed above: temporary properties and temporal boundness are much the same thing. Many presumed syntactic differences between stage and individual level predicates are explained by the presence of an additional spatio-temporal argument in the valence of stage level predicates. The logical consequence of these two assumptions is to try to explain the first (the special relationship between stage level predicates and theticity) with the second (the spatio-temporal argument of stage level predicates). This step was taken almost simultaneously by a number of people – e.g. Erteschik-Shir (1997), Jäger (1997, 2001), Pinto (1997), McNally (1998), Tortora (1999), Longobardi (2000), etc. I shall only describe the common kernel of these approaches, but the reader should keep in mind that there are considerable differences between some of them.

Individual level predicates are characterized by the fact that their subjects must be topics. With stage level predicates this is not the case: their subjects may, but need not be topical. If the subject of a stage level predicate is not topical, then the role of the topic is played by the event (or spatio-temporal) argument these predicates always have in their valence frame. In this case, the verb-subject complex plays the role of the focus, or comment. This is at the same time the definition ofthetic sentences: inthetic sentences, the event (as expressed by the verb-subject complex) is asserted of a certain spatio-temporal frame. Syntactically, this is achieved by the scrambling of the spatio-temporal argument to the position outside the VP, i.e. to the canonical position of topical subjects in categorical sentences. The non-topical subject remains within the VP, which is a necessary syntactic prerequisite for the formation of a pragmatically assertive verb-subject complex. In this way, the effect of spatio-temporal boundedness ofthetic sentences in Russian and Japanese (as well as in comparable constructions in other languages, notably in English prosodic inversion), is explained as the consequence of a pragmatically driven syntactic movement of the spatio-temporal argument.

If I use a spatio-temporal argument (coindexed by default with *hic et nunc*) as the topic about which I want to convey new information, then it is only logical that the information conveyed is understood as pertaining to a certain space and time. Thus, the non-episodic reading of (2-40) (*Idet DOŽD*, 'goes rain') is easy to account for: the assertion that it is raining is made with respect to a *hic et nunc*-indexed covert event argument, so that the only possible reading is the one related to the present situation. The contrast between the imperfective verb

in (a) and the perfective one in (b) in (2-43), although somewhat more complex, is also deducible from the existence of a covert event argument. This argument in (2-43) refers to the contextually given section of time, that of Pjotr waking up. If the imperfective verb is used to assert an event of this section of time, the interpretation is that the whole section of time is ‘covered’ by the event, and no claim is made as to the beginning of the event, since imperfectivity signals that the left and the right borders of events are not specified. This allows for the pragmatically plausible interpretation that the phone started ringing before Pjotr woke up. If, on the other hand, it is the perfective verb that is the means of asserting the event about the covert spatio-temporal argument, then the speaker asserts that the beginning of the event took place at the given section of time, since perfective verbs conceptualize events together with their left or right borders. The resulting interpretation is that at the moment of Pjotr’s waking up the phone started to ring.

The status of overt spatio-temporal adverbials inthetic sentences is unclear. Some people (Rosengren 1997) claim that they are not topics but rather some kind of restrictive modifiers of the event; some are, again, convinced that they play essentially the same, topic, role as the covert event arguments (Erteschik-Shir 1997). Upon this interpretation, if I say *In the garden was a rabbit*, I assert of the garden (at a particular time) that the event of a rabbit being there obtains with respect to it.

Note that Lambrecht’s notion of sentence focus is here simply embedded in a predicative structure: the verb and the subject form a focus domain which represents an assertion about an external argument, which means that the subject-verb complex serves as a kind of predication (in Sasse’s terms) about a predication base in the form of an event argument. A similar analysis had already been proposed by Fuchs (1976, 1980; see 2.4), only that she was more interested in the semantic relationship between the verb and the subject, whereas the proponents of the event argument interpretation mostly deal with the relation between this presumed argument and the verb-subject complex as a whole.

Attractive as it is, this interpretation encounters the same problem as all approaches with lexical basis: counterexamples. Although many – perhaps the majority – of ‘thetic’ sentences do seem to be somehow spatio-temporally bound, there are also some clear instances of the ascription of permanent properties occurring inthetic constructions. In some cases, one is entitled to speak of lexical reinterpretation – an individual level predicate is interpreted as a stage level predicate, so that the idea of the topical event argument may be salvaged:

(2-44) German, prosodic inversion

What's the matter with you? *PETER ist tot.*

'PETER is dead.'

In this example, the individual level predicate 'be dead' may be understood as reinterpreted to mean more or less the same as 'to die', so that one can speculate about the introduction of a covert event argument in its valence. However, not all instances of 'thetic' individual level predicates are liable to this kind of double reading. Consider the following examples and their English translations:

(2-45) Serbo-Croat, VS order

What's wrong with my looks? *Velika ti je GLAVA.*

big to-you is head

'Your HEAD is big.'

(2-46) German, prosodic inversion

Why don't you like Cologne? *Die STRAßEN sind häßlich.*

'The STREETS are ugly.'

I cannot imagine an episodic, stage-level reading for predicates like 'be big' or 'be ugly' with subjects like 'the head' and 'the streets'. Nevertheless, 'thetic' sentences (2-45) and (2-46) are perfectly grammatical and felicitous. Furthermore, even when the predicate in a 'thetic' construction is of the stage level type, the impression of spatio-temporal boundness is not equally strong in all 'thetic' utterances.

(2-47) Modern Greek, VS order

Tu arese i MARIA.

to-him liked the Maria-NOM

'He liked Maria.'

(a) He spent his whole life alone. *Tu arese i Maria*, but he never had the courage to approach her.

(b) He was kind of restless the whole evening: *tu arese i Maria*, but he didn't have the courage to approach her.

If (2-47) is used in context (b), the verb has a punctual reading, and the impression of spatio-temporal boundness is fairly strong: he started to fancy Maria within the section of space and time determined by the context. On the contrary, in context (a), the verb has a complexive reading, and the feeling that the event of liking Maria is asserted with respect to a particular stage is absent.

I will have more to say on this and similar issues in the empirical part of this study. For the time being, suffice it to note that the type of phenomena I adduced here to illustrate the difficulties of the approach based exclusively on the assumed event argument (heterogeneous predicates, cancellability of the boundness effect) point out to pragmatics rather than to syntax or semantics. The intuition that VS sentences and related constructions are in many cases somehow deictically related to the present moment and the current scene, or to the contextually determined stretch of time and space, is doubtlessly correct. I am not sure, however, that this intuition is satisfactorily captured by the postulation of additional arguments and by the restriction of theticity to certain predicate classes.

A final note on Sections 2.5 and 2.6. As the reader will have noticed by now, the prevailing attitude to VS order and related constructions in the last twenty years or so has been the one I call thetic: it is widely believed that there is a primitive, not further reducible category situated on the level of cognition, semantics or pragmatics, and this category is directly expressed in grammar through a certain number of syntactically, intonationally or morphologically marked constructions.

2.7. Skeptics

The growing awareness of the importance of naturally occurring language data, characteristic for the Nineties, has brought some fresh air into the research of VS order. This is not to say that everything written in the past was based on three or four invented examples (though the number of studies using this method is considerable): what is meant is rather that the large stock of factual knowledge acquired in the past hundred years or so, together with innumerable contradictory interpretations, is now felt to be ripe to be tested on the language material itself. Some work in this direction was conducted by Ulrich for Rumanian (1985) and by Myhill for a number of unrelated languages (see 2.3) in the Eighties; Jacennik and Dryer's study of the Polish VS order (1992) represents an interesting application of Myhill's methodology. However, the real work on the assessment of the theticity issue and its relationship to VS order and similar constructions was undertaken by the creator of the universal thetic category himself: in the first half of the past decade, Sasse initiated a research of the VS construction in eight European languages (Latin, Italian, Spanish, Russian, Modern Greek, Hungarian, Romani and Turkish), the results of which are published in Matras and Sasse (1995).

The idea underlying the VS study was simple: a number of parameters usually associated with theticity and related notions was checked on the naturally occurring language material from the languages in which VS order is claimed to be the principal means of expressing theticity. The results were twofold: first, a couple of fine accounts of different types of VS constructions both for the languages in which the structure had been examined many times before (Russian, Maslova 1995) and for those for which there had been no comparable specialized studies (Romani, Modern Greek, Hungarian, Turkish – Matras 1995, Sasse 1995a, Schroeder 1995), and second, a general cross-linguistic account of formal, semantic and pragmatic features of VS order and its relationship to the notion of theticity, based on the material collected in the VS study (Sasse 1995b, 1996).

One of the major results of Sasse's crosslinguistic generalization is that there is no such thing as the unitary VS structure, or, the other way around, that VS order is employed in a number of semantically and pragmatically unrelated constructions. The most important VS types attested in more than one language are: inversion (a syntactic phenomenon of subject postponement due to the presence of another preverbal element), polarity focus, narrow subject focus, 'connective' VS type (the status of which is unclear), and finally a class of sentences labeled 'thetic'. Moreover, it is demonstrated that the existence and frequency of particular VS types is a language-specific matter, dependent on the overall grammatical structure in a manner more complex than Myhill's Greenberg-based generalization predicts. Thus Hungarian cannot have a VS narrow subject focus construction, since all narrow foci must occur preverbally in this language (Sasse 1995a); Turkish has only the 'connective' VS type, because it lacks inversion contexts and, more importantly, because the only postverbal elements allowed are 'given', referentially continuous referents (Schroeder 1995), etc.

Once the non-thetic verb-subject constructions are sifted out, Sasse goes on to explore the assumed universal properties of the 'thetic' VS type, which, although partially overlapping in various languages, turn out to be subject to dramatic crosslinguistic variation. Thus the monoargumentality constraint, the restriction of the number of arguments in a thetic sentence to one, seems to be absolute in Italian and Spanish, but virtually absent in Modern Greek. The restriction on definite subjects in VS constructions (known in English linguistics as definiteness effect in *there* sentences) seems to be relatively strong only in Russian, where definiteness is strongly associated with the preverbal position. In Hungarian, for instance, exactly the opposite holds, namely definite/referential subjects, if they constitute a part of the broad focus, have to be postverbal. Other languages are somewhere between these two extremes. A similar variability is also observed for other presumably universal features like

inanimateness and non-agentivity of subjects. The next step is to enumerate the functions thetic constructions typically assume in discourse: thetic sentences introduce new referents (introductory function), give background descriptions (descriptive), encode sudden events interrupting the main narrative line (interruptive), serve for utterances made out of the blue (annunciative) and identify presupposed but unknown states of affairs (explanative). The only common feature of these contexts is, according to Sasse, the low presuppositionality of the arguments, i.e. subjects. Even the presence and the grade of obligatoriness of these discourse functions is subject to cross-linguistic variation, with respect both to the construction type used in a language and to the idiosyncratic semantic and pragmatic restrictions a language displays. To illustrate the cross-constructive variation, I reproduce here a table from Sasse (1996: 43, slightly modified) which summarizes the distribution of three ‘thetic’ constructions (VS order, prosodic inversion and split constructions) over the five discourse functions mentioned above:

(2-48) *Functional ranges of alternative ‘thetic’ constructions*

	(X)VS order	prosodic inversion	split constructions
Introductory	+	–	–
Descriptive	+	(+)	–
Interruptive	+	+	+
Annunciative	+	+	+
Explanative	+/-	+	+

VS order seems to cover the widest range of discourse functions (although it is as a rule less automatically triggered by explanative contexts than the other two constructions, therefore the +/- sign in the table), while split constructions are restricted to more eventive types of ‘theticity’. If theticity were a unitary category, one would expect ‘thetic’ constructions to cover the whole of its functional area; since this is not the case, its categorial status is doubtful.

This becomes even more patent when one compares the lexical features of the predicates appearing in VS constructions. There is indeed something like a common core of ‘thetic’ verbs, comprising existentials and cryptoexistentials, verbs denoting sudden events, etc., or, more generally, verbs with stage level semantics and less agentive subjects. Verbs of this kind tend to occur in ‘thetic’ constructions in most languages, a fact derived from the discourse functions of these constructions: the enumerated semantic fields are the typical lexical solutions for presentation, description, interruption, etc. The problem is that there are dramatic differences in other semantic areas: some languages seem to be able to employ virtually every predicate in VS structures (Greek, Serbo-Croat), whereas in others the range of VS predicates

seems to be restricted only to a subset of the ‘thetic’ lexical core (Polish, Czech). Furthermore, some subject-verb combinations seem to be quasi-lexicalized in VS order in some languages, but not in others.

The practical consequence of these considerations is that the presumed universal thetic category turns out to be to a large extent dependent on the lexical structure and the type of interaction between lexicon and grammar in a language. It is possible for a state of affairs to be obligatorily expressed by a ‘thetic’ construction in one language and to be, due to the lexical semantics of the predicate expressing it, completely excluded from ‘thetcity’ in the other.

Research based on naturally occurring language data thus renders results much more complex than the ones that reveal themselves when one casts a deep intuitive glance on one’s own linguistic competence. The conclusion Sasse draws is worth quoting:

Is theticity a category? The answer is clearly no. It is a conglomeration of similar presuppositional/assertional conditions prevailing in similar semantic areas, which are frequently expressed by comparable constructions in different languages. (Sasse 1996: 51)

This loss of the categorial status, however, does not mean that the phenomenological area of ‘thetcity’ is a mere phantasm and as such best forgotten as soon as possible. In the concluding remark of his 1996 paper, Sasse notes that verb-subject combinations described as thetic in the previous research seem to represent ‘only one of several subpatterns of a more general pattern, which occurs under similar conditions with similar discourse-pragmatic effects’ (1996: 52). This more general pattern seems to refer to the mechanisms of creating semantically and pragmatically unified verb-noun combinations marking a broad scope of assertion.

These are also the general postulates on which I am going to base my study of VS order in the languages of the Balkans: first, a skeptic attitude towards the categorial status of theticity, and second, the assumption that the whole issue has something to do with the very general principles of presuppositional/assertional behavior and with the ways languages formally react to these principles.

Appendix: VS order in generative grammar

The systemic attitude to VS order, i.e. its reducibility to other principles, is – not unexpectedly – the most common departure point in generative grammar, though there are

some exceptions (e.g. Drubig 1992, Basilico 1999). Since the main point of this type of research is the question of syntactic derivation, the postulated principles underlying this order are formal in nature. Accordingly, the problems to be solved concern the formal restrictions posited in the current version(s) of the theory, and do not arise if one does not work within this framework, which is the case in the present study. For this reason, i.e. since the problems the generative descriptions try to solve are generally not the problems I am interested in, the overview of the extremely rich and ramified generative literature on VS order that follows is separated from the rest and is short, non-exhaustive and rather rudimentary. (For more detail see Drubig 1992, Bresnan 1994, Pinto 1997, Eguzkitza and Kaiser 1999, and Culicover and Levine 2001.) The few generative studies on the VS construction that take a more meaning-oriented perspective have been mentioned elsewhere.

Since generative grammar in its early decades was marked by a strong interest in English and a strong lack of interest in other languages, it is no wonder that the VS-structure, being only a minor phenomenon of the English grammar, received little or no attention at that time. Few early transformational accounts on English stylistic inversion (e.g. Emonds 1970, Aissen and Hankamer 1972) have left no visible trace in the subsequent theorizing on the topic. It was only with the increasing interest in Romance languages in the late Seventies and early Eighties, together with Perlmutter's Unaccusative Hypothesis (Perlmutter 1978) and with some theory-internal developments (Chomsky 1981), that brought the 'subject inversion' into the spotlight. The first really influential work on the topic, Kayne and Pollock's article on French stylistic inversion (1978), already addresses some of the issues that were to dominate the generative VS-research in the next decades, notably its relationships with the pro-drop parameter and with the wh-movement. In the subsequent years, the inventory of the problematic issues increased considerably, coming to comprise phenomena like case-assignment, unaccusativity, the presence of a separate focus (or topic) projection in a language, heavy NP shift, etc. The prominence given to one or the other issue has led to the crystallization of the two basic types of interpretation of the VS-phenomenon in generative grammar: the unaccusativity/case-assignment and the focus projection approach.

The former interpretation is concentrated on the question of the obligatory presence of the structural subject in the clause, i.e. the obligatory occupance of Spec,IP (Extended Projection Principle, Chomsky 1981) and on the way the subjects receive case-assignment, since verb-subject sentences (mostly in Romance) seem to violate both the EPP and the case-assignment rules. The task of the theory is to account for these violations by reducing the properties of the inverted structures to the postulated rules and thus fit them into the architecture of the

assumed universal grammar. The early work on these topics (Saltarelli 1981, Burzio 1981, 1986, Rizzi 1982, Belletti 1988) tried to solve the problem by capitalizing on the analogy of expletives in Germanic with subject inversion in Romance. Roughly, it is the rich morphology of many Romance languages that is considered to be responsible for VS order. The inversion is claimed to be only apparent, the real structure being the same as in English or German expletive constructions. Namely, expletives (*there* in English, *es* in German, *er* in Dutch, *il* in French) are assumed to satisfy the EPP and thus force the subject not to raise to Spec,IP. In pro-drop languages, Infl has pronominal properties and therefore allows for pro-dropping. It is assumed that Italian and Spanish also have an expletive in the form of a dropped *pro*, i.e. an element without morphological substance, but with all the syntactic properties of an expletive. These properties enable it to satisfy the EPP and license the non-raising of the subject, just like the overt expletives in Germanic and French do. The subject itself is somehow right-adjoined to the VP and gets its case assigned by a special case-transmission from the pro-dropped expletive.

Safir (1985) extends the analogy: the definiteness effect, i.e. the exclusion of definite subjects from expletive constructions in English and French (Milsark 1974, 1977), is supposed to be present in Romance inversion, too. Burzio (1981, 1986) refines the analysis by using the Unaccusative Hypothesis: there is not only one, but two positions of the subject in covert expletive constructions. With unergatives and transitives, the subject can only occupy the commonly assumed right adjoined position; with unaccusatives, it can be both in its base-generated position and right-adjoined. An influential paper by Belletti (1988) represents the final point: both definiteness effect and unaccusativity are accounted for by one single syntactic device, namely by the stipulated capability of unaccusative verbs to assign case to its sole, VP-internal arguments. This case is called partitive and restricted to indefinites (the nomenclature reveals a strong influence of the Finnish morphology on this analysis). This line of research has one serious flaw: empirical inadequacy. It is not only languages with rich morphology that license inversion, the definiteness effect simply does not exist in the Romance inversion of the Italian and Spanish type and in English locative inversion¹, and the restriction to unaccusative verbs seems to be unwarranted (for a more detailed critique, see Levin and Rappaport Hovav 1995, Pinto 1997; for a more fundamental critique, see Bresnan 1994). Nevertheless, the research conducted in the last years shows that the basic idea has still not been given up: different movement analyses have been proposed (Kayne and Pollock

¹ In fact, the definiteness effect has only a limited holding even in English *there*-existentials and in French stylistic inversion: see Ward and Birner 1995, Ward 1998, and Etchegoyhen and Tsoulas 1998.

1998, Eguzkitza and Kaiser 1999), and a subtype of subject inversion, the so-called light inversion, has been singled out, in which the restrictions proposed by Burzio, Belletti and others are said to be valid (Culicover and Levine 2001).

The second approach to the VS-phenomenon is strongly influenced by the notion of discourse configurationality, developed during the Eighties as a method of giving a configurational analysis for languages like Hungarian and Basque (Horvath 1985, É.Kiss 1987, 1994, 1998a, Brody 1990, Ortiz de Urbina 1989). The basic idea is that certain syntactic phenomena are sensitive to pragmatic/semantic features carried by constituents, notably [+focus] and [+topic]. If a constituent is marked by one of these features, it has to move – either to an independently defined position (e.g. Spec,CP in Brody 1990 and Drubig 1992), or to a special functional projection designed to satisfy the feature in question (e.g. FP in É. Kiss 1994, Tsimpli 1995, 1998, Barbosa 1998, TP in Basilico 1999).

There is a number of ways discourse configurationality, thus defined, is supposed to interact with the VS order. If one wants to make the focus feature responsible for the movement, one can say that the subject carries the feature [+focus], which then forces it to move to the appropriate position in order to be checked (Tsimpli 1995, Barbosa 1998); or one can postulate complicated movements of other constituents which then allow the subject to be placed in its feature-checking position, which results in the surface VS order (Ambar 1992, 1998, Samek-Lodovici 1998). If one wants the topic feature to trigger inversion, then topical subjects move to a preverbal position where this feature is checked (TP or Spec,IP), whereas the subjects that are not topical (whether [+focus] or not) stay in the base-generated position or move somewhere within the VP, rendering thus a VS order (Drubig 1992, Basilico 1999).

A somewhat different approach partly based on discourse meanings operates with the notion of the Nuclear Stress Rule (NSR): the nuclear sentence stress is assigned to the most embedded node and can project to all nodes to the left of it in a right-branching structure. The impact this kind of rule can have on the interpretation of the VS-structure is predictable: in order to comply to the demands of the NSR, the subject can remain in its base-generated position, provided that the EPP is satisfied in some other way – by a covert or overt spatio-temporal argument or otherwise. Different versions of this type of explanation are to be found in Pinto (1997), Costa (1996, 1998), and Zubizarreta (1998). In all of these interpretations, discourse notions like topic and focus seem to play the decisive rule. On closer inspection, it becomes clear that most of them have very little to say about the meaning of the inverted structures they deal with: topic and focus are reduced to abstract features, stipulated only in order to rectify the postulated syntactic operations through the checking theory or the NSR.

In sum: no matter whether the purely formal or the pragmatically motivated formal approach is assumed, most of the generative research on the phenomenon of VS order has one feature in common, namely the attempt to reduce the inversion to one or the other syntactic regularity holding also in other constructions. The meaning of the construction is in general either completely neglected or very roughly defined only to be used as a basis for syntactic generalizations. I shall therefore make use of the results of this line of research only to a very limited extent.

3. Aims and announcements

Now that the excessively long review of the results and the problems of the previous research on the topic has hopefully fulfilled its purpose, that is, to create a background knowledge on the part of the reader regarding the questions of VS order and theticity, I may try to depict the purposes of this study in some more detail. This explains the first part of the extravagant title of this section, ‘aims’, which will, for the sake of dramatic presentation, be formulated as questions. The second part, ‘announcements’, refers to the second component of the section: apart from questions, I shall also give approximate, rough and of necessity apodictic answers. That is, I shall formulate the problems connected with VS structures in the Balkans and, in a manner of introduction to my own way of thinking on these problems, give a rough sketch of my solutions to these problems. The latter primarily in order to render the reading of the present study easier for the reader, but also in order to avoid the irritating compositional principle of suspended solutions: I do not think that a linguistic study should look like a thriller.

3.1. Forms and functions

The underlying axiom of my approach to VS order is that word order phenomena in free word order languages, when they carry meanings (which is not always the case), are to a great extent driven by discourse-pragmatics. This is tantamount to saying that the phenomena under discussion will be explained in terms of information structure, with presupposition, assertion, topic and focus as the basic notions (cp. Section 4).

The primary question I ask is:

(3-1) How many VS constructions are there in the languages under consideration?

A construction is a mapping of a particular form and a relatively general meaning. This general meaning may have more interpretations, or the other way around, it can represent the common denominator of a number of more specific meanings. This is the reason why I shall not start from meanings but rather from forms. My first step will be to establish how many forms with VS structure exist in the languages of the Balkans. In order to achieve this, I shall use intonational patterns, pragmatic and semantic properties of the elements of the constructions, differences in syntactic restrictions holding in various constructions, etc. After

formally distinct VS categories have been established, I shall represent their general meanings and the mechanisms through which specific meanings, or interpretations, are deduced from the basic, general meaning.

There are three major VS constructions, represented in all three languages I deal with. Capitals in my designations of the constructions mark the position of the sentence accent. The major constructions are:

(i) **vS-construction**

(3-2) Albanian (Kadare, *Ëndrra martuese*, p. 146)

Në jetën tonë po hynte DËBORA.

in life-the our PTC came-in show-the

‘The SNOW was entering our lives.’

The sentence stress is placed on the subject; the verb is deaccented. The number of elements allowed between the verb and the subject and after the subject is restricted. This construction marks the focus domain comprising minimally the subject and maximally the verb-subject complex. Whether other elements may be included in the focus domain is a question I shall try to answer in Chapter 11.

(ii) **VsX-construction**

(3-3) Modern Greek (Kapandai, p. 48)

Ke PROSPAΘUSE o Petros mes ston polemo na vri XRONO ...

and tried the Petros middle in-the war to find time

‘And Petros tried to find time in the middle of the war ...’

The subject is obligatorily deaccented, the verb and the element X carry the sentence stress. The presence of further element(s) after the subject is not obligatory, though common. The subject tends to be directly postverbal, but there is some variation among the languages in this respect. The subject codes a continuous, ratified topic referent. The assertional structure is either broad focus on the verb and the X element or polarity focus on the verb.

(iii) **Inversion**

(3-4) Serbo-Croat (Andrić, p. 114)

Gdje mi je BILA pamet?

where to-me:CLIT is been reason

‘Why didn’t I think of it?’

Inversion is triggered by a mixture of discourse-pragmatic and purely syntactic constraints. As a discourse-pragmatic device, it occurs in presuppositional contexts and is designed to distinguish between different types of presupposed information. As a purely syntactic matter,

it is triggered by the existence of certain sentence elements which have to be both clause-initial and verb-adjacent. As a result, everything else, including subjects, has to be placed after the complex formed by these elements and the verb. Note that the term ‘inversion’ is often used in a broader meaning, to denote all instances of postverbal subjects; the restricted use of the term in this study is a terminological convention borrowed from Sasse (1995a,b).

This classification is of course only the first step, since it raises more questions than it is able to answer. The first group of questions pertains to the formal peculiarities of Balkan languages which allow for the existence of apparently completely unrelated constructions with VS order:

(3-5) How do different types of VS order fit into the word order systems of Balkan languages, or, the other way around, what does the exploration of VS structure contribute to the general understanding of the word order systems in these languages?

(3-6) How are formal and statistical differences among different VS patterns in the three languages under consideration to be explained with respect to the differences in their overall structure?

Though almost synonymous at the first sight, these two questions demand different answers. I shall try to account for the apparent polysemy of VS order by resorting to a kind of item-and-arrangement sentence model, i.e. by postulating a sentence structure consisting of a number of pragmatically defined slots. This sentence structure will be represented as a template comprising all possible slots. Very roughly, my basic findings are that the modern Balkan languages have postverbal focus domains with sentence stress on the last argument of the verb, which accounts for the vS-construction, and at least two topic slots, one preverbal and one directly postverbal, the latter of which is the basis for the VsX-construction. Inversion is explained by two factors: by a similarity between the markings for integrated and non-integrated readings in assertional and presuppositional contexts, and by a purely formal adjacency constraint holding for certain types of sentence elements.

Question (3-6) presupposes that there are formal and statistical differences in VS structures of the Balkan languages, an assumption abundantly confirmed by my data. Some of these differences are to be accounted for by minor variations in sentence templates, but this is only a part of the story. It is for this reason that I used the phrase ‘general structure’ instead of ‘general grammatical structure’ in formulating (3-6). Apart from syntactic parameters, there are parameters of variation which have to do with the lexical structure and with referentiality phenomena. Lexical structure, the way I shall use the term, refers both to the internal semantic structure of lexemes and to the phenomena of argument selection. For instance, all three

modern languages have two basic ways of expressing existence, a construction with impersonal ‘have’ and one or more verbs meaning ‘to be, exist’, but the distribution of these two expression types is different in all three, with the use of the ‘have’ construction being most constrained in standard Serbo-Croat (chiefly with mass nouns and plurals) and virtually semantically unconstrained in Albanian. This has a bearing on the frequency and semantic types of subjects licensed in the vS construction in these languages, since existence is one of the prototypical semantic fields of this construction. The number of similar lexical idiosyncrasies is rather high; I hope to be able to capture at least some of them. The same holds for various referentiality effects derived from word order phenomena.

3.2. VsX-construction

At first sight, this construction is less problematic than the other two, an impression certainly having to do with the fact that it was, probably due to its complete absence from languages like English, not subject to so many interpretations. Apart from the commonplace claim in Slavonic studies that this VS construction sounds somehow ‘folkish’ (e.g. Adamec 1966, Krylova and Khavronina 1986), Myhill and Givón’s (2.3) theory that the function of VS order is to express narrative continuity is the only interpretative attempt aiming at some generality. It is not per chance that more fine-grained studies of the construction, like Matras (1995) for Romani and Valioui (1994) and Sasse (1995a, 1996) for Modern Greek, deal with Balkan languages: VsX is, in at least one of them, a very prominent and frequent sentence type. The first question, given the pragmatic orientation of the present study, is:

(3-7) What presuppositional/assertional structure does VsX-construction encode?

The *raison d’être* of this construction is a postverbal topic slot in the languages of the Balkans, which is the preferred position for referentially continuous, ‘ratified’ (Lambrecht and Michaelis 1998) topics occurring in a changed discourse frame. The rather undetermined term ‘change of discourse frame’ may refer to the change of scene, of spatio-temporal frame, perspective, discourse universe, etc. This kind of topic – referentially continuous, discontinuous with respect to the discourse frame – is compatible with two assertional structures: broad focus, comprising the verb and optionally one or more of the elements occurring postverbally, and polarity focus structure, in which the positive or negative polarity of the whole proposition is asserted via narrow focus on the verb. Obviously, I take the

discourse structure to be crucial in the definition of the meaning of VsX-construction. The following question is the logical consequence of this attitude:

(3-8) What discourse functions does VsX-construction perform?

The two assertional structures behave differently in discourse. The broad focus VsX structure is generally employed in narrative texts and has four basic functions. The first is to give background identificational or descriptive information on major participants in the narrative, the second, to mark a change of spatio-temporal frame within the referentially continuous topic chain. The third function is to resume the interrupted narrative chain if the basis of this chain is a major participant in the narrative. Finally, in interactive contexts, it may be used to produce what I shall call reactive effect. If VsX-construction is coupled with polarity focus, this structure either fulfills the two typical functions of this focus type, the corrective and the cataphorically adversative one, or covers a range of illocutionary acts, from reassurance to threat. Only the first and the last functions seem to be exclusively linked to VS order, all others being expressible with other constructional types of polarity focus as well. This brings me to the next question:

(3-9) Are there syntactic contexts characteristic of VsX-construction?

The answer is yes. Both focus types occur almost exclusively in main clauses.

(3-10) Are there lexical, syntactic, etc. restrictions to VsX-construction, comparable to vS-construction?

At the first sight, nothing of the kind is the case: VsX-construction seems to allow for a range of lexical items and argument structures as wide as that of the canonical SV sentences. On closer inspection, however, a strong variation among the languages may be established, with MG using VsX with all predicate types and all kinds of subject referents, Alb. and SC mostly with the copula and the pronominal subjects.

As indicated above, one of the major concerns of the present study is to determine language-specific discourse preferences which are not triggered by the formal properties of languages. In the case of VsX-construction, this task is intermingled with sociolinguistic considerations. In one of the languages of my sample (Modern Greek), this construction is considered to be a feature of colloquial style, and as such avoided in all kinds of official writing. In others (Albanian, Serbo-Croat), it is felt to be a feature of chosen style. There are also some differences in the frequency patterns and in the sociolinguistic status between the broad focus and the narrow focus types. The question is thus formulated as follows:

(3-11) What preferences do the languages under consideration display as to the use of broad and polarity focus VsX-constructions? In what relationship do the sociolinguistic factors stand to these preferences?

The answer to this question is somewhat surprising. In Alb. and SC, the polarity focus VsX clauses seem to be a productive pattern in discourse, whereas the broad focus clauses seem to be in the process of dying out. In MG, both types are equally frequent and productive.

3.3. Inversion

Inversion is the only phenomenon described in the present study which has attracted the attention of generative linguists working with Balkan languages, who tend to explain it in terms of adjacent functional projections on a deep level of derivation (cf. e.g. Tsimpli 1995, 1998, Motapanyane 1997, Alexiadou 2000, and the references quoted therein). Since I work with a flat, template-based model of syntax, this kind of argumentation will find little echo in the present study. Now, inversion is, as I defined it in 3.1, partly a formal matter having to do with positional idiosyncrasies of certain types of sentence elements, and partly a construction triggered by discourse-pragmatic considerations. This is a challenging constellation: for a full understanding of the interaction between pragmatics and syntax, one has learn how to keep apart the meaningful and the purely formal variation. In order to achieve this, one has to find criteria to discern between these two types of variation. The first task in my dealing with inversion is thus to describe it:

(3-12) Which sentence elements trigger inversion, how obligatory is it with various elements, and do all inversion triggers behave in the same way?

There is a great deal of cross-linguistic variation here, regarding both the identity of inversion triggers and the obligatoriness of inversion. The only common inversion-trigger in all three languages are fronted quotes. The rest is a heterogeneous group comprising interrogative words, fronted focus expressions, some (but not all) subordinators and relative pronouns. Clauses with the first three elements are classified to a larger group labeled *focal inversion*, those with the last two to a group named *subordinate inversion*. Languages also vary as to the extent to which inversion is obligatory with similar inversion triggers. For instance, some the subordinator *ja na* in MG triggers inversion much more regularly than its Alb. counterpart *që të*. Finally, there seem to exist differences in syntactic behavior of certain inversion triggers even within one language: for instance, sentence-initial focus expressions in Modern Greek do

not cause exactly the same type of inversion as, say, the subordinator/modal particle *na*. All this means that the principal diagnostic test for inversion is the presence of certain elements, which have to be defined separately for every language, in the clause-initial position. But this is not enough. Inversion often occurs, especially in SC, but also in Alb., and, seldom, in MG, when there is not obligatory adjacency between the clause-initial element and the verb. The reason for VS order in such cases is presumably discourse-pragmatic in nature. The question is:

(3-13) What discourse-pragmatic principles regulate inversion in addition to inversion-triggers?

I shall try to demonstrate that the occurrence of inversion in the syntactic contexts in which the inversion-trigger does not require the adjacency of the verb has to do with distinguishing degrees of the ratified status of the sentence elements. The opposition ratified vs. non-ratified, i.e. marked as easily retrievable for the hearer vs. marked as potentially contestable, is characteristic for presuppositional contexts. One of the discourse-pragmatic rules operative in this type of context defines the clause consisting only of ratified elements as the one in which the sentence stress is placed on the verb which is not in the clause-final position. Inversion occurs when the subject is postposed in order to ‘cover’ this position and thus prevent the verb from ending up there. The presupposed material in a sentence thus seems to follow rules similar to the assertional material, although there are differences, as will become patent in the course of the present study.

3.4. vS-construction

As has probably become clear by now, vS-construction, which extensionally covers sentence types usually labeled ‘thetic’, is by far most controversial, most complex and potentially most interesting. In what follows, I list the major questions arising both from the relevant literature (cf. Section 2) and from my corpus data, and give the first approximate answers.

(3-14) Is theticity a category?

My answer to this question will be unequivocally no. The reasons for this are too complex to be elaborated upon here (for a full account, see Chapter 11, especially 11.6.), but some of them are indirectly mentioned in the following question-answer pairs.

(3-15) On what level of linguistic description is vS-construction to be accounted for? Is it a pragmatic, semantic, lexical, or purely formal phenomenon?

As indicated above, I believe vS-construction to be primarily a pragmatic phenomenon, or more precisely, a phenomenon of assertional structure. The verb and the subject are either construed as a unified broad focus construction, or the subject is narrowly focused. The assertion conveyed by the verb-subject complex or by the subject focus is either related to a non-subject topic, a direct or indirect object frequently coded as a pronoun, or to an indirect topic, which may, but need not be overtly coded (for the notion of indirect topic, cf. 4.3.3. and 4.5.2.). On the other hand, as has probably become clear from the state of the art report in Section 2, vS-constructions show cross-linguistically a non-trivial affinity to certain lexical classes of predicates, to a certain field of sentence semantics, to certain types of argument structure, etc.

(3-16) What lexical types of predicates are prototypically employed in vS-construction, and why?

(3-17) What kind of semantic relationship between the predicate and the subject prevails in vS-construction, and why?

(3-18) Is there really a restriction as to the number of arguments allowed within vS-construction, and if so, why?

(3-19) What is the preferred semantic role of the subject? If there is such a role, why is it preferred?

(3-20) Is there a tendency for the subject to be new to discourse? If so, why?

The reader probably recalls that existential semantics (in broader sense of the word) has played an important role in the discussion of theticity, that subjects were claimed to be obligatorily new to discourse and preferentially non-agentive, that verbs and subjects often stand in some sort of ‘lexical solidarity’ relationship (Coseriu 1967), that it has been repeatedly claimed that only one argument (i.e. the subject) is allowed in vS-clauses, etc. These and similar claims, amply evidenced in Section 2, have in the last decades led to the postulation of separate predicate classes with specific argument structures that somehow seem to trigger ‘thetic’ constructions (I mean unaccusative verbs and stage level predicates). I could not agree less. In fact, I shall try to demonstrate that all formal and semantic restrictions and tendencies observable in vS-construction are derivable from its basic pragmatic meaning, that of conveying an assertion in which the assertional material is composed of the predicate and the subject, or of the subject alone. I shall also try to show, along the lines proposed in Sasse (1996), that the impression that theticity is a category is only an impression, caused by a cross-linguistic tendency to use certain types of predicates with certain presuppositional/assertional structures in certain discourse surroundings. I shall push this line even harder than

Sasse, and claim that what has been considered to be restrictions of vS-structures are in fact merely pragmatic tendencies which can be cancelled under appropriate pragmatic conditions. It will be shown that the question of assertional scope *within* the vS-construction, i.e. that of the broad or narrow focus reading of this construction, sometimes claimed to be lexically determined (e.g. Pinto 1997), is in fact a matter of the frequent use of certain predicates, or certain topic types, with broad or with narrow focus, i.e. a statistical, not a categorical matter. All these questions taken together bring me to another, more complex one:

(3-21) What kind of unity does the vS-construction represent?

Recall that many scholars have put forward a claim that ‘thetic’ constructions somehow represent a unified piece of information, though there is no consensus as to the nature of this unity – it is sometimes considered to be a semantic, sometimes a pragmatic phenomenon, but the whole issue is usually dealt with in a highly vague manner. From the viewpoint of vS-construction, the unity question clearly concerns only one of its readings, the broad verb-subject focus one. This is the starting point for my answer: I shall elaborate upon Fuchs’s idea that the relationship between the subject and the verb has to be in a way non-controversial for the given context (cf. 2.4.), and try to embed this idea into a wider context of broad focus domains’ formation.

Oversimplifying significantly, I shall claim that for a noun and a verb to be able to form a broad focus, independently of the grammatical relation carried by the noun, the speaker has to consider the internal relationship between them non-controversial. Or, from a different point of view, speakers use as the assertional material in broad assertions only those verb-noun combinations that are less probable to be internally contestable. The internal non-controversality of the assertional material enables the hearer to process the external relationship of this material to the topic referent instead of concentrating on the internal structure of the assertive material itself. In other words, the unity that exists between the verb and the subject in the broad focus reading of vS-construction is triggered by pragmatics, but this pragmatic motivation has some important consequences for the lexical choices and the overall semantic structure of the proposition which speakers choose to convey.

It has often been noted (cf. 2.5.2 and 2.6.3) that ‘thetic’ sentences are somehow deictically bound to a stretch of space or time determined by the utterance itself or by context, that they are based on perception, or similar. The question is:

(3-22) Where does the effect of spatio-temporal boundness, or the perceptual effect, carried by many vS-sentences, come from?

I should once more like to make the presuppositional/assertional structure of vS-construction responsible. Many vS-utterances carry assertions about indirect topics in the form of situations or locations (cf. 4.3.3. and 11.6.). These implicit superordinate indirect topics are more often than not deducible from the context only by relating them to a particular time or space, or both. Yet another phenomenon of time reference has been often associated with vS-construction and often treated simply as the other side of its assumed time-space boundness:

(3-23) Why are generic or habitual statements, or simply statements referring to lasting properties of individuals, seldom conveyed by vS-sentences?

Generic, etc., statements are not absolutely banned from vS-construction, as I indicated in 2.6.3, but they doubtlessly very seldom occur in this syntactic disguise. I think that this issue should be interpreted separately from the frequent space-time boundness of vS-sentences. The latter is, as indicated, a by-product of the topic types often employed in these sentences, whereas the avoidance of generic, etc., statements has rather to do with the general principles of communicative behavior: If something is topicworthy enough for the speaker to make a statement concerning its permanent properties, then this something is generally topicworthy enough to be coded as direct topic. Since direct topics are almost as a rule subjects, and since direct topicality in the languages under consideration is coded primarily with sV order (or, under certain circumstances, with VsX construction, see below), vS construction is in most cases not the option speakers choose to express this kind of statement. With this pragmatic explanation, I am approaching the last question I have to ask about vS-construction:

(3-24) What discourse functions does vS-construction fulfil?

I have already announced that I shall contend that the proper location of ‘theticity’ is the interface between lexicon, information structure and discourse. In other words, discourse surroundings in which vS-constructions occur are the ultimate answer to the question why such a construction should exist at all. The basis of my work in this direction is Sasse’s (1996) classification of ‘thetic’ contexts, which I shall reformulate so as to show two basic types of discourse contexts for broad assertions conveyed by vS-sentences. According to different topic types, I shall classify five functions identified by Sasse into two groups. In the first group, comprising Sasse’s introductory and descriptive functions, the assertions are made with respect to an indirect locational topic. In the second, comprising the remaining functions (annunciative, explanative, interruptive), a relation is established between the asserted material and a situation.

What locations and certain types of situations have in common (for definitions of these terms see 11.6.) is that they are not, or only seldom, coded as subjects in the languages of the

AME type (for locations cf. Bresnan 1994). This means that, in order to make an assertion about these discourse referents, speakers have to use constructions in which subjects are included into focus domain and are not under the presupposition of topicality – in the languages of the Balkans, it happens to be the vS-construction, due to language-specific properties of their word order systems.

But there is more to this: speakers generally use non-subject referents as topics only when subjects are less topicworthy than these referents (see 4.3.1 and 4.3.2 for some elaboration of the notion of topicworthiness). This is typically the case when the subject is non-agentive and/or inanimate, when it is indefinite, or when it for any other reason does not make the intentions of the speaker in the given stretch of discourse optimally transparent. Thus, many lexical and sentence-semantic properties of vS-sentences are derivable from the basic discourse functions they perform. However, languages may differ with respect to the preferred information structure in certain contexts: thus it may be the case that, say, an animate subject, even though it is not the optimal topic of the sentence, is coded as one, simply because the language in question has a strong conventionalized tendency to use animate referents as topics rather than anything else. This is the most interesting part of my work: to try to find microstructural discourse preferences of the languages I deal with.

Despite my aversion towards the compositional principles of a detective novel when they are applied to writing linguistic studies, I feel it would be premature to give an outline of the concluding remarks of the present study here. Suffice it therefore to say that the major general outcome of my research is that the word order systems of the languages of the Balkans are a product of an interplay of pragmatic, semantic and purely formal factors. The fact that in certain discourse contexts certain assertional structures and certain semantic classes are often used may lead to the formation of semantic and syntactic restrictions in some languages, and thus suggest the conclusion that we are dealing with specific cognitive categories which are directly reflected in syntax and semantics. But this is only an impression.

4. Information structure: presupposition, assertion, topic and focus

“Terminological profusion and confusion, and the underlying conceptual vagueness, plague the relevant literature to a point where little may be salvageable.” (Levinson 1983:x)

My line of explanation of VS order in the Balkans will be, as indicated in Section 3, phrased out in terms of communication strategies and information structure. In other words, I shall be working with the notions of presupposition, assertion, topic, focus, contrast, etc., all of which are notorious for having meanings fluctuating from one scholar to another, to the extent that the opinion that one should abandon the whole research field, as exemplified in the introductory quote, is in no way exceptional. However, despite my sympathy for the air of despair in Levinson’s characterization of the topic-focus articulation (based on my own personal experience with the relevant literature), I do not share his pessimism. In order to avoid terminological confusion and conceptual vagueness, I shall dedicate the present chapter to clarifying the basic concepts which will be used in the rest of the study. A brief sketch of the pragmatically rooted model of human communication I subscribe to is followed by a concise discussion and explanation of the principal notions of my descriptive model, which draws heavily on the ideas of Stalnaker (1974, 1978) and Lambrecht (1994).

4.1. The incremental model of communication

Let me start with some truisms (which have become truisms only after the publication of Grice 1975 and 1978). Communication is an activity, and as such goal-oriented and intention-driven. It is a cooperative enterprise, demanding at least two parties willing to participate in it and contribute to it. Its principal purpose is to achieve a consensus between the interlocutors on how the world is. Each of these points will be explained in what follows.

First, the last point. Utterances carry propositions, conveniently described by Stalnaker (1978: 316) as ‘representation[s] of the world as being a certain way’; these representations, if accepted by the interlocutors, reduce the number of ways the world can be (possible worlds), and contribute thus to the creation of the mutual understanding of how the world actually is (for the interlocutors). In a way, utterances eliminate uncertainty, or ignorance, by eliminating alternatives: each utterance you accept from me adds a new partial description of the world to your stock of descriptions, removes your previous conceptions of the world which are at

variance with this description, and brings us a step closer towards identical representations of the world.

Now, people generally do not enter communication as *tabulae rasae*: participants in a communication have already gathered, from the previous communicative acts they were involved in, a vast amount of representations of the world, or propositions, which they bring to the communicative act as a dowry. This is called encyclopedic knowledge, or world knowledge. Furthermore, the communicative act does not take place in an abstract, isolated space: the immediate environment supplies us with further representations, providing the interlocutors with situational knowledge. Let us call these two types of knowledge, both independent of the current communicative act, background knowledge. The background knowledge stocks of the interlocutors, due to the huge but limited number of experiences human beings are capable of having, usually partially overlap, meaning that a considerable number of the representations of the world is shared by the speaker and the hearer. Let us call this shared subset of the set of background knowledge propositions common ground. The greater the common ground, the smaller the set of possibilities, i.e. of the ways the world could be. Propositions accepted during the communication are added to the common ground, i.e. they become part of the background knowledge.

The definition of the purpose of communication can now be reformulated: Interlocutors aim at enlarging the common ground to the extent that they reach the level of having the same set of propositions, or entertaining the same set of beliefs, with respect to a particular section of the universe.

To say that every new accepted proposition is added to the common ground is simply another way of saying that the state of knowledge of the interlocutors is constantly changing during the communication, i.e. that the level of the mutual consensus on the way the world is grows gradually with every utterance. This is why this model of communication is called **incremental**: it defines communication as a step-by-step progress towards mutual knowledge. The details of the actual process will be dealt with in the following sections.

Before I turn to these (essential) details, however, a short assessment of two further truisms mentioned at the beginning of this section in the light of the incremental model is in order. The idea that communication is goal-oriented and intention-driven means within the incremental model that the speaker confronts the hearer with ever new representations of the way the world is (propositions) with the intention to form her/his background knowledge in a certain manner. It is self-evident that the choice of the communicated propositions is determined by this goal. But it is not only what is said that makes the hearer accept the desired

and abandon the undesired representations of the world: it is also how it is said, and, no less, what is not said.

This is where the notion of cooperativity (first formulated, once again, by Grice 1975) enters the scene. What is said – explicitly expressed propositions – is only the tip of the iceberg (Sperber and Wilson 1986, Carlton 1998, 1999), or, better, the starting point, both in encoding and in decoding processes. The speaker starts with the assumptions about the state of the hearer’s knowledge, i.e. with the assumptions about the scope of the common ground, and with the assumption that both s/he and the hearer are aware of a number of social principles which define the cooperative communicative behavior, the main ones being to communicate no more and no less than necessary¹. Now, this means that the proposition to be added to the common ground must contain the right measure of descriptive content: to assert what is already assumed to be known or not to assert what is assumed not to be known either leads to misunderstandings or urges the hearer to pragmatically enrich the descriptive content of the communicated proposition so as to maximize its cognitive effects (Sperber and Wilson 1986), i.e. so as to make it capable of eliminating as many possible worlds from her/his background knowledge as possible. Thus the increase of mutual consensus has to be a cooperative effort of the interlocutors. The fact that the way the propositions are encoded plays an essential role in this process will be demonstrated in the following sections.

4.2. Presupposition and assertion

4.2.1. Basic issues

The importance of the notion of common ground is not exhausted by the description given above. Recall that the aim of communication is to achieve the level of having the same set of propositions *with respect to a particular section of the universe*. This means that the participants in a communicative act do not simply discard any old background knowledge propositions which do not correspond to the accepted new representation of the way the world is, but only those which might be a relevant alternative to the accepted proposition. In other

¹ The number of principles varies from author to author. Grice’s original four maxims and almost a dozen submaxims (Grice 1975) have been frequently subject to reduction, the most important being Horn’s Q- and R-principles (Horn 1984, 1989) and Sperber and Wilson’s (1986) monistic Principle of Relevance.

words, the set of possibilities among which the speaker wants the hearer to distinguish is not infinite but rather limited by the assumed common ground. Consider the following example:

(4-1) *He was sorry his father was ill.*

If a person utters this sentence, s/he instructs the hearer to discard the potential propositions ‘he was glad his father was ill’, ‘he did not care for his father’s illness’, etc., but not the potential propositions ‘he was glad to see his father’, ‘he was sorry his father was not there’, etc. Why is this so? The first set of propositions is judged against the same background knowledge as the uttered proposition, the second is not: ‘his father is ill’ belongs to the common ground of the interlocutors, ‘he saw his father’ and ‘his father was not there’ do not. In uttering a proposition, the speaker signals to the hearer to look for potential propositions which are based on the same common ground propositions as the uttered proposition and eliminate them, if they are incongruent with the new proposition. The domain of assertion is thus explicitly limited by the extent of the assumed common ground: the hearer has to process only that information which is relevant to a particular propositional content, or, as Stalnaker (1978:321) puts it: ‘(the set of the potential propositions determined by the assumed common ground) is the set of possible worlds recognized by the speaker to be “live options” relevant to the conversation’.

Common ground, i.e. the background knowledge assumed by the speaker to be shared by her/him and the hearer, is made up of propositions, i.e. of representations of the way the world is. To assume that a proposition is a part of the common ground means to presuppose it: in order to assert (4-1), the speaker first has to presuppose the proposition ‘his father is ill’, i.e. to presuppose that the hearer is somehow already familiar with the fact that his father is ill. If this is not the case, the hearer will not, or only with great difficulties, be able to judge the assertion ‘he was sorry about it’, since the domain of the assertion remains unspecified. Now, recall that communication is a goal-oriented, cooperative enterprise. If I, owing to an unsuccessful or underspecified presupposition, do not succeed in forming the background knowledge of my hearer, or if s/he has to invest to great an effort to adjust her/his representations of the world to mine, then the communication failed. In order to avoid this, I have to make my presuppositions accessible and as explicit as required (according to the principles of saying no more and no less than necessary). This has some nontrivial consequences for the form of the sentences I utter in order to convey my propositions: the proposition does not contain only the descriptive material which the hearer has to decode in order to understand my assertion, but also the descriptive material which shows to which part

of the common ground the assertion is to be related. So natural language sentences generally contain both presupposed and asserted material¹.

The notion of presupposition based on the common ground, however, is not without problems². For the sake of simplicity, I have treated it simply as the assumption on part of the hearer that a proposition is shared by the interlocutors. However, in a number of linguistic constructions unequivocally identified as presupposition-carriers (definite descriptions, factive verbs, clefts, etc.), it is not only propositions which can be reasonably assumed to be shared by the interlocutors that occur, but also propositions that are somehow ‘new’ (Prince 1978a, Lewis 1979, Grice 1981, Reinhart 1982, Lambrecht 1994, Delin-Oberlander 1995, Dryer 1996, Abbott 2000, etc.). Compare the following example from Grice (1981: 190). While discussing a concert, I can say:

(4-2) *My aunt’s cousin went to that concert.*,

though I know perfectly well that the person I am talking to is very likely not to know that I have an aunt, let alone that my aunt has a cousin. In spite of the hearer’s unforgivable ignorance, I use the definite description ‘my aunt’s cousin’, which is generally agreed to be a marker of presupposed propositions. In other words, I presuppose that my hearer knows the propositions ‘there is a person who is my aunt’, and ‘there is a person who is my aunt’s cousin’, although I know that s/he does not know them. There are, as far as I know, two lines of explanation for this.

¹ This functional explanation for the ‘old-new’ character of utterances was first explicitly put forward by Stalnaker (1974, 1978), and had an immense influence on such important approaches as that of Lewis (1979), Heim (1988), Vallduví (1992), Lambrecht (1994), etc.

² As has probably become clear by now, I use the term ‘presupposition’ as a shorthand for ‘pragmatic presupposition’, as defined by Stalnaker (1974), Horn (1986, 1989, 1996), Lambrecht (1994) and Abbott (2000), among others. The problem I am concerned with is the relationship between this, ‘pragmatic’ presupposition and the notion of common ground. What I am *not* going to deal with is the endless debate on whether presuppositions are pragmatic or semantic in nature. ‘Semantic presuppositions’ were very popular in the Seventies (for an excellent survey, see Levinson 1983), somewhat forgotten in the Eighties, and revived by Neo-Strawsonians like Burton-Roberts (1989), descriptive linguists like Mohanan and Mohanan (1999) and by people allowing for both pragmatic and semantic presuppositions like McCawly (1993) and Lambrecht (1994) in the Nineties. Consequently, the debate goes on with an undiminished fervor, as witnessed by some recent polemics like those between Carston (1998, 1999), Burton-Roberts (1999) and Seuren (2000), and between Mohanan and Mohanan (1999) and Simpson (1999). I also have little to say about the treatment of presupposition in formal pragmatics (see e.g. van der Sandt 1988, 1992). The prehistory and the early history of the notion of presupposition are nicely summarized in Donnellan (1981), Levinson (1983) and Horn (1996).

The first one operates with the notion of pretense and was first conceived by Stalnaker (1974): the speaker can pretend that the auditor already knows a proposition in order to achieve certain rhetorical effects; the auditor recognizes the speaker's intention to presuppose this proposition, and the communication is successful¹. The pretense account was further refined by Lewis's principle of pragmatic accommodation:

If at time *t* something is said that requires presupposition *P* to be acceptable, and if *P* is not presupposed just before *t*, then – *ceteris paribus* and within certain limits – presupposition *P* comes into existence at *t*. (Lewis 1979: 340)

That is, in uttering (4-2), although I know that the hearer does not share with me the propositions 'there is a person who is my aunt' and 'there is a person who is my aunt's cousin', I use a presupposition-carrying definite description, because I expect her/him to use the principle of pragmatic accommodation and somehow treat the two propositions as belonging to the common ground, i.e. to act *as if* s/he has already been aware of them. Heim (1988 [=1982], 1983) tried to state more precisely the part of Lewis's definition which refers to the limits of presupposition, relying on Clark's (1977) notion of bridging: 'When a new file card is introduced under accommodation, it has to be linked by crossreferences to some already-present file cards' (Heim 1988:373). The only propositions outside the common ground the speaker is allowed to presuppose are the ones that are indirectly linked to the propositions which were directly evoked in the previous discourse. Now, as is the case with the whole familiarity theory, the constraint seems to be too strong: I can start a text with a sentence like:

(4-3) *The first person to sail to America was an Icelander.* (Abbott 2000:1427),
and my auditor (or reader) will understand the presupposed proposition 'someone sailed to America first' even though it is not cross-referenced to any previously evoked proposition. The latest important extension of the pretense account is that of Lambrecht (1994: 65-73): as the starting point he uses the fact that certain presupposition-carrying constructions are used with pragmatic accommodation with a regularity which surpasses the frequency of *ad hoc* rhetorical effects. Thus *it*-clefts are often used as text-opening devices to refer to propositions which cannot be assumed to be partially shared by the interlocutors, and, more importantly, which do not seem to yield readily to the accommodation analysis:

¹ Stalnaker is to my knowledge the only theoretician to extend the pretense account to the cases of phatic communication: in saying 'filthy weather, isn't it', I pretend that the auditor knows less than s/he actually does, in order to establish a contact, or for any other reason.

(4-4) *It was George Orwell who said that the best books are those which tell you what you already know.* (Lambrecht 1994:71)

In order to explain this and similar cases, Lambrecht invokes the concept of conventionalized presuppositional structures: certain constructions are frequently used in certain text surroundings, so that they eventually become conventionalized in these text surroundings, even in the cases when one cannot speak of accommodation in Lewis's sense.

The pretense account has been criticized on two points (see especially Dryer 1996 and Abbott 2000). First, for its lack of comprehensiveness: not all instances of excessive presuppositions are used for rhetorical purposes, and not all of them have the air of pretense (a point already noted, but not explained, by Stalnaker 1974): thus, (4-2) does not seem to be rhetorically marked in any relevant way, and it does not demand any particular additional processing effort on part of the hearer. Secondly, for its potential vacuousness: if there are no clearly set limits to accommodation (Heim's constraint is, as indicated, too restrictive), then every counterexample to the common ground view of presupposition can be treated as accommodated without further discussion, and the theory is not falsifiable *qua* vacuous.

This is why the second line of explanation was devised: it was first sketched informally by Jerrold Sadock (quoted by Stalnaker 1974:202 as p.c.) and Grice (1981), then definitely shaped as a theory by Abbott 2000. I shall call it a nonassertional account. The basic idea is that presuppositions are not derivable from the common ground but rather from the decisions the speaker makes as to what s/he considers newsworthy and therefore important enough to be asserted, and what is noncontroversial and therefore not asserted. Thus, the speaker in (4-2) considers the existence of 'my aunt's cousin' noncontroversial, i.e. s/he expects the auditor to take it from her/him, as the auditor has no reason to doubt it; that someone sailed to America first (4-3) is not a proposition the hearer would be likely to doubt, given the fact that our encyclopedic knowledge tells us that for every repetitive state of affairs there is a time it takes place for the first time.

So far, so good. But why do the speakers presuppose in the first place, if presuppositions have nothing to do with the common ground and are thus unrelated to the way the propositions are added to the knowledge stock? Abbott (2000:1431) explains this apparent paradox in terms of the quantity of information which speakers are allowed to convey in one utterance: since human processing resources are limited, there is only one assertion allowed per clause, so speakers have to decide what is going to be the main point, which is then what they assert, and what is less newsworthy and more likely to be taken for granted by the hearer, which is what they presuppose. Since the propositions contained in the common ground are

most likely to be taken for granted, etc., they are typically presupposed, but the relationship between presupposition and common ground is a statistical, not a categorical one. On this account, even (4-4) receives an explanation in terms of the presupposition-assertion pair: from the two possible informational points, the propositions ‘the best books are those which tell you...’ and ‘it was said by George Orwell’, the speaker decides to start the conversation making the latter the main point and treating the former as something the auditor is to take for granted. Presuppositions are thus merely nonassertions.

The nonassertional account has a number of advantages: it offers a straightforward explanation for the presuppositions which do not belong to the common ground, it takes into account the processing capabilities of language users and emphasizes the creative side of communication with the speaker leading the hearer through communication by creating the main points, i.e. deciding what is to be asserted (see Section 4.5.). The deficiency of this approach in comparison to the pretense account (and the common ground account in general) is that it does not specify how exactly the presupposed propositions are incorporated into the hearer’s model of the world and how the asserted propositions are processed with respect to the background knowledge.

In other words, I would like to have both the incremental model of communication based on the notion of common ground and the nonassertional account of presuppositions. My way of achieving this compromise is based on certain elements of Wilson and Sperber’s idea of ordered entailments (1979) and Heim’s notion of context change (1983: 117-119). Presupposed propositions contained in utterances are indeed propositions we have to mention but choose not to assert, since our asserting capacities are limited. In uttering a sentence which conveys both a presupposed and an asserted proposition (as the overwhelming majority of natural language sentences do), I give a signal to the auditor to construct his background knowledge so, that s/he *first* incorporates the presupposed proposition into it, and only *then*, in relation to the background knowledge enriched in this way, which is now our common ground, add and evaluate the asserted proposition. When the presupposed proposition is already contained in the background knowledge, i.e. when it is already a part of the common ground of the interlocutors, the hearer simply has to recognize it as such and to relate the asserted proposition to the common ground it defines, eliminating those potential propositions which contradict it. When the presupposed proposition is not in the common ground, an additional effort is necessary: the hearer first has to add this proposition to his background knowledge. The common ground between the interlocutors is thus enlarged, and the asserted proposition is added to and evaluated against this newly reached common ground.

Thus, when a speaker utters (4-2), s/he signals to the auditor to add the propositions ‘there is a person who is my aunt’, and ‘there is a person who is my aunt’s cousin’ to her/his background knowledge and then relate the assertion to the field of possible worlds in which these presupposed propositions hold. If the auditor does not already have them in her/his background knowledge, s/he incorporates them into it (if they are accepted), and creates a new common ground between her/him and the speaker. The next step is to eliminate all the possible worlds in which ‘my aunt’s cousin’ did not go to that concert, whereby the asserted proposition ‘my aunt’s cousin went to that concert’ is also added to the common ground (of course, if it is accepted), whereby the mutual consensus on the way the world is increased.

To sum up: *to presuppose means to determine the order of assessing propositions conveyed by the utterance.* In this light, it is only natural that the propositions which are already contained in the common ground are usually chosen to be presupposed: the fact that they are already there makes it unnecessary or even counter-productive with respect to the goals of communication to assert them (recall the principles of cooperative communication mentioned in Section 4.1.). It is even more natural that the propositions which are judged noncontroversial – like the fact that my aunt has a cousin, or the fact that someone was the first to sail to America – are more readily relegated to presupposition than the ones whose content may be contested: if both the presupposed and the asserted propositions can be accepted or rejected (and this is one of the consequences of the model outlined above), it is more economical to process the one that is more likely to be accepted first, and judge the one that could be disagreed upon in relation to the common ground thus obtained.

Finally, if both types of propositions may be contested (which is the case in (4-4)), the speaker may either split the utterance in two or more assertions (e.g. if s/he considers a certain proposition to be too complex to be efficiently processed as a presupposition), or decide which of the two or more propositions s/he wants to go on record as contributing to discourse, meaning that s/he decides which one of the propositions is most newsworthy or relevant relative to the his/her communicative intentions and accordingly asserts only this proposition, presupposing the rest.

All this means that speakers’ behaviour with respect to presupposing vs. asserting is indeed free, i.e. dependent on the issues of intention and assumed relevance, but that this freedom is conventionally restricted by the principles of cooperative communicative behavior, more specifically by speakers’ assumptions about hearers’ processing capacities.

From this perspective, assertion may be defined in two ways. First, on the ordering view of presupposition, it is the last in the series of propositions added to the common ground by a

single utterance, and as such the most prominent one. Second, more content-oriented, it is a new relation added to the common ground so as to reduce the number of possible representations of how the world is, a relation established by adding a proposition to the common ground as formed by the background knowledge of the interlocutors, by the previous communication, and by the presuppositions evoked.

Unlike Abbott (2000:1432), I think that assertions have to be new (apart from rhetorically marked cases of phatic communication mentioned in note 1). However, ‘new’ should be understood in a somewhat broader sense here: if the speaker asserts something of which s/he knows that it is already in the common ground, s/he instructs the hearer to exploit the principles of communicative behavior and to draw inferences, or, in terms of Relevance theory, to pragmatically enrich the proposition, so as to *make* it ‘new’. To use the ever-quoted example:

(4-5) *It is so hot in here.*

If a person utters this sentence in an overheated room, s/he can be pretty sure that the auditor is also aware of the proposition ‘it is hot’, i.e. that the proposition is a part of their common ground. The auditor is also aware that the speaker has to be aware of the latter fact, and, relying on the principles of saying no more and no less than necessary, tries to figure out why the speaker should utter a familiar proposition. One of the most common solutions is that s/he is indirectly asking her/him to open the window, turn down the heating, or something similar; in any case, s/he is conveying a new proposition.

What is the use of distinguishing between presupposed and asserted propositions if in both cases the propositional content is or may be added to the common ground? Though I have already partially answered this question, I consider it worth while to elucidate it in more detail. Suppose you and I speak of a concert, and you state that you did not go there because you had some appointments you were unable to cancel. In response to your statement, I say:

(4-6) It’s a pity! *My aunt’s cousin went to the concert* and found it astonishingly good.

(4-7) Don’t be too disturbed. *The only one to go to the concert was my aunt’s cousin,* who is ready to cancel everything on short notice just to hear a good concert.

The propositional content and the quantity of added information (of eliminated possible worlds) is in both cases roughly the same, but the assertional structure is different. In (4-6), the presupposition is ‘there is a person who is my aunt’s cousin’, and what is asserted is ‘he went to the concert’; in (4-7), I presuppose ‘someone went to the concert’ and assert that it was my aunt’s cousin. According to the ordering account of presupposition, what I do in these two utterances is to give you two different sets of instructions in which order you are to add

the propositions to the common ground. In (4-6), you are first to accept the proposition about the existence of a person corresponding to the description ‘my aunt’s cousin’, and then to add the proposition ‘he went to the concert’ to the common ground already enriched with the representation of my aunt’s cousin. In (4-7), the order is reversed: you should first accept that someone went to the concert, and then add the identification of the variable ‘someone’ with ‘my aunt’s cousin’ to the common ground. These different orderings correspond to the different intentions I have in both cases: in (4-6), I want to contrast your not going to the concert with somebody else’s going there (where this person only accidentally happens to be my aunt’s cousin), and then increase your disappointment by quoting this person’s positive judgment of the concert; in (4-7) my intention is to single out a person who did go to the concert and, having done this, ascribe this enterprise to his extraordinary qualities as far as concerts are concerned, offering you a sort of consolation for not going there. In both cases, it is the asserted proposition that corresponds to my intentions, whereas the presupposed one is only a necessary prerequisite to process the assertion. Thus, presupposing and asserting propositions fulfills the function of *hierarchically representing conveyed information* so as to render speakers’ intentions as transparent as possible: asserted propositions are pushed into foreground as carriers of relevance, i.e. as carriers of information because of which the utterance was made at the first place.

Of course, the hierarchical representation of information need not be triggered by the issues of relevance and intentions, though I assume this to be its primary purpose. The following example illustrates how this primary use can be easily overridden by considerations of politeness:

(4-8) *We regret to inform you that your insurance policy is hereby cancelled.* (Abbott 2000:1430)

Although the insurance agent writing this sentence actually wants to convey that the insurance policy is cancelled, s/he treats this proposition as presupposed, asserting instead her/his regret because of this fact. The unpleasant character of the main point makes the indirect expression more polite and thereby more desirable than the direct one; giving the main point a status of a backgrounded, presupposed proposition is one of the ways to do it, just like (4-5) is a more appropriate way to ask someone to open the window (see Leech 1983 for a detailed discussion of indirect speech acts). Both the insurance agent and the unhappy addressee know perfectly well what is at stake, but they are ready to accept the conventionalized game of pretending that they do not, using the language as a means of alleviating potentially conflictive situations. Nevertheless, it is important to keep in mind that we are dealing with the *exploitation* of

linguistic means, i.e. with the use of the primarily intention-driven linguistic device of presupposing and asserting for other means.

4.2.2. Two notes on presupposition

The notions of presupposition and assertion as defined above need some refinement in order to be operative in concrete language descriptions. First two minor precisions.

As Horn (1996:317) notes, in the immense literature on presupposition, it is generally unclear who or what presupposes: it is sometimes sentences, sometimes propositions themselves that presuppose other propositions; utterances and words are also mentioned in this context, etc. I once again follow Stalnaker (1974), and claim that it is *speakers* who presuppose. In order to *mark* their presuppositions, however, as has been most consequently claimed by Prince (1978b, 1985, 1986, 1998) and Lambrecht (1994, 1997, 2000, 2001), they use certain linguistic *forms* which are conventionally (i.e. language-specifically) and systematically connected to certain presuppositional structures. This is the point which makes presupposition and assertion relevant for linguistic descriptions.

Secondly, presuppositions have often been defined in terms of beliefs on part of the speaker and/or the auditor: in order to presuppose a proposition, it is said, the speaker has to believe, or to assume that the auditor believes, that the proposition is true (Rooth 1996, 1999, Dryer 1996). I believe that this claim, rooted in the philosophical tradition, interested in truth values, is far too strong for natural languages: as demonstrated by Karttunen (1976), Lambrecht (1994:74ff.) and Behrens and Sasse (2003:12ff.), it is enough for the speaker and the hearer to entertain a propositional thought, i.e. to have a mental representation of the world as described by the proposition, without necessarily committing themselves to its truth. More precisely, it is enough for the hearer to accept the existence of a proposition *relative to* the ongoing discourse. In this way, even the notorious bald king of France and all the unicorns can fall under the scope of presupposition and become discourse referents, provided the hearer accepts the game and, in a manner similar to Coleridge's 'postponement of disbelief', accepts the proposition of the existence of these controversial entities *for* the given discourse. The distinction between discourse world and 'real world' is linguistically irrelevant. (See Gundel 1985: 101ff. and Abbott 1994 for excellent refutations of presuppositions as beliefs.)

4.2.3. Two kinds of presuppositions

The distinction between the **existential** and the **relational presuppositions** I am going to draw is of utmost relevance for the present study¹. Under the label ‘existential presupposition’ I lump together different kinds of presuppositions described in the relevant literature, from existential presupposition proper as exemplified by definite descriptions (see above) to factive presuppositions (‘I regret that I shouted at him’) and presupposed complements of so-called aspectual verbs (‘She stopped beating her husband’), etc. The reason for this promiscuous use of the term ‘existential presupposition’ is simple: in all the types quoted, what is presupposed is the *existence of an entity or of a state of affairs*, and in all of them the entity or the state of affairs plays *the role of a term* (as opposed to predicate) on the propositional level. Or the other way around: I contend that the use of a denotatum as a term implies that the existence of that denotatum is presupposed. The practical, linguistic side of this claim is poignantly expressed by Abbott (1994:478):

[T]he use of any NP seems to bring along a presupposition to the effect that the associated CNP has a non-empty denotation, unless of course that is being explicitly denied or questioned in the utterance. Consider the utterance schemata in (20):

- (20) a. Mary verbed several glarphy blinks.
b. Some glarphy blinks were / weren’t predicate.

I claim that unless the addressee had some reason to object, the existence of glarphy blinks would be automatically accommodated ... on the occasion of utterance of this sort.

So, on the level of linguistic form, everything that is coded as an argument or adjunct of a verb, be it as NP, PP or subordinate clause, carries an existential presupposition. Or, the other way around, in order to encode something as non-main-verb, the speaker has to presuppose its

¹ A similar distinction is drawn by Lambrecht (1994, 2001; cp. also Lambrecht-Michaelis 1998), who distinguishes between the knowledge and consciousness presuppositions (\approx my existential presupposition) and the topicality presupposition (\approx my relational presupposition). I have chosen the Strawsonian term ‘existential’ rather than the more hearer-oriented terms ‘knowledge’ and ‘consciousness’ in order to capture the nature of the proposition thus presupposed and to include the cases of indefinite specific reference (which are hardly covered by the term ‘knowledge’). The first point is of some representational relevance: I shall represent every existential presupposition as ‘there is X’ or ‘it is the case that X’ in order to emphasize the uniform nature of this phenomenon, which is often obscured by *ad hoc* paraphrases like ‘I have a car’ as the presupposed proposition conveyed by ‘my car’. In my notation, the presupposition here is ‘there is an entity which corresponds to the description “my car”’. ‘Relational’ is used instead of ‘topical’ in order to avoid confusion with the closely related but distinct notion of topic.

existence. Of course, there are exceptions to this, i.e. contexts where the presupposition is cancelled. Apart from the self-evident cases of negation and question mentioned by Abbott, there are predicates and modal contexts which more or less systematically cancel existential presuppositions. For instance, some verbs which take propositional complements generally cancel existential presuppositions due to their lexical semantics. (One of) the lexical meaning(s) of ‘to think’, namely ‘not to be sure whether something is the case’, automatically precludes existential presupposition of its complements, as in ‘I think you are wrong’: to utter this sentence in the context where the complement is presupposed would simply be contradictory, since the meaning of the verb explicitly states that the state of affairs it is applied to is not presupposed. For a discussion of the cancellation of existential presuppositions in modal contexts see Stalnaker (1974).

The other type, the relational presupposition, is more difficult to capture intuitively. I shall call relationally presupposed those propositions which are *presupposed relative to the utterance*. What does this mean? Consider once more (4-6): the proposition ‘someone went to the concert’ is unequivocally existentially presupposed, meaning that the speaker instructs the hearer to add the representation of the world in which it is the case that someone went to the concert to her/his background knowledge before processing the assertion. There is, however, another kind of presupposition associated with this proposition: the speaker presupposes that the auditor will optimally decode her/his communicative intentions at the given moment in discourse by relating the asserted proposition to this rather than to any other proposition. Let me make this cumbersome definition somewhat clearer. Recall that the speaker’s intention in uttering (4-6) is to single out a person who did go to the concert and, having done this, ascribe this enterprise to his extraordinary qualities as far as concerts are concerned, offering to the hearer a sort of consolation for not going there. Now, the best way to achieve this goal is to find a proposition against which the hearer can easily single out the person the speaker has in mind: Obviously, this proposition is the one which designates the property on the basis of which the person is singled out in the first place: ‘someone went to the concert’. When the speaker utters a sentence, s/he signals to the hearer that the relationally presupposed proposition is the optimal background against which the assertion is to be assessed. It is in this sense that relationally presupposed propositions are presupposed relative to the utterance.

Existentially and relationally presupposed propositions are in an asymmetric relation: all relationally presupposed propositions are existentially presupposed, but not all existentially presupposed propositions are relationally presupposed. So (4-6) also conveys the proposition

‘there is a person who is my aunt’s cousin’, which is existentially presupposed (the hearer has to add it to her/his background knowledge in order to process the assertion), but not relationally: the speaker does not presuppose that this proposition is the optimal background against which the assertion is to be assessed¹. Thus, relationally presupposed propositions constitute a proper subset of the existentially presupposed ones. Why should this complicated distinction be necessary? First and foremost, the expressions conveying relational presuppositions tend to be formally differentiated from the rest of the sentence. We have already seen that existential presuppositions are marked by the non-verbal (or non-predicate) status of the expressions that encode them; relational presuppositions get additional markings, like the subject-status of their carrying expressions, special sentence positions, relative-clause status in clefts, etc. Secondly, this distinction should help avoid misunderstandings of the sort exemplified by Dryer’s (1996) critique of Lambrecht’s (1994) definition of sentences like *my CAR broke down* as sentences with no presuppositions: Dryer considers this definition unsatisfying, as there is at least one obvious presupposition here, namely ‘I have a car’ (I would rather say: ‘there is an entity which satisfies the description “my car”’, see note 1). What Lambrecht had in mind was of course ‘with no relational presuppositions’: existential presuppositions are present in virtually every natural language utterance. (For further elaborations of the notion of relational presupposition see 4.3.1 and 4.3.2.).

4.2.4. Sources of presuppositions

One of the vexing questions of the presupposition research are the clues from the context licensing presuppositions, which I am going to call sources of presuppositions. There is much literature on this. Although scholars tend to be extremely creative as to the terminology they use, there appears to be a sort of consensus that what is presupposed is somehow ‘old’ (‘familiar’, ‘given’, ‘activated’, ‘lit up’, etc.). If it is not ‘old’, then it has to be at least ‘accessible’ (‘anchored’, ‘linked by bridging’, ‘crossreferenced’, etc.). Now, the ordering view of presuppositions I adopt operates on a different level: presuppositions are not automatically triggered by the proposition being ‘old’, etc., nor is everything that is ‘new’ automatically asserted. It is rather the speaker who makes the decision how s/he is going to hierarchically present the information s/he intends to convey according to her/his ideas of

¹ At this point, my analysis in terms of existential vs. relational presupposition is reminiscent of Lambrecht’s analysis in terms of semantic vs. pragmatic presupposition (1994:51-65, 270-277).

newsworthiness. This means that there is no categorical relationship between being ‘old’ and being presupposed. However, there is a statistical relationship between the two categories: it is typically the case that what is supposed to be somehow known to the interlocutors is not asserted, since being known normally implies not being newsworthy. Furthermore, as indicated above (Section 4.2.1), presuppositions have a crucial role in interpreting assertions, since they determine the particular section of the background knowledge assertions are to be related to. Consequently, not all of the ‘old’ knowledge is equally relevant to the discourse. Thus, I would like to reformulate the question of the sources of presuppositions as follows:

(a) What context clues *facilitate* (not trigger) presuppositions? (On basis of what evidence can I assume that you are ready to take my proposition for granted?)

(b) What context clues make one section of the background knowledge more relevant than the others?

The answer to both questions is the same: it has to do with the assumptions of the speaker on the momentary state of consciousness of the auditor, as most clearly shown by Chafe (1980, 1987, 1992, 1994). Since I will deal with the sources of presuppositions in greater detail in the empirical part of this study, here I shall only list them and briefly comment upon them.

(1) Physical context. Let me use an example from Stalnaker (1978:323). If a goat walked into the room, a normal reaction would be to ask ‘How did that thing get in here?’, presupposing ‘that thing got in here (somehow)’. The physical context of communication makes it not only easy to present this proposition as presupposed, it also makes it relevant, i.e. lit up (the term is from Chafe 1987) in the consciousness of the interlocutors.

(2) Textual context. If I tell you the story of the goat trotting into the room which happens to have two doors, and say ‘And then a goat came in through a door’, you may ask me ‘Which door did it come in through?’, presupposing ‘the goat came in through one of the doors’. The previous text, where the existence of two doors and the entrance of the goat through one of them was asserted, licenses your presupposing of this proposition and makes it acutely present in our respective minds. It is important not to mistake textual context for context in general, since the former is only a subset of the latter. Though there were some ingenious terminological solutions to this potential ambiguity in the past (‘co-text’ for textual context in the Prague School – e.g. Adamec 1966), I shall refrain from multiplying technical terms and use the transparent opposition ‘textual context’ vs. ‘context’.

(3) Frames. Now suppose I tell you the story about the goat without explicitly mentioning that the animal came in through a door: I simply say ‘And then a goat came in’. You are still entitled to ask me ‘Which door did it come in through’ and to presuppose that the goat entered

through one of the doors. What licenses this presupposition and makes it relevant is our encyclopedic knowledge, which contains the stereotype that one typically enters rooms through doors. I shall call this type of knowledge *frame*¹: situations are abstracted from their concrete manifestations and reduced to stereotypic schemes which contain the typical participants and the typical actions taking place in a situation. Thus the frame ‘entering the room’ minimally contains the one who enters, the room, the door, and the action of walking in; the frame ‘tea party’ (van Oosten 1986) contains the hosts and the guests, tea-cups, tea-pots, cakes, the actions of people coming and going, drinking and eating, etc. The way frames are relevant for creating presuppositions is obvious: If the speaker assumes that the auditor shares a certain frame with her/him, s/he can presuppose the propositions contained in that frame more easily than if the frame were not present. Furthermore, the mention of one part of the frame makes the other parts potentially relevant for further communication. There is a number of ways frames can be evoked in discourse:

(a) Permanently available frames. Some basic concepts like causation or ordering constitute frames which are inherently evocable in every moment of communication, irrespective of the current topic. Thus the speaker can always, in any context, presuppose a proposition of the type ‘this was caused by someone’, or ‘this was first done by someone’ (cp. (4-3)). Less spectacularly, the existence of certain discourse referents is inherently noncontroversial and thus inherently presupposable: ‘I exist’, ‘you exist’, ‘there is a person who is my mother, my father, my aunt’ et sim. are cases in point. Finally, noncontroversial states of affairs like ‘the sun rose yesterday’ are inherently easy to presuppose, though perhaps not so automatically equipped with relevance as the previous ones.

(b) Discourse-evoked frames. The type of frames most often dealt with in the literature are the frames explicitly evoked in the course of communication. If the speaker speaks of a goat entering the room, s/he explicitly evokes the frame of entering the room; if the topic of a conversation is a tea-party, the tea-party frame becomes automatically lit up in the consciousness of the interlocutors.

(c) Lexically evoked frames. Lexical frames are sometimes difficult to distinguish from discourse-evoked frames (and are probably not ontologically different from them), but I

¹ The term is borrowed from Fillmore (1982). Alternative names for similar phenomena are *script* (Shank and Abelson 1977), *scenario* (Brown and Yule 1983), *scene* (van Oosten 1986), etc. Within the familiarity theory of presuppositions, what I call frames has been understood dynamically, as a process of stipulating familiarity on the basis of encyclopedic knowledge: Clark (1977) speaks of *bridging*, Prince (1981) of *anchoring* or *stereotypic assumptions*, Heim (1988) of *cross-referencing*, etc.

consider it useful to treat them as a separate category for descriptive purposes, as will become clear in the course of this study. The use of a word or phrase may activate a frame in the interlocutors' consciousness in (at least) three ways. First, the existential presupposition of certain entities can contain information that is more precise (or less general) than is the mere knowledge of existence. The existential presupposition of 'house' is not only 'there is a house', but also 'a/the house was built'; to existentially presuppose 'river', speakers have not only to presuppose 'there is a river', but also 'a/the river flows', etc. (see Allerton-Cruttenden 1979, Goldberg-Ackerman 2001). The second type is somewhat more general than the first: some entities regularly appear in a limited number of states of affairs, which makes the relationship of the words denoting these entities and these states of affairs highly inferable from each other (more often from the entity to the state of affairs than the other way around). Thus, if I mention a book, I may count on it that my interlocutor can easily construe the reading frame from this mention; 'door' will typically evoke the frames of entering and leaving, etc. The third type is purely lexical: in idiomatic or quasi-idiomatic collocations consisting of a predicate and an argument, the mention of the argument facilitates or even automatically triggers the mental representation of the predicate and consequently of the state of affairs denoted by the collocation. The mention of 'attention' certainly makes it easy for speakers to presuppose 'pay' and to treat the state of affairs 'someone pays attention' as something their auditors are ready to take for granted and consider relevant in the given section of discourse; 'duty' will probably evoke 'perform', etc.

Presuppositions facilitated by frames are linguistically the most interesting of the three basic types listed. On one hand, they display the strongest variation from speaker to speaker, but on the other, there seems to be a tendency to conventionalize certain types of frame-facilitated presuppositions in certain kinds of contexts. Since languages may differ with respect to these conventionalizations, this can be used as a good starting point for comparison of discourse strategies crosslinguistically.

A note on terminology is in order here. As indicated above, *context* is often used in its narrow sense, as 'textual context'; on the other hand, Stalnaker uses it in its broadest sense, as the set of propositions assumed to be shared by the speaker and the auditor irrespective of the sources of these propositions. My use of the term is somewhere between these two extremes: *context* is in this study the sum of all sources of presuppositions (as defined above) active in the current discourse. Propositions shared by interlocutors (Stalnaker's *context*) may, via exploitation of pragmatic principles, differ from those explicitly licensed by context (in my use of the term). So I draw a line between *context* and *presupposition*, whereby the latter is

typically licensed, or better facilitated by the former, but not necessarily so. This distinction will play a role in my explanation of the VS order.

4.2.5. Definitions

PRESUPPOSITION is a communicative act whereby the speaker instructs the hearer to add a proposition to her/his background knowledge and thus expand the common ground between the interlocutors **prior** to adding another (asserted) proposition, in order to delimit the field of relevance of the asserted proposition. Presupposed propositions have the air of being taken for granted, old, given, etc. not because these are their inherent properties, but because speakers typically choose to presuppose propositions which are likely to be taken for granted, i.e. the noncontroversial, less newsworthy ones. More precisely, although virtually everything can be presupposed, the most frequent targets of presupposition are the propositions mediated by the physical or textual context of communication or by frames. Two types of presupposition are distinguished: the existential presupposition, by virtue of which the existence of entities and states of affairs is presupposed, and the relational presupposition, i.e. the presupposition of a proposition with respect to the utterance, which identifies a proposition as the optimal basis for the recognitions of speakers' intentions. Finally, it is the speakers who presuppose, not sentences. Nevertheless, sentences do carry presuppositions, or presuppositional structures, in that speakers use specific linguistic forms to mark the presupposed descriptive material.

ASSERTION is a communicative act whereby the speaker instructs the hearer to add a proposition to her/his knowledge **after** all the presupposed propositions have been processed, and to eliminate all the possible representations of the world which do not hold if the asserted proposition is accepted. The asserted proposition is judged against the background of the common ground between the interlocutors, i.e. against the section of the common ground delimited by presupposed propositions. Thus, to assert means to establish a relationship between a proposition (a representation of the way the world is) and a certain section of the background knowledge. Speakers choose to assert those propositions which are judged by them to be most newsworthy, i.e. the potentially most disputable of all the propositions carried by utterances. This is why asserted propositions are so often felt to be new and unpredictable: they represent the main point of the utterance, what the speaker intends to go on record as having contributed to the communication.

CONTEXT of a communicative act is a set of indications drawn from the physical or textual surroundings or from the cognitive frames activated in discourse which facilitate the creation of presuppositions. The set of actual presuppositions employed in the communicative act does not necessarily correspond to the set of those facilitated by the context: speakers may presuppose more or less than the context suggests. Context and actual presuppositions have thus to be strictly kept apart.

4.3. Topic

4.3.1. Basic issues

Communication is, to repeat the mantra, goal-oriented and cooperative. Accordingly, the information transmitted has to be somehow organized, so as to be optimally understood, stored and, when necessary, retrieved. The intuition that the knowledge achieved in communication is organized (as opposed to chaotic) was with all vehemence expressed by Reinhart (1982) and Heim (1988) and has been subject to a great number of formalizations ever since (Vallduví 1992, 1994, von Stechow 1994, van Kuppevelt 1995, 1996, Tomlin 1995, Vallduví and Engdahl 1996, Roberts 1996, Erteschik-Shir 1997, Portner and Yabushita 1998, Grabski 2000, Jacobs 2001, Hendriks 2002, to mention just a few). The basic idea hidden behind all the various approaches is that a certain type of presupposition serves as the organizing principle for the ingoing information: during communication, we use these presuppositions to attach the asserted information to them, so it remains stored and organized in locations which they determine. The impression is that the utterance conveys information *about* this type of presupposed material, that this material is in a way ‘the peg on which the message is hung’ (Halliday 1970:161) or the ‘hitching post for the new knowledge’ (Chafe 1987:28). I shall call the part of the proposition which carries this kind of presupposition **topic**.¹

¹ The idea that propositions have two parts, one of which is the basis on which the new information is added while the other represents the new information, is very old: Aristotle’s division of judgment into subject and predicate (*hupokeimenon* and *katēgoroumenon*, ‘the underlying’ and ‘the predicated’) is the first formal expression of this intuition, which had dominated the Western philosophical and linguistic thought until Frege recast the notion of proposition in a radically semantic manner. The idea soon found its way back to linguistics and philosophy, first in the Prague School theory of actual sentence bipartition, then as a part of Strawson’s efforts to refute Russell’s theory of definite descriptions. Mathesius (1939) introduced the term *theme* instead of

Thus, if I say *Peter is a smart fellow*, then ‘Peter’ is (in the majority of contexts) the topic of the utterance: I tell you about Peter that he is a smart fellow, and in doing this I instruct you to store the information that someone is smart under the entry ‘Peter’. The information is not equally distributed but rather attached to the mental representation of only one element of the proposition conveyed.

Intuitively, this idea is appealing, but extremely difficult to formalize. The most frequent expression of the intuition that topics are a kind of information center in utterances is the notion of aboutness: topics are what utterances are about or what propositions are construed about (Strawson 1964, Chafe 1976, Reinhart 1982, Gundel 1985, 1988, Lambrecht 1994, Molnár 1993, 1998, Jacobs 2001, etc.). Sometimes this difficult notion is made easier to grasp through metaphors: in Reinhart (1982), knowledge acquired through communication is compared to a library catalogue, with topics serving as subject entries; Vallduví (1992), following Heim (1988), treats knowledge as a file, and topics as file-cards into which assertions are entered; Hendriks (2002) invokes the picture of a shoe box. Yet another metaphor (if it is a metaphor at all) is the definition of topic as an open question: topics determine open questions in the communication which they help resolve (von Stechow 1994, van Kuppevelt 1995, Roberts 1996). In whatever way we chose to describe topics, there are at least two (nonmetaphorical) questions that arise in this connection: First, is there any evidence that the information exchange really functions via topics, and second, how does the idea of

subject, or *logical subject*, in order to avoid confusion with the label for the primary grammatical relation; Strawson (1964) and Hockett (1958) used *topic* to the same effect. In the course of time, however, the equation “logical subject = topic/theme” was partly lost out of sight, and the newly introduced term started to be used in a number of meanings which are only vaguely, if at all, related to its original sense. In order to preclude misunderstandings, I list here the most common meanings given to the term *topic* (for more detail, see papers in Dittmar 1992, especially Schlobinski and Schütze-Cobrun 1992):

- (1) topic = ‘logical’ subject: *topic* is what the utterance/proposition is about.
- (2) topic = old information: *topic* is the part of the proposition which conveys old, already known, ‘contextually bound’, etc. information.
- (3) topic = discourse topic: *topic* is the discourse referent which displays great referential persistence and represents an important protagonist in a text. Thus, it is not defined on the sentence level and is rather a gradual notion, with different referents having different grades of topicality.
- (4) topic = constituent: *topic* is a constituent on the left periphery of the sentence, usually not valence-bound. Pragmatically, it is similar to setting adverbials of space and time, since it delimits the domain in which the proposition expressed by the sentence holds.

My use of *topic* is closest to the one described by the equation (1).

asymmetric information storage fit into the incremental model of communication outlined above?

Most of the evidence offered for the existence of topics in communication is indirect and intuitive in nature. Strawson (1950, 1964) and Reinhart (1982) claim that in assessing the truth of a proposition, we are likely to search our knowledge of the topic, and not of the rest of the proposition: if I want to know whether ‘Peter is a smart fellow’ is true, I shall rather search my knowledge of Peter and see if he corresponds to the description ‘being a smart fellow’ than search my knowledge of smart fellows and see if Peter is among them¹. In a similar vein, Portner and Yabushita (1998) notice that certain discourse phenomena are better understood if we assume some sort of asymmetric information storage: if a discourse referent appears in a text twice, once as a topic and once as non-topic, then it is more felicitous to resume it with the description given to it in the utterance whose topic it was. Here is the English translation of one of the examples they adduce. The text is a description by a waiter at Good Eats Café of one evening’s events:

- (4-9) A woman with a small child came in first, and she ordered chicken fried steak. Next, a young man holding a tennis racket came in. He handed the racket to her, and went to the bar to get a beer. Another man and woman, who were late, came in. They seemed to have been to a movie.
- (a) The woman who ordered a chicken fried steak left first.
- (b) ??? The woman who the man handed a racket to left first.

Sentence (b) is, according to Portner and Yabushita, less felicitous than (a), because the information that the woman ordered chicken fried steak, conveyed in the sentence where she is the topic, is stored under the entry ‘the woman with a child’ and thus directly retrievable; the information that the young man handed the racket to her is stored under the entry ‘the young man’, since he is the topic of the sentence conveying this, and is thus less accessible when the woman is referred to. Chierchia (1992) claims that topicality influences proportion readings of conditional donkey sentences. In addition, Grabski (2000) demonstrates that elaborative text sequences work better when the topics of elaborative sentences are in some sort of set-relationship with the topic of the sentence which is to be elaborated upon. Psycholinguistic experiments like those of Itagaki and Prideaux (1985) and Tomlin (1995),

¹ Strawson (1964) claims that the presumed lack of truth value of ‘the King of France is bald’ is due to the auditor’s incapability to search her/his knowledge about the King of France, since there is no such person. The problem of the King of France and his baldness is elegantly solved by Karttunen’s notion of discourse referent (see Section 4.2.2; see also Horn 1996, Behrens and Sasse 2003).

though not absolutely conclusive, seem to point out that there is some sort of correspondence between topics and attention states of interlocutors, etc.

Now, evidence of this kind is admittedly not overwhelming, though it unequivocally shows that there is a consensus among scholars on the existence of aboutness topics. The most convincing evidence are thus still some formal properties of languages: a great number of constructions, lexical and grammatical, are best explained as somehow related to topicality. Many languages have morphological means of marking phrases which carry the descriptive material conveying the topical part of propositions, the best known cases being Japanese, Korean and, probably, Tagalog (see, e.g., papers in Hinds, Maynard and Iwasaki 1987, Schachter 1977, Shibatani 1991); certain intonational contours (e.g. L+H* in English, Pierrehumbert and Hirschberg 1990) and certain sentence positions (see e.g. Gundel 1988, É.Kiss 1998a, Matić 2002) seem to be consistently applied in marking topic expressions, etc. Even more striking is the close connection between topicality (as conceived above) and such core elements of language as the grammatical relation of subject, which is nowadays widely accepted to be a mixed category, the result of a grammaticalization of the pragmatic relation of topic and the semantic role of agent (see papers in Li 1976, Sasse 1982, 1995c, Comrie 1988). To sum up: speakers of languages have an intuition that certain constructions are there in order to mark parts of propositions as what these propositions are about – and this is up to now the best piece of evidence we have for the existence of an organizational principle based on aboutness.

The question of applicability of the notion of topic to the incremental model of communication outlined in 4.1 is intended as a convenient introduction to a more precise description of how I think aboutness is to be explained cognitively; my account is based mostly on Reinhart (1982), Lambrecht (1994), and Portner and Yabushita (1998). I have already mentioned that not every element of the common ground is relevant for the processing of assertions: only those the speaker chooses to presuppose (and mark as such) play a certain role in the process of communication, since presupposed propositions *delimit* the part of the common ground which the auditor has to activate.

Recall that ignorance, i.e. the lack of consensus between interlocutors on the way the world is, is defined as the existence of more than one possible ways the world is (more possible worlds) in the knowledge of interlocutors, and that communication is the elimination all the potential propositions which do not fit the representation of the world given by the asserted proposition. Now, if to presuppose means to delimit the relevant part of the common ground, then by presupposing a proposition the speaker actually instructs the auditor approximately as

follows: concentrate only on those propositions (possible representations of the way the world is) in your background knowledge which contain the elements which I have marked as presupposed. Consider the following sentence, uttered in answer to the question “What about your friends?”:

(4-10) *Peter works in the University library.*

There are two explicitly marked existential presuppositions here: ‘there is a person named Peter’ and ‘there is (one) thing corresponding to the description «University library»’. The auditor is expected to add these two propositions to her/his background knowledge, if they have not already been there. So far, so good, but we are still not any closer to the notion of topic. This is why I introduced the second type of presupposition, the relational one, in 4.2.3: the speaker relationally presupposes those propositions which s/he thinks represent the optimal background for processing the assertion, i.e. those, which make her/his communicative intentions optimally transparent to the hearer. Relationally presupposed propositions are thus in a way presupposed relative to the utterance, since the speaker gives the hearer the sign to process the assertion conveyed by the utterance relative to the relationally presupposed proposition.

How can this be applied to (4-10)? The context of the utterance reveals that the most probable communicative intention of the speaker is to increase the auditor’s knowledge of her/his friends. So s/he chooses, among the two possibilities (‘Peter’ and ‘the University library’), the one which will make this intention most transparent, ‘Peter’, and represents it as relationally presupposed (formally, by coding it as the subject and by using the L+H* contour). In this way, s/he signals to the hearer that, among all the potential representations of the world in which Peter is involved, s/he should eliminate all those in which the relation ‘works in the University library’ does not hold. In contrast, the speaker does not directly commit her/himself to the relationship between the University library and Peter working there: the hearer is simply not directly instructed to eliminate the possible representations of the world pertaining to the University library in which the relation ‘Peter works in’ does not hold, because s/he is not instructed in the first place to search all the representations in which the University library is involved.

Thus, *relational presuppositions are instructions to the auditor to concentrate on a sequence made of a discourse referent and the possible worlds containing it and to add and evaluate the asserted proposition relative to this rather than to any other sequence. The topic of an utterance is that element of the proposition which is under the scope of the relational presupposition, i.e. the discourse referent which defines the sequence of the possible worlds to*

*be searched*¹. All the phenomena enumerated above as indirect evidence for the existence of topics (verification, resumption, etc.) are easily explained by this definition.

The most striking consequence of this model of information transmission, in which assertions are in a way ‘clustered’ around certain discourse referents, is the asymmetry of knowledge. By this I mean the fact that, upon this interpretation, one would have to assume that on hearing (4-10) the hearer knows that one of the properties of Peter is his employment in the University library, but s/he does not know that one of the properties of the University library is Peter’s employment in it: the common ground increases only with respect to topics. This is clearly counterintuitive: the conversation in (4-10) can be easily continued by an ironic remark like ‘oh, the University library must be extremely happy about that!’, where the knowledge that Peter’s employment is one of the properties of the University library is clearly present. If I say ‘I saw your cat’, you know not only about me that I saw your cat, but also about your cat that it has the property of having been seen by me.

There have been a couple of proposals how to solve the apparent contradiction between the idea that information is given only about certain discourse referents and the obvious increase of knowledge about all discourse referents. In one of them, the latter fact is simply denied, directly or indirectly (Kuroda 1972, 1990, Ladusaw 2000, Sasse 1987; see Sasse 1996 against this view). Others have postulated some sort of accommodation mechanism in the vein of Lewis (1979). Thus Vallduví (1992, 1994) claims that the asserted proposition is directly entered only on the file-card of the topical item, whereas the other elements of the proposition are cross-referenced to this item, so that in order to retrieve the asserted proposition from a non-topical item, one first has to go to the topical file-card (for the problems of this approach see Hendriks 2002).

My proposal is once again, as in the definitions of presupposition and assertion, based on the fact that communication takes place in time. The speaker gives the hearer the instruction to add and evaluate the asserted proposition with respect to the relationally presupposed element, the topic (by concentrating on the possible worlds containing the topic, etc.). This is the *explicit* assertion, the representation of the world the speaker commits her/himself to; the hearer is expected to follow this instruction first. Once the assertion is assessed with respect to the topic, the new information obtained in this way simply spreads to all other elements involved, the knowledge being rather a network than a catalogue.

¹ The definition of topic as a sort of searching instruction neatly explains the effects of informational separation, addressation and frame-setting, which Jacobs (2001) singles out as the prototypical topic features. The question of predication will be addressed below.

Thus, apart from one explicit assertion, every utterance can be said to convey a number of derived, *implicit* assertions stemming from the explicit one: in uttering ‘Peter works in the University library’ I commit myself to the claim that one of Peter’s properties is his work in the University library and imply that the auditor is free to conclude that one of the properties of the University library is Peter’s work there. The difference between the former and the latter lies both in the commitment on the part of the speaker and in the temporal relationship: implicit assertions are processed after the explicit one¹.

I should like to distinguish three phenomena along this line: first, the explicit assertion, which is the act of relating the asserted proposition to the topic; second, the implicit assertions, which are the extensions of the explicit assertion by the hearer; and third, the state of knowledge achieved after both types of assertions have been processed.

Since implicit assertions play no role in my description of VS order, I shall refer to the explicit assertion simply as **assertion**. Since the final state of knowledge reached by processing the whole utterance (i.e. my knowledge of both Peter and the University library) closely resembles the Fregean predicative scheme, I shall call this state **predication**.

Thus, in describing (4-10), I shall say that the speaker asserts about Peter that he works in the University library, i.e. that there is an assertional relationship between ‘Peter’ and ‘works in the University library’, and that there is a relation of predication both between ‘Peter’ and ‘works in the University library’ and between ‘the University library’ and ‘Peter works in’.

Schematically:

$$(4-11) \quad F(x,y) \wedge x = \text{topic}$$

	x (topic)	y (non-topic)
assertional relationship	F(y)	/
predicational relationship	F(y)	F(x)

In this way, both the asymmetry of assertion and the symmetry of the resulting knowledge are salvaged².

¹ I assume that, in postulating the subject-predicate (later topic-comment, theme-rheme, etc.) division of judgment, Aristotle had the explicit assertion in mind, and that, in rejecting this dichotomy and creating the modern symmetrical predicate calculus, Frege tried to capture the intuition triggered by the sum of the explicit and implicit assertions.

² Note that the terms *predication* and *assertion* are usually used promiscuously, as quasi-synonyms (see 2.5.1). There are to my knowledge two attempts to keep them apart explicitly, and both differ from my proposal in a significant way. The first one is that of Sasse (1987): for Sasse, predication is but one type of assertion (the other being a thetic assertion); for me, predication is the indirect consequence of every assertion. The second, advocated by Lambrecht (1994), is based on the distinction between semantics and pragmatics: predication is a

4.3.2. Discourse properties of topics

One discourse referent is chosen as topic rather than the others because the speaker believes that her/his intentions at the given point of communication are best made transparent by relating the assertion to that discourse referent. The fact that the topical material is of necessity presupposed and that presuppositions are very frequently equated with ‘old’ or ‘believed’ information led many people to posit more or less severe restrictions on topics. According to the stricter variant of this view, topics can be only those referents which are definite or generic (e.g. Kuno 1972 Kuroda 1972, É.Kiss 1987, 1994, Portner and Yabushita 1998). In the more lenient version, the topics only have to be referential (Lambrecht 1994), which is often assumed to mean that quantified phrases are never topical (e.g. Davison 1984). The linguistic reality does not seem to care about these restrictions: indefinite nonspecific, quantified and nonreferential phrases carelessly appear in many constructions which are widely assumed to mark topicality (see Ward and Prince 1991, Birner and Ward 1998, Lötscher 1992):

- (4-12) ... consider a person who knows arithmetic, who has mastered the concept of numbers. In principle, he is now capable of carrying out or determining the accuracy of any computation. *Some computation he may not be able to carry out in his head.* Paper and pencil are required ... (N. Chomsky, *Rules and representations*, 1980:221; from Birner and Ward 1998:80)
- (4-13) Did she buy a whole new wardrobe for school? “Not really. I have a great deal of clothes ... *Most of my stuff, my mom gets at Alexander’s.*” (Birner and Ward 1998:81)
- (4-14) (a description of the demonstrations of the Kurds in Berlin) *Organisiert hat die Demos die KPD.* (Deutschlandfunk, 28.08.2000)

There have been attempts to explain away some of these cases as instances of ‘focused topics’ (Molnár 1998, Büring 1999, 2000), or to attribute to nonreferential phrases some kind of referentiality (Lambrecht 1994:76). The incremental model outlined above does not pose any

semantic notion referring to the fact that certain parts of propositions are there in order to ascribe properties; assertion is a pragmatic act whereby a new relation is being added to the stock of knowledge of the interlocutors. For me, property-assigning is one of the effects of assertion (see below 4.4.1), whereas predication is one of its consequences.

restrictions to topicality and does not encounter this kind of problem: anything that can be present in one's representations of the way the world is can be chosen to be topic. Thus, in order to interpret (4-12), the hearer simply has to activate all the potential propositions in which 'he' stands in some relationship to an unspecified type or quantity of computing, without necessarily having the concrete mental representation of 'some computing'¹. However, this should not mean that there is no relation at all between definiteness, referentiality, etc., on one hand, and topicality on the other. However, this relation is rather statistical than categorical: The overwhelming majority of topics in naturally occurring texts are indeed definite, etc., but this alone does not make them topical. Referentiality and definiteness (or identifiability, if you like) make the relational presupposing of a discourse referent, i.e. its choice for the topic, easier, but they do not determine it.

Now, it is not only reference-related features that seem to stand in some kind of statistical relationship to the choice of the topic. It has been known for a long time that certain ontological properties play a certain role as well: thus entities are more often used as topics than states of affairs, animate entities more often than inanimate ones, the first and the second person rather than the third, basic level objects more often than super- and subordinate level objects (in the sense of Mervis and Rosch 1981), etc. Furthermore, the role played by a term in the proposition may be relevant as well: Agents are more often topics than, say, recipients, recipients more often than patients or themes, etc. These, and various other, features have often been represented as topicality hierarchies (for a good summary, see Sasse 1982:269; see also Lambrecht 1995). All these tendencies are easily explainable as consequences of the function of topics, namely to render the intentions of the speaker at the given moment in discourse as transparent as possible: we are mostly concerned with ourselves and with the things similar to us, i.e. with human, agentive, etc. phenomena, the fact poignantly subsumed under the term 'egocentric principle' (Sasse 1982, van Oosten 1986).

However, it is important to keep in mind that we are dealing with statistical tendencies, i.e. with paradigmatic values, not with syntagmatic laws. For this reason, I shall call this type of tendency towards topicality **statistical topicworthiness**. The concrete decision of the speaker to choose the one instead of the other element of the proposition for the topic of the utterance may run in the opposite direction: certain intentions are better decoded if states of affairs rather than entities are chosen as topics, or a thing rather than a human being, etc. The actual

¹ The examples with nonreferential, usually predicative, expressions in topic positions like (4-14) may need a more complex explanation; see Büring (1999), Matic (2003).

estimation of the appropriateness for the topic role relative to the current utterance will be called **actual topicworthiness** throughout this study.

One of the properties commonly associated with the notion of topic since Weil (1844) is that it represents old or familiar information. As I hope to have shown in 4.2., this is in the case of both existential and relational presuppositions only a statistical tendency explainable by the nature of presuppositions. In spite of this, topics, even if they are not contextually given, i.e. old or familiar, often have the air of having been somehow present in the minds of interlocutors prior to the utterance. Consider the following examples:

(4-15) Gestern gab Peter ein gelungenes Fest.

(a) *Dieses Mal hatte [er]_{Top} bemerkenswerterweise sein GANZes Geld für alkoholische GeTRÄNke ausgegeben.*

(b) *Dieses Mal hatte [er]_{Top} [sein ganzes Geld]_{Top} bemerkenswerterweise für alkoholische GeTRÄNke ausgegeben.* (Grabski 2000:181-182)

(4-16) One morning I came downstairs to breakfast, and my mother, who had gotten up earlier and listened to the news, announced to me: [*Truman*]_{Top} *DIED*. (Schmerling 1976:41)

Only ‘he’ is the topic of (a) in (4-15), whereas in (b) both ‘he’ and ‘all of his money’ are topical, as witnessed by the position of the sentence adverbial *bemerkenswerterweise* and by the intonation. The effect of the topical status of ‘all of his money’ in (b) is that of familiarity, as if the interlocutors have already been thinking of what Peter had done with all his money, although the contextual clues for this are weak or nonexistent. In (4-16), which is a perfect example of a contextless, i.e. out-of-the-blue, utterance, the topical status of ‘Truman’ somehow seems to imply that Schmerling’s mother assumed Schmerling had been thinking about Truman’s fate before.

How is this effect to be accounted for? Recall that by relationally presupposing something and making it a topic, speakers signal that the assertion is to be processed with respect to this element, meaning that the hearer has to concentrate only on those possible propositions which contain that element. All this takes place in time, meaning that you first process the presupposed information, and only thereafter that what is asserted. Recall also that speakers mark as topics those elements which make their intentions transparent. Taken together, these facts mean that the speaker creates *expectations* on the part of the hearer: if I tell you that this is the topic of my utterance, and you invest some effort to choose the possible propositions containing that topic, then you are entitled to expect some further specification of the topic. It is these expectations which come into being before the assertion is processed that create the

impression of oldness, or familiarity. When the speaker in (4-15) explicitly marks ‘Peter’ and ‘all of his money’ as topical, the hearer develops the expectation that there is some relevant relationship between Peter and all of his money in the context of a good party, the expectation which is satisfied by the assertion ‘he gave it away for alcoholic drinks’. We tend to understand this expectation as a sign that the interlocutors had already thought of the topical element(s) and considered it/them relevant before the utterance was produced. I shall call this phenomenon the **effect of anaphoric relevance**, or simply **anaphoric relevance**. Let me note that there is some irony in play here: instead of topics being inherently old or given, the speaker *makes* them old by investing them with relational presupposition.

Somewhat less conspicuous, but still present is the converse phenomenon of **cataphoric relevance**: topics tend to evoke expectations on the part of the hearer that the communication to follow will somehow pertain to them or reveal some further facts about them. Thus (4-16) seems to imply that the utterances which are to follow will also pertain to Truman, or to his death, or will be in any way related to them. The speaker can exploit these expectations (which are sometimes more and sometimes less present, presumably depending on the nature of assertion related to the given topic) to represent the interpretation of the subsequent discourse as relevant to the topic of one utterance, even if there is no overt relationship:

(4-17) “What about the people we used to go to school with?”

“Remember that old pedantic bore, Peter? [*He*]_{Top} *works with JOAN now. You can imagine how orderly her office is nowadays.*”

In interpreting the clause ‘how orderly her office is’, the hearer is more likely to think of Peter as the principal cause of the new orderliness than of Joan: the speaker simply used the cataphoric expectations associated with topics and indirectly gave another piece of information on the topic of the previous assertion.

The final point I want to address in this section is the question of the number of topics allowed per assertion. The debate on this issue, starting, as far as I know, with Reinhart’s (1982) apodictic claim that propositions contain only one topic, and continuing with low intensity over the greater part of the Eighties and Nineties, seems to have resulted in some sort of consensus on pragmatic and syntactic admissibility of multiple topics¹. The examples of multiple topics adduced by now are (4-12) [‘he’, ‘some computation’], (4-13) [‘most of the stuff’, ‘my mum’] and (4-15)(b) [‘he’, ‘all of his money’]. The particular thing about the

¹ An excellent defence of multiple topics in pragmatic terms is given by Lambrecht (1994); syntactic aspects are nicely presented in É. Kiss (1987) and Erteschik-Shir (1997), syntax and pragmatics of multiple topics in Nikolaeva (2001). For a short and intelligent critical assessment see Jacobs (2001).

utterances with more than one topic is the impression that they are not only about the topics themselves, but also about the relationship holding between them: in (4-12), for instance, Chomsky asserts something not only about ‘he’ and ‘some computation’, but also, and first of all, about the relationship between these two elements. One has the feeling that the actual topic is ‘his relationship to some computation’. The extended incremental model explains this straightforwardly: when the speaker encodes a proposition with more than one topic, s/he signals to the hearer to concentrate only on those potential propositions in which the two or more topical elements stand in some relation to each other, and to evaluate the assertion only against the background of thus defined possible worlds.

4.3.3. Some formal properties of topics

There is no one-to-one correspondence between information structure and linguistic form. Topics are not exempted from this fate: in probably every language there is more than one formal marker for topics, none of which is there only in order to mark topics. For this reason it is difficult to generalize on the universal formal properties of topics (*pace* Gundel 1988) and downright wrong to generalize on discourse properties of topics in general by concentrating on only one type of topic marking in just one language¹: the best result to be achieved with the latter procedure is a more or less precise description of one topic construction in one language, not universal properties of topics (see Prince 1998). So I shall not try to describe the prototypical formal devices for topic marking. In fact, this section deals with only one point of topic marking: with the difference of what I shall call directly and indirectly marked topics, or simply **direct** and **indirect topics**.

At least since Chafe (1976), two types of topics have been differentiated: in Chafe’s terminology, they are called Chinese-style and English-style topics. The former are extraclassical and determine the ‘frame within which the sentence holds’; the latter are within the clause and represent ‘what the sentence is about’ (Chafe 1976:51). The notion of extraclassicality has been defined somewhat broader since, comprising all the elements outside

¹ This tendency has been gaining momentum in the last years, as a research program in McNally 1997, and practically in Vallduví 1992 (Catalan, sentence initial and sentence final expressions), Pierrehumbert and Hirschberg 1990, Hendriks 2002 (English, L+H* intonational contour), Frey 2000, Grabski 2000 (German, elements in the *Mittelfeld* between the verb and sentence adverbials), Portner and Yabushita 1998 (Japanese, *wa* phrases).

the verb valence – thus the clause-initial temporal and local adverbials are treated as external topics by Nikolaeva (2001). She uses a mixture of formal and discourse criteria to define the two types, labeled external and internal topics: the former are outside the verb valence and serve as frame-setting devices, which is why they ‘tend to correspond to adverbials’ (Nikolaeva 2001:11); the latter are valence-bound and represent what the sentence is about. A similar conglomerate of syntax and pragmatics is to be found in the proposal by Gasde (1999), who distinguishes between frame-setting and aboutness topics along the same lines as Chafe and Nikolaeva, only that a more detailed list of the types of frame-setting topics is given: spatial and temporal topics, Chinese-style topics proper, PP individual frames, conditionals. Jacobs (2001) supplements this list with topicalized manner adverbials.

In my model there is no room for pragmatic, i.e. ontological, differentiation of the two types, since I understand aboutness as frame-setting (and frame-setting as aboutness): the speaker uses topics to signal to the auditor against which possible worlds s/he is to evaluate the asserted proposition. Intuitively, however, there is a noticeable difference between the topics in (4-18) and (4-19) on one hand and (4-20), (4-21) and (4-21) on the other:

(4-18) *Peter is smart.*

(4-19) *Peter habe ich gestern gesehen.*

(4-20) *Shuiguo, wo xihuan pingguo.* (‘fruit, I like apple’; Chinese; Xu 2000:23)

(4-21) *Körperlich geht es Peter gut.* (Jacobs 2001:655)

(4-22) *In the yard was an old bicycle.*

Sentences (4-18) and (4-19) seem to be about their topics (‘Peter’). This is not the case with sentences (4-20), (4-21) and (4-21): ‘fruit’, ‘physically’ and ‘in the yard’ are indeed felt to be only the frame within which the sentence is to be interpreted. Thus, one could paraphrase (4-21) as ‘with respect to his body, [Peter]_{Top} is well’. I shall contend, as indicated above, that this apparent difference in interpretation is not ontological, i.e. that we are in both classes of cases dealing with the same unified phenomenon of topic (aboutness topic, if you like), but that it is rather formal in nature. The topics straightforwardly interpreted as standing in the aboutness relationship with their sentences are my direct topic, the ones which look like mere frame-setting expressions are my indirect topics.

The idea underlying this division is almost trivial: direct topics are only those which are presented as entities and coded as direct arguments of the predicate, i.e. as subjects and objects (or ergatives and absolutes, or whatever primary grammatical relations the language has); all other types are indirect. It is only in the former case that both the relevance relationship between the asserted proposition and its part under the relational presupposition,

the topic, is marked explicitly and the denotatum of the topic expression is presented as an entity. In (4-18), ‘Peter’ is directly presented as an entity and the relevance relationship between ‘Peter’ and ‘Peter is smart’ is explicitly marked (the grammatical relation of subject, concord); the same holds in (4-19) and, for that matter, (4-6), where ‘the person who went to the concert’ has the same properties as ‘Peter’. The instances of indirect topics lack at least one of the two properties mentioned above, or both. The extraclausal expression ‘fruit’ in (4-20) is presented as an entity, but its relevance relationship to the assertion is not explicitly marked. The relevance relationship of the locative PP ‘in the yard’ in (4-21) is expressed explicitly, but its denotatum is not an entity but a location. Finally, in (4-21), the topical expression ‘physically’ has neither of the two properties.

What does all this have to do with the impression of an adverbial-like, frame-setting interpretation? The speaker marks these expressions as relationally presupposed, i.e. topical, but they lack some of the properties which enable the hearer to choose the set of possible worlds related to the denotata of these expressions. So s/he has to invest an additional effort: where the explicit relevance relationship is not marked, it has to be construed pragmatically; where the topic is not encoded as an entity, the existence of an entity carrying the existential presupposition which approximately corresponds to the description given by the expression has to be inferred. Both processes are nicely illustrated by (4-21): ‘physically’ is marked as topical (initial position, fall-rise contour), but neither the relevance relationship nor an entity-like description are given. The auditor has to construe an existentially presupposed entity corresponding to ‘physically’, something like ‘the body’, and then to establish a relationship between this entity and the direct topic, ‘Peter’. It is only now that s/he can search her/his background knowledge and concentrate on all the possible representations of the world, i.e. potential propositions, in which there is a relationship between ‘Peter’ and ‘his body’, and eliminate all those in which this relationship is not ‘feels well’. In other words, I claim that (4-21) can be paraphrased with the cumbersome but useful ‘as for’-paraphrase: ‘as for his body, as for Peter, the relationship between them is »feeling well«’. Similar analysis can be applied to other examples of this kind. The impression of the mere frame-setting – note that I claim that it is only an impression – stems from the obliqueness of the indirect topics and from the additional effort needed to process them¹. The other possible source of this impression is the fact that indirect topics more often than not occur together with direct topics (as in (4-20) and (4-21)), in which case they are of necessity interpreted as part of the presupposition ‘the

¹ A comparable account of indirect topics, with different terminology, is given by LaPolla (1995) in his analysis of Chinese locative phrases.

relationship of the direct topic to the indirect topic' and are thus felt to be somehow backgrounded or 'secondary' (Nikolaeva 2001), serving to highlight an aspect of the direct topic.

The most important practical consequence of this purely formal approach is that, in contrast to the ontological interpretations listed above, direct and indirect topics represent a language-specific matter. As Lambrecht (1995) demonstrated on the example of English and French psych-verbs, the same state of affairs with the same presuppositional structure can be expressed with different types of topics in different languages, depending on the overall grammatical structure of the given language and on the idiosyncratic properties of predicates:

(4-23) *My FOOT hurts.*

(4-24) *J' ai mal au PIED.* ('I have pain in-the foot')

In both cases the topic on the propositional level is 'I', with the assertion 'the foot hurts' pertaining to me and giving new information about me. However, in English, this topic is indirect, coded as a possessive pronoun, so that the hearer has to construe its entity status and the relevance relationship between 'I' and the asserted proposition, whereas in French it is direct, coded as a subject, with the relevance relationship and the entity status explicitly marked on the grammatical level. The notion of indirect topic plays an important role in my interpretation of the VS order, so that it will be further elaborated upon in Sections 4.5.2, 11.2.1., and 11.6.

4.3.4. Definitions

TOPIC is the element of the proposition under the scope of the relational presupposition, i.e. the discourse referent which defines the sequence of possible worlds the hearer is to search in order to assess the asserted proposition. Thus, assertion are realized and the mutual consensus on the way the world is increases with respect to topics. As a result, the utterance is felt to be about the topical discourse referent. The choice of the topic depends on the considerations on the part of the speaker concerning the way s/he will make her/his intentions optimally transparent to the hearer. Statistically, human, agentive, entity-like, familiar, etc. discourse referents are the most frequent topics; the actual decision on the choice of the topical referent, however, may run contrary to these statistical tendencies if the speaker's intentions demand it. There are no restrictions as to the referential or cognitive status of topics, though they typically have definite or generic reference. By using a discourse referent as the topic of an utterance, the speaker creates expectations on the part of the hearer that the assertion to come

conveys information relevant to that discourse referent. If these expectations are interpreted as pertaining to the previous discourse, I speak of the effect of anaphoric relevance; if they pertain to the following discourse, the cataphoric relevance is at stake. On the level of linguistic form, topics may be direct, if they are coded as direct arguments of the predicate, or indirect, if they are coded in any other way. In the latter case, the effect of oblique topicality is due to the fact that the hearer has to pragmatically construe the entity status of the topic and/or the relevance relationship between the topic and the asserted proposition.

PREDICATION is the state of the knowledge of the interlocutors indirectly resulting from the assertion. Whereas the assertion is the speaker's explicit instruction to the hearer to eliminate the possible worlds which do not fit the description given by the asserted proposition *within* the sequence of possible worlds defined by the topical discourse referent, predication is the result of the secondary spread of the elimination process over the sequences defined by the non-topical discourse references contained in the proposition. Thus, although the speaker explicitly commits her/himself only to the relationship between the topic and the asserted proposition, the final result of the assertion is the increase of mutual consensus on all the possible relationships holding within the asserted proposition. Since the relationship established by the assertion is the only one the speaker commits her/himself to, it is foregrounded and relevant to the discourse; the knowledge resulting from predication is backgrounded, less easily retrievable, and less relevant for the progress of communication.

4.4. Focus

4.4.1. Basic issues

Upon the definition of assertion as a new relation established against the common ground, the assertion conveyed by an utterance obviously corresponds to the *whole* proposition conveyed. In uttering (4-18), for instance, I assert the whole proposition 'Peter is smart', by relating the property of being smart to my mental representation of Peter: both 'Peter' and 'is smart' are necessary for the assertion to be realized. This is tantamount to saying that there is no part of the proposition which could be equated with assertion. On the other hand, it is clear that there is a significant informational difference between 'Peter' and 'is smart': 'Peter' is the topic of the utterance, meaning that the speaker relationally presupposes it, whereas there is no

relational presupposition connected with ‘is smart’. This element contains the descriptive material which is to be assessed with respect to the topic: it is, to paraphrase Lambrecht (1994:213), *the part of the proposition whereby the assertion differs from the relational presupposition*, or, somewhat simpler, *the part of the proposition which is not presupposed with respect to the utterance*, i.e. not under the relational presupposition. I shall, in accordance with the current usage, call this element **focus**. Assertion can thus be defined as the establishment of a relation between the relationally presupposed topic and the relationally nonpresupposed focus. Focus is not the same as assertion but rather only a necessary part of it, the point which should not be obscured by the fact that I shall throughout this section, for the sake of brevity, speak of the assertional definition of focus and of focus as the assertional part of the proposition. Focus is, after all, the carrier of the assertion proper.

This property inspired many metaphorical descriptions: focus is the ‘information point’ of the utterance (Bolinger 1954), it is the element of the proposition which is ‘highlighted’ (Chomsky 1972 and the bulk of the subsequent generative literature), to mention just a few. The feeling that focus is somehow the most prominent part of the sentence is a direct consequence of its carrying the assertive descriptive material: The asserted proposition is the point the speaker wants to make by uttering a sentence (4.2.1.), so that the part of the proposition which distinguishes the assertion from the relational presupposition is felt to be the carrier of the main point.

The fact that focus represents the *relationally* nonpresupposed part of the proposition does not mean it cannot contain *existentially* presupposed material, or even entirely consist of it (see 4.2.3). Consider the following sentences:

(4-25) *I am sorry that I missed the CONCERT*, not that I spent time with YOU instead.

(4-26) Can you imagine what happened to me yesterday? *I bumped into PETER!*

The focus in (4-25) is ‘that I missed the concert’, with the propositional function ‘I am sorry that X’ under the relational presupposition. The whole focal part – the complement clause serving as a term of the predicate ‘be sorry’ – is existentially presupposed: in order to utter (4-25) felicitously, the speaker has to presuppose the proposition ‘it is the case that I missed the concert’. However, this does not prevent this proposition from playing the role of the focus, as it is not presupposed relative to the utterance: the speaker instructs the hearer to search all the possible worlds in which s/he is sorry for some reason (relational presupposition) and to eliminate all those in which this reason is not her/his missing the concert (assertion). *Mutatis mutandis*, the same can be said of (4-26) (the existentially presupposed referent ‘Peter’ is a part of the focus ‘bumped into Peter’). Existential presuppositions are situated on a level of

information processing different from that on which assertion takes place, i.e. on which the notion of focus is operative: as I hope to have demonstrated in 4.2., the processing of existential propositions is the first step in the decoding of utterances, whereas relational presuppositions and assertions are processed afterwards, on the basis of the common ground reached by the acceptance of the existentially presupposed descriptive material.

This relationship between focus and existential presuppositions explains why the widespread idea that focus necessarily represents ‘new’ information is, if taken as a categorical property, not unproblematic. As indicated in 2.2.4, ‘new’ has been used as a label for at least two phenomena: as a property of the denotata of linguistic expressions and as a type of relationship between a linguistic expression and the utterance. If, by saying that focus is new, the former definition is meant, i.e. newness as the criterial property of focal denotata, then this claim is empirically and theoretically untenable, as witnessed by the examples and the discussion above. ‘Old’, existentially presupposed denotata freely occur as foci. If the latter definition is meant, newness as a relationship between the focus and the utterance, then this claim is, albeit with certain reservations, essentially correct, as I demonstrated in the section devoted to assertion (4.2.1). In terms of topic and focus, the newness of assertive material may be phrased out as follows: for an utterance to be informative, the relationship between the topic and the focus must be new, i.e. it must not already belong to the common ground of the interlocutors. Of course, the speaker may choose to assert a relationship that is already known to both interlocutors, but this is only an invitation to the auditor to exploit her/his knowledge of communication principles and *make* the ‘old’ assertion ‘new’. In sum: it is the relationship of the focus to the topic that has to be unknown and new, not its denotatum.

Apart from newness, focus has often been associated with contrast and with ascription of properties. The former feature gave rise to the theory of focus based on alternative semantics (Rooth 1992, 1996, 1999, Roberts 1996, 1998; papers in Bosch and van der Sandt 1999), the latter to the structured meaning theory and related approaches (Atlas and Levinson 1981, Löbner 1990, von Stechow 1991, Krifka 1991, Jacobs 1991, von Stechow 1994)¹.

The alternative-semantics account of focus operates with the notion of alternative propositions: focus evokes alternatives and rules them out. Thus, if I say:

(4-27) *John introduced Bill to [SUE]_{Foc}.* (Rooth 1996:275),

¹ My account of the two currently most prominent theories of focus is of necessity short and extremely nontechnical. For more details and more formulas, see von Stechow (1991), Rooth (1996), É.Kiss (1998a: 707ff.) and Devine and Stephens (2000: 72ff.).

I evoke the set of propositions of the form ‘John introduced Bill to X’ and rule them out by asserting that the only true representation of the world is the one in which X = ‘Sue’. The structured meaning theory in its various versions tries to capture the feeling that whatever the syntactic structure of the sentence, in its semantic / pragmatic structure the semantic value of the focus is ascribed as a property to the topic referent. Oversimplifying a little, one can say that in (4-27), the semantic value ‘Sue’ (focus) is ascribed as a property to the individual who has the property of being the person John introduced Bill to (topic).

The intuitions underlying both of these two approaches seem to be somehow right. I shall argue that my definition of focus as the part of the proposition not relationally presupposed in relation to utterance, i.e. as the assertional part of the proposition, has the advantage of incorporating both the intuition of contrast and that of the ascription of properties.

Recall that assertion is the last in the series of propositions added to the common ground by an utterance, and that it is processed so that the auditor searches all the potential propositions containing the denotatum of the topic and eliminates those which do not fit the description given by the asserted proposition. Focus is that part of the asserted proposition which has to be assessed with respect to the topic. The explanation for the impression of contrastiveness of focus lies at hand: potential propositions in which the topic is involved, i.e. the background against which the assertion is assessed, differ from each other in one point, namely in that part which is not the denotatum of the topic, i.e. the focus. In assessing (4-27), the auditor concentrates on the propositions which contain the relationally presupposed material ‘John introduced Bill to someone’ and eliminates those in which the person to whom Bill is introduced is not Sue, i.e. which do not correspond to the description given by the asserted proposition. The eliminated propositions differ from the asserted one only in their focal parts, i.e. in the identification of the variable ‘someone’. The scope of alternative semantic values is thus determined by the relational presupposition rather than by the focus – the hearer chooses all those propositions which contain ‘John introduced Bill to X’. The role of the focus is to instruct the hearer which alternative propositions evoked by the relational presupposition are to be upheld and which eliminated from the discourse model.

The notions of direct and indirect topic, described in Section 4.3.3, may help us understand the property-assigning feature of focus. First the simpler case, that of direct topic. Consider once again (4-10), repeated for convenience as (4-28):

(4-28) *Peter works in the University library.*

Upon hearing this sentence, the auditor searches all the potential propositions in which Peter is involved and discards those in which he does not work in the University library: ‘Peter’ is

the topic, ‘works in the University library’ is the focus of the utterance. After s/he had accepted the asserted proposition, the auditor’s knowledge of the discourse referent ‘Peter’ increased, since s/he related the asserted proposition to background knowledge defined by this discourse referent. This new relation is felt to be a new property which has been added to the mental representation of ‘Peter’, whereas the assertive part of the asserted proposition, i.e. the part denoting the relation proper, the focus ‘works in the University library’, is felt to denote this property.

So far, so good. But how can a sentence like (4-27) be analyzed in this way? ‘John introduced Bill to someone’ is hardly to be understood as a discourse referent to which a property is ascribed by relating a proposition to the background knowledge defined by it. The following paraphrase may be helpful as a first approximation: ‘the entity who has the property of being the person to whom John introduced Bill has a property of being identical to Sue’. I shall contend that the subject of this almost amusingly cumbersome sentence, together with the restrictive relative clause, is the *indirect* topic of the utterance. Indirect topics are characterized by the fact that the hearer has to pragmatically *construe* either their entity status or their relevance relationship with the asserted proposition. It is the former feature that is of interest in this context: from the relationally presupposed propositional function (or open proposition) ‘John introduced Bill to X’ the auditor pragmatically construes an existentially presupposed mental representation of an entity, ‘a person John introduced Bill to’, giving it the status of a discourse referent. Once this accommodating action is accomplished, the process of relating focus to the background knowledge defined by this construed discourse referent is the same as with direct topics. Accordingly, the feeling that a new property has been assigned to a discourse referent by the focus comes into being just like in the simpler case of direct topics. The ‘search and eliminate’ principle of the incremental model and the assertive nature of focus are thus a sufficient explanation for the property-assigning feature of focus, too.

Focus is the assertive part of the asserted proposition, i.e. that part which is not under the relational presupposition. Four features repeatedly associated with the notion of focus – prominence, newness, contrastiveness and ascription of properties – are deducible from its assertive nature. Some approaches to information structure, like those of Molnár (1993, 1998), Rosengren (1997), Drubig (1992), etc., differentiate between two or three informational articulations on the basis of these features. Thus Molnár and Rosengren speak of the theme-rheme, topic-comment and focus-background structures, whereby the first is meant to represent the feature of newness, the second the ascription of properties, and the

third the prominence of focus. Upon the assertive definition of focus, this differentiation turns out to be superfluous: focus is ‘new’, highlighted, and ascribes properties due to its one criterial property, that of being the assertive part of the proposition. For this reason I shall not use the terms ‘comment’ and ‘rheme’: what is meant by them is already contained in my ‘focus’.

Finally, a rather uncontroversial issue (*pace* Krifka 1991): in contrast to topic, there can be only one focus per utterance. This naturally follows from the assertional definition of focus: if the asserted proposition is defined as the last in the series of propositions added to the common ground by an utterance, it is logical that there is only one proposition within the utterance which fulfills this criterion, and consequently only one unit of descriptive material carrying this proposition, i.e. only one focus.

4.4.2. Types of foci

The idea that there are two ontologically, qualitatively different focus types, although relatively old (it was to my knowledge first expressed by Halliday 1967), has gained momentum in the recent years, primarily through the publications of Szabolcsi (1983, 1994) and É. Kiss (1998a: 707ff., 1998b). These two types are called **information focus** and *identificational focus* by É. Kiss (I shall use the term **operator-like focus** for the latter). This is how they are defined:

An identificational (i.e. an operator-like, D.M.) focus represents a subset of the set of contextually or situationally given elements for which the predicate phrase can potentially hold; it is identified as the exhaustive subset of this set for which the predicate phrase actually holds. Semantically, the constituent called identificational focus represents the value of the variable bound by an abstract operator expressing exhaustive identification. ... If a sentence part conveys new, nonpresupposed information marked by one or more pitch accents – without expressing exhaustive identification performed on a set of contextually or situationally given entities, it is not an identificational focus but a mere information focus. (É. Kiss 1998b: 245-6)

É.Kiss’s information focus resembles (though defined in a rather rudimentary way) my focus, and does not need further specification; it is the operator-like focus, or rather the idea that it is ontologically different from the ‘information focus’, that I find problematic. First one Hungarian example from É.Kiss (1998a:710):

(4-29) Who did you meet at the concert?

(a) *Erzsivel találokztam.* ('Erzsi-with I-met') – 'I met ERZSI' (and no-one else)

(b) *Találokztam Erzsivel.* ('I-met Erzsi-with') – 'I met ERZSI' (among others)

Sentence (a) contains an operator-like focus: the focus-operator 'Erzsi' is said to introduce the set of relevant persons and asserts that it was Erzsi I met, and that I did not meet anyone else from this set. The information focus in (b) does not have this exclusive effect.

Against this (influential) view, and following Lambrecht (1994: 286ff.), Vallduví (1992) and Vallduví and Vilkkuna (1998), I consider the operator-like exhaustive focus to be simply a pragmatically conditioned variant of the normal, i.e. 'information' focus. There are two reasons for this. First, the exhaustive interpretation is in most (all?) cases due to **contrast**. Now, I have already mentioned that contrast is one of the general focus features. How can, then, a feature that is always present in focusing be responsible for one particular reading of focus? The answer is simple: it is a gradient notion. The principal feature of contrast is the elimination of alternatives. But the number of alternatives may vary from infinite to two: the smaller the number of alternatives to be eliminated, the stronger the contrast. Applied to focus processing, this means that, if the speaker limits the number of the possible worlds which the hearer is to search in her/his common ground, the feeling of contrast is stronger.

In a broad sense, every semantic peak is contrastive. Clearly in *Let's have a picnic*, coming as a suggestion out of the blue, there is no specific contrast with *dinner party*, but there is a contrast between picnicking and anything else the group might do. As the alternatives are narrowed down, we get closer to what we think of as a contrastive accent. (Bolinger, quoted after Lambrecht 1994: 290)

The stronger the contrast, the stronger the exhaustiveness effect. This is again explained pragmatically: if I assert one of the two alternatives, my auditor will, relying on the communicative principles of telling no more and no less than necessary (4.1), conclude that I intend to eliminate the other, even if it is not logically incompatible with the asserted one; if I assert one of many alternatives, the auditor is free to eliminate only those which are logically incompatible and to feel no commitment as to the others. Thus, the exhaustiveness effect (as demonstrated by Horn 1981) is simply a conversational implicature which naturally arises in absence of a block. Contrast and exhaustiveness are pragmatic values independent from focus: foci associated with explicit contrast are not ontologically (qualitatively) different from those which are not.

What I aim at is that all foci are basically the same. This is not meant to imply that contrastiveness and exhaustiveness cannot be formally marked: in many languages, explicitly contrastive foci have a marking different from those which are not explicitly contrastive (cp.

(4-29)). This brings me to the second reason I do not accept the ontological differentiation of operator-like and information foci. As in the case of topic (4.3.1), in the focus research there has been a tendency to concentrate on one type of focus marking in one language and to proclaim the results thus reached as being universal. Surprisingly, the deepest impact on the general theory of focus was not borne out by the research of the English focus constructions, but rather by the work done on Hungarian (Szabolcsi 1980, 1981, 1983 Horvath 1985, É. Kiss 1987, 1994, 1998b, Brody 1990, etc.). The Hungarian focus construction is characterized by a specific feature, namely by the fact that apart from marking the assertive part of the proposition, focus also marks this part as exclusively holding in the given context. In other words, in accordance with the principle that there is no one-to-one correspondence between form and content, what is coded by the linguistic form called ‘focus’ in Hungarian are two things: assertiveness (focus proper) and exhaustiveness, i.e. a pragmatic feature paraphrasable with ‘only’, closely related to but different from focus. This is an idiosyncratic, language-specific fact of Hungarian grammar, not a universal feature of focus: in other languages, the linguistic form specialized for marking focus does not carry this second meaning (though it can be partly united with other meanings, see É. Kiss 1998b, Matić 2002). I shall have more to say on the topic of differentiation of pragmatic values like topic and focus on the conceptual level on one hand and their concrete encodings in languages on the other below. For now, suffice it to conclude that the distinction between information foci and operator-like foci is due to an independent pragmatic feature, which limits the number of alternatives to be eliminated; there is no qualitative, ontological difference between the two types.

Even though there are no qualitatively different types of focus, there is a significant variation as far as quantitative properties are concerned. The famous Vallduvian (1992, 1994) broccoli examples may serve as the first approximation for this:

(4-30) *The boss* [*hates* [*BROCCOLI*]_{Foc}]_{Foc}

(a) What did you find out about the company?

(b) Is it true that the boss hates peanuts?

In context (a), the most probable interpretation of (4-30) is the one with ‘hates broccoli’ as the focus: the hearer is instructed to search all the potential propositions pertaining to ‘the boss’ and to eliminate those in which he does not hate broccoli. In the context (b), the preferred interpretation is the one with ‘broccoli’ as focus: the hearer concentrates on the potential propositions in which the boss hates something and eliminates those in which this something is not broccoli. I shall call the former type **broad focus**, the latter **narrow focus** (only partly corresponding to the predicate and argument focus in Lambrecht 1994). Broad focus is the

focus comprising the main predicate and optionally some of its arguments or adjuncts. Narrow focus is confined to one argument or adjunct. The difference between these two types is purely quantitative: if an utterance contains a broad focus, then the greater part of the propositional material is not under the scope of the relational presupposition, i.e. the greater part of the proposition is assertive; if a narrow focus is used, the greater part of the proposition is relationally presupposed, i.e. the smaller part is assertive.

Broad and narrow focus have been frequently equated with information focus and operator-like focus, respectively (e.g. É. Kiss 1998b). It is true that the exhaustive interpretation is more frequently found with the narrow focus, but this is again a purely pragmatic issue: it is generally the case that the number of alternatives is smaller when the greater part of the propositional material is relationally presupposed. Thus exhaustiveness (and the explicit contrastivity) is only a statistical, not a categorical property of narrow foci.

A more challenging hypothesis associated with the broad/narrow focus distinction is Lambrecht's (1994) implicit claim that narrow foci function through identification, whereas this is not the case with broad foci, which are thought to ascribe properties. In uttering (4-30) in the context (b), with narrow focus, the speaker instructs the hearer to identify the variable contained in the relationally presupposed proposition 'the boss hates X'; the assertion is simply the identification of 'X' with the denotatum of 'broccoli'. If (4-30) is uttered with broad focus, in the context (a), the assertion is not identificational, but 'predicational', i.e., in my terminology, property-assigning: the speaker adds a new property, 'hates broccoli' to the hearer's mental representation of 'the boss'. If this analysis turned out to be true, we would have to admit two ontologically, qualitatively different types of assertion, i.e. of foci, which are in some nontrivial way connected with the scope of assertive material in a proposition.

To express my opinion in advance, I think that the two alleged types of foci (of assertion), the **identificational focus**, in which the new asserted relation is established via identifying a variable, and the **additive focus**, in which the asserted relation is achieved by adding new descriptive material, are, similar to the operator-like and information focus, only pragmatically conditioned variants of the same unitary phenomenon of focus. Furthermore, I do not think that there is more than a mere statistical connection between broad and additive foci and between narrow and identificational foci. That is, narrow foci tend to be identificational, but need not be that way; broad foci are more often additive than not, but this is merely a question of frequency. The difference between identificational and additive foci is again rather quantitative than qualitative in nature, though in a way differing from the broad/narrow focus distinction: it is based not on the bare proportion of the relationally

presupposed material in a proposition, but on different types of expectations built into relational presuppositions. Consider the following pieces of discourse:

(4-31) Peter wanted to go home. But *he had lost his KEYS* and he knew it would take hours to find them.

(4-32) Peter wanted to go home. But he couldn't get in: *he had lost his KEYS*, and now he was standing helplessly in front of the door.

I should like to argue that *he had lost his KEYS* in (4-31) and (4-32), though formally identical and with identical presuppositional structures ('had lost his keys' is a broad focus in both cases), conveys an additive assertion in (4-31) and an identificational assertion in (4-32). In (4-31), there is no particular expectation connected with the relationally presupposed 'he' apart from the expectation normally associated with topics that the utterance to follow is somehow relevant with respect to the denotatum of the topic expression. It is this simple relevance expectation that is satisfied by the broad focus 'had lost his keys'. In (4-32), on the contrary, there is, apart from this, an additional expectation: the textual context signals to the auditor that there is a certain state of affairs in which the relationally presupposed discourse referent 'he' is involved, and that the assertion to follow will identify this state of affairs. The broad focus in (4-32) thus not merely adds a new relation to the common ground, it does it in a particular way, by identifying the state of affairs the existence of which is presupposed. The difference between the two foci can be schematically represented as follows:

(4-31') [he]_{Topic, relationally presupposed} [had lost his keys]_{Focus, asserted}

(4-32') [[he]_{Topic} takes part in X]_{relationally presupposed} [X = had lost his keys]_{Focus, asserted}

Consider now (4-33) and (4-34), both with a narrow focus:

(4-33) I went to see 'The Lord of the Rings' with the kids yesterday. As usual, the girls found the male actors the most interesting part of the film. *Mary fell in love with FRODO*, Joan with Aragorn.

(4-34) I went to see 'The Lord of the Rings' with the kids yesterday. You know, the girls are now at the age when they constantly fall in love with film characters. *It is FRODO* *Mary fell in love with* this time.

Here the forms are not identical (the point I am going to return to in 4.4.3), but the presuppositional structure is: 'Frodo' is the narrow focus, 'Mary fell in love' is relationally presupposed. Nevertheless, the assertions conveyed are different. In (4-33), the speaker instructs the hearer to search all the potential propositions in which Mary falls in love after watching 'The Lord of the Rings' and to eliminate all those in which the object of her love is not Frodo: the new relation is simply added to the relationally presupposed material. In (4-

33), the process is somewhat more complex: the hearer is instructed by the clues from the textual context to search all the potential propositions in which Mary falls in love *with one of the characters from the film*, and to add a new relation to her/his knowledge by identifying this character with Frodo. Schematically:

(4-33') [Mary fell in love]_{relationally presupposed} [Frodo]_{Focus, asserted}

(4-34') [Mary fell in love with X]_{relationally presupposed} [X = Frodo]_{Focus, asserted}

Note that in both cases it is not the way the assertion works that is different (the 'search and eliminate' principle is universally applied): it is simply the quantity of the implicit knowledge on the part of the hearer regarding the *type* of relation which is to be expected that makes the identificational and the additive focus distinct. In using the additive focus, the speaker does not presuppose that the hearer can identify the type of the relation s/he is about to assert; in using the identificational focus, s/he does. In a way, identificational foci are similar to the contrastive, or operator-like, foci, in that the number of alternatives to be eliminated is narrower here than in the case of additive foci. They differ from explicitly contrastive foci in that the number of alternatives is not restricted by the contextually given set of possible alternatives, i.e. by the knowledge of possible tokens for the focus value, as in the (b) sentence in (4-30), for instance, but by the knowledge of the type to which the focus value belongs. In this light, it is self-evident why narrow foci are more often identificational than the broad ones: when the greater part of the proposition is relationally presupposed, it is much more probable that the hearer will have enough context clues to infer the type of relation which is to be asserted than in the case only one discourse referent is presupposed.

Given the relevance of the number of alternatives for two of the three divisions proposed, namely for the information-focus vs. operator-like focus distinction and for the identificational vs. additive focus distinction, I propose to reduce these two distinctions to a single scale:

1. **additive focus** ('information focus', 'neutral focus'): the set of alternatives is unlimited.
 2. **identificational focus**: the set of alternatives is limited by the knowledge of the type the focus belongs to.
 3. **explicitly contrastive focus** ('contrastive focus', 'operator-like focus', 'exhaustive focus'): the set of alternatives is limited by the knowledge of the possible tokens for the focus value.
- The number of alternatives is highest with additive foci and is gradually reduced towards the lower end of the scale (a similar proposal was put forth by Sasse 1996 concerning broad foci only).

Let me now summarize this winding section. Focus is a *unitary phenomenon*, based on eliminating alternatives and ascribing properties. Since the number of alternatives to be eliminated can be reduced through contextual clues, and since this reduction can influence the *interpretation* of the focus, though *not* the principles according to which it works, three types of focus are differentiated regarding this feature – additive, identificational and explicitly contrastive focus. Yet another distinction is based on the proportion of the assertive material with respect to the proposition: broad focus, in which the predicate is a part of the assertive material, and narrow focus, in which the assertive material consists of only one argument or adjunct. There is a statistical correlation between the narrow focus and the identificational and explicitly contrastive foci on one hand, and between the broad focus and the additive focus on the other, but this correlation is pragmatically conditioned and is not categorical.

4.4.3. Some formal properties of focus

As indicated in 4.4.2, I consider it important to clearly distinguish between focus as defined in notional terms and focus as marked in particular languages. This is not due to any methodological purism but has, as I hope to have shown, important descriptive consequences. Let me repeat that I do not believe that there is a possibility for a one-to-one relationship between linguistic form and content. A construction which is primarily used as a focus marker may encode some other pragmatic or semantic value together with it; or it can encode only those foci which are accompanied by an additional pragmatic or semantic value. In order to understand what focus is, one has to abstract away from these additional features, i.e. one has to abstract away from the concrete formal marking and work with a purely notional apparatus¹. Of course, this is only the first step. The second step is to investigate the interplay of assertiveness and various other pragmatic and semantic features in determining the formal structure of a language as well as the specifiedness degree of the assertional structure in a language.

¹ This is also the reason why gross comparisons of ‘focus’ between languages are not only very coarse-grained, but also frequently downright wrong: not only that one focus construction in one language generally does not encode the same cluster of features as another focus construction in another language, but it is usually very different even from other focus constructions in the same language (Rooth 1996, É. Kiss 1998b; see also Section 5.3.).

It is the latter question this section is intended to shed some light on. I shall contend that it is a language-specific matter to what degree the assertional (i.e. focus) structure, as depicted in the foregoing chapters, is explicitly formally marked. More specifically, languages differ as to the degree to which the quantitatively determined focus types described in 4.4.2 are formally distinguished. In order to demonstrate this, I need one further distinction (first introduced by Lambrecht 1994), namely that between focus structure and focus construal. **Focus structure** is a property of constructions, or sentences: it represents the number of possible assertional interpretations of a construction, or of a sentence. Thus, the sentence ‘Peter saw your CAT’ can be used to assert that Peter [saw your cat], not [destroyed your car], that Peter saw [your cat], not [your car], etc. Furthermore, it covers the whole scale as to the number of alternatives to [saw your cat]: the broad focus may be asserted against the background of an unlimited number of possible propositions as well as against the background of only one other possible proposition. Which one of these interpretative possibilities is realized in the concrete communication is the matter of **focus construal**: the speaker uses a particular focus structure and the clues from the context to instruct the hearer which particular interpretation is to be chosen. Consider the following conversation:

(4-35) A: You know, I’ve heard that Peter saw a tiger in front of his house.

B: *Peter saw your CAT.*

The speaker B, who wants to make an explicitly contrastive assertion which identifies the variable in the relationally presupposed proposition ‘Peter saw something’, chose the focus structure which, among other interpretations, allows for the narrow focus construal on ‘your cat’ and for the explicitly contrastive focus. The textual context helps the hearer to *construe* this assertive type from the underspecified focus structure of the sentence *Peter saw your CAT*.

Where a focus structure allows for only one focus construal, I shall speak of an **unequivocal focus structure**; where one particular focus construal is only a subset of the interpretations licensed by the focus structure, we are dealing with an **underspecified focus structure**. Languages differ as to the predominance of the one or the other focus structure type, the borders drawn between different specification fields, and the obligatoriness of the use of unequivocal focus structures where they are at disposal. Let me illustrate this with an example.

The important difference between broad and narrow focus is marked by an unequivocal focus structure in a number of languages with specific word order systems, in recent years labeled discourse configurational, like Hungarian (Behrens 1982, É. Kiss 1987, 1994, 1998b), and in languages with morphological focus systems, like Yukagir (Fortescue 1996, Maslova

1997, 2003) or Somali (Saeed 1984, Tosco 2002). Witness the Hungarian sentences in (4-36) and the Kolyma Yukagir examples in (4-37)

(4-36) (a) *János* [*MEGHívta* *Évát*]_{Foc}

János preverb-invited Éva

(b) *János* [*ÉVÁT*]_{Foc} *hívta meg*.

János Éva invited preverb

‘János [invited [*Éva*]_{Foc}]_{Foc}.’

(4-37) (a) *tudel* [*end’o:npele alhudo:l lebieget mid’um*]_{Foc}

he animals lower earth-from took

(b) *tudel* [*end’o:npelek*]_{Foc} *alhudo:l lebieget minmele*.

he animals-FOC lower earth-from took-OBJ.FOC

‘He [took [animals]_{Foc} from the lower earth]_{Foc}’ (Maslova 1997: 464)

The (a) sentences contain broad foci: *meghívta Évát* ‘invited Éva’ and *end’o:npele alhudo:l lebieget mid’um* ‘took animals from the lower earth’ mark the assertive parts of their respective propositions. The broad focus scope is marked by the sentence accent on the preverb and the postverbal position of the argument in Hungarian, and by the absence of the focus marker on the object and the focus conjugation on the verb in Kolyma Yukagir. The (b) sentences represent narrow focus counterparts of the sentences in (a), with *Évát* and *end’o:npelek* as argument-focus expressions. In Hungarian, the narrow focus is marked by the immediately preverbal position of the focused argument and by the resulting postposition of the preverb, as well as by the position of the sentence accent on the focused argument. In Yukagir, the ‘predicative’ ending *-lek* on the focused argument and the object-focus conjugation are the markers of narrow focus. With respect to the broad/narrow focus distinction, then, Hungarian and Yukagir dispose of unequivocal focus structures: one focus structure allows for only one focus construal. In contrast to them, a great number of languages display only underspecified focus structures, at least as far as the broad/narrow focus distinction is concerned¹. Consider the English translations of (4-36) and (4-37). *János invited ÉVA* allows for at least two focus construals: the broad focus construal ‘invited Éva’ and the

¹ Focus structures underspecified as to the broad/narrow focus distinction have attracted a great deal of attention in the last twenty or so years, mostly under the label *focus projection* (for a good overview of the relevant literature see Kennedy 1999): roughly, the idea is that the single constituent which can be focused in the narrow focus construal projects the focus feature to the predicate in the broad focus construal, whereby the focus feature is generally understood as an abstract, more or less syntactic, phenomenon and the projection as a derivational process.

narrow focus construal ‘Éva’. The same holds true for three of the languages which are the object of the present study – Albanian, Modern Greek and Serbo-Croat:

(4-38) *Gjoni [ftoi [EVĒN]_{Foc}]_{Foc}* (‘Gjon-the invited Evë-the.’, Albanian)

(4-39) *O Janis [kalese [tin EVA]_{Foc}]_{Foc}*. (‘the Jannis invited the Eva.’, Modern Greek)

(4-40) *Jovan [je pozvao [EVU]_{Foc}]_{Foc}*. (‘Jovan AUX invited Eva.’, Serbo-Croat)

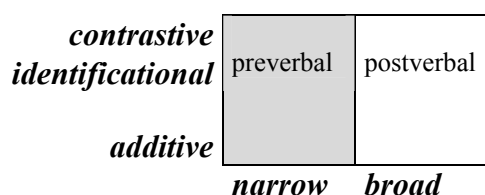
In all three cases, the focus structure is underspecified between the narrow (*Evën, tin Eva, Evu*) and the broad (*ftoi Evën, kalese tin Eva, je pozvao Evu*) focus interpretation. In other words, English, Albanian, Modern Greek and Serbo-Croat dispose of underspecified focus structures with respect to the broad/narrow focus distinction, since one focus structure allows both for the broad and for the narrow focus construal (though see Section 8.3.).

Broad and narrow foci, however, are not the only relevant variants of focus structure: orthogonal to this distinction is the one based on the number of alternatives against which the focus is assessed, with additive, identificational and explicitly contrastive foci as the relevant landmarks on the scale from an infinite number of alternatives to the contextually given set of only a few. Furthermore, some of the restricted number of alternatives can be conventionally associated with further semantic effects, like exhaustiveness (see 4.2.2.). Now, this distinction represents a variable independent of the broad/narrow distinction outlined above. This means that a language may possess unequivocal focus structures relative to the broad/narrow focus distinction on one hand, and only underspecified focus structures relative to the number of alternatives on the other; or the underspecifiedness with respect to the broad/narrow focus distinction is paired with unequivocal focus structures with respect to the number of alternatives, etc.

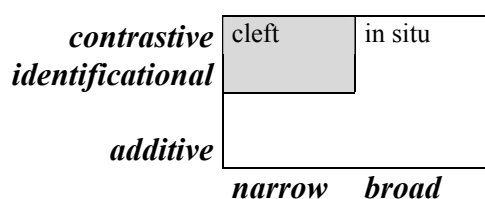
Let me adduce one example to render this point clearer. As demonstrated above, Hungarian is unequivocal as to the broad/narrow focus distinction, but underspecified as to the number of alternatives involved, only with a conventionalized exhaustiveness interpretation (É. Kiss 1998b), which can nevertheless be cancelled under the appropriate pragmatic conditions (Roberts 1998). Thus, the narrow focus sentence (b) in (4-36) can be uttered both in the context where there is only one possible alternative to Éva being invited by János, e.g., Erzsi, and in the context in which there are no restrictions to the number of possible candidates for an invitation. On the other hand, English, which is underspecified as to the broad/narrow focus construction, has a focus structure which is unequivocal with respect to the number of alternatives, at least as far as the narrow focus is concerned. If the narrow focus is nearer to the additive part of the scale, the prosodic focus in situ, as exemplified in the translation of (4-36), will be used. If the narrow focus is identificational, or

explicitly contrastive, the use of the cleft construction is preferred or obligatory: *It was EVE John invited*. When the speaker uses broad focus, however, s/he does not have any unequivocal focus structure at her/his disposal, so the underspecified structure like that in the translation of (4-36) is the only option. The only exception to this is Hiberno-English, in which the cleft construction can be used for identificational, etc. foci even when the focus is broad: *It is inviting EVE (that) he was* (Filppula 1999). A very simplified picture of this is given in the following diagrams (a continuous field in the diagram denotes one marking strategy):

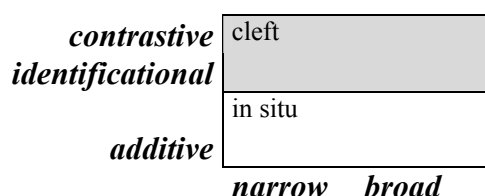
(4-41) (a) Hungarian



(b) English



(c) Hiberno-English



The relevance of the distinction between unequivocal and underspecified focus structures for the present study is primarily methodological in nature. While describing the interplay of information structure and grammar, one has to keep the notion of focus structure and focus construal clearly apart. The former is a fact of grammar and as such language-specific; the latter is a matter of interpretation of grammatically encoded pragmatic values, and probably subject to more universal rules than the former.

4.4.4. Definition

FOCUS is that part of the proposition whereby the assertion differs from the relational presupposition, i.e. that part which is not relationally presupposed. It is thus the carrier of the assertion in utterances, which is why it is felt to be the most prominent element of every utterance. Though relationally nonpresupposed, focus can contain existentially presupposed material, so that it is not necessarily referentially ‘new’. However, the relationship between the topic and the focus in an utterance has to be ‘new’, or at least processed so as to become such. Two basic properties of focus – exclusion of alternatives and attribution of properties – are reducible to its assertive function. There can be only one focus per utterance. With respect to the proportion of assertive material in a proposition, two types of focus are distinguished – broad and narrow focus; with respect to the contextually licensed number of possible worlds to be eliminated, a scale may be established, with additive, identificational and explicitly contrastive foci as relevant landmarks. An important feature of the formal marking of focus is the distinction of focus structure and focus construal: focus structure is a name for the set of assertional interpretations licensed by a linguistic form; focus construal is a concrete assertional interpretation chosen on the basis of focus structure and contextual clues for a particular utterance. Focus structures can be unequivocal, if they correspond to only one focus construal, or underspecified, if the number of focus construals they license is larger than one.

4.5. Terminological issues

In order not to fall under the scope of Levinson’s critique concerning conceptual vagueness and terminological confusion dominating the field of topic-focus research, as expressed in the motto of this chapter, I shall devote a separate section to clarifying the issues of terminology. Much of what appears here has, in one form or the other, already been said; only some of it is new. The section should therefore be read as a summary of the conclusions arrived at under the terminological aspect.

4.5.1. Four levels

I strictly distinguish between four levels, the propositional level, the level of expression, the syntactic level, and the level of pragmatic inference. Although this sometimes leads to a rather

cumbersome way of expression, the distinction is essential: much of the confusion referred to above stems from the joyfully careless way in which many scholars jump from one level to another.

First the propositional level. **Propositions**, understood as relational entities consisting of a certain number of terms and a predicate, have a **presuppositional structure** and an **assertional structure**. Both terms refer practically to the same thing, viz. to the fact that propositions are informationally structured; the difference is only in the emphasis given to one of the two basic notions of information structure. Consequently, my choice of the term ‘presuppositional’ or ‘assertional structure’ will depend solely on the relevance of the one or the other notion in the given context.

The element of the proposition which carries the relational presupposition will be called **topic**; the one representing the assertive material is **focus**. In other words, the terms ‘topic’ and ‘focus’ refer *only* to mental representations, referents or denotata of expressions, *never* to the expressions themselves. Both topics and foci are treated in this study as ontologically unitary phenomena, though this is not to say that there are no differences between different topic and focus types. With respect to their assumed status in the mind of the hearer, topics may be **ratified** or **non-ratified**, the former treated as easily accessible to the hearer, the latter less so (more on this in Section 6.2.). With respect to the way they are encoded, topics may be **direct** or **indirect**, a distinction about which I shall have to say more in Section 4.5.2. Furthermore, the topical material may comprise only one or two terms in a proposition, or it may equal the whole proposition minus one focused term. Generally, I shall speak of both as ‘topic’; however, when the broad scope of relational presupposition is to be emphasized, the term **presupposed propositional function** will be used. Thus, in *It is PETER I saw yesterday*, the denotatum of the string ‘I saw (someone) yesterday’ will be called presupposed propositional function. The distinctions pertaining to focus have been abundantly explained and illustrated in 4.4.3. Depending on the scope of assertion, focus may be **broad** and **narrow**. The number of alternatives evoked by focus varies from two to infinite; if a focus is closer to the former end of the scale, I shall speak of **contrastive foci**, if to the latter, of **additive foci**. The term **identificational focus**, which shall play an important role in the present study, denotes the kind of focus in which the alternatives are restricted as to the ontological class they belong to.

The second level is that of expression. When referring to concrete lexical material, I speak of **topic expressions** and **focus expressions**. If, for any reason, the aspect of

presuppositional/assertional structure is to be brought to the foreground, the terms **presupposed material** and **assertional material** will be used instead.

On the level of syntax, the relevant notions are **topic structure** and **focus structure**, denoting properties of syntactic structures and intonational patterns. They refer to the number of possible presuppositional and assertional interpretations of a construction. Or, the other way around: The topic structure and the focus structure of a construction represent the sum of all topic-focus configurations, or presuppositional/assertional structures, with which a construction may be invested. When topic/focus constructions are defined within the word order system of a language, which is the case in the languages under consideration in the present study, I speak of **topic positions** and **focus positions**, referring to the slots in the sentence template of a language reserved for topic or focus expressions. Both topic structures and focus structures may be **unequivocal**, if they allow for only one topic-focus configuration, or **underspecified**, if the number of possible interpretations is larger than one. A construction may be unequivocal or underspecified with respect to all the differentiations of topic and focus enumerated above.

Finally, the level of pragmatic inference. In the case of underspecified topic/focus structures, the topic-focus configuration of the utterance, i.e. its presuppositional/assertional structure, has to be inferred by the hearer through the clues from the context, the frames evoked by the lexical material used, etc. In other words, confronted with an underspecified topic/focus structure, both the speaker and the hearer have to *construe* that presuppositional/assertional structure which is most plausible, choosing from the interpretations licensed by the given topic/focus structure. The terms used for this are **topic construal** and **focus construal**. A construction with an underspecified topic or focus structure has more topic or focus construals paradigmatically, but a concrete utterance can have only one construal. In constructions with unequivocal topic or focus structures, the paradigmatic value of the construction is equal to its syntagmatic value, i.e. one.

Instead of a summary, here is an example designed to illustrate how my analytical apparatus works. The analysis is performed in detail, i.e. with a lot of redundancies:

(4-42) [How did your colleagues spend their holidays?]

Peter, he went to ITALY.

1. level: Proposition. [Peter went to Italy]

presuppositional/assertional structure:

[Peter], [Italy] – existentially presupposed

[Peter] – relationally presupposed

[went to Italy] – asserted

topic: [Peter]

focus: [went to Italy]

2. level: Expression. *Peter, he, went, to, Italy*

relationally presupposed material: *Peter, he*

assertive material: *went, to, Italy*

topic expression: *Peter, he*

focus expression: *went, to, Italy*

3. level: Syntax. *Peter, he went to ITALY.*

topic structure: left dislocation, unequivocal

focus structure: postverbal focus domain, underspecified as to broad/narrow focus

(both [went to Italy] and [Italy] possible foci)

4. level: Pragmatic Inference.

topic construal: unnecessary, since unequivocal topic structure

focus construal: broad focus [went to Italy]

Only the 4. level deserves some additional comment: in the context of the question “How did your colleagues spend their holidays?”, the broad focus construal [went to Italy] is more probable than the narrow focus construal [Italy]. However, it is not inconceivable that the latter variant is more plausible in certain cases, e.g. in lists: *Peter, he went to ITALY, Mary went to BRAZIL, and Ann went to BAHAMAS*. Here, the speaker seems to presuppose ‘everybody went somewhere’, where ‘everybody’ = {Peter, Mary, Ann}. The narrow focus construal is thus more probable: [Peter went to X], [X=Italy]. The level of pragmatic inference represents a problem for linguists, who (wish to) work with clear-cut notions (see Section 11). The speakers of natural languages seem to come to terms with this indeterminacy without difficulties.

4.5.2. Focus domains and indirect topics

The terms *focus domain* and *indirect topic* denote informationally underspecified structures of major relevance for the present study.

Focus domain is the term I shall use for that part of the sentence which represents the maximal scope of assertion in constructions with the focus structure underspecified as to the

feature broad/narrow focus¹. Consider example (4-42): as shown above, there are at least two focus construals for *he went to ITALY*, [went to Italy] and [Italy]. On the propositional level, the terms used to describe this are *broad* and *narrow* focus. The terms *broad* and *narrow focus construal* describe the process on the level of pragmatic inference. On the level of expression, the focus expressions in the broad focus construal are *went, to, Italy*, in the narrow one only *Italy*. What we do not have is a term which would describe this kind of situation on the level of syntax: *focus domain* is intended to mend this problem. Thus, in the sentence *he went to Italy*, the focus domain of the sentence is *went to Italy*, meaning that only this part of the sentence, but not the subject *he*, may be interpreted as focus; which one of the possible construals will be used is a matter of pragmatic inference, not of syntax.

Focus domain is consequently a part of grammar, i.e. a conventional pairing of a certain form with a certain meaning. The meaning in question is approximately ‘the maximal scope of assertion’; the form varies from language to language (cp. Matic 2002). Many European languages, including Alb., MG and SC, have a postverbal focus domain, with the verb as the left, and the nuclear stress as the right border. The nuclear stress (more on the notion in 5.3.) generally falls on the last non-verbal part of the focus domain. All this is illustrated by the sentence *he went to ITALY*: the focus domain is *went to ITALY*, the left border of the domain being the verb, the right one the constituent carrying the nuclear stress. If the sentence is extended to *he went to Italy YESTERDAY*, the focus domain is *went to Italy YESTERDAY*; if the extension is left unaccented, the result, *he went to ITALY yesterday*, will have the same focus domain as the original sentence, *he went to ITALY*, the temporal adverbial being excluded from the focus domain by being placed after the nuclear stress, i.e. the right border of the domain.

The principles of focus domain formation in Alb., MG and SC are identical to those described for English. For instance, the MG sentence *o Petros pije stin ITALIA* (‘the Peter went in-the Italy’) has as its focus domain *pije stin ITALIA*. The extended version, *o Petros pije stin Italia XTES* (‘the Peter went to Italy yesterday’), if it is accented on the last element, has the focus domain *pije stin Italia XTES*. If the temporal adverbial is void of accent, as in *o Petros pije stin ITALIA xtes*, the focus domain is *pije stin ITALIA*. The focus domain in the Balkan (and many other) languages, being a construction, may be represented as a template (‘x’ denotes any element, ‘X’ any element carrying the nuclear stress):

(4-43) **[[verb] [x] [X]]_{Focus Domain}**

¹ Both my terminology and my notional apparatus are strongly influenced by Lambrecht (1994).

As Lambrecht (1994) has noted, focus domains, as syntactic entities, may be discontinuous, with the topical material intervening between the verb and the nuclear stress, as in *I saw him YESTERDAY*, where the maximal scope of assertion is defined by the sentence fragment *saw him YESTERDAY*. Without an accent, *him* cannot but be topical. One of the presuppositional structures is thus [I in relation to him]_{Topic} [the relation = saw yesterday]_{Focus}. In spite of this, the topic expression *him* is placed within the focus domain, due to an independent rule of the English grammar. It is in this sense that focus domains may be discontinuous. This phenomenon will be of some importance in the description of the syntax of the Balkan languages (cp. especially Chapter 6).

The notion of **indirect topic** has been dealt with in some detail in 4.3.3. In the light of the division of labor between the four levels proposed above, the term presents us with a slight terminological problem, since it extends over more than one level. Recall that indirect topics are those topics which are either not presented as entities or not explicitly marked for the relationship with the asserted proposition. The first property has to do with the propositional level; the second with both the propositional and the syntactic levels. Obviously, some precision is needed.

When speaking of the propositional level, the term **indirect topic** will be used for those elements which are either not terms of the predicate, or not its direct arguments, or not present in the expressed proposition at all, but subsequently construed in order to process the assertion. The first case is illustrated by (4-21) (*Physically, he feels fine*), the second by (4-22) and (4-23) (*In the yard is a bicycle, My FOOT hurts*): ‘physically’ is not an entity; ‘in the yard’ and ‘my’ are not direct terms of the predicate. The third case, an indirect topic not present at all in the expressed proposition, is illustrated by *It is RAINING*: in order to process this termless proposition, the hearer has to construe an entity with respect to which it is relevant: ‘the place where we are now’, ‘the place and the time where the hero of the novel is at the moment’, ‘London’, ‘Paraguay’, or whatever the context clues are. This construed entity is then treated as an indirect topic, on a par with ‘body’, derived from ‘physically’, ‘the yard’ derived from ‘in the yard’, and ‘I’ derived from ‘my’.

The level of expression is less complex: in the examples quoted above, *physically*, *in the yard*, and *my* will be called **indirect topic expressions**. In the case of *It is raining*, I shall speak of sentences without indirect topic expressions.

Syntactically, indirect topics are encoded as non-subjects and non-objects, i.e. without an explicit marking of the relevance relationship to the predicate. Since they are not encoded as topics, or are not encoded at all, but the hearer has to construe them on the level of pragmatic

inference, indirect topics are **underspecified topic structures**. Whereas in *Peter went to ITALY* the topic of the proposition is expressed unequivocally, as an unaccented subject preceding the focus domain, in *My FOOT hurts* there is no unequivocal topic: the hearer uses her/his world knowledge and the clues from the context in order to interpret the proposition conveyed by the sentence as an assertion about the referent 'I'.

Unlike direct topics, where the relationship between form and function is straightforward, indirect topics are thus subject to **topic construal** on the fourth level, that of pragmatic inference.

4.6. Why is information structure formally marked?

I should like to conclude this chapter with a few remarks on the function of the formal expression of information structure. It is generally assumed that it is the principle of cooperative communicative behaviour that is responsible for the fact that topic and focus are probably universally somehow formally marked across languages: the speaker makes it easier for the hearer to decode the message by giving her/him instructions how to process the information contained therein. This is doubtlessly true, but I think that there is more to the coding of information structure than simply being nice to one's interlocutors. Recall that communication is, apart from being a cooperative action, also an intention-driven and goal-oriented enterprise, and that one of the principle reasons for the choice of presupposed and asserted descriptive materials is to make one's intentions transparent. Recall also that the context does not determine the presuppositional and assertional structures, but merely makes the one or the other structure more or less probable.

These three facts point to the explanation I want to offer (for a similar approach see Keijsper 1985:65ff. and Fuchs 1980). Speakers mark the information structure not only in order to ease the processing, but also in order to express their intentions: as Keijsper (1985:76) puts it, speakers are not only altruistic, but also, and primarily, egoistic. This is why the idea of the partial context-independence of information structure is of importance here: in the given context, the choice of one information structure rather than the other is a signal to the hearer that the speaker has one intention rather than the other. In a way, the speaker uses information marking to lead the hearer through the discourse.

Let me illustrate this with an example from a story by Šuškin (taken from Keijsper 1985:66; only the English equivalent of the Russian original is given). All the inhabitants of a

village gather near the local club, where a travelling amateur theatre is expected. A person, a new-comer in the village, tries to attract attention by starting a fight. But nobody is paying attention. The story ends with the words:

(4-44) *Just at that moment the theatre group arrived.* And everyone went to watch the theatre group.

The first sentence may be read with two different accentuations, i.e. with two different information structures. In the first reading, it is the verb that carries the sentence accent: *just at that moment the theatre group ARRIVED*; ‘the theatre group’ is marked as the relationally presupposed topic, ‘arrived’ as the focus. In the second, the accent is carried by the subject: *just at that moment the THEATRE group arrived*; the subject-verb complex is marked as the broad focus. What is the practical difference between these two readings? In the first reading, with the accent on *arrived*, the speaker referentially presupposes ‘the theater group’, meaning that s/he creates the expectation on the part of the hearer that the new assertion will pertain to this discourse referent. This expectation is reinterpreted as anaphoric relevance of ‘the theatre group’(see 4.3.2.): the knowledge that the theatre group is to appear is evoked in the form of an expectation which is satisfied by the assertion ‘arrived’. The message is roughly ‘the theatre group, whose arrival you, reader, as well as the villagers, had (more or less impatiently) expected, finally appeared (and put an end to the unpleasant scene)’. In the second reading, where the subject-verb complex ‘the theatre group arrived’ is focused, there is no effect of anaphoric relevance, since there is no relational presupposition, so that no expectations are built into the presuppositional structure. Therefore, the utterance has a note of suddenness, of an event abruptly interrupting the ongoing state of affairs (see Sasse 1996 for the interruptive function of the verb-subject foci). In the former case, with the focus on ‘arrived’, the impression is that the arrival of the theatre group is presented from the perspective of the villagers who are annoyed by the aggressive behavior of the intruder and are only waiting for the theatre group to appear; in the latter, with the verb-subject focus, the event is rather presented from a quasi-objective point of view, as if someone were observing the whole situation from outside. These two effects, which reveal two different intentions on the part of the speaker, are due to different presuppositional structures, i.e. different knowledge frames these presuppositional structures evoke.

It is not only different presuppositions, but also different (types of) assertions that can be used by the speaker to reveal her/his intentions and thus lead the hearer through the information flow. The relevant point here are the alternatives, i.e. the propositions excluded by the assertion, with respect to both their scope and number. An example (modified from

Keijsper 1985:70) should illustrate this. An apartment is being renovated, which disrupts the tenor of daily life. Remembering the good old times before the renovations, a member of the family says:

(4-45) *We used to drink tea in the evening.*

Again, this sentence can be pronounced with at least two different accentuation patterns and three information structures:

(a) *We [used to drink tea in the [EVENING.]_{Foc}]_{Foc}*

(b) *We used to drink [TEA]_{Foc} in the evening.*

If the underspecified focus structure (a) is construed as the broad focus on ‘used to drink tea in the evening’, the speaker relationally presupposes only the discourse referent ‘we’. The number of alternatives to the asserted proposition is, as is more often than not the case with broad focus construals, practically unlimited. These two facts – the narrow relational presupposition and the additive broad focus – determine the interpretation of this pragmatic construal: the expectation of the hearer is to be informed about ‘we (at the time before renovations)’; the additive focus, which is, due to the great number of alternatives generally not interpreted as exhaustive (see 4.4.2.), leaves most of the alternatives simply unmentioned upon. This leads to the expectation that the asserted proposition is only one of the assertions which hold true of ‘we (at the time before renovations)’. The ideal textual context for this construal is thus the one in which a number of propositions is asserted about the topical discourse referent in order to give a comprehensive information about it – say: *we used to drink tea in the EVENING, and we watched TELEVISION, talked to EACH other*, etc. The broad focus construal thus reveals the speaker’s intention to give a (more or less) complete picture of a certain discourse referent (which is, by the way, the reason why broad focus is the basic means of producing longer stretches of narrative texts).

If the same focus structure is construed as the narrow focus on ‘in the evening’, the number of alternatives is drastically reduced: only those potential propositions in which ‘we’ used to drink tea at a certain time are the relevant background here. Although the reduction of the number of alternatives may lead to the exhaustiveness effect, this need not be the case; what is certain is that the alternatives excluded through assertion evoke a sort of contrastive interpretation, since the number of times at which one can drink tea during the day is limited. The hearer is thus likely to interpret the utterance as a statement in which the speaker intends to contrast the previous state, when we had our tea at the evening, with the present state, when we drink tea on irregular basis, or at absolutely impossible times, or only in the morning, etc. An appropriate textual context would be, e.g. *we used to drink tea in the EVENING, and look at*

us NOW: FIVE o'clock in the morning, and we are sitting in this CHAOS and hastefully sipping our TEA. Giving a narrow focus interpretation to the adjunct *in the evening* thus instructs the hearer to interpret the utterance as a complaint about the loss of regularity in the everyday life.

Finally, if the unequivocal focus structure (b) is chosen, the same principle of limited alternatives as in the case of the narrow focus on 'in the evening' is at work. But the alternatives are different here, since it is the beverage which we used to consume that is focused now. The alternatives are, as is typical for narrow foci (though not necessary, see 4.4.2), limited: it is either the case that we drink other, presumably less tasty, things instead of tea, or we do not drink anything at all. In both cases, the alternatives against the background of which the speaker asserts the proposition (b) suggest that our drinking habits have deteriorated. A good textual context would be *we used to drink TEA in the evening; the bloody renovations cost so much MONEY that the only thing we can afford in the evening now is a glass of WATER.* Narrow focus on 'tea' thus suggests that the utterance is to be understood as a complaint about the material deterioration of the family due to great costs of the renovations.

The two examples adduced above have hopefully clearly demonstrated what I mean by saying that speakers, in formally marking information structure, lead their audience through information flow. The fact that they are to a certain degree free to choose the information structure independently of the clues from context, or, the other way around, that every context allows for a certain number of information structures, is crucial in this respect: one of the points I intended to emphasize in this chapter is this **creative nature** of the discourse-pragmatic notions like presupposition, assertion, topic and focus. Keijsper's examples have also shown that information structure is, like everything else explicitly coded in language, only a starting point in both encoding and decoding utterances. The rest is the work of pragmatic principles, primarily the ones of saying no more and no less than necessary. To paraphrase Carston (1999:377): formally encoded meanings are only a thin icing on the substantial pragmatic cake.

5. Preliminaries

5.1. Preliminary notes on Albanian, Modern Greek and Serbo-Croat

In this section, a number of peculiarities of Alb., MG, and SC relevant for the present study will be addressed.

As is well known, Alb., MG, and SC are Balkan languages, Alb. and MG belonging to the core of the *Sprachbund* and SC being located somewhere on its margins. In practice, this sometimes simply implies that some typical Balkanisms are present in Alb. and MG, but not in SC, which is the case with most morphological areal features. Sometimes, however, as will become patent in the course of this study, especially with respect to the syntactic areal features, the core-margin opposition is reflected in different sociolinguistic statuses of certain features in Alb. and MG on one hand, and in SC on the other, the common pattern being the folkish/colloquial character of a Balkanism in standard Alb. and standard MG versus poetic/archaic note carried by the same feature in SC, though there are exceptions to this (see especially Chapter 6). These ‘stylistic’ differences presumably stem from dialectal and historical divergences in the development of the modern standard languages on the Balkans. Since the purpose of my study is not primarily sociolinguistic, nor diachronic, but rather descriptive, I shall confine myself to registering the facts of this kind and giving a tentative explanation, leaving deeper insights into the complicated problems of the Balkan diachrony and sociolinguistic variation to the future research.

All three languages have a relatively rich morphology and are roughly characterizable as SVO. In the descriptions of the MG¹ and Alb.² word order systems, however, it has been repeatedly stated for almost a century that the basic word order is SVO in the main clauses, but VSO in the dependent ones, which is partly confirmed by the statistics presented in 5.5.3. Since this is obviously a question of some relevance for the present study, it will be dealt with in detail in Chapters 9 and 10. Note that no such claims have been put forward concerning SC embedded clauses.³

¹ Thumb 1910: 192, Tzartanos 1963: 276, Mackridge 1985: 237, Lascaratou 1989, 1998.

² Haebler 1957, Buchholz and Fiedler 1987: 557.

³ This seems to be the right place to give a short assessment of the relevant literature on word order in Alb., MG and SC. There is to my knowledge only one monograph on Alb. word order, Skënderi (1997), which contains a fine bibliography and a nice list of word order patterns based on Rushi (1983), but unfortunately gives neither interesting descriptive details nor deep theoretical insights. Some interesting small-scale work was done by Floqi

Apart from being SVO (with or without shades of VSO), Alb., MG and SC are often characterized as free word order languages, which is tantamount to saying that much of their word order is pragmatically determined. The main purpose of my work is to find out what kind of pragmatic motivation is responsible for the word order flexibility, so that this topic will receive due attention in the course of the present study. Suffice it for now to say that all three languages have two topic positions, one clause-initial and one postverbal, at least two narrow focus positions, a clause-initial and a clause-final one, and the typical A(verage) M(odern) E(uropean) focus domain, with the verb as the left and the accented constituent as the right border of the domain. Very roughly, the Balkan sentence can be represented as follows:

(5-1) **[Topic] [Narrow Focus] [[Verb] [Topic] [Narrow Focus]]_{Focus Domain}**

(1969, 1976), and Rushi (1983, 1984, 1985, 1988), who tried to apply the Praguean model on Alb. The only study of verb-subject order I am aware of is an article by Haebler (1957), in which it is claimed that Alb. is verb-initial in embedded clauses, and that MG displays the same syntactic behaviour. The best description of Alb. word order is thus still to be found in the standard work on Alb. grammar, Buchholz and Fiedler (1987: 536-561). As far as MG is concerned, the situation is only slightly better. There are two book-length studies: Lascaratou 1989 (cp. also Lascaratou 1998), in which it is claimed, in the spirit of Simon Dik's LIPOC and comparable principles, that the complexity and the length of constituents are the major factors influencing word order, and Tzanidaki 1996 (known to me only from the summary in Tzanidaki 1998), where some well-known facts are described with the technical apparatus of Hudson's Word Grammar. The number of papers dealing with MG word order is considerable, with almost all of the work conducted in some version of the Chomskyan variant of generative grammar. Although most of this work is of no interest to those who do not adhere to this particular linguistic theory, I should like to single out the long-lasting debate within the British school of MG studies: Initially, Horrocks (1983) and Philippaki-Warbuton (1985) expressed opposed opinions on the basic word order in MG, the former allowing for two basic WOs (SVO and VSO), the latter arguing that VSO is the only basic word order, SVO being derived through topicalization. This soon developed into a debate on configurationality, with Catsimali (1991) and Horrocks (1994) treating MG as a non-configurational or semi-configurational language, and Philippaki-Warbuton (1987), Tsimpli (1995), and others, ascribing it configurational or discourse-configurational features (see Horrocks 1994 and Alexiou and Anagnostopoulou 2000 for details). There are three studies devoted specifically to VS: Valioui (1994), Sasse (1995) and Alexiadou (1996, 2000), all of which will be (or have been) commented upon in due course. The most informative account is again to be found in a section on word order within a standard grammar: Mackridge (1985: 233-249). Finally, the situation in SC is much better than in the other two languages, due to the existence of one excellent book: Popović 1997 (written in 1972), which gives plenty of naturally occurring material classified in a fine-grained fashion, with a very good feeling for detail. Although the theoretical framework within which the book is written (the Praguean theme-rheme dichotomy) is somewhat obsolete, this does not make the sound descriptive work conducted less useful and informative. Since Popović gives a detailed overview of the (not very rich and not very helpful) literature up to 1997, I refer the interested reader to the pages 13-18 of this study.

Of course, there are also some relevant differences in the syntactic structures of the Balkan languages. Some of them, like those pertaining to relative elements and subordinators, will be handled *suo loco*, in the sections devoted especially to these topics. Some, however, are ubiquitous in the present study, so they will be dealt with in this, introductory, section.

First the clitics, which are in all three languages subject to special positioning rules. These words display different syntactic patterns of behaviour and comprise slightly different word classes in Alb. and MG on one hand, and in SC on the other. In Alb. and MG, all clitics are directly preverbal in finite clauses; in SC, they occupy the Wackernagel position, i.e. the position after the first word/phrase of the intonation unit:

(5-2) Alb.: *E pashë, i dhashë një lule.*

her:CLIT I-saw, to-her:CLIT I-gave a flower

MG: *Tin iða, tis eðosa ena luluði.*

her:CLIT I-saw, to-her:CLIT I-gave a flower

SC: *Videh je, dadoh joj cvet.*

I-saw her:CLIT, I-gave her:CLIT flower

‘I saw her, I gave her a flower.’

In the interlinear gloss, clitic words will be marked with :CLIT, as the issue of clitics has some relevance for my purposes. Namely, clitics may, due to their strictly grammaticalized position, apparently break some of the adjacency rules I posit. For instance, I shall claim (Chapter 8) that verbs of saying have to be immediately adjacent to the preceding quote. In Alb. and MG, this constraint can be apparently violated if the sentence contains a clitic. However, in view of the fact that clitics conform to independent word order rules and form a phonological word with the verb, this violation is indeed only apparent, the [clitic+verb] complex counting simply as [verb]. In contrast, due to the Wackernagel position of the SC clitics, in this language the question of adjacency does not arise at all, as is visible from the following example:

(5-3) Eng: *You needn't stray off too far ..., his partner admonished.* (London, p. 18)

Alb: [quote], *e porositi Henriku.* (p. 18)

him:CLIT commanded H.-the

MG: [quote], *ton simvulepse o sindrofos tu.* (p. 24)

him:CLIT advised the colleague his

SC: [quote], *opomenu ga drug.* (p. 22)

admonished him:CLIT friend

As far as word classes the clitics encompass are concerned, the main difference lies in the status of the auxiliaries and the copula. In Alb. and MG, the auxiliaries and the copula are not clitics, apart from the future particles *do të* and *tha*. In SC, the past tense auxiliaries and the present tense forms of the copula are Wackernagel clitics.

(5-4) Alb.: ***Kemi*** *thënë* | *se ne jemi të lumtur*.

AUX:1PL said | that we are the happy

MG: ***Exume*** *pi* | *oti emis imaste eftixismeni*.

AUX:1PL said | that we are happy

SC: ***Rekli smo*** | *da smo mi srečni*.

said AUX:1PL:CLIT | that are:CLIT we happy

‘We said that we were happy.’

Vertical strokes in (5-4) mark intonation units. What the example illustrates is that auxiliaries (*kemi*, *exume*) and the copula (*jemi*, *imaste*) in Alb. and MG behave like other verbs, whereas in SC, due to the Wackernagel rule, they must not occupy the first position in the clause (*Rekli smo*, not **Smo rekli*), and are confined to the second position only (*da smo mi*, not **da mi smo*). The consequences of this difference for the description of verb-subject order will be dealt with later (see 6.3.2.3., 8.1.1.3., 8.2.2.4. and 11.2.1.).

Finally, let me note in passing that Alb. and MG, but not standard SC, have a clitic doubling construction, in which the referents of direct or indirect objects are represented simultaneously by a clitic pronoun and a full noun phrase.

The second relevant point of divergence is the position of modal adverbials. Adverbials, like all other elements, may of course be placed almost anywhere in the sentence, if certain discourse-pragmatic conditions are met. Nevertheless, they have a special syntactic status, being verb modifiers in the narrower sense of the word. Thus, when modal adverbials are neither under narrow focus nor in any relevant way topicalized, but form a constituent-like syntactic and semantic unity with the verb instead, they tend to occupy a relatively firm position in the clause. In MG and Alb. this position is directly after the verb (see, e.g., Mackridge 1985: 239ff, Holton et al. 1997: 358, Buchholz and Fiedler 1987: 553ff), in SC directly before the verb (Popović 1997: 139ff):

(5-5) Eng.: ... *his teeth* ... *sank easily into the yielding flesh* ... (London, p. 134)

Alb.: ...*çatajtë* ... *nguleshin lehtë në trupin e epur* ... (p. 131)

jaws-the stuck easily in body-the the bent

MG: ...*ta ðondia tu viθizondan efkola sti malaki sarka* ... (p. 163)

the teeth his sank easily in-the soft flesh

SC: ... *njegovi zubi ... su lako zapadali u meko meso ...* (p. 150)

his teeth AUX:CLIT easily sank in soft flesh

This difference in the default position of modal adverbials will play some role in the description of inversion (Chapters 8 and 9) and of the vS-construction (Chapter 11).

The third issue which will be briefly addressed here is the expression of the subject. Alb., MG and SC are what is misleadingly called pro-drop languages, i.e. their subjects do not have to be explicitly expressed, presumably because they are already contained in the verb morphology. Now, the issue of expressing subjects (as opposed to not expressing them) is a very complicated matter, and I do not hope to adequately describe it in this short note. However, there seem to exist some clearly recognizable tendencies (which can be numerically verified; cp. 5.5.). Namely, MG omits its subjects more often than the other two languages, and Alb. more often than SC. Thus, a scale can be established, with MG on the left end, SC on the right, and Alb. somewhere in between, where left means ‘frequent subject omission’ and right ‘frequent subject expression’. Some statistical tendencies which the first sight may seem surprising can be explained by resorting to this provisional scale.

And finally, one lexico-grammatical difference. Diverging preferences appear to exist with regard to encoding states of affairs in argument positions, with Alb. tending to use embedded clauses more often than abstract nouns, MG and SC the other way around. Although the choice of the one or the other strategy for expressing states of affairs in argument positions is by and large dependent on the register in all three languages, the speakers of Alb. seem to consistently use the embedded-clause strategy more frequently than the speakers of the other two languages, across registers. Thus, in the parallel translations of Jack London’s *White Fang*, in the Alb. text 49 English abstract nouns are translated as embedded clauses, in the MG one 15, and the SC one only 10. In a sample of 1000 clauses from an Alb., a MG and a SC newspaper (‘Koha ditore’, ‘Elefterotipia’, ‘Vesti’), Alb. has 188 embedded complement clauses vs. 92 abstract nouns, the ratio in MG is 198 vs. 182, in SC 174 vs. 165. The relevance of this phenomenon for the present study lies in the fact that embedded clauses are one of the prototypical syntactic environments of verb-subject order, so that their stronger presence in one language may imply the greater frequency of verb-subject clauses in that language (cp. e.g. 8.2.4).

5.2. Heaviness and cataphoric relevance

The heaviness principle and the principle of cataphoric adjustment are the labels for two positioning tendencies (one may call them iconic) which are shared by all three languages and which are responsible for a great deal of word order variation occurring among the elements with the same informational status, e.g. among those which are parts of a presupposed propositional function, or parts of an extended focus domain.

The principle of **cataphoric adjustment** may be defined as follows: If a sentence is followed by a subordinate or coordinate structure which shares one of the elements with the sentence, and if no discourse-pragmatic factors are at work which would regulate the relative positions of the sentence elements, then the shared element will more probably than not appear at or towards the end of the sentence. The structures triggering this scheme are, among others, appositive depictive predicates, gerunds, embedded and coordinate clauses sharing the subject with the previous sentence/clause, etc. The principle of cataphoric adjustment is illustrated by the following examples:

(5-6) Alb (Kosovarja, p. 6/48)

[quote], *fliste si i humbur Noli dhe më shikonte pandërprerë.*

spoke as the lost N. and me looked unintermittently

‘[quote], Noli spoke confusedly and looked at me intently.’

(5-7) SC (Vesti, p.9)

[quote], *kaže za “Vesti” Slavko Štrbac, dodajući da ...*

says for “Vesti” Slavko Štrbac adding that AUX:CLIT

‘[quote], says Slavko Štrbac to “Vesti”, adding that ...’

Both examples are instances of quotation inversion. As I shall try to show in Section 8.1., the material after the quote, i.e. the verb, the subject, and the additional elements, represents a ratified presupposed propositional function, meaning that all the elements have the same informational status. This makes quotation inversion an ideal material for illustrating the two principles under discussion, so that it will be used for the heaviness principle, too.

Let us now look at the examples themselves. In (5-6), the quotation inversion clause is followed by a coordinate clause with the same subject, in (5-7) by the gerund (*dodajući*) with the same feature. This brings the subject, i.e. the shared element, to the final position in these clauses, with the additional elements, the depictive predicate *si i humbur* and the addressee expression *za “Vesti”* preposed with respect to the subject. Of course, we are dealing with a tendency here, not with a rule: the subject–depictive predicate (*fliste Noli si i humbur*) and the

subject–addressee (*kaže Slavko Štrbac za „Vesti“*) orders would be perfectly grammatical and felicitous, though perhaps less stylistically polished.

The second tendency observable in my corpus seems to be somewhat stronger than the first: it is the **heaviness principle**, according to which the relatively heavier elements tend to end up after the relatively lighter elements¹. In the case of quotation inversion, which is our guinea pig in this section, this means that the relative order of the subject and the elements other than the verb partly depends on their respective lengths and grades of morphological complexity: the heavier the element, the greater the probability that it will be postposed. A beautiful illustration is given in the sentences (5-8) and (5-9), both appearing on the same side of an Albanian journal (vertical strokes mark constituents):

(5-8) Alb. (Koha Ditore, p. 2)

[quote], *u tha ai | gazetarëve | në kryeqytetin kosovar, në Prishtinë.*

them:CLIT said he to-journalists in capital the of-Kosovo, in Prishtina

‘[quote], said he to the journalists in the capital of Kosovo, in Prishtina.’

(5-9) [quote], *u tha gazetarëve | në Sofje | Arbën Xhaferi, kryetar*

them:CLIT said to-journalists in Sofia A. Xh., president

i PDSH-së, e cila merr pjesë në koalicionin qeveritar në Maqedoni.

of-the PDSH, the which takes part in coalition-the governing in Macedonia

‘[quote], said to the journalists in Sofia Arbën Xhaferi, the president of PDSH, which partakes in the coalition government of Macedonia.’

Orders [subject] [addressee] [setting] in (5-8) and [addressee] [setting] [subject] in (5-9) are unequivocally triggered by the relative heaviness of the elements. This tendency is weak in the cases where the difference in the grade of heaviness is small, and becomes stronger exponentially with the growing complexity.

The principle of cataphoric adjustment and the heaviness principle are tendencies rather than rules. They work independently of the semantic or syntactic statuses of the expressions whose position in the clause they influence, and seem to be operative in the same way in all three languages.

Before closing this section, one point should be emphasized. Despite Hawkins (1994) and his adherents, there is no evidence in my corpus that heaviness plays the decisive role in determining word order in Alb., MG and SC. It is rather a minor factor which can always be overridden by stronger factors, the most important one being discourse pragmatics. Or the

¹ For the notion of heaviness and its syntactic consequences see Arnold et al. (2000) and the literature cited therein.

other way around: if a discourse-pragmatic rule is at work in a clause, the heaviness principle is irrelevant for its word order. This is important in so far as I shall argue that verb-subject order is basically triggered by discourse-pragmatically motivated word order rules, meaning that it is never triggered by the heaviness principle. Heavy subjects and light verbs, or, for that matter, heavy verbs and light subjects, may occur in both SV and VS clauses, depending solely on the discourse pragmatics of the given clause.

5.3. Some technical details

My syntactic representations are simple to the point of being self-explanatory: sentence structures and constructions are viewed as flat configurations consisting of discourse-pragmatically and syntactically defined slots. For a particular construction, the maximal projection of the slots potentially occurring in that construction, arranged in the proper order, is the **construction template**; for a language, the maximal projection of the slots potentially occurring in its sentences is the **structural sentence template** for that language.

Templates are graphically represented as ordered sequences of syntactic positions, each position marked with square brackets: thus, in (5-1), the sentence structure of the three languages under consideration was roughly delineated as [Topic] [Narrow Focus] [[Verb] [Topic] [Narrow Focus]]_{Focus Domain}.

The model of intonation embraced here deserves some more explanation. Basically, my attitude to the tricky question of intonation is a mixture of Lambrecht's (1994) functional account and Pierrehumbert's autosegmental model (Pierrehumbert and Hirschberg 1990) in Ladd's (1996) metrical reinterpretation.

Tunes carried by utterances are described as sequences of discrete tones, which may be high (H) and low (L). There are two basic kinds of tones, **pitch accents** and **boundary tones**. It is the first type, pitch accents, that are of particular interest for the present study. Pitch accents – some prefer the term '**intonation peaks**' – are tones mapped on the accented syllables of the words occurring in sentences: A particular tone cannot be arbitrarily aligned with any old syllable in the word, but occurs only on that syllable which carries the lexically determined accent for that word. In Pierrehumbert's notation, pitch accents are marked by an asterisk – H*, L*, H+L*, H*+L, etc. Consider the following MG example (simplified from Ladd 1996: 214):

$$\begin{array}{cc} H^* & L^* \\ | & | \\ (5-10) & Xo.'re.vi & a.'pop.se? \end{array}$$

‘Is she dancing TONIGHT?’ (as opposed to some other time)

Intonation peaks are aligned with the accented syllables, H* appearing on *-re-* and L* on *-pop-*, not on *xo-* or *-vi*, and not on *a-* or *-se*. Not every word in the sentence must be assigned a pitch accent, but it may be so.

Perceptually most prominent intonation peak is called **nuclear stress** or **primary accent** or **sentence accent** – in (5-10) it is the L* on *apopse*. Its function is most frequently that of focus marking, but, as I shall try to demonstrate (Chapters 8-10), not necessarily so¹.

All other intonation peaks are thus non-nuclear, **secondary accents**. According to their position, they may be divided into **prenuclear** and **postnuclear accents**. These accents are of interest here because they, if perceptually prominent, may play a ratifying function, serving to draw the hearer’s attention to a particular referent and ‘ratify’ it as the sentence topic (see Lambrecht 1994, Lambrecht and Michaelis 1998; cp. also Sections 6.2. and 8.2.2.) Thus, the prenuclear accent on *John* and *Bill* in (5-11) ratifies the referents ‘John’ and ‘Bill’ as contrastive topics, whereas the nuclear accent on *Mary* and *Jane* marks their referents as focal (adapted from Ladd 1996: 87):

$$H^*+L \quad H^* \quad H^* \quad H^*$$

(5-11) *John studied Mary, and Bill studied Jane.*

I shall use only a very rudimentary intonational notation. Small capitals mark all kinds of relevant sentence accents (primary/nuclear and secondary/non-nuclear, high and low pitch accents); double underscore indicates the primary, single underscore the secondary accent. Separate intonational phrases are marked by a single vertical stroke. Rendered in this notation, (5-11) would look as follows:

(5-12) JOHN studied MARY | and BILL studied JANE

Sentence and construction templates may also be intonationally specified; for instance, (5-1):

(5-13) [TOPIC] [NARROW FOCUS] [[Verb] [Topic] [NARROW FOCUS]]_{Focus Domain}

The other tone type, boundary tones (‘edge tones’, marked with %), serve as boundary markers of intonation units and are not aligned with particular syllables, but may stretch across different domains, from one syllable to more words, depending on the material

¹ Contrary to the frequently expressed opinion, the nuclear stress cannot be universally defined as the last intonation peak in the utterance, but only as the most prominent one, for the simple reason that, unlike English, many languages do allow for postnuclear accents (cp. Ladd 1996: 212ff.).

available. These tones will be largely ignored in the present study, since of no immediate relevance to its topic. Nevertheless, it is important to keep in mind that they do exist, since in certain cases their semantic import makes them subjectively more prominent than pitch accents. This is often the case in questions (see Ladd 1996: 172ff.): the pitch accent may appear to be less prominent than the boundary tone because the latter often marks the illocutionary force. For instance, Hungarian belongs to the group of languages which in yes/no questions place the primary accent on the verb, but because of the prominence of the edge tones in this question type, one may fail to notice this fact (Ladd 1996: 116):

L* H L%

(5-14) *Beszél a tanár?*

‘Is the teacher talking?’ (lit. talks the teacher)

In my notation, which, as indicated above, does not take notice of edge tones, (5-14) would be transcribed as BESZÉL a tanár, despite the fact that for many people the edge sequence H L% is more prominent than the low pitch accent on the accented (first) syllable of the verb. Pitch accents are placed on a different level on analysis than edge tones, so that the question of the relative prominence of the ones or the others simply does not arise.¹

5.4. Corpus

Introspectional data (including elicited speakers’ judgments), although sometimes very useful, are in no way sufficient for the kind of investigation I intend to conduct in the present study: not only do they frequently reflect the speaker’s idea of how a language should behave rather than how it really behaves, they also do not allow for a fine-grained comparison of discourse conventionalizations across languages, and, not least, for a diagnostics based on statistical data. For these reasons, the study is based on a relatively large corpus of Alb., MG and SC texts – approximately 15.000 clauses for each language. The language variant chosen is in each case the standard language, to the exclusion of dialectal and strongly diverging substandard forms.

Although it is probably illusory to think that the corpus for a syntactic investigation (as opposed to a lexicological or morphological one) may as yet be large enough to be truly representative, I considered it important to make my corpora as extensive as possible, in the

¹ I tacitly pass over innumerable problems every intonational analysis is intermingled with, since a detailed account of them would fill more than one book. I refer the reader to the works mentioned above.

hope to achieve something like approximate representativity, and to be able to draw conclusions based on numerical relationships. This had some consequences for the composition of the corpus.

The most important one is the exclusion of spoken language data. I had originally collected a number of spoken texts and partially transcribed them, but it turned out that a two-hour conversation contained less clauses than forty pages of a book, and took at least ten times more time to process: obviously, even with much more spoken material than a couple of two-hour conversations, the spoken data would represent only an insignificant fraction of the corpus and would thus perform merely an ornamental function. Considerations of economy thus led me to the decision to exclude the spoken language from the numerically processed corpus altogether. Spoken language data are, however, not completely absent: I have used the material from my recordings if they offer insights not derivable from the written material, as well as the material unsystematically collected while taking part in or listening to conversations in Alb., MG and SC.

I have tried to compose a balanced corpus for all three languages, so that at least the most relevant registers and text types are well represented. In particular, my choice looks as follows: First, for each language, a representative of a ‘higher’ and of a ‘lower’ journalistic style, respectively – for Alb., an edition of the daily journal *Koha Ditore* and of the ladies’ magazine *Kosovarja*, for MG, an edition of the journal *Elefterotipia* and of the teenagers’ magazine *Ciao*, for SC, an edition of the ‘serious’ daily *Vjesnik* and of the less prominent journal *Vesti*. Second, for each language, at least one instance of a predominantly narrative prose writing, and at least one of a more or less expository prose, with little narrative material; at least one of these must contain larger stretches of retold dialogues. For Alb., these are the science-fiction novel *Udhëtim i jashtëzakonshëm* (‘A strange journey’) by Astrit Bishqem and small portions from a more intellectually challenging work, *Kronikë në gur* (‘Chronicle in stone’) by Ismael Kadare; for MG, five stories from the collection of short stories *Efta fores to ðaxtiliði* (‘Seven times the ring’) by Ismini Kapandaj, and portions of two novels: *I mitera tu skilu* (‘Dog’s mother’) by Pavlos Matesis and *I mejali prasini* (‘The big green’) by Evgenia Fakinou; for SC, the collection of short stories *Grobница za Borisa Davidoviča* (‘A tomb for B.D.’) by Danilo Kiš and the short novel *Prokleta avlija* (‘A Devil’s yard’) by Ivo Andrić.

Although each of these sources deserves a detailed comment, I shall confine myself to the most important issues. Both the Alb. journal and the magazine are published in an area where the Gheg dialect is spoken, Kosovo, but the language used is Standard Albanian, based on the Tosque dialect; neither my informants nor I were able to detect any relevant dialectal traits.

The Alb. ladies' magazine *Kosovarja* has a literary section with a number of short stories, which, being love stories, mostly add to the bulk of non-narrative, expository prose, and carry some features of the 'folkish' Alb. style. Because of the specific sociolinguistic situation of standard MG, the literary sources have been chosen not only so as to represent different text types, but also the three social strata in the language: Kapandai's short stories are written in the classical demotic style, which is partly based on the oral folklore tradition; Matesis' novel is a first-person confession with a lot of retold dialogues and is therefore a good example of the MG conversational style; Fakinou's novel is written in the modern, urban variety of the standard language. The SC literary sources display the same kind of diversity: while Kiš's postmodern prose is the perfect example of the SC as spoken and written in large urban centers, Andrić's novel, staged in the time of Turkocracy on the Balkans, has a number of linguistic features of the SC popular orally transmitted tales.

It is all but accidental that for all three languages at least one source was taken which is to a certain extent close to the oral folk tradition: the linguistic communities of the Balkans started to be urbanized only in the second half of the 19th century, so that folklore traditions are much more present in the consciousness and the everyday language behaviour than is the case in the Western societies.

Despite my efforts to make comparable corpora for each language, some unbalanced features necessarily remain. In order to make up for this, at least partially, three translations of the same text have been added, so that in this part of the corpus the context variability is reduced to the minimum. The original text is Jack London's novel *White Fang*, more precisely its parts chosen so as to contain the right proportion of narrative, expository, and conversational material. Now, as is always the case with translations, Alb., MG, and SC versions of this text (named *Dhëmbi i bardhë*, *O Asprođondis* and *Bijeli Očnjak*, respectively) vary along some additional parameters apart from the language itself, which should be kept in mind in analyzing these texts. The MG translation is probably best suited to my purpose, being true to the original but not slavish, so that the MG idiomaticity is not diminished through the influence of the English text. This is the main trouble with the SC translation, which is very literal, so that in some instances one has the impression of reading an English text with SC words. On the other end of the scale, the Alb. translator sometimes does not understand the meaning of the original, and is from time to time very creative – not a bad thing in itself, but not exactly ideal for the analysis of the kind envisaged here.

Although only the above mentioned texts are subject to statistical evaluation, my material is not confined to them: this narrower corpus is regularly supplemented by data from other

sources, both spoken and written. Spoken language data are, as indicated above, adduced every time a phenomenon described is typical for this language variety, or even virtually absent from the written language. Additional written language data are used when the narrower corpus does not contain examples which I would consider beautiful enough.

Finally, a word on elicitation and introspection. As has been said at the outset, I consider these two methods insufficient, actually even rather ill-suited for the investigation of the kind conducted here. This attitude, however, does not imply that I do not use them as *supplementary methods*: when the corpus does not give an unequivocal answer, or does not contain the form-meaning pair one could expect on the basis of theoretical predictions, or simply when I want to confirm a conclusion drawn on the basis of the corpus data, I consult both my native speaker intuitions of SC and the intuitions of my Alb., MG, and SC informants.

5.5. Preliminary statistics

In order to give the first, rough impression of the frequency of VS order and its relationship to the two major alternative strategies, SV order and the zero-subject (henceforth zeroS) strategy, some basic statistical data are presented in this section.

5.5.1. General statistics and cross-linguistic variation

The general distribution of the three strategies (VS, SV and zeroS) within the whole corpus (original texts and translations) is presented in Table (5-15):

(5-15) *VS, SV and zeroS in Alb., MG and SC*

	zero subject	subject-verb (SV)	Verb-subject (VS)	Σ
Alb.	7834 – 49.5%	5754 – 36.4%	2225 – 14.1%	15813 – 100%
MG	7822 – 54.1%	3968 – 27.4%	2673 – 18.5%	14463 – 100%
SC	5718 – 40.3%	6393 – 45.1%	2065 – 14.6%	14176 – 100%

Some tendencies are recognizable even from this basic statistics: as indicated above (Section 5.1.), MG uses zeroS most frequently, Alb. somewhat less, and SC significantly less; the reverse order is found with the SV strategy, which is most prominent in SC, least so in MG. As far as VS order is concerned, two points are worth mentioning. First, it is a minor, though not marginal, strategy in all three languages, making up between one seventh and one fifth of all clauses, in contrast to the other two strategies, which cover between one third and one half

of all clauses. Second, it is in MG that VS order is most prominent, with some 18%, whereas Alb. and SC are approximately on the same level, with about 14%. These numerical relations are confirmed by the analysis of the three parallel translations of Jack London's novel, with the minimal context and text-type variability:

(5-16) *VS, SV and zeroS in the translations of Jack London's White Fang*

	zero subject	subject-verb (SV)	Verb-subject (VS)	Σ
Alb.	1755 – 48.0%	1480 – 40.5%	419 – 11.5%	3654 – 100%
MG	2075 – 58.8%	944 – 26.7%	511 – 14.5%	3530 – 100%
SC	1380 – 42.3%	1543 – 47.3%	337 – 10.3%	3260 – 100%

The predominance of the zeroS strategy in MG is even stronger here, as well as the predominance of the SV strategy in SC. The fact that VS order is somewhat less frequent in all three languages has to do both with the text type and, presumably, with the influence of the source language.

Before turning to finer statistical evaluations, a word on zeroS is in order, especially because the clauses containing it will be left out of consideration in further calculations. The frequency with which subjects are not expressed in MG is visible already from the two tables adduced by now. It becomes even more prominent when the first and the second person subjects are considered separately from the third person subjects. The former are, as is well known, almost always omitted. Now, since the Alb. corpus contains more first and second person subjects than MG and SC (because of the love stories in the magazine *Kosovarja* and the first-person novel *Udhëtimi i jashtëzakonshëm*), the number of zeroS clauses in this language automatically rises.

In particular, the relationship between 1./2. person and 3. person subjects in the three languages looks as follows: In Alb., the ratio of 1./2. vs. 3. person subjects is 3516 : 12297, i.e. 22.2% : 77.8%, in MG, 2029 : 12434, i.e. 14.0% : 86.0%, in SC, 1534 : 12642, i.e. 10.8% : 89.2%. That 1./2. person subjects are indeed generally not expressed is visible from the following figures: In Alb., out of 3516 1./2. person subjects, 3088, i.e. 87.8% are not expressed; in MG, the percentage is 89.2% (1810 out of 2029), in SC, 80.4% (1234 out of 1534).

Once the 1./2. person subjects are excluded, the differences between the languages concerning the expression of the subject are somewhat clearer (languages are ordered according to the frequency of subject omission):

(5-17) *Subject omission vs. subject expression with 3. person subjects*

	subject omission (zeros)	subject expression (SV/VS)	Σ
MG	6012 – 48.4%	6422 – 51.6%	12434 – 100%
Alb.	4746 – 38.6%	7551 – 71.4%	12297 – 100%
SC	4484 – 35.5%	8158 – 74.5%	12642 – 100%

While MG omits its 3. person subjects in almost one half of all clauses, the percentage of zeroS clauses in Alb. and SC only slightly exceeds one third. Although this represents an important fact about the three Balkan languages in itself, it is relevant in the context of the present study only because of the influence the presumably independent parameter of subject omission has on the statistical relationships between the percentages of VS clauses in the three languages: the greater the percent of zero subjects, the smaller the percent of VS clauses in a language. In other words, in order to calculate the actual frequency of VS clauses, the parameter of subject omission has to be kept constant, which is best done if zeroS clauses are simply left out.

Tables (5-18) and (5-19) give the numerical relationships of SV and VS clauses in the whole corpus and in the parallel translations, respectively:

(5-18) *SV and VS orders in Alb., MG, and SC*

	subject-verb (SV)	verb-subject (VS)	Σ (SV/VS)
Alb.	5754 – 72.1%	2225 – 27.9%	7979 – 100%
MG	3968 – 59.8%	2673 – 40.2%	6641 – 100%
SC	6393 – 75.6%	2065 – 24.4%	8458 – 100%

(5-19) *SV and VS orders in the translations of Jack London's White Fang*

	subject-verb (SV)	verb-subject (VS)	Σ (SV/VS)
Alb.	1480 – 77.9%	419 – 22.1%	1899 – 100%
MG	944 – 64.9%	511 – 35.1%	1455 – 100%
SC	1543 – 82.1%	337 – 17.9%	1880 – 100%

The picture is quite clear now: in MG, the number of VS clauses is significantly higher than in Alb. and SC, and it is slightly higher in Alb. than in SC. With respect to the frequency of VS order, Alb., MG and SC are thus to be ordered along a scale:

(5-20) *Frequency of VS clauses in Alb, MG, and SC*

MG >> Alb. > SC

5.5.2. Intralinguistic variation

The corpus on the basis of which the above scale is established is of course not absolutely homogenous – as a matter of fact, it has been composed the way it is in order to capture the differences between the most relevant registers of the written language. It is therefore interesting to see how VS and SV clauses are distributed across the sources. The breakdown of the data is given in Table (5-21); the first row, which contains the numerical relationships obtaining in the corpus of the given language as a whole, i.e. the mean value for all sources, serves as the *tertium comparationis*.

(5-21) *SV and VS orders in Alb., MG and SC sources*

	SV	VS		SV	VS		SV	VS
Alb.	72.1%	27.9%	MG	59.8%	40.2%	SC	75.6%	24.4%
Bishqemi	59.5%	40.5%	Kapandai	34.3%	65.7%	Andrić	77.1%	22.9%
Kadare	69.2%	30.8%	Matesis	53.8%	46.2%	Kiš	77.6%	22.4%
Koha Dit.	74.4%	25.6%	Fakinou	78.7%	21.3%	Vjesnik	63.7%	36.3%
Kosovarja	73.7%	26.3%	Elefterot.	67.1%	32.9%	Vesti	69.4%	30.6%
			Ciao	65.0%	35.0%			

Most of the sources for all three languages are close to the mean value within tolerable limits of deviation, confirming thus indirectly the relative representativity of the corpus.

There are three exceptions, however: Bishqem's science-fiction novel *Udhëtimi i jashtëzakonshëm* has some 13% more VS clauses than the mean value for Alb.; The same phenomenon is observable in Kapandai's short stories *Efta fores to ðaxtiliði*, where the deviation with respect to the mean value for MG is even higher, amounting to some 25%; On the other side of the scale, Fakinou's novel *I meyalitë prasini* contains more than 19% less VS clauses than the mean value for MG. Each of these exceptions deserves a separate explanation.

Bishqem's novel consists for a greater part of retold dialogues, and contains therefore a host of quotation inversion clauses of the type *[quote], said his friend* (see Section 8.1), which obligatorily display VS order, as well as quite a high number of two typically conversational clause types, interrogative clauses, which are often formed with VS (Section 8.2.), and clauses with fronted narrow focus, also usually VS (Section 8.3.). Thus it is the discourse type – conversation and the narrator's intervention in it – that is responsible for the high ratio of VS clauses in this case.

A similar explanation may be offered – although this time in the opposite direction – for the low percentage of VS clauses in Fakinou's novel: the novel is a specimen of modern prose

writing, which disposes not only of dialogues, but also of descriptions and similar ornaments, which are, as will be seen in the course of the present study, one of the prototypical discourse surroundings for VS clauses.

The case of Kapandai's short stories with the disproportionately high number of VS clauses is somewhat more complex. As indicated in 5.4., the stories are written in traditional demotic, partly in a style close to the MG oral folk tradition. One of the typical traits of this sociolinguistic or stylistic variant of MG is the frequent use of VS order in narration, more precisely of the VsX construction (Sections 3. and 6.), which is used in a more limited set of contexts in the more official and more 'urban' MG writing (see Section 6.6.2. for details).

5.5.3. Clause types and VS order

The last important point I should like to investigate in this preliminary statistical evaluation is the question of the preferential formally definable contexts for VS order, or, more simply, of the syntactically definable clause types in which VS order appears with a frequency higher than average.

The first parameter is that of the syntactic status of the clause and the grammatical person of the subject. The categories investigated are on one hand main, embedded, and relative clauses, and on the other, orthogonally to the former division, the 1./2. person and the 3. person subjects. The categories are thus defined as, for instance, 'main clause with a 3. person subject', 'embedded clause with a 1./2. person subject', etc. The one exception where the differentiation with respect to the person is not made are relative clauses, because the number of 1./2. person subjects in this category is negligibly small – 7 in Alb., 3 in MG and SC, respectively, all zeroS – so that these 13 clauses are simply added to the category 'relative clause, zeroS'.

On the other hand, relative clauses are split into two categories: the sum of all relatives (marked simply as 'relative') and **oblique relative clauses**. By the latter term only those relative clauses are meant in which the relative element is *not* the subject of the relative clause (i.e. *the man with whom I spoke, the man which we saw* etc., but not *the man who sold the world*; see Section 9.1. for more details).

Two kinds of measurements will be performed: first, the presence of VS order within a category (e.g., how many 3. person main clauses are VS, SV, etc.), and second, the distribution of VS order across categories (e.g., how many VS clauses are 3.person main clauses, how many 3. person embedded clauses, etc.). In both cases, the numerical relations

obtaining for VS order will be compared to the mean value for the whole corpus, i.e. for all clause types (zeroS + SV + VS). The categories in which VS occurs more often than average are then considered the typical environments for this order.

The data pertaining to the first measurement are presented in the following tables; the first rows contain the mean values for the whole corpus of the given language. For instance, the first row of the table devoted to Alb. data reads as follows: zeroS clauses make up 49.5% of all clauses in Alb., SV clauses 36.4%, etc. The rows that follow the first one are dedicated to individual clause types, the second row reading as follows: zeroS clauses make up 28.0% of all 3. person main clauses, SV clauses 51.2%, etc.

(5-22) *Presence of VS, SV and zeroS in formally definable clause types*

	zero S	SV	VS	Σ
Alb.	49.5%	36.4%	14.1%	100%
main 3. p	1940 – 28.0%	3546 – 51.2%	1435 – 20.8%	6921 – 100%
main 1./2.p	1943 – 84.7%	272 – 11.9%	79 – 3.4%	2294 – 100%
embed. 3. p	1952 – 60.7%	770 – 24.0%	493 – 15.3%	3215 – 100%
embed. 1./2.p	1145 – 93.7%	53 – 4.3%	24 – 2.0%	1222 – 100%
relative	125 – 8.7%	1113 – 77.7%	194 – 13.6%	1432 – 100%
oblique relat.	125 – 30.7%	89 – 21.8%	194 – 47.5%	408 – 100%

	zero S	SV	VS	Σ
MG	54.1%	27.4%	18.5%	100%
main 3.p	2790 – 40.6%	2367 – 34.5%	1707 – 24.9%	6864 – 100%
main 1./2. p	1231 – 87.5%	123 – 8.7%	53 – 3.8%	1407 – 100%
embed. 3. p	2590 – 68.1%	548 – 14.4%	665 – 17.5%	3803 – 100%
embed. 1./2.p	579 – 93.1%	26 – 4.2%	17 – 2.7%	622 – 100%
relative	221 – 16.3%	904 – 66.7%	231 – 17.0%	1356 – 100%
oblique relat.	221 – 44.5%	45 – 9.0%	231 – 46.5%	497 – 100%

	zero S	SV	VS	Σ
SC	40.3%	45.1%	14.6%	100%
main 3. p	1957 – 27.5%	3721 – 52.3%	1434 – 20.2%	7112 – 100%
main 1./2. p	811 – 74.6%	258 – 23.7%	19 – 1.7%	1088 – 100%
embed. 3. p	1617 – 52.4%	1101 – 35.7%	367 – 11.9%	3085 – 100%
embed. 1./2.p	423 – 94.8%	23 – 5.2%	0 – 0%	446 – 100%
relative	331 – 17.7%	1290 – 69.1%	245 – 13.1%	1866 – 100%
oblique relat.	331 – 41.1%	229 – 28.5%	245 – 30.4%	805 – 100%

The tables reveal a number of more or less predictable facts. When the subject is the first or the second person, and is not omitted (which is seldom the case), the order is between three and fifteen times more often SV than VS. In other words, the 1./2. person subject is, when exceptionally overt, almost always placed in front of the verb. Since these data are so

unequivocal, the first and the second person subjects shall not be dealt with in this statistical overview any more.

Further, it is clear that, in relative clauses, the relative element is very frequently the subject of its clause; since the position of the relative element is not subject to variation (it has to be the first word in the clause), this fact distorts the picture of the distribution of the VS, SV and zeroS strategies in relative clauses.

The third point readily observable is the variability of the zeroS strategy in cases in which the subject is the third person: in all three languages, it is significantly less frequent than the mean value in main clauses (28.5% vs. 49.5% in Alb, 40.6% vs. 54.1% in MG, 27.5% vs. 40.3% in SC), but significantly more frequent than average in embedded clauses (60.7% vs. 49.5% in Alb., 68.1% vs. 54.1% in MG, 52.4% vs. 40.3% in SC). This variability also significantly distorts the picture of the distribution of SV and VS order, this time in main and embedded clauses.

Thus, in order to get a more conclusive impression of the distribution of VS and SV orders, both the relative clauses in which the relative element = subject and the clauses with omitted subjects should be left out of consideration. The results achieved in this way are presented in the following tables:

(5-23) *Presence of VS and SV orders in formally definable clause types*

	SV	VS	Σ
Alb.	72.1%	27.9%	100%
main	71.2%	28.8%	100%
Embedded	60.9%	39.1%	100%
oblique relat.	31.4%	68.6%	100%

	SV	VS	Σ
MG	59.8%	40.2%	100%
Main	58.1%	41.9%	100%
Embedded	45.2%	54.8%	100%
oblique relat.	16.3%	83.7%	100%

	SV	VS	Σ
SC	75.6%	24.4%	100%
Main	72.2%	27.8%	100%
Embedded	75.0%	25.0%	100%
oblique relat.	48.3%	51.7%	100%

In all three languages, main clauses display a surprising congruence with the mean value, with only an insignificantly higher variance in SC. Embedded clauses seem to be one of the

syntactic surroundings in which VS order occurs significantly more frequently than average in Alb. and MG (39.1% vs. 27.9% in Alb., 54.8% vs. 40.2% in MG), but not in SC (25.0% vs. 24.4%). The state of affairs obtaining in oblique relative clauses is even more convincing: in all three languages, VS order occurs significantly more often than average (68.6% vs. 27.9% in Alb., 83.7% vs. 40.2% in MG, 51.7% vs. 24.4% in SC), again more so in Alb. and MG than in SC. These results are, with insignificant deviations due to the influence of the original text, confirmed by the data from the parallel translations of London's novel:

(5-24) *Presence VS and SV orders in formally definable clause types in the translations of Jack London's White Fang*

	SV	VS		SV	VS		SV	VS
Alb.	77.9%	22.1%	MG	64.9%	35.1%	SC	82.1%	17.9%
main	77.7%	22.3%	main	65.5%	34.5%	main	75.2%	24.8%
embed.	55.7%	44.3%	embed.	36.8%	63.2%	embed.	88.4%	11.6%
obliq.rel.	25.7%	74.3%	obliq.rel.	11.4%	88.6%	obliq.rel.	64.7%	35.3%

The picture obtained by the first measurement is confirmed also by the second measurement undertaken, than of the distribution of VS order across categories. The data are presented in the following tables:

(5-25) *Distribution of VS order compared to the mean value in Alb., MG and SC*

	main 3	main 1./2.	embed. 3	embed. 1./2.	relative	Σ
Alb. Σ	71.6%	5.0%	18.2%	1.1%	4.1%	100%
Alb. VS	64.5%	3.5%	22.2%	1.1%	8.7%	100%

	main 3	main 1./2.	embed. 3	embed. 1./2.	relative	Σ
MG Σ	70.5%	3.0%	21.0%	0.7%	4.8%	100%
MG VS	63.9%	2.0%	24.9%	0.6%	8.6%	100%

	main 3	main 1./2.	embed. 3	embed. 1./2.	relative	Σ
SC Σ	69.7%	3.8%	19.8%	0.3%	6.4%	100%
SC VS	69.4%	0.9%	17.8	0.0%	11.9%	100%

The tables read as follows: the first rows as 'of all the clauses in the given language, 3. person main clauses make up 69.7%, 1./2. person main clauses 3.8%, etc.', the second rows as 'of all the VS clauses in the given language, 3.person main clauses make up ...'. The results are obtained through comparison: if in a category the value for VS clauses is higher than the mean value, then that category is a syntactic context with an increased frequency of VS clauses. As already indicated, the results are in full compliance with the results of the first measurement. Main clauses correspond to the mean value and oblique relative clauses display

a significantly higher percentage of VS clauses than average in all three languages. With respect to embedded clauses, the languages diverge from each other, with a higher ratio of VS clauses in this context in Alb. and MG, but not in SC.

(5-26) *Syntactic contexts with a higher incidence of VS clauses than average*

Alb., MG: – 3. person embedded clauses

– oblique relative clauses

SC: – oblique relative clauses

5.4.4. Summary

Statistically, VS-clauses represent a minor, but not marginal phenomenon in the languages under consideration, with the generally frequency ranging from one seventh to one fifth of all clauses. When only clauses with overt subjects are taken into consideration, the cross-linguistic variation becomes more conspicuous: MG has significantly more VS clauses than Alb. and SC, Alb. somewhat more than SC. Although most sources in the corpus are numerically very close to the mean value, a few display a disproportionately high or low percentage of VS clauses; this seems to be related to the variation across text-types and registers. As far as syntactic contexts in which VS clauses occur are concerned, some clear tendencies are observable: first, the first and the second person subjects very infrequently occur in VS clauses; second, main clauses correspond very closely to the average value; third, VS order seems to be the preferred order in oblique relative clauses in all three languages; fourth, in Alb. and MG, embedded clauses display a disproportionately high ratio of VS clauses, but in SC they do not.

6. VsX construction

The description of VS order in the Balkans begins with VsX construction, for the simple reason that it is probably the least complex of the three constructions identified, its functions being easiest to grasp intuitively.

The label ‘VsX’ is borrowed from Jacennik and Dryer (1992); the presence of ‘X’ refers to the fact that more often than not an additional element apart from the verb and the subject is present; capitalized ‘V’ and ‘X’, opposed to small ‘s’, are meant to give a rough indication of the intonational pattern characteristic of the construction, with accented verbs (and additional elements) and unaccented subjects. The label should, however, not be understood as the precise description of VsX construction, as divergences from this basic pattern are abundantly attested as well, but merely as a convenient shorthand. Here are some examples:

(6-1) Alb. (Bishqemi, p. 33)

“Ne nuk jemi këlyshë ariu në kopshtin zoologjik!”, ia këputi.

we not are young of-bear in garden zoological to-him-it he-tore

“Filloi TERS kjo punë”, mendova.

began bad this thing I-thought

‘»We are not baby bears in the zoo!«, he couldn’t suppress it any more.

»This adventure hasn’t begun very successfully.«, I thought.’

(6-2) MG (Kapandai, p. 138)

... vriskete tora ðesmios ke anisxiros, jatri ke nosokomes pernun, ... pai pja o

is-found now captive and weak doctors and nurses pass-by goes already the

Filipos, ðEN ine pja aftos EðO, ke mono mesa apo to ðomatjo pernun ... jatri

Ph., not:CLIT is more he here and only within of the room pass-by doctors

‘[My father] is now held captive, weak, doctors and nurses are passing by, ...

Philippos is disappearing slowly, *he is not here anymore*, it is only doctors passing by through the room...’

(6-3) SC (Andrić, p. 64)

“... to je istorija, nauka, a od nauke ne može biti štete.” ...

it is history, science, and of science not can be harm

“NEĆU ja, efendija, da lupam glavu o tom. Ja istoriju ... ne znam.”

not-want I efendi to break head about that I history not know

‘»...But it is history, a science, and science cannot do any harm!’ ...

»*Efendi, I don't want to rack my brains with it. About history, I know nothing...*«?

The basic function of VsX construction is to encode propositions whose subject topics are continuous over larger stretches of discourse, but which themselves are discontinuous as to the discourse frame. According to the assertional structure used, two basic types can be differentiated, the polarity focus type, in which the positive or negative polarity of the proposition is asserted, and the focus domain type, in which the scope of assertion is marked by a focus domain, which, in turn, may be construed as broad (predicate +/- arguments) or narrow (argument) focus, more often the former than the latter.

In the interpretations of VS order proposed by Givón, Hopper, Myhill and others mentioned in Section 2.3., the expression of continuous topicality figures as its most prominent function. In the account that follows, I shall try to evaluate this claim with respect to the Balkan VsX construction.

6.1. Statistics

Although VsX construction exists in all three languages under consideration, there is a strong crosslinguistic variation with respect to its frequency: in contrast to being a rather minor phenomenon in Alb. and SC, it seems to be a very prominent discourse device in MG. The relevant data are presented in Tables (6-4) and (6-5), the former giving the absolute numbers for original texts and the translation of Jack London's *White Fang*, the latter the share of VsX clauses within the class of VS clauses, within the class of clauses with overt subjects (SV/VS), and the percentage within the whole corpus (Σ):

(6-4) *VsX clauses in the corpus: absolute values*

	original texts	Translation	total
Alb.	107	13	120
MG	505	52	557
SC	99	17	116

(6-5) *VsX clauses in the corpus: percentages*

	VsX vs. VS	VsX vs. SV/VS	VsX vs. Σ
Alb.	120 vs. 2225 = 5.4%	120 vs. 7979 = 1.5%	120 vs. 15813 = 0.8%
MG	557 vs. 2673 = 20.8%	557 vs. 6641 = 8.4%	557 vs. 14463 = 3.9%
SC	116 vs. 2065 = 5.6%	116 vs. 8458 = 1.4%	116 vs. 14176 = 0.8%

Unlike Inversion (see Chapters 7. – 10.), and far more than vS construction (see Chapter 11.), the use of VsX construction is a matter of the speaker’s choice, in which, as I shall try to show, the considerations of sociolinguistic nature play a significant role. In other words, VsX is also subject to intralinguistic variation. Consider the following data (the first column in each table refers to the absolute number of VsX clauses, the second to its share in the group of VS clauses, the third in the group of clauses with overt subject; the first rows indicate the mean value for the given language):

(6-6) *VsX in Alb., MG, and SC sources*

	num	VS	SV/VS		num	VS	SV/VS		num	VS	SV/VS
Alb.	120	5.4%	1.6%	MG	557	20.8%	8.4%	SC	116	5.6%	1.4%
Bishq.	41	6.3%	2.5%	Kapand.	296	36.2%	23.8%	Andrić	39	9.4%	2.1%
Kadare	3	3.1%	0.6%	Matesis	77	19.2%	8.8%	Kiš	21	4.3%	0.9%
Koha Dt.	16	3.0%	0.8%	Fakinou	21	13.4%	2.9%	Vjesnik	15	7.2%	2.6%
Kosovar.	47	7.3%	1.9%	Elefterot.	63	13.2%	4.3%	Vesti	24	3.9%	1.2%
translat.	13	3.1%	0.7%	Ciao	48	15.6%	5.5%	translat.	17	5.0%	0.9%
				translat.	52	10.2%	3.6%				

First the cases in which VsX is less frequent than the mean value. As could be expected, all three translations of an English novel belong here, which is presumably due to the influence of the language of the original, which does not have anything similar to this construction. Journalistic registers, represented by *Koha Ditore* in Alb., by *Elefterotipia* in MG, and by *Vesti* in SC, also seem to disfavor VsX. The same holds for the representatives of the modern, ‘urban’ belletristic style (Kadare, Fakinou, Kiš). In sum: The more official and/or ‘urban’, ‘westernized’ the register, the smaller the ratio of VsX clauses.

The instances of overrepresentation of VsX construction are more variegated. In MG, it is either prose writings which imitate the colloquial style or contain a lot of quoted speech (Matesis) or those which have a traditionalist, ‘folkish’ tone (Kapandai’s short stories) that tend to have more VsX than the mean value. In Alb. and SC, VsX is especially often used in love stories from *Kosovarja*, by Bishqemi and by Andrić, who belong to the latter group (traditionalist writing); in SC it also occurs very frequently in a pathetic, sublime style, as represented by the funeral speeches for the Croatian president Tudjman printed in *Vjesnik*. The distribution of VsX across registers may be roughly summarized as follows (for a more fine-grained picture, see 6.6.):

(6-7) *Distribution of VsX across registers*

	official	‘urban’	colloquial	traditional	pathetic
Alb.	–	–	+/-	+	–
MG	–	–	+	+	–
SC	–	–	+/-	+	+

6.2. Ratified topics, discourse frames, and the Balkan syntax

The basic function of VsX construction is, as indicated, to encode assertions which change the discourse frame but keep the topical referent constant. I shall, following Lambrecht and Michaelis (1998) call this type of topic, often referred to as ‘constant’ or ‘continuous’ (Matić 2003), *ratified topic*, in order to emphasize the negotiating character of topic determination.

Topics of propositions are in the present study treated as a kind of instruction to the hearer, defining the set of possible worlds to be searched in order to assess the assertion. The speaker chooses as topic that element of the proposition which s/he assumes to be (a) less probable to be contested by the hearer, and therefore easily presupposable, and (b) the optimal background for the hearer to decode the intentions of the speaker (see 4.3. for details).

In the incremental model of communication, in which the mutual consensus on the way the world is increases step by step, with every utterance, the status of topics with respect to this consensual knowledge may be twofold. First, the interlocutors may differ as to the topic, i.e. the section of the universe, they posit as relevant for the ongoing communication. More precisely, it may be the case that the speaker determines as the relevant section of the universe¹ a topic which the hearer is not entitled to expect on the basis of the previous communication, or which may be expected, but it stands in the contrastive relation to some previous discourse referent. These topics are thus *non-ratified*, i.e., at the moment of the utterance, the speaker and the hearer have not yet come to an agreement as to the topic of the further communication. Non-ratified topics are typically expressed lexically, often with rich descriptive content; as a rule, non-ratified topic expressions carry a secondary sentence stress (called *activation accent* by Lambrecht 1994, *ratification accent* by Lambrecht and Michaelis 1998). Syntactically, these expressions tend to be placed clause-initially; some types are often extraposed. Note that my non-ratified topics are not equal to contrastive topics, which have been often dealt with in the relevant literature: the latter are rather a subtype of the former (see Matić, 2003, for more detail).

It is the second topic type that is of interest here. When the speaker assumes that the hearer expects a certain section of the universe to be the object of the increment of knowledge, i.e. when s/he assumes that the hearer expects assertions about a certain topic, it is *ratified topics* that are used: the interlocutors have reached the mutual consensus on the further increment of knowledge before the moment of the utterance. Thus, in a sequence like *I saw Eve yesterday*.

¹ The one with respect to which s/he intends to change the hearer’s mental representations of how the world is.

She is fine, the discourse referent ‘Eve’ is a ratified topic in the second clause, in which it is encoded as *she*; after the context has been created by the first clause, the speaker is entitled to assume (on the basis of the principles of good communicative behaviour) that the hearer expects some information on ‘Eve’.

The example given above perfectly illustrates two important features of ratified topics. First, the clues that enable the speakers to treat a topic as ratified. It is often the referential continuity of the topic referent in the ongoing discourse, i.e. its continuous presence in the immediate context, as in the case of ‘Eve’. However, the topic referent does not have to be contextually given; it may be, for instance, inherently easy to activate (‘I’, ‘you’), or of continuous interest for the given interlocutors (‘the daughter’ for the parents, ‘the father’ for the sons, ‘the boss’ for the employees, etc.), or a major character of a longer stretch of discourse which has not been mentioned for some time (‘Bilbo Baggins’ in the second book of *The Lord of the Rings*), etc. In other words, ratification is only statistically connected with contextual givenness: a given discourse referent is easier to treat as ratified than the one which is not, but it does not have to be ratified, and, the other way around, a ratified topic is often given, but it need not be, depending on the intentions and assumptions of the interlocutors.

The second point of interest is the encoding of ratified topics. When treated as ratified, ‘Eve’ is pronominalized and unaccented, which is, according to Lambrecht (1994: 172ff.), the prototypical way of expressing this topic type. This claim is certainly correct as far as languages like English and French are concerned, but not, or not entirely so, when the language type misguidingly called pro-drop languages (Alb., MG, and SC belonging here) is taken into consideration. In these languages, the normal way to express ratified topics is not to express them (especially if they are subjects, which is most often the case). Thus, in Alb., the sequence with ‘Eve’ would sound as follows: *Pashë Evën dje. Ĕshtë mirë.* (Ø saw:1SG Eve yesterday. Ø is:3SG fine). In this light, Lambrecht’s claim must be reformulated for the three languages under consideration: in Alb., MG, and SC, ratified topics, if they are subjects, are preferably not expressed at all, i.e. they are encoded by the zero-marking.

How is VsX construction, then, to be linked to ratified topics? (Recall that I contend that the subjects in this construction are discourse-pragmatically ratified topic expressions.) This is where the notion of discourse frame enters the scene. Under the notion of discourse frame I shall subsume the space and time within one discourse universe, i.e. scene, or the one discourse universe among other possible discourse universes, i.e. perspective.

Assertions are made with respect to topics which are placed on a certain spatio-temporally defined scene: In uttering the sequence about Eve, the speaker invites the hearer to assess the

focus ‘is fine’ not with respect to all possible times and spaces in which the referent ‘Eve’ may exist, but only in those which are defined by the proposition ‘I saw Eve yesterday’. Furthermore, assertions are made within discourse universes defined by the speakers: ‘I saw Eve yesterday’ is placed in a discourse universe in which the speaker partakes as one of the participants; ‘She is fine’ is an assertion made within a discourse universe in which the speaker is a judging subject, not a participant. When I refer to the existence of the topic on a certain scene, I shall speak about *spatio-temporal frames*; when the choice of the discourse universe is meant, the term *perspective* will be used. Both are subsumed by the notion of *discourse frame*.

Back to VsX. When the speaker chooses to encode the topic referent as a ratified subject topic, s/he may, as indicated, use zero-marking. This is indeed the most common strategy, both when the discourse frame remains the same with respect to the previous utterance, and when it changes, as illustrated by the following variations of the Eve-example, all in Alb.:

(6-8) *Eva nuk është e sëmurë. Është mirë.* (Eve not is the ill. Is fine.)

(6-9) *Eva qe e sëmurë. Tani është mirë.* (Eve was the ill. Now is fine.)

(6-10) *Pashë Evën dje. Është mirë.* (I-saw Eve yesterday. Is fine.)

In sequence (6-8), the discourse frame remains unchanged between the two utterances; in (6-9), the spatio-temporal frame changes, in (6-10), the perspective (the spatio-temporal frame remaining roughly the same, with ‘yesterday’ defining a broadly understood present). In all three cases, the zero-strategy is the usual variant in all three languages under consideration.

However, in certain discourse surroundings, speakers may decide to explicitly mark the change of the discourse frame, while keeping the topic ratified. This is the proper place of VsX construction: Postverbal topical subjects encode ratified topics belonging to propositions in which the discourse frame changes, and which occupy such a position in the discourse, that the speaker considers this change worth marking. Take, for instance, (6-3): the speaker (an official reacting to the pleas of an old man to release someone from the jail) wants to deny the implicit assumption of the petitioner that he wants to think over the whole issue. To this end, he uses VsX construction: *NEÇU ja da lupam glavu...* (not-want I to break head...) ‘I don’t want to rack my brains with it’. What is marked here is, of course, the ratified topicality of ‘I’ and the change of perspective, from the world of your judgment to the world of my judgment. The speaker might have decided not to mark the latter point explicitly – in this case, he would have used the zero strategy.

In terms of markedness, the whole situation depicted in this section may be analyzed as follows: both postverbal and zero topics are marked for the feature [+ratified], whereas

preverbal topics carry the feature [–ratified]. As to the continuity of discourse frames, both preverbal topics and zero topics are neutral as to the feature [+/-continuous]; postverbal topics are marked for [–continuous]. Table (6-11) summarizes this:

(6-11) *Meanings of topic expressions in Alb., MG, and SC*

	ratified topic	continuous discourse frame
preverbal topics	–	/
zero topics	+	/
postverbal topics	+	–

The simplified template of the Balkan sentence structure proposed in (5-1) and (5-13) – [TOPIC] [NARROW_FOCUS] [[Verb] [Topic] [NARROW_FOCUS]]_{Focus Domain} – finds its natural explanation in this distribution of topic types: preverbal topic expressions, carrying the secondary stress, encode non-ratified topics which are not marked for the continuity of the discourse frame; the essential part of VsX construction, postverbal topic expressions, are unaccented since they encode only ratified topics, and are marked for the discontinuity of the discourse frame. In what follows, I shall try to show that this sentence pattern may need some further elaboration.

6.3. Formal properties of VsX construction and information structure

6.3.1. Length and intonation

Unlike Inversion (Chapters 8. and 9.) and vS construction (Chapter 11.), VsX clauses are more often than not long, meaning that they contain at least one more element apart from the verb and the subject. Thus, out of 120 VsX clauses in my Alb. corpus, only some 28, i.e. 23.3% have the form Vs, consisting only of a verb and a subject. In MG, the ratio is 557 : 143, i.e. 25.6%, in SC, 116 : 32, i.e. 27.5%. The frequency of additional elements in VsX clauses becomes even more conspicuous when they are compared to other constructions: for instance, in a subtype of Inversion, quotation inversion, short clauses (V+S) make up between 73.7% and 77.8% of all inverted clauses (see 8.1.1.2. for more detail). The prototypical VsX clause is thus long, consisting of a verb, a subject, and at least one element besides them.

The most prominent feature of VsX construction is certainly its intonational contour. The subject is always unaccented. Thus, if the clause is short (only the verb and the subject) the nuclear accent invariably falls on the verb¹:

(6-12) Alb. (Koha Ditore, p. 11)

NUK MJAFTOJNË këto ndryshime.

not:CLIT suffice these changes

‘These changes are not ENOUGH.’

In long clauses, the distribution of accents, or intonation peaks, depends on the focus structure. If the polarity focus structure is used, only the verb is accented, the rest of the clause being marked by the fall of fundamental frequency:

(6-13) MG (Kapandai, p. 37)

ËEN EPERNE i kubura tu fotia

not:CLIT took the pistol his fire

‘His pistol WOULDN’T shoot.’

In focus domain structures, with broad and with narrow focus construal, both the verb and the additional element carry an accent; the one on the non-verbal element (X), as in all kinds of focus domains in the Balkan languages (cp. 4.5.2.), is the nuclear accent, while the verb carries only the secondary stress:

(6-14) SC (Andrić, p. 22)

IMAO sam ja ženu MISIRKU.

had AUX:CLIT I wife Egyptian

‘I was married to an Egyptian woman.’

The intonational patterns for VsX construction may thus be represented as follows:

(6-15) short clauses: [VERB][subject]

polarity focus: [VERB][subject][rest] / [VERB][rest][subject]

focus domain: [VERB][subject][REST] / [VERB][REST][subject]

A word on the quality of intonation peaks is order here. VsX clauses in Alb., MG, and SC have, for both a native and a non-native ear, a kind of ‘singing’ intonation. In the focus domain type, this is doubtlessly due to the multiple sequence of accented and unaccented strings: accented verb, unaccented subject, accented X (as, e.g., in (6-14)). In the polarity

¹ Clitics in Alb. and MG stand invariably in front of the verb (5.1.) and form a phonological word with it. When the accent falls on the clitic, the verb remains unaccented, but the phonological word consisting of the clitic and the verb counts as accented. This accent shift on the clitic regularly occurs with the negative particles *nuk* (Alb.) and *ëen* (MG), and only occasionally with MG pronominal clitics.

focus type, it is rather the nature of the accent itself that produces this impression. Namely, in this focus structure, especially if any kind of contrast is involved, the nuclear stress often has the form H+L*, i.e. there is a rapid fall from high to low on the accented syllable. For instance, in (6-3), the accented syllable *ne-* (in the verb *neću*, ‘I-not-want’) is aligned with the tonal sequence H+L*, as well as the syllable *đen* (in the phonological word *đen eperne*, ‘not it-took’) in (6-13).

6.3.2. Elements of the construction

6.3.2.1. Verb and preverbal elements

Unlike Inversion, in which an element standing in front of the verb represents the defining part of the construction, and vS type, which is very frequently introduced by an element other than verb and subject, VsX clauses are in the overwhelming majority of cases verb-initial in the narrower sense of the word. In the Alb. corpus, in 76.7% of all VsX clauses the verb is in the absolute initial position (92 out of 120 clauses); in MG, the percentage is 85.5% (476 out of 557)¹, and in SC 82.7% (96 out of 116). The preferential absolute initial position of the verb is illustrated by all the examples adduced in this section by now. Nevertheless, the statistical data show that we are not dealing with a syntactic rule here, since exceptions do occur, but rather with a tendency, which presumably has a functional explanation.

VsX clauses mark referential continuity within a context of the discontinuity of the discourse frame, either in terms of the spatio-temporal coordinates or of the perspective. The fact that a ratified subject topic is overt (as opposed to the default zero strategy) is a sufficient sign for the change of the discourse frame. This makes the explicit marking of the change by a non-ratified preverbal element (setting adverbial or similar) superfluous. In other words, in most cases, it is enough to put the ratified subject topic in the postverbal topic slot in order to signal to the hearer that the discourse frame has changed with respect to the previous utterance. For instance, in (6-14), the speaker saying ‘I was married to an Egyptian woman’ (*Imao sam ja ženu Misirku*) changes the spatio-temporal coordinates with respect to the previous utterance, which deals with his present miserable state (he is in prison). To say ‘Many years ago, I was married...’ (*Pre mnogo godina, imao sam ja...*) would be, although not ungrammatical, somewhat awkward: the fact that ‘I’ is overt (*ja*) and postverbal is a sufficient instruction to the hearer that the scenery has changed, although the topic remains the same.

¹ As indicated in Section 5.1. and in the note on page 164 above, proclitics in Alb. and MG form a phonological word with the verb, so that the fact that they appear in front of the verb does not count for our purposes.

However, in the case of the change of scene, when the speaker wishes to precisely state in which way it changes, non-ratified preverbal elements may occur:

(6-16) MG (Kapandai, p. 37)

Tin ali mera ton ESTILE o Petros piso ston PASA...

the next day him:CLIT sent the P. back to-the Pasha

‘[Ali-Pasha sent a certain Ali-Beqir to catch the Greek rebel Petros for him, but Petros overcame Ali-Beqir.] The following day, Petros sent him back to the Pasha.’

It is of some relevance for the narration that the main character of the story, Petros, sent Ali Beqir to Pasha on the day that followed the skirmish between them. Consequently, both the fact that the scene changed (postverbal ratified topic) and the way it changed (preverbal non-ratified setting adverbial) are marked.

When it is the perspective that changes, preverbal elements are even more infrequent, but not impossible; when the speaker for some reason wants to precisely state the source of the new perspective, s/he may use some kind of extraposed expression (‘as-for-construction’) or a contrastive preverbal topic:

(6-17) MG (Elefterotipia, p. 23)

Oso ja to mitropolititi Pireos Kaliniko ... ðEN ton afora to thema tu Timiu Stavru ...

as for the metropolitan of-P. C. not him:CLIT concerns the theme of precious cross

‘[In an article about a piece of the Holy Cross the Greek Church is to exhibit in Athens:] As for the metropolitan of Piraeus, Callinicos, ... the whole issue of the Holy Cross does NOT concern him.’

Examples of this kind are very rare: I have two instances in Alb., seven in MG, one in SC. Obviously, the VsX construction itself is felt to be more at home in the contexts in which the hearer is assumed to be able to figure out the source of the change of the discourse frame solely on the basis of the fact that the construction has been used. As a corollary to this, the verb occupies the absolute initial position in the majority of cases.

6.3.2.2. Subject and other postverbal elements

While the formal features of the construction described by now are common to all three languages, the issue of the relative order of the subject and the additional element(s) appearing after the verb is connected with some clearly divergent tendencies.

The question is simple: in what order do the postverbal elements occur? The simplicity reoccurs in the answer only in the case of SC, where the subject almost without exceptions precedes all other elements. The situation in Alb. and MG is somewhat more complex.

6.3.2.2.1. Serbo-Croat

Let us begin with SC. There are 84 long VsX clauses in the corpus; out of this number, 12 are left out of consideration, since the additional element is preverbal; so we are left with 72 potentially relevant instances. In 66 cases, i.e. 91.7%, the order is [verb][subject][x] (cf. (6-3) and (6-14)); in only 6 clauses (7.3%) is the order the other way around, i.e. [verb][x][subject]. The tendency is easy to formulate: In SC, the subject tends to follow the verb directly, with everything else coming afterwards. The six attested exceptions to this, completely grammatical, are a proof that we are dealing with a tendency (as opposed to a rule) which can be overridden by other tendencies or rules. Indeed, in all six cases, it seems to be the heaviness principle (see Section 5.2.) that is responsible for the postposition of the subject:

(6-18) SC (Vjesnik, p. 4)

NE BI inače imala pokrića tvrdnja da dr. Franjo Tujman, ostavljajući nas,

not:CLIT would otherwise have cover claim that dr. F. T. leaving us

odlazi u legendu.

goes into folk tradition, into legend

‘[President Tujman has died, but he will live in the memory of all citizens, for he has accomplished incredible deeds (a list follows).] Otherwise, the claim that dr. Franjo Tujman, who has left us, is passing into the legend, WOULDN’T be true.’

The extremely long subject (which is, nevertheless, treated as presupposed and topical, judged both by the context and by the H+L* tone on the auxiliary, indicating polarity focus) is placed after the object, *pokriće* (‘cover’), because of its length. There is, however, yet another possible reason for the postposition. The verb-object complex *imati pokrića* (‘to have cover’) is an idiom, meaning roughly ‘to be true, not to be vacuous’. There is a tendency in all three languages to keep close-knit lexical combinations (not necessarily idioms) together (see below, and especially 8.1.1.2.4. for more details): in other words, even if the subject were not so long, it would probably be placed after the object, in order to keep the parts of the idiom together. Be it how it may, these are only exceptions: the basic order of postverbal elements in SC is for the subject to precede everything else.

6.3.2.2.2. Albanian and Modern Greek

Alb. and MG data are less equivocal: In Alb., the number of long clauses is 92, 79 of which are relevant for the order of postverbal elements. Out of this number, 36 (45.6%) display the order [verb][subject][x], 43 (54.4%) the order [verb][x][subject] (cp. (6-1)). In MG, 389 out of 415 long clauses are of interest. Preposed subjects ([verb][subject][x]) are found in 210 clauses, i.e. 53.9% (cp. (6-2) and (6-13)), the postposed ones ([verb][x][subject]) in 179, (46.1%). All my attempts to find relevant discourse-pragmatic differences between the directly postverbal and clause-final subject topics failed. Therefore I tested a number of parameters not directly connected with the notion of topicality, and reached a rather unspectacular result: all of them play a certain role in triggering the variation, but none of them the decisive one. Here are the details.

If the additional element is a modal adverbial, it is more probable that it will stand directly after the verb, with the subject following: in Alb., the order [adverbial][subject] is attested in 9 instances, [subject][adverbial] in 3 (75.0% : 25.0%), while the ratio in MG is 56 : 19 (74.7% : 25.3%). The following example illustrates the most common order (v-adv-s):

(6-19) MG (Kapandai, p. 47)

Xamojelase amixana to palikari

smiled perplexedly the youngster

‘The youngster smiled, embarrassed.’

Other syntactic/semantic categories (objects, setting adverbials, predicatives, etc.) do not display any clear preference for the one or the other position.

Heaviness and cataphoric relevance (see Section 5.2. for the notions) seem to influence the relative order of postverbal elements in Alb. and MG in the same way as in SC: the heavier the element, the more probable its clause-final position. Here is an example of a heavy additional element, a predicative, (6-20) and of a heavy subject (6-21), both clause-final:

(6-20) Alb. (Kosovarja, p. 6)

Është ky i pari ministër i jashtëm që e viziton këtë shtet.

is this the first minister the outer who it visits that state

‘This is the first minister of foreign affairs to visit this country.’

(6-21) MG (Fakinou, p. 13)

... itan ekso ap' ta siniθismena afto pu zituse i Eleni tu.

was out of the usual that what asked the E. his

‘...what his daughter Eleni was asking was beyond his daily routine.’

As far as additional elements are concerned, this is not a rule, but merely a tendency: the examples of non-postposition are also attested:

(6-22) Alb. (Bishqemi, p. 62)

... *nuk kanë ç'na duhet aeroplanë të tillë*

not:CLIT have what us is-needed aeroplanes the such

‘... such aeroplanes do not have what we need.’

Heavy subjects, however, seem to stick more strictly to the heaviness principle: there is not a single example of a non-postposed heavy subject in my corpus. On the other hand, light, i.e. pronominal, subjects display a relatively strong tendency to be directly adjacent to the verb: in Alb., 26 pronominal subjects are found in the order [v][s][x] versus 3 in the order [v][x][s] (89.7% : 10.3%); in MG, the ratio is 46 : 16 (74.2% : 25.8%). The following sentence is an illustration of the dominant order:

(6-23) Eng: *I thought he must be somebody.* (Jack London, p. 142)

MG: *To katalava ejo oti prokite ja spuđeo atomo.* (p. 173)

it understood I that it-is-about for important person

Yet another factor influencing the relative order of postverbal elements is the semantic relationship between the verb and the element X: as in SC, if they form a kind of lexical unit larger than word, then it is more probable that the order will be [verb][x][subject] than [verb][subject][x], close-knit verb-noun combinations tending to be immediately adjacent, as shown in example (6-24):

(6-24) MG (Kapandai, p. 236)

Ke tendone t' aftia tu o Petros...

and stretched the ears his the P.

‘And Petros pricked up his ears (= listened attentively)...’

The relatively firm position of adverbials next to the verb (see also 5.1. and 8.1.1.2.3.), the heaviness principle and lexical units larger than words are the factors that may trigger minor word order variation in most constructions in the Balkan languages (see 5.1. and 5.2. for a general overview, and 8.1. for Inversion). In the Alb. and MG VsX construction, however, there also exist two construction-specific factors: focus type and the semantics of the subject.

There seems to be a slight preference for the subjects of polarity focus clauses to be placed directly after the verb, whereas the opposite holds for focus domain clauses (with broad or narrow focus construal). Thus, in Alb., there are 18 polarity focus clauses with [verb][subject][x] order as opposed to 13 such clauses with the order [x][subject], the ratio being 58.0% : 42.0%; in contrast, with focus domain clauses the order [s][x] is found in 18

clauses, [x][s] in 30 (37.5% : 62.5%). In MG, in polarity focus clauses the ratio of [s][x] and [x][s] is 55 : 39 (58.5% : 41.5%), in focus domain clauses 140 : 155 (47.5% : 52.5%). The statistical differences between the two focus types are admittedly far from significant, but they should not be underestimated either. Furthermore, there are two contexts in which the polarity focus seems to attract the subject as close to the verb as possible with a much greater regularity: when the verb is modal ('can' and 'must', but also propositional attitude verbs, phasal verbs, desideratives, and similar), or when it is negated. In MG, out of 19 long polarity focus clauses with a focused modal verb, 17 (89.5%) have the order [s][x]; out of 37 long negated clauses, 25 (67.7%) have this order; in Alb., the results are 6 out of 7 for modal verbs (85.7%) and 10 out of 14 for negated verbs (71.4%). For MG, the position of the subject immediately after the negated verb in polarity focus clauses is illustrated by (6-3) and (6-13). Here are two Alb. examples:

(6-25) Alb. (Kosovarja, p. 4/30)

NUK SHKOJNË gjërat ashtu ...

not:CLIT go things so

'It simply doesn't work that way ...'

(6-26) Alb. (Bishqemi, p. 78)

MUNDIM ne ta kthejmë këtë alamet anije...?

can we to-it turn that giant of-ship

'CAN we turn that giant ship?'

The last factor to be mentioned here (but not the last one conceivable) is the mode of reference of the subject. Epithet subjects (sometimes difficult to distinguish from apposition) refer anaphorically to a discourse referent but ascribe it a sense different from the previous mention: they are instances of an anaphor with the same extension and a different intension¹. This subject type is relatively frequent in VsX construction (see 6.4.1.), and it tends to appear clause-finally, i.e. after all other postverbal elements: in Alb., the ratio between clause-final and directly postverbal epithet subjects is 3 : 0, and in MG it amounts to 27 : 11. A nice illustration is given in (6-27):

(6-27) MG (Matesis, p. 32)

... ði then pijenan ekðromi na kseskasun i murles.

and-so went excursion to relax the crazy

¹ The typical examples are those of negative characterization: *I met Peter_i yesterday. **The idiot_i** didn't recognize me.*

‘[The women from the neighborhood bring bombs to the guerilla around the town pretending to go on a picnic; this time they took food, blankets, even some instruments with them], and so the crazy cows went on a picnic to relax!’

Obviously, I am not able to offer the definite rule for the placement of postverbal elements in Alb. and MG, because there is probably no such a thing at all. Instead, a number of stronger or weaker tendencies may be observed. The subject will most probably appear directly after the verb if it is light, especially if it is pronominal, and if the focus structure used is polarity focus, especially with modal and negated verbs. The clause-final position of the subject is preferred if the other postverbal element is a modal adverbial or an element in the relationship of lexical solidarity with the verb, if the subject is heavy, or if it belongs to the class of epithet subjects.

6.3.2.3. Complex predicate forms

A specific problem in the syntax of VsX construction is the behavior of complex predicate forms, more precisely of the tense and mood forms formed with an auxiliary and a participle and of the nominal predicates consisting of the copula and a predicative. The data from SC are of some interest only in those tense forms in which the copula/auxiliary is not clitic, which leaves us practically only with the past tense of nominal predicates and with the negative forms of the auxiliary (cp. Section 5.1.).

What is specific about complex predicate forms is the possibility of inserting the subject between the copula/auxiliary and the predicative/participle (hereafter simply *insertion*). There seems to be some significant variation here, both across and within languages, as the following table shows:

(6-28) *Insertion of subject between [aux/cop] and [ptcp/pred]:*

	[aux][ptcp][s]	[aux][s][ptcp]	[cop][pred][s]	[cop][s][pred]
Alb.	11 – 100%	0 – 0%	26 – 61.9%	16 – 38.1%
MG	25 – 92.6%	2 – 7.4%	42 – 77.8%	12 – 22.2%
SC	2 – 18.2%	9 – 81.8%	1 – 3.1%	31 – 96.9%

The columns to the left contain data on non-insertion, those to the right on insertion. The general impression is clear: Alb. and MG tend to keep [aux/cop] and [ptcp/pred] together and place the subject only after them, although there is some variation here; SC inserts the subject between the two parts of the predicate with some regularity.

(6-29) MG (Ciao, p. 24)

Exun maθi i iθopji oti ta orea ine ta ðiskola.

have learnt the actors that the beautiful are the difficult

‘The actors have learnt that only difficult things are really good.’

(6-30) Alb. (Bishqemi, p. 111)

Nuk ishte detyra ime kjo.

not was duty-the my this

‘This was not my duty.’

(6-31) SC (Andrić, p. 32)

... *nije on nevin*

not-is he innocent

‘... he is not innocent.’

This is roughly in accordance with the positioning tendencies observed in these languages with respect to the behavior of the subject and other postverbal elements in the previous section: the immediately postverbal position of the subject in SC and the clause-final position with a lot of variation in Alb. and MG; even the independent ordering principles, like heaviness, seem to work in the same fashion (cf. (6-20) and (6-21)). In fact, if we treat the copula/auxiliary as the ‘verb’ and the predicative/participle as a lexical complement, then the special section devoted to complex predicate forms is superfluous.¹

But there is more to it. First, the relationship between the auxiliary and the participle in Alb. and MG is not precisely the same as that between the copula and predicative: the former are almost never separated, the latter may be so, and with a not insignificant frequency. Actually, my Alb. informants were decidedly against the variants with the subject inserted between the auxiliary and the participle I presented them with: the string *?kemi ne thënë* (‘have we said’), for instance, is judged very strange. The speakers of MG claimed that the only two examples of insertion from my corpus (both from one source, Kapandai’s collection of short stories) sound very archaic: thus *exi o θios tu fiji* (‘has the uncle his gone’, Kapandai, p. 121) could have been said by someone in the eighteenth century, but not by a person in today’s Athens, as one of my informants put it. All this points to the advanced grammaticalization of complex tense forms in the two languages, with the two building blocks of the tense form, the auxiliary and the participle, being practically inseparable.

¹ This is the analysis Fillmore (1999) proposes for the auxiliary-participle complex in English; for more details, see Section 8.1.1.3.

Nominal predicates, on the other hand, are not subject to this kind of restriction, which is why they are found with the variation typical for the two languages in question.

The second relevant point is the fact that the possibility of inserting the subject within the complex predicate form – regularly used in SC, possible with nominal predicates in Alb. and MG – is a construction-specific matter. It may occur in some types of Inversion (especially quotation inversion and wh-inversion, see Sections 8.1. and 8.2.), but it is almost completely absent from vS construction (see 11.2.1.). Inserted subjects are thus, apart from the absolute initial position of the verb and the intonational pattern, the third formal feature which separates VsX construction from other types of VS order.

6.3.2.4. Formal properties and information structure: A summary

In this section, the structural properties of VsX construction will be summarized, with considerable simplification, in the form of sentence templates. The notational apparatus used is the one introduced in Section 5.3.; # marks the clause boundary.

(6-32) *Sentence templates for VsX construction:*

Alb., MG and SC

short clauses, polarity focus type and focus domain type

#[VERB] [subject]#

Alb. and MG

long clauses, polarity focus type

#[VERB][subject][rest]#

#[VERB][rest][subject]#

long clauses, focus domain type

#[VERB][REST][subject]#

#[VERB][subject][REST]#

clauses with nominal predicates

#[copula][PREDICATIVE][subject]#

#[copula][subject][PREDICATIVE] #

SC

long clauses, polarity focus type

#[VERB][subject][rest]#

long clauses, focus domain type

#[VERB][subject][REST]#

clauses with nominal predicates

#[copula][subject][PREDICATIVE] #

The three distinctive features of VsX construction – the clause-initial position of the verb, the intonation pattern, and the possibility to insert the subject between the copula and predicative – are represented in the templates. The differences between the three languages are made somewhat more prominent than they really are by simplification: the prototypical VsX clause in SC has a subject that is immediately adjacent to the verb; in Alb. and MG, the position of the subject is influenced by a number of competing factors, so that both variants, the immediately postverbal and the clause-final subject, are almost equally frequent.

6.4. Semantic and informational features of subjects and verbs

6.4.1. Subject

Subjects of VsX clauses are often pronominal, or at least modified with a pronoun, as witnessed by (6-1) – (6-3), (6-12), (6-13), (6-20) – (6-22), (6-26), (6-30) and (6-31).¹ Here is some statistics:

(6-33) *Pronominal and lexical subjects in VsX construction*

	pronominal S	lexical S	Σ
Alb.	62 – 51.7%	58 – 48.3%	120 – 100%
MG	145 – 26.0%	412 – 74.0%	557 – 100%
SC	69 – 59.5%	47 – 40.5%	116 – 100%

In Alb. and SC, clauses with pronominal or pronominally modified subjects make up more than a half of all VsX clauses; in MG, they represent less than a third. There are two reasons for this asymmetry: First, as shown in 5.1. and 5.5., MG tends to leave its subjects unexpressed, i.e. to use the zero topic encoding, significantly more often than Alb. and SC. Thus, when no descriptive content is felt to be necessary for the hearer to identify the referent, the speakers of MG are much more likely not to express the subject than is the case with the speakers of the other two languages. Second, as will become patent in the course of this and the following chapters, VsX construction is for a greater part a moribund category in Alb. and SC, kept alive in a number of semi-petrified textual contexts, most of which demand a pronominal subject. In MG, the construction is very much alive, the number of contexts in which it occurs being much larger; consequently, lexically expressed subjects are immeasurably more frequent in this language than in the other two, so that the share of the pronominal ones automatically sinks.

¹ The same phenomenon has been observed in Polish by Siewierska (1987) and Jacennik and Dryer (1992).

The high or the relatively high ratio of pronominal subjects has to do with the fact that the subjects in VsX construction encode ratified topics: when a referent is ratified as the topic, it often need not be further identified via lexical description. However, even when lexically encoded, ratified topics usually display a high degree of referential continuity, meaning that they have already played a role of a discourse referent in the (more or less) immediately preceding text; or they are somehow present in the physical context (deictically given); or they are, even though not immediately present in the textual and physical context, major participants of the larger stretch of discourse; or they are easily made relevant through different semantic frames. Note that one discourse referent may be both immediately present in the textual context and play the role of the major discourse referent, or similar. Consequently, the fact that a referent is referred to in one row of the following table does not imply that it cannot appear in one or more other rows. This is why the total number of instances could not be given.

(6-34) *Contextual boundness of the 3rd person subjects in VsX clauses*

	Alb.	MG	SC
(1) 1 st and 2 nd person deictic	12 – 10.0%	20 – 3.6%	12 – 10.3%
(2) 3 rd person deictic	3 – 2.5%	6 – 1.1%	4 – 3.4%
(3) mentioned in the last 1–5 claus.	48 – 40.0%	267 – 47.9%	52 – 44.8%
(4) mentioned in the last 6–10 claus.	3 – 2.5%	49 – 8.8%	5 – 4.3%
(5) mentioned in the last 10–20 claus.	0 – 0.0%	5 – 0.9%	0 – 0%
(6) inferable	27 – 22.5%	169 – 30.3%	25 – 21.6%
(7) major participant	28 – 23.3%	194 – 34.8%	30 – 25.8%

Deictic categories (1) and (2) are better represented in Alb. and SC than in MG, which is probably to be explained by their being inherently pronominal, with MG tending not to express pronominal subjects more often than Alb. and SC. The nature of the corpus (more first and second person subjects in Alb. and SC than in MG, see 5.4. and 5.5.) may play a certain role as well.

The ratio of categories based on referential continuity, (3) through (5), is almost equal in all three languages if the subject referent has been mentioned up to five clauses before the VsX sentence. The greater the distance between the last mention of the subject referent and the clause, the greater the difference between MG on one hand and Alb. and SC on the other. Namely, MG seems to use VsX construction more readily than Alb. and SC in those cases in which the subject referent has been mentioned quite some time ago.

The greater tolerance of the MG VsX construction towards less immediately textually or physically present subject referents is also confirmed by the category (6): if the subject

referent is not directly mentioned in the preceding text or not directly present in the physical context, but has to be inferred from the semantic frames invoked in the discourse instead, MG allows for a VsX clause significantly more often than Alb. and SC. Let me illustrate what I mean by inferring from a semantic frame by discussing the following example: In the first paragraph of an article in the Greek daily *Eleftherotipia*, a crisis of the Greek Government provoked by the refusal of the minister of culture to fly to the opening of the Olympics in Sidney is described, and the conjecture is expressed to the effect that this is a sign of deeper misunderstandings within the government; the second paragraph is opened with the following sentence:

(6-35) MG (*Eleftherotipia*, p. 4)

Veveos, ðen perimene o kirios Simitis na erthun ola mazemena...

of-course not expected the mister S. to come all gathered

‘Of course, Mr. Simitis has not expected that all the problems will surface at the same time...’

Now, Mr. Simitis, the prime minister, has not been mentioned at all in the preceding text; the author of the article, however, relies on the frame ‘the present Greek government’, which he assumes to be shared by him and the readers, and expects the readers to infer the immediate relevance of the whole story for the prime minister, whose identity is, of course, known to the readers of *Eleftherotipia*. It is in this sense that the ratified status of the subject in VsX clauses can be inferred from the semantic frames. And it is this inferencing device that is used more often in MG VsX clauses than in the Alb. and SC ones.

Finally, the row (7) shows that the subjects of MG VsX clauses are significantly more often major participants in the narrative than is the case in Alb. and SC. This difference has to do with the difference in the discourse functions fulfilled by VsX clauses in MG on one hand and in Alb. and SC on the other, so that it will be explored in some detail in Section 6.5., devoted to the role of VsX construction in discourse.

The properties of subjects dealt with by now can be subsumed under the label ‘informational’ or ‘discourse properties’. As far as the semantics of these phrases in the narrower sense is concerned, only two features seem to be of some relevance. Namely, the fact that some 75.8% (91) subjects in Alb., 91.0% (507) in MG, and 64.7% (75) subjects in SC refer to humans or personified animals (mostly in the translations of Jack London’s *White Fang*) is not typical only of VsX construction, but represents a characteristic of the subjects in most other kinds of constructions (except for vS construction; cf. 11.3.2.). What is interesting is rather that among non-human subject referents, states of affairs play a rather prominent

role, especially in SC: thus, in SC, subject expressions referring to a situation amount to 36 (31.0%), in Alb. to 21 (17.5%), and to 24 (4.0%) in MG. The prototypical mode of reference to a state of affairs is via pronoun, in SC always in neuter singular, in Alb. and MG either feminine/neuter singular or agreeing with the predicative nominal in gender and number. I shall call this type of pronominal subject ‘non-syntactic pronoun’ (NSP), referring to the fact that it does not take its agreement features from an antecedent nominal to which it is anaphorically bound.¹ Here are two examples illustrating the two types of NSP:

(6-36) SC (Kiš, p. 93)

Bilo je to jednog užasno studenog popodneva...

was AUX this:NEUT:SG on-one terribly cold afternoon

‘[After a description of a funeral] This happened on a terribly cold afternoon....’

(6-37) Alb. (Kosovarja, p. 5/47)

Ishte ky kontakti i parë i nuses me djalin ...

was this:MASC:SG contact:MASC:SG the first-of-the bride with the boy ...

‘[After a description of the first encounter of a young bride and a boy she falls in love with] This was the first time the bride met the boy...’

NSPs as subjects of VsX clauses are very frequent in SC (26 instances, i.e. 22.4%), less so in Alb. (12, i.e. 10.0%), and rather infrequent in MG (7, i.e. 1.3%). In fact, it is this frequency relationship between NSPs that is responsible for the difference in the frequency of subjects referring to states of affairs in the three languages: when expressed lexically, they are approximately equally frequent (10 in SC, 9 in Alb., 17 in MG). The point is that MG, and, to a certain extent, Alb., very often use zero encoding to refer to previously described states of affairs; in SC, this is one of the few contexts where zero encoding is not grammatical. The result is the high frequency of VsX clauses with a NSP subject in SC, and a relatively low one in MG and, less dramatically, in Alb. The typical behavior of the three languages in regard to NSP subjects referring to states of affairs is neatly illustrated by the following example, defining the relationship between a dog and its master:

(6-38) Eng: *It was a pain and an unrest ...* (London, p. 154)

Alb: *Ishte dhimbje dhe shqetësim ...* (p. 151)

was pain and unrest

MG: *Itan ponos ki anisixia ...* (p. 186)

was pain and unrest

¹ The term represents a slight adaptation of Gensler’s (1977) *non-syntactic anaphor*.

SC: *Bio je to neki bol i neki nemir...* (p. 173)

was AUX this some pain and some unrest

More about NSP subjects in Section 6.5. Let us now turn to the second interesting semantic feature of VsX subjects. As mentioned in 6.3.2.2.2., VsX clauses seem to be a perfect context to use an epithet subject: If one defines epithets rather narrowly, as expressions which ascribe an utterly new sense to the referent invoked before (as in (6-27), where ‘the women’ mentioned in the previous clause are reinterpreted as ‘crazy cows’), their number is relatively small in Alb. (3, i.e. 2.5%), but not insignificant in MG (38, i.e. 6.9%) and SC (7, i.e. 6.0%). In view of the fact that such subjects are rather infrequent in general (0.7% in the whole Alb. corpus, 1.1% in MG, 0.7% in SC), this is a rather impressive ratio. The reason for the relatively frequent use of epithets in VsX clauses is the fact that the uses of the two categories, epithets and postverbal topical subjects, are a perfect match: VsX construction marks referential continuity and various types of frame discontinuity, the change of the perspective being the relevant one here; epithets represent an evaluative device, by means of which the judgment of the speaker is presented without explicit assertion, simply by giving the hearer the instruction to find the referential antecedent of the presupposed nominal expression in the previous text. The connection is clear: to give an evaluation means, as a rule, to change the perspective, usually from the narrative world to the world in which the speakers judges the qualities of the participants; the referential continuity is a part of the definition of VsX construction. One more example in addition to (6-27), this time with a somewhat more affirmative judgment:

(6-39) MG (Kapandai, p. 137)

Mia zoi olokleri itan o babas_i, i sxesis tus safos kaθorismenes mesa se plesia

a life complete was the father the relations their clearly defined within in frames

me periyramata kaθara ..., ke istera arostise. AROSTISE o babakis tis_i ke ...

with contours pure and later he-got-ill got-ill the father:DIMIN her and

i Andiyoni ... parakaluse to Θεο na ton sosi ...

the A. prayed the God to him save

‘Her father was a life contained in itself, their relationship always clearly defined within the frames with sharp contours ..., and then he got ill. Her dear old daddy got ill, and Antigone prayed to God to save him....’

There is a fine change of sense, and a fine change of perspective in this piece of discourse. The first part, in which ‘the father’ is simply called *babas*, ‘father’, is a story of the past of the two characters, the father and the daughter. The second part, which begins with a VsX clause

arostise o babakis tis, is a story of the father's last days. The main topic, the discourse referent 'the father' remains the same, but the way he is referred to is slightly different: now he is *o babakis tis*, 'her (dear, little) daddy'. By the use of a VsX clause the hearer is instructed to seek for the old topical referent in a changed perspective (from the ancient history to the current agony), so s/he pairs *o babakis tis* with *o babas* straightforwardly; the change in the emotional perspective is encoded additionally, by slightly changing the sense of the referring expression: 'the father' of the author's old days is now 'the dear daddy'.

To sum up: In Alb. and SC, the subjects in VsX clauses are pronominal in about a half of all instances; in MG, this is the case in less than a third of all cases. The difference between Alb. and SC on one hand, and MG on the other, is further confirmed by the measurements of the degree of the current relevance of the subject referents in the discourse: whereas Alb. and SC seem to prefer their VsX subjects to be immediately given in the physical or textual context, MG allows for a much greater range of the kinds of relevance carried by the subject referent. As far as semantic features are concerned, VsX construction sides with most other clause types in having human subjects most frequently. A specific feature of the construction is the relatively high frequency of the subjects referring to a state of affairs, especially in SC, where the preferred mode of reference are NSPs, which occur less often in Alb., and especially in MG. Finally, if the speaker of Alb., MG, or SC intends to use an epithet subject, s/he is, especially in MG, most likely to use VsX construction.

6.4.2. Verb

Although all kinds of predicates may occur in VsX clauses, there are some clearly recognizable tendencies, which are, interestingly, not absolutely identical in the three languages under consideration. Five classes of predicates appear in the construction with a frequency higher than average: nominal predicates with a copula (cp., e.g. (6-20), (6-21), (6-37), (6-38), (6-30), (6-31), etc.), modal verbs (in the broader sense of the verb: 'can', 'must', but also 'begin', 'try', etc. – see (6-1), (6-3), (6-26), etc.), verbs denoting cognitive activities ('know', 'remember', 'speak', etc. – cp. (6-23), (6-24) and (6-29)), verbs denoting emotions (e.g. (6-17) and (6-19)), and verbs of (non)existence and (dis)appearance (cp. (6-2), (6-16), (6-18), etc.). The following table contains data on the absolute and relative frequency of each of these classes in Alb., MG and SC:

(6-40) *Semantics of the predicates of VsX clauses*

	Copula	modal	cognitive	emotion	existence	Σ
Alb.	42 – 35.0%	9 – 7.5%	11 – 9.2%	10 – 8.3%	16 – 13.3%	88 – 73.3%
MG	55 – 9.9%	43 – 7.7%	83 – 15.1%	74 – 13.3%	59 – 10.5%	339 – 56.4%
SC	42 – 36.2%	10 – 8.6%	13 – 11.2%	6 – 5.2%	14 – 12.1%	85 – 73.3%

The first difference between the three languages to be noted is the one seen in the last column: whereas the five predicate classes identified as prototypical cover some two thirds of all instances of VsX construction in Alb. and SC, they occur only about a half of the VsX clauses in MG, meaning that the range of verbs used in VsX construction in MG is larger than in the other two languages, a fact pointing out to the greater productivity of the construction in this language.

The second important difference is the frequency of the use of nominal predicates with the copula. In Alb. and SC, this is the major lexical filling of the verb slot in VsX construction, with more than one third of all instances; in MG, the frequency of this predicate type amounts to only about one tenth. This is partly due to the fact that other predicate types are better represented in MG, so that the percentage of copular clauses automatically sinks. In part, however, there are certain features of the use of VsX construction in discourse in MG which strongly disfavor the frequent use of nominal predicates; the details will be presented in Section 6.5.

Modal and verbs of existence seem to be equally well represented in all three languages; as for the remaining two categories, verbs of cognition and emotion, they are slightly more frequent in MG than in the other two languages, with SC having a very low ratio of the latter.

There is also some variation as concerns the composition of the predicate classes themselves. In the group of nominal predicates, Alb. and SC seem to prefer the stative copula (*ështëë, biti* – ‘to be’), with only one example of the dynamic one in each (*bëhet, postati* – ‘to become’); in MG, the dynamic copula *jinome* is used in 11 out of 55 copula clauses. Furthermore, only in MG do other copular verbs, like *meno* ‘remain’, *pao ja* or *lojizome*, roughly ‘count as’, ‘be considered’, etc. occur as well.

The situation with other predicate classes is comparable, the general pattern being greater variety in MG, and only a limited number of verbs in Alb. and SC. For instance, in Alb. and SC, among modal verbs, only those meaning ‘can’ (*mund, moçi*), ‘want’ (*do, hteti*) and ‘begin’ (*fillon/zë, početi*) are attested, whereas MG has, apart from the above mentioned (*boro, thelo, arxizo*), various other verbs as well, like *anangazome* ‘be forced’, *prepi* ‘should’, *prospaθo* ‘try’, *kano* ‘make an attempt’, *kataferno* ‘succeed’, etc. Similar differences appear in all the

predicate classes. It should, however, be noted that it is only in the class of copular sentences that the greater variety of verbs in MG cannot be ascribed to the greater number of examples, since with all other groups the MG corpus contains four to eight times more VsX clauses than the other two languages.

The second interesting semantic feature of the predicates used in VsX clauses is the frequency with which they occur with negation. In Alb., 27 out of 120 clauses are negated (22.5%), in MG 74 out of 557 (13.3%), in SC 26 out of 116 (23.2%). This is significantly more than the mean value for negation in all the sentences in the corpus, which amounts to 8.3% (1312 clauses) in Alb., 8.5% (1229) in MG and 7.7% (1091) in SC. Negated predicates are well represented in the examples adduced by now, e.g. in (6-1) – (6-3), (6-12), (6-13), (6-22), (6-30), (6-31), (6-35), etc. The fact that the percentage of the clauses with negative polarity is higher in Alb. and SC than in MG is to be understood as a consequence of the greater frequency of polarity focus VsX clauses in these two languages, negated verbs being one of the prototypical forms of focusing the polarity of the proposition (cf. 6.5.1.).

The last feature of the VsX verbs worth mentioning is a discourse-pragmatic one. Verbs used in the construction are more often ‘old’, i.e. either directly mentioned in the immediately preceding context or inferable from semantic frames, than is the case in other constructions. Thus, Alb. has 20 ‘given’ verbs (16.7%), MG 75 (13.3%), SC 12 (10.4%). Almost all of the clauses containing such verbs are instances of polarity focus: to assert merely the fact that a state of affairs occurs or does not occur means to presuppose the whole proposition, ‘oldness’ being one of the important factors facilitating such extensive presuppositions (see 4.2.4.). Consider the following sentence:

(6-41) SC (Andrić, p. 20)

»... što ti nisi najurio tog svog otrova ...?« ... »Najurio, najurio!

why you not:AUX chase-away that your poison chase-away chase-away

Nije to tako lako...« - »*Ah šta! NAJURIO bih ja nju...*«

not-is that so easy oh what chase-away would I her...

‘[A conversation about an evil wife] »...why didn’t you simply chase away that snake?« ... »Chase away, chase away! It is not that easy....« - »Bullshit! I would chase her away...!«

The proposition ‘I would chase her away’ is presupposed, i.e. the speaker assumes that the hearer is able to create the mental representation of the world as being that way, which does not necessarily mean that he assumes that the hearer believes that the proposition is true, but merely that s/he is able to accept the possibility of the existence of such a proposition; what is

asserted is the fact that the speaker believes that the proposition is true, i.e. its positive polarity – ‘I WOULD chase her away...’. The speaker’s assumption that the hearer is ready to take such an extensive presupposition from him is based on the fact that ‘chasing away the evil woman’ has been repeatedly mentioned in the immediately preceding textual context. This is the reason ‘old’ verbs occur so often in VsX construction with polarity focus structure.

To conclude: Five predicate classes are identifiable as typical of VsX construction, nominal predicates with the copula, modal verbs, verbs of cognition, emotion, and existence. Copular sentences are significantly more frequent in Alb. and SC than in MG, modal and verbs of existence are approximately equal, whereas verbs of cognition and emotion occur somewhat more often in MG. In all these verb classes, MG disposes of a greater range of verbs than the other two languages. The other two features of the verbs in VsX construction, frequent negative polarity and the ‘givenness’ of the verb, are connected with one of the two focus structures compatible with the construction, the polarity focus structure.

6.5. Discourse functions of VsX construction

The informational and semantic properties of the elements of the construction are, as I shall try to demonstrate in this chapter, determined by its basic function, to encode ratified topics in a changed discourse frame, and by the functions it performs in the development of discourse.

As I have repeatedly pointed out throughout this chapter., VsX construction can be used with two focus structures, the polarity focus and the focus domain structure, the latter being construable as a broad or a narrow focus. The polarity focus structure is formally characterized by a specific intonational contour (see 6.3.1.) and, especially in Alb. and MG, by the tendency of the subject to appear directly after the verb; as shown above, some types of verbs and of subjects tend to be used more frequently with the polarity focus. For a demonstration of how polarity focus structures work, see sentence (6-41).

The focus domain type has an intonational contour of its own (6.3.1.), and is, just like the polarity focus, particularly frequent with certain subject and predicate types. Sentence (6-1) is a nice example of a focus domain with a narrow focus construal: *filloi ters kjo punë* (‘began bad this thing’) has as its ratified topic the subject phrase *kjo punë*, with the verb and the adverbial *filloi ters* as a syntactic domain which may be interpreted as asserted, i.e. focal; the fact that the hearer can infer from the previous text that ‘this thing’ has already started makes the narrow focus construal on *ters*, ‘bad’ more probable. The resulting assertional structure is

thus $[[\textit{filloi}]_{\text{Presupposed}} [\textit{ters}]_{\text{Focus}}]_{\text{Focus Domain}} [\textit{kjo punë}]_{\text{Ratified Topic}}$. However, if the hearer would for any reason be assumed not to be able to accept the predicate as presupposed, a broad focus construal would be possible as well: $[[\textit{filloi ters}]_{\text{Focus}}]_{\text{Focus Domain}} [\textit{kjo punë}]_{\text{Ratified Topic}}$.

Let us now look at some statistics on the frequency of the two focus types in VsX clauses:

(6-42) *Polarity focus and focus domain VsX clauses:*

	polarity focus	focus domain (broad/narrow)	Σ
Alb.	64 – 53.3%	56 – 46.7%	120 – 100%
MG	168 – 30.2%	389 – 69.8%	557 – 100%
SC	63 – 54.3%	53 – 45.7%	116 – 100%

Alb. and SC, as will be the case in the greater part of this chapter, side together, with polarity focus clauses amounting to more than 50% of all VsX clauses; in MG, they make up less than a third of the total number. The explanation of these statistical facts will be offered in what follows.

6.5.1. Polarity focus

Before the position of polarity focus VsX clauses in discourse is discussed, a note on terminology and ontology is in order. Most scholars writing on this kind of focus (Gussenhoven 1983a; Dik et al. 1981 [*polar focus*], Höhle 1992 [*verum focus*], Adamec 1966 [*verificative focus*]) emphasize that the notion of polarity focus can be applied only to those cases in which only the truth or falsity of a proposition is asserted; when it is the TAM component that the speaker asserts (as in *I am not ill; I WAS ill*), one should not speak of polarity focus. This purism stems from an implicit or explicit belief that what is focused in the case of polarity focus is a covert illocutionary operator (labeled *VERUM* by Höhle); it follows naturally that, when tense, aspect or mood are asserted, some other covert operator has to be focused. *Ergo*, polarity focus is not the same kind of thing as TAM focus.

There are two problems with this claim. First, it is often impossible to distinguish between the assertion of the commitment of the speaker to the truth value of a proposition and the assertion of TAM. Second, the multiplication of entities in the already overpopulated world of language by postulating the existence of covert operators should *a priori* be avoided (*entia non sunt multiplicanda praeter necessitudinem*).

My approach to the problem is therefore based on the radical pragmatic ideology (see Goldberg and Ackerman 2001 for a similar proposal). The polarity (or TAM) focus interpretation comes into being when the verb is focused, but the hearer does not find its

lexical or informational content sufficient to justify the assertion of the content itself: in the context of the utterance *I am not ill*, the sentence *I WAS ill* cannot be meaningfully interpreted as the assertion of the descriptive content of *was*, although it is formally marked as focus – the verb is both ‘given’ and void of descriptive content. The hearer has to enrich the message by searching the appropriate object of assertion within the semantics of the predicate: in this particular case, it will probably be the verb’s temporal reference. If the context were *You were not ill*, the same sentence, *I WAS ill*, would be interpreted as the assertion of the positive polarity of the predicate, i.e. as the ‘true’ polarity focus.

The consequences of this approach are predictable: since the same principle is at work in both the ‘TAM’ focus and the ‘real’ polarity focus, and since the interpretation of the assertion as being of one or the other kind is merely the matter of pragmatics, I do not see any compelling reason to separate the two types. Actually, as will become patent briefly, there are more than just two interpretations of this type of focus. The term *polarity* has been kept simply because it is already an established term in linguistics.

Back to the point: There are two types of use of polarity focus: the illocution-oriented and the text-cohesion-oriented one. In the description that follows, they will be dealt with separately.

6.5.1.1. Polarity focus in illocutionary contexts

Not surprisingly, polarity focus used in illocutionary contexts is a feature of the spoken language; most examples from my corpus are therefore taken from retold dialogues, and in a number of cases informally collected spoken data had to be adduced.

(1) **Confirmation and denial of beliefs.** Polarity focus VsX clauses may be used in conversation to confirm or deny both one’s own and the beliefs of the interlocutor(s), both the implicit and the explicitly expressed ones.

The confirmation and denial of one’s own beliefs typically occur with cognitive verbs (including verbs of saying), and, of course, with first person subjects:

(6-43) Eng.: *I tell you right now, Henry, that critter’s the cause of all our trouble.*
(London, p. 18)

Alb: *Ta THEM unë, Henrik, për të gjitha kusuret tona fajin ajo e ka.* (p. 18)
to-you-it:CLIT say I H. for the all troubles-the our guilt she it has

- (6-44) Eng. *I thought he must be somebody*. (Jack London, p. 142)
 MG: *To KATALAVA eyo oti prokrite ja spuđeo atomo*. (p. 173)

it understood I that it-is-about for important person

- (6-45) SC (Andrić, p. 10)

KAŽEM ja uvijek: *Nisi ti Rastislav, nego Raspislav!*

say I always not-are you R. but R.

‘[After criticizing a young monk for being extravagant, an old monk concludes:]

»As I always say: You are not Rastislav, but Raspislav (“a squanderer”)«’

In all three examples, the speaker uses polarity focus to confirm the beliefs he entertained before; especially with verbs of saying, the combinations like the above (*e them unë, to leo eyo, kažem ja*) are semi-idiomatic expressions used when the constancy of one’s claims and beliefs is to be emphasized: the message reads as ‘I do say (think, believe) now as you know that I have always said (thought...)’. Of course, ‘I’ is a ratified topic in such contexts (it is all about *my* beliefs), and the perspective changes from the world of uncertainty to the world in which the knowledge is achieved via confirmation.

If it is the beliefs of the interlocutor(s) that are to be confirmed or denied, the predicates and the subjects used vary depending on the belief assumed or expressed. Here are two examples:

- (6-46) Alb. (Kosovarja, p. 4/30)

NUK SHKOJNË gjërat ashtu si mendon ti.

not:CLIT go things so as think you

‘[You try to keep your husband by making scenes. Be careful:] It simply doesn’t work the way you think it does.’

- (6-47) MG (Ciao, p. 42)

»*Ola afta pu mas ðijjiste, ðen sas tromazun otan simvenun?*«

all that REL us you-told not you they-scare when they-happen

»*Oxi. ... Mu ARESI i epikinonia afti.*«

no me:CLIT likes the communication that

‘[An interview with a woman who receives signs from beyond from her late husband.] »All the things you told us, doesn’t it make you shudder when it happens?« - »No. ... I LIKE that communication between us.«’

In (6-46), the implicit belief of a reader that she may keep her husband by being jealous is refuted by a polarity focus VsX clause: the presupposed proposition ‘things work that way’ is denied: ‘things DON’T work that way’. The subject *gjërat*, ‘the things’, refers quasi-

anaphorically to the ratified topic ‘keeping one’s husband’; the change of perspective takes place between the perspective of the reader herself and that of the writer. Example (6-47) is adduced to show that denials do not have to be negative: the speaker assumes that the hearer assumes ‘you do not like that’, and denies this belief with ‘it is not the case that I do not like it’, i.e. ‘I do like it.’

In straightforward answers to explicitly asked yes-no questions, VsX construction is usually not used, the zero-coding strategy being preferred instead, presumably because questions do not imply beliefs, but rather uncertainty. However, if the speaker feels that the question is rhetorical, i.e. an expression of a belief, not of uncertainty, s/he may use VsX construction in order to emphasize the change of perspective, from ‘your’ belief to ‘mine’:

(6-48) SC (Kiš, p. 7)

»Zar može jedna ... kokoška koja vredi ... pet červonaca da bude ravna jednom

QUEST can one hen which is-worth five coins to be equal to-one

*tvoru koji ... smrdi nadaleko?» – »Ne, ... **ne može se ravnati kokoška koja vredi***

polecat which stinks afar no not can REFL be-equated hen which is-worth

***pet červonaca sa smrdljivim tvorom.**«*

five coins with stinking polecat

‘»Can a ... hen, which is worth five coins, be considered equal to a polecat, which ... stinks awfully?« - »No, ... a hen which is worth five coins cannot be compared to a stinking polecat.«

The question is in fact a statement; the answer is encoded as a VsX clause, with the repeated postverbal subject, because it is not a genuine answer, but rather a confirmation of the interlocutor’s belief expressed in the form of a question.

The use of polarity focus VsX clauses as confirmations or denials of beliefs is equally productive and frequent in all three languages.

(2) **Promises, reassurances, and threats.** The speaker may make a promise, or reassure the apparently uncertain hearer, or threaten her/him, by pretending that the hearer has a proposition in mind, which s/he then affirms or refutes with a polarity focus VsX clause. Which one of the three interpretations will be chosen depends both on the context and on the content of the presupposed proposition. First a promise, naturally from a mouth of a politician. In an interview devoted to the tasks the future Serbian government has in clarifying the political and other crimes committed by the previous government, the prospective minister of internal affairs uses the following two sentences almost one after another:

(6-49) SC (*Glas javnosti*, 20.12.2000, p. 1)

DOBIĆEMO mi policiju koja će biti maksimalno efikasna. ...

will-get we police which will be maximally efficient

USPOSTAVIĆEMO mi red u ovoj zemlji.

will-establish we order in this country

‘We shall have a police which will be maximally efficient. ... We shall establish order in this country.’

The ratified topic in these clauses is ‘we, the future government’. The speaker acts as if the readers of the daily in which the interview is published had the propositions ‘we shall have a police...’ and ‘we shall establish order’ in mind, but were not sure of their truth value. The polarity focus assertion confirms the that they are true indeed: the perspective changes from the presumed uncertainty of the readers to the certainty of the politician. In VsX clauses used as promises, it is usually the future tense that is used; the subjects are often first or second person pronouns.

When the speaker has the reason to believe that the hearer is worried because of a certain state of affairs, s/he can utilize the same device, this time for reassurance. A reader complains to a fortune-teller that her child suddenly stopped talking. The fortune-teller offers comfort:

(6-50) Alb. (Kosovarja, 4/37)

Do t’i KTHEHET të folurit shpejt.

will *të*-to-him:CLIT return the speaking soon

‘His ability to speak WILL return soon.’

The proposition is presupposed, the assertion confined to the positive polarity, the scene changed from the reader’s uncertainty to the fortune-teller’s confidence, the topic being held constant: ‘his ability to speak’.

If the content of the proposition is not considered pleasant to the hearer, the principle according to which promises work may be used as a threatening device. In fact, what is a promise for one, may be a threat to the other: thus (6-49) may be read as a threat, if the readers are the members of the previous government. The following example is more unequivocal in its threatening quality:

(6-51) MG (Matesis, p. 65)

θa su to KAPSO eyo to spiti, exo ton tropo!

will:CLIT to-you:CLIT it:CLIT burn I the house I-have the way

‘I shall burn your house down, I’ll find the way!’

The speaker acts as if the hearer were not sure if she is going to burn his house down; in order to ascertain that her intentions are serious, she uses a polarity focus VsX clause.

The use of polarity focus VsX clauses in promises, threats, etc., resembles the use of particles in some other languages, like *schon* in German (*Ich KRIEG' dich schon*) or *már* and *még* in Hungarian¹. The construction is productive and frequent in all three languages.

(3) **Unknown reality: questions, wishes, orders, and alternatives.** The two uses of polarity focus VsX clauses described by now are based on the assumed ignorance of the hearer. When polarity focus is used to mark the ignorance of the speaker her/himself, I shall speak of the VsX clauses marking unknown reality.

The most prominent context for this assertional structure are yes-no questions: the speaker treats the proposition as presupposed, and by focusing the polarity of the predicate signals that s/he is not familiar with the truth value of the proposition; the boundary tones (H L%) on the right periphery of the clause are in this context interpreted as an illocutionary marker, i.e. as a request for information.²

(6-52) Alb. (Kosovarja, p. 16)

Tash je fytyrë e njohur. A të PENGON kjo?

now you-are face the known QUEST you:CLIT oppress you

‘You are famous now. Does it bother you?’

(6-53) MG (Kapandai, p. 469)

To onoma tu itan Fileteros ... ðen to ksexase i Andiyoni - KSEXNIETE tetio onoma?

the name his was Ph. not it forgot the A. is-forgotten such name

‘His name was Philetairos ... Antigone has not forgotten it – can one forget such a name at all?’

Note that, in contrast to languages such as English, verb-subject order is not obligatory in yes-no questions in the Balkan languages: it is used only then when the speaker wants to keep the referential continuity, i.e. to use a ratified topic, and intends to mark the change of discourse

¹ In fact, Hungarian also has a comparable construction with a clause-initial verb and a specific intonational contour. In the Hungarian linguistic tradition, however, this construction is not interpreted as a carrier of polarity focus, but rather as a syntactically encoded aspect, which is called *existential aspect* by É. Kiss (1987: 72ff.) and Kenesei, Vago and Fenyvesi (1998: 306ff.), *experiential aspect* by Kiefer (1994: 420ff.).

² For a discussion of the assertional structure of questions, see Section 8.2., devoted to wh-inversion; for the function of boundary tones in questions, see 5.3. A similar analysis of accented verbs in yes-no questions in terms of polarity focus is proposed by Höhle (1992).

frame – in (6-52), from a scene-setting utterance to a request for information, in (6-53) from a narration to a general statement – i.e. in those contexts in which VsX construction is at home. If these conditions are not fulfilled, the subject is either not expressed, or it is, if non-ratified, placed preverbally. In Alb. and MG, VsX construction is very often used in yes-no questions: in Alb., out of 62 yes-no questions with overt subjects, 34 (54.8%) are VsX; in MG, out of 38 such sentences, 17 (44.7%) are VsX; in SC, they are rather infrequent (13 out of 66, i.e. 19.7%).¹

Mutatis mutandis, the analysis of VsX clauses as polarity focus structures marking the unknown reality of the proposition expressed may be applied to optative and imperative clauses like the following:

(6-54) SC (Kiš, p. 113)
ŽIVELI mladenci! ŽIVELA Crvena flota!
 live:OPT bride-and-bridegroom live:OPT red fleet
 ‘Long live the newly weds! Long live the Red fleet!’

(6-55) Alb. (Bishqemi, p. 67)
QESHNI ju, qeshni ...
 laugh:IMP you laugh:IMP
 ‘[Partin and Arbin cannot stop laughing.] »You two just laugh...«’

Finally, in expressing indifferent alternatives, SC, but not Alb. and MG, has a special construction formed with the conditional auxiliary in which VsX order is practically obligatory (Popović 1997: 149):

(6-56) SC (Vesti, p. 12)
BIO Milošević izabran ili ne...
 AUX:COND M. elected or not
 ‘No matter whether Milošević is elected or not’

In all four types (there are perhaps some more), the ignorance of the speaker is what triggers the polarity focus. Postverbal subjects are used when the conditions for the use of ratified

¹ The situation in SC is somewhat complicated by the existence of a question particle. Namely, the particle *li*, being a Wackernagel clitic, has to occupy the second position, with the verb preceding it. The verb-initial position of the verb is thus highly grammaticalized. However, since *li* practically does not exist in colloquial and in the ‘normal’ written language, being an archaism reserved for highly pathetic styles, I feel justified in ignoring it here (cp. Popović 1997: 147ff. for a sound philological analysis and Rivero 1993 for a generative interpretation). The Alb. question particle *a*, although very often used with verb-subject order, does not syntactically require it (cp. Buchholz and Fiedler 1987: 556).

topics are fulfilled. Note that languages diverge with respect to the use of VsX construction in yes-no questions, wishes, orders, and in expressing alternatives: MG and Alb. seem to employ the construction in all but the last context rather regularly, whereas SC seems to avoid it in questions, but uses it the remaining three.

(4) **Intensifying.** The last illocution-oriented discourse function of VsX clauses I am going to discuss is that of marking the degree of the fulfillment of the action described by the predicate to an extent which surpasses what is expected. This use is, both in my corpus and in informally collected data, attested only in SC. Furthermore, my Alb. and MG informants, when confronted with constructed examples of this kind, showed little enthusiasm. Although the failure of my elicitation technique is in no way the definite proof that VsX construction in these two languages cannot be used to intensify predications, I shall refrain from using unconfirmed constructed examples and confine myself only to SC.

(6-57) SC (Andrić, p. 112)

Vidim, ODULJILO se ovo moje tamnovanje...

I-see got-long REFL this my imprisonment

‘And so, this imprisonment of mine seemed to have no end...’

The nuance achieved by the use of the polarity focus variant of VsX construction is best described as ‘more than expected’: the whole proposition is presupposed, so that the focus on the verb can only be interpreted by an enrichment of its content. With predicates which denote gradable states of affairs, like ‘get long’ in (6-57), one logical enrichment is ‘more than expected’. This is how the intensifying interpretation comes into being. Of course, the subject has to be a ratified topic in the changed discourse frame for the use of a VsX clause to be felicitous.

A particularly interesting point is that verbs which in themselves signify a high grade of the fulfillment of an action often occur in this construction, marking the ‘more than expected’ nuance redundantly, both by the semantics of the verb itself and constructionally. Namely, in SC there is a highly productive mechanism of deriving imperfective verbs with the meaning ‘to do X a lot, up to and beyond saturation’ with the preverb *na-* combined with reflexivization: thus there is *najesti se* ‘eat enough’ from *jesti* ‘eat’, *naspavati se* ‘sleep enough’ from *spavati*, etc. These verbs are often found in intensifying VsX clauses, as in the following example from a conversation in which I myself took part. S. is baking a cake; B. is his girlfriend, notorious for being such a bad cook that she can’t make a sandwich for herself; D. (me) is a silent guest who is trying to be helpful:

- (6-58) S.: *Jel' imamo još praška za pecivo?* B: *Nemam pojma, pogledaj u orman.*
 QUEST have-we more powder for baking I-not-have notion, look in cupboard
 S: *Kvasca ima na tone, jer ga Boba manijakalno kupuje u želji da mesi hleb.*
 yeast it-has on tons since it B. maniacally buys in wish to knead bread
NAMESILA se Boba, ali praška za pecivo uopšte nemamo.
 a-lot-kneaded REFL B. but powder for baking at-all we-not-have
 'S.: »Have we got any more baking powder?« – B.: »I haven't got the faintest idea, check in the cupboard.« – S.: (checking in the cupboard) »We have tons of yeast, 'cause Boba buys it obsessively, planning to bake bread once. (looking at D., ironically) Our Boba's baked a whole lot of it, but we haven't got any baking powder at all.«' [07.06.2001]

To sum up: Polarity focus VsX clauses marking a specific illocutionary force of the utterance are all based on the notions of knowledge, certainty, and expectation. No matter which one of these cognitive phenomena is decisive in the given utterance, the change of perspective is so to say built in the polarity focus, by turning ignorance into knowledge or by surpassing expectations. The clauses used in the situations described above are responsible for the frequency of negated, 'old', and cognitive verbs in VsX construction (see 6.4.2.), as well as for the relatively high percentage of first and second person subjects, and pronominal subjects in general.

6.5.1.2. Polarity focus as a text-cohesion device

Polarity focus VsX clauses may, by creating or responding to expectations via contrast, be used to establish logical relations between sentences or paragraphs. Depending on the direction of the contrast, they may receive a concessive interpretation, if the contrast projects cataphorically, or an adversative one, if it is anaphoric.

- (1) **Concessive link.** As indicated above, when the contrast evoked by focusing the predicate is meant to evoke alternatives in the following discourse, VsX construction is understood as a concessive link, often paraphrasable with 'although':

(6-59) MG (Ciao, p. 24)

KRINI i kiria Papađopulu, ala ðen krini emena.

judges the Mrs. P. but not judges me

‘[Reacting to a critical statement by an actress about her colleagues, an actor says:] Mrs. Papadopoulou JUDGES people, but she does not judge me.’

(6-60) SC (Mandeljštam, I, p. 241)

ZNALA sam ja to, ali sam ipak bila spremna da bežim s njim

knew AUX I that but AUX nevertheless was ready to flee with him

‘[That miserable creature treated his girlfriends brutally.] I KNEW it, but I was still ready to run away with him...’

A very frequent type of concessive VsX clauses is formed with modal verbs meaning ‘can’:

(6-61) MG (*Elefterotipia*, p. 6)

BORI stin Turkia o kosmos na ine siniθismenos ston iθietero afto rolo ton

can in-the Turkey the world to be used to-the special that role of-the

enoplon θinameon, ala i Turkokiprii exun tis iθies andilipsis me tus Elinokiprius.

armed forces but Turkish-Cypriots have the same opinion with the Greek-Cypriots

‘It is perhaps true that people in Turkey are used to the special role of the Armed Forces, but the Cypriots of Turkish origin share the views of the Cypriots of Greek origin.’

These verbs are used as hedges: instead of only saying ‘I know that it is the case that X, but...’, the speaker says ‘It may be the case that X, but...’.

Whereas my MG corpus contains a wealth of examples for concessive VsX clauses, SC has only a couple, and Alb. only one. The speakers of MG tend to express concession using VsX clauses more often than the speakers of the other two languages, or, more cautiously, the rhetoric of the written language allows for this type of VsX more often in MG than in Alb. and SC, which rather use an explicit marking of concession, with subordinate clauses introduced by a concessive subordinator. It should be emphasized, however, that the concessive interpretation of polarity focus VsX clauses is possible, if infrequent, in these two languages as well.

It is interesting to note that the zero subject strategy is almost never used in concessive contexts, although subject topics are almost always referentially continuous, and therefore instances of ratified topics: the explicit marking of the change of perspective appears to be an important feature of such contexts.

(2) **Adversative link.** If the contrast evoked by polarity focus on the verb is a response to a proposition in the preceding discourse, the interpretation is adversative:

(6-62) MG (Kapandai, p. 118)

ΘΙΜΟΤΑΝ ομος ο Jorjis tis manas tu ta lojia ke kitakse ton θio tu ...

remembered though the G. of-the mother his the words and looked the uncle his
‘[Giorgis is told by his mother not to leave the room while his uncle is there; he
feels very uncomfortable and would like to go out.] However, Giorgis THOUGHT
of his mother’s words and looked at his uncle...’

The frequency relationships with adversative polarity focus VsX clauses are the same as in the case of concessive focus: some 23 examples in MG, one in SC, one in Alb. And again, it is rather the rhetorical structure of the written language than the structural differences that is responsible for this.

Let me now conclude the section on polarity focus VsX clauses as a text-cohesion device: Both types identified, the concessive and the adversative link, are based on contrast, which is an inherent feature of focus, especially of narrow focus (see 4.4.), of which polarity focus is a subtype. In both types, it is the change of perspective characteristic for concessive and adversative contexts, paired with the ratified status of subject topics, that triggers the use of VsX construction. Unlike illocution-oriented polarity focus clauses, which are highly productive in all three languages, the text-cohesion-oriented VsX types are a common discourse device only in MG, with Alb. and SC tending to resort to other kinds of constructions in concessive and adversative contexts.

6.5.2. Focus domain: broad and narrow foci

As has been explained in some detail in 4.5. and repeated in the introduction to 6.5., focus domain is a term marking (in the three languages under consideration) the part of the sentence stretching to the right of the verb and marking the potentially assertive material. The left border of the domain is the verb, the right border the non-verbal element carrying the nuclear accent. The domain – and this is of some relevance here – may be discontinuous, meaning that elements not being interpretable as assertive, i.e. as parts of the focus, may intervene between, say, the verb and the carrier of the nuclear accent, as in the following example: A: *What’s the matter with Peter?* B: *I* [[saw]_{Focus} *him* [YESTERDAY]_{Focus}]_{Focus Domain}, where the focus domain is *saw him YESTERDAY*, while the actual focal material, the utterance being an instance of a broad focus construal, is only ‘saw yesterday’.

It is a matter of context and intentions whether a focus domain will be interpreted with a broad or narrow construal. In those VsX clauses in my corpus which may be interpreted as having a focus domain (as opposed to polarity focus), the broad focus interpretation seems to be the prevalent one. Although the distinction between broad or narrow scope of assertion is not always clear-cut, so that one has to admit for a certain degree of arbitrariness, the numerical values look (approximately) as follows: in Alb., out of 56 focus domain clauses, the narrow focus construal is probable in 25 (44.6%); in MG, the ratio is 389 to 59 (15.2%), in SC 53 : 11 (20.6%). The reason for the relatively high percentage of narrow foci in Alb. will become patent in the course of this chapter. Let us now look at the discourse functions of focus domain VsX clauses more closely.

(1) **Digression in narration.** When a narrative text is told, the speaker may wish to interrupt the narrative chain by giving a descriptive statement about one of the participants or about the whole situation, or by offering an explanation for a state of affairs just evoked, or in order to identify one of the participants. This done, the narrative chain is resumed, and continued there where it was interrupted by a digression. VsX clauses are an ideal means to introduce a digression of this kind: The topic is ratified, being either a participant of the state of affairs described in the preceding narrative portion of the text or the state of affairs itself, and the scene changes, from the narrated world to the world of the omniscient narrator, from the event to the comment. Alb. sentence (6-37) (*ishte ky kontakti i parë i nuses me djalin...*, ‘was this contact the first of-the bride with boy-the’) is a nice example of a descriptive digression: an encounter of a young woman and a young man at her wedding is described; then (6-37) is inserted, as a comment of the narrator; then the description of the wedding is continued. A description of a participant is given in the following example:

(6-63) SC (Kiš, p.55)

*Avram Romanič ... se sav predao poslu ... **Bio je to čovek svog POSLA!***

A. R. REFL whole gave to-work was AUX it man of-his work

Ne samo što je napravio od mene pravog ... protojereja, nego mi je ... stavio i

not only that AUX made of me real archpriest but to-me AUX put also

lažan trbuh.

artificial belly

‘[A report about the preparations for a theater show] Avram Romanič was concentrated on his work ... It was a man of his trade! Not only did he make a real priest out of me, but he also ... put me an artificial belly.’

The judgment of the speaker – A. R. was a good craftsman – is inserted into a narrative text describing the course of events before a theater performance. An explanatory digression is exemplified by (6-64):

(6-64) MG (Kapandai, p. 44)

Ki otan sto ðeftero to xrono pali ðen epjase peði, ksafniastike o Kalomiras,
and when in-the second the year again not took child was-surprised the Lucky
ðen ixē SINIΘISI aftos ... pos ine boreto na tu ARNIETE i tixi tu kati.

not AUX was-accustomed he that is possible that him denies the luck his anything
[The leader of the rebels, ‘Lucky’ (Kalomiras) and his wife want a child, but seem to have no success; the years are passing by...] ‘And when the second year passed, and she did not get pregnant, Lucky was surprised – he had not yet experienced ... that his luck denied him anything.’

Descriptive, explanatory, and identificational digressions occur in all three languages with a similar, relatively high, frequency. The typical predicate in such clauses is the copula (description and identification being the proper semantic fields for nominal predicates), which partly explains the large share of copular sentences within the sum of VsX clauses. Pronominal subjects prevail, as witnessed by the examples, since the discourse referent referred to is usually mentioned in the immediately preceding clause. A very frequent use is made of NSPs, as in (6-37) and (6-64), when the inserted description is made with respect to a state of affairs. As shown in 6.4.1., this kind of pronoun cannot be left unexpressed in SC, unlike Alb. and MG, which is the reason why clauses like (6-64) are very frequent in this language, but less so in Alb. and, especially, in MG, where the zero strategy is very prominent.

At least in MG, this infrequent use of VsX clauses with NSP subjects is richly compensated for by a type of digression which is encoded by VsX only in this language. Namely, when a narrative chain is interrupted by a description of the emotional of one of the participants in the event, or by an identification of the source of the information on which the narration is based, it is only in MG that the VsX strategy is used, Alb. and SC usually employing the canonical subject-verb sentences. Consider the following example:

(6-65) MG (Kapandai, p. 120)

... tora pu tin alaksan tin isoðo 'jinan oyðonda enea ta skalopatia
now when it changed the entrance became eighty nine the stairs
ke ti θimate tora o Joryis tin KAPETANISA pos ta anevokatevene,

and her remembers now the G. the captain's-wife how them up-down-went

petuse i Afrodi ton peðikon tu xronon, ma tora ðen boruse pia

was-flying the A. of-the childhood- his years but now not could more

‘... now that the entrance has been changed the staircase has eighty-nine steps, and Giorgis remembers now how the lady captain used to go up and down the stairs, she was almost flying, the Aphrodi of his childhood, but now she could not walk any more...’

The shift in the temporal perspective and the continuous topic – ‘Giorgis’ is the main character of the story – seem to be the sufficient trigger for VsX in MG, but not so in Alb. and SC. The predicates frequently found are those of emotion and cognition, which in part accounts for their relative frequency in MG and for the lower percentage in Alb. and SC.

One special case of identificational digressions has to be mentioned yet. Alb. and, to a much lesser extent, MG, have developed a kind of cleft construction based on VsX: a copular sentence with a postverbal subject, either in the form of a pronoun modified by a relative clause, or in the form of a headless relative clause:

(6-66) Alb. (Kosovarja, p. 5)

Nuk jam UNË ai që kam të drejtë të thërras dikë për patriotizëm...

not am I he who have the right to shout something about patriotism

‘I am not the one who should dare say anything about patriotism...’

(6-67) Alb. (Kosovarja, p. 6)

Është KY që bëri më shumë për 143 bashkëqytetarët e tij në kazamatet serbe.

is he who did more much for 143 fellow-citizens the his in prisons Serbian

‘It is he who did most for 143 fellow-citizens imprisoned in Serbia.’

The former sentence is only a semi-cleft, whereas the latter seems to me to be a full-fledged cleft construction. Namely, the former is still an instance of VsX construction (narrow focus construal of the focus domain *nuk jam unë*, postverbal unaccented subject *ai që...*, the ratified status of the subject); the latter is further grammaticalized. As indicated above, both semi-clefts and real clefts are usually used to express exhaustive identification functioning as a digression in narration. MG does not use this kind of clefting very often, but it does exist in the language; in SC, nothing comparable could be traced down.

(2) **Left and right border of a paragraph.** The beginning and the end of a paragraph are the two places where a change of scene is most likely to occur: at the beginning, because a new paragraph is often separated from the previous by a new spatio-temporal frame, by a new

perspective, or both; at the end, because many paragraphs end with a sort of conclusion, which implies the shift of perspective.

In both cases, provided that the topical referent is not changed with respect to the preceding text, or at least that it is a major participant in the discourse, or inherently ratified in some way, VsX construction may be used. The most frequent type, attested in all three languages under consideration, are descriptive or identificational statements encoded as copular sentences, with all the characteristics mentioned in connection with VsX used in digressions (pronominal subjects, often NSPs, the copula, etc.). Consider the following examples:

(6-68) Alb. (Koha Ditore, p. 3/14)

Është vendi dhe RASTI kjo të themi se është për të ardhur keq që rezultatet

is place and occasion this to we-say that is for the future bad that results-the
e studiuesve tanë ... nuk i kemi bërë sa duhet të njohura në botë...

of-the scholars our not them we-have made as-much should the known in world

[In a speech devoted to the 80th birthday of the Albanian linguist Shaban Demiraj: ‘In the Balkan studies, ... Sh. Demiraj has succeeded in creating a more complete picture of linguistic Balkanisms from the point of view of Albanology.] This is the right place and the right occasion to say that it is bad for the future that we have not made the results of our scholars known in the world as much as they should be...’

(6-69) SC (Vijesnik, p. 3)

Bila je to kruna svega njegovog NASTOJANJA da se ... svijetu predoči

was AUX that crown of-all his effort that REFL to-world be-shown

što je jedan mali narod sposoban stvoriti ...

what is one small nation capable to-make

[In a funeral speech for the Croatian president Tudjman: ‘Dr. Tudjman returned from Rome tired, but proud of the exhibition which was a confirmation of the Croatian identity, which is more than thousand years old.] It was the culmination of his efforts ... to show the world what a small nation is able to create... [‘Dr. Tudjman returned from Rome, and it was his last journey...’]

The Alb. example is an instance of a paragraph-opening VsX clause; the SC one, for a paragraph-closing type. The former opens a new scene, from a laudation to Sh. Demiraj to the general problem of Albanian linguistics, and uses an inherently ratified topic, the present place and time, referred to with a NSP. The topic of the latter is the referentially continuous

‘exhibition’, referred to with a pronoun, and the perspective shifts from the narration of the president’s last journey to the estimation of the value of that journey. This is, in its turn, also the final clause in the paragraph; the one following returns once again to the events connected with the journey.

MG uses copular clauses in these two functions somewhat less frequently than Alb. and SC, partly because of its general tendency not to express subjects when no descriptive content is felt to be needed, partly because the latter strategy, that of closing a paragraph with a concluding remark in the form of a VsX copular clause, seems not to be employed at all in the contemporary written MG.¹

However, the paragraph-opening VsX clauses do exist in MG, both with copula and with other predicates. This is where the productivity of VsX construction in MG in comparison to Alb. and SC is clearly seen again: whereas in the latter the paragraph-opening VsX is virtually restricted to the copular clauses (only 2 examples with another verb in SC, 1 in Alb., all judged as extremely pathetic and archaic by the native speakers), MG freely uses all kinds of predicates in this context (out of 77 clauses with the opening function, 25 are with a copular verb, 20 with ‘to be’, 5 with ‘to become’, and 52 with other predicates):

(6-70) MG (Elefterotipia, p. 13)

Ipoqramise episis o ipuryos pos o OTE xriazete aksiopisto SIMAXO...

emphasized also the minister that the OTE needs reliable ally

[A report from the press conference at which the minister of economy, Mr. Papandoniou, explains why the government wants to keep its share in the national telecommunications company (OTE): ‘He mentioned the example of the National Bank, where the state has 0% of the actions, but still exercises some influence.]

The minister also emphasized that OTE needs a reliable partner...’

The subject is a ratified topic, ‘the minister’; the VsX construction is used to open a new paragraph in order to clearly mark the shift of interest in the minister’s speech, from the fact that the state can influence important companies even when it has no share in them, to the fact that the telecommunication company needs a reliable partner, i.e. the state. The predicate used is *ipqramizo*, ‘emphasize’, i.e. a non-copular verb. In Alb. and SC, a VsX clause with this verb in this context would be inappropriate, being reserved only for archaizing or pathetic registers. The same analysis can be applied to the following example, too:

¹ The strategy was very much alive, indeed quite prominent, in Ancient Greek – see Matić 2003.

(6-71) MG (Kapandai, p. 48)

Oso ke ðiskola na itan ta praymata ja tus xristianus, ama polemuse konda tus
how and difficult that were the things for the Christians when fought near them
o kapetan Kalomiras_i epernan ðaros, ðen lipopsixusan... Ke prospaθuse o Petros_i
the captain Lucky_i they-took courage not they-lost-hope and tried the Petros_i
mes ston polemo ton astamatito na vri XRONO ... na ksanapai sta kalivia ...
in-middle in-the war the incessant to find time to again-go to-the cottages
‘No matter how difficult the situation was for the Christians, when Captain Lucky_i
fought on their side, they drew courage from him and did not lose their spirits ...
Petros_i tried, in the middle of this incessant war, to find time to go to the cottages
once again...’

This passage illustrates the use of paragraph-opening VsX clauses in a narrative text: when the speaker wants to mark a spatio-temporal or a shift in the perspective (here from the general description of the war to the personal story of the main character) and open a new textual unit in this fashion, focus domain VsX clauses are the ideal means. Again, it is only in MG that this use is abundantly attested.

(3) **Resumption of an interrupted narrative thread.** One further function of VsX in texts, especially in narrative texts, is to resume an interrupted narrative thread. The indication of the continuity of the reference of the subject topic, paired with the indication of discontinuity of the discourse frame, makes them theoretically an ideal means for this. However, this discourse function is attested only in MG. Instead of VsX, clauses with preverbal, non-ratified subject topics are met in Alb. and SC.

Here is an example to illustrate the way VsX clauses resume interrupted narration: In a book on the Second World War in Greece, a story of the hungry inhabitants of a Greek provincial town is told; they try to steal some potatoes from a store-house; the occupation force, German soldiers, having heard of this, arrive promptly and force the Greeks to give the potatoes back. A brave woman, Mrs. Kanello, shouts at them: »You are going to pay for this! ‘Russians are coming’ – do you know that song? You will pay for everything!« After this quote, the scene suddenly changes: some thirty years after the war, two women, the brave one and the narrator, while drinking their coffee, angrily conclude that nobody has paid for anything. Then, the interrupted scene with potatoes, hungry Greeks and orderly Germans is resumed with the following sentence:

(6-72) MG (Matesis, p. 68-9)

Tu jirizi tis PATATES tu Jermanara i kiria Kanelo ke kinai kata to meros mas.

to-him returns the potatoes to-the German the Mrs. K. and moves to the part our

‘Mrs. Kanello gives the German the potatoes back and starts moving towards us.’

The effect of a VsX clause in this context reads as ‘back to the point’, as if the actual, main point of the discourse were returned to after a less important excursus. If subject-verb order were used, the effect of returning the camera to the scene previously shot would not be present.

This has to do with the ratified status of postverbal subject topics: if the speaker marks a topic as ratified, s/he instructs the hearer to search for its relevance in the previous text; if it is marked as non-ratified (which would be the case if the subject were preverbal), the instruction is simply to activate a previously irrelevant discourse referent. It is for this reason that in this use of VsX with broad focus the subject topics are almost exclusively those discourse referents which are the major participants over larger stretches of discourse, i.e. those whose relevance is not only local.

The fact that only MG, but not Alb. and SC, employs its VsX clauses as a resuming device explains the high percentage of major participants among the referents of the subjects of VsX clauses in MG, in contrast to the other two languages (cp. Section 6.4.1., column 7 in Table (6-34))¹.

(4) **Reactions.** The last discourse function to be described here is the major point of divergence between the three languages, being present, indeed extremely productive, in MG, and almost nonexistent in modern Alb. and SC.

This function has been identified by Matras (1995) in Romani, and by Sasse (1995a, 1996) in MG, who label it *connective* and tentatively propose its classification as ‘thetic’. Since all instances adduced by Sasse (the same holds for Matras’ Romani examples) and all the (numerous) instances in my corpus formally belong to the VsX type, and since the subjects in these clauses are unequivocally topical and ratified in that function, I suggest that this classification be revised: verb-subject order used in reactive contexts cannot be counted as ‘thetic’, however one chooses to define the latter term, but rather as an instance of VsX construction, which is the formal reflex of the existence of a postverbal topical slot.

Consider now the following examples:

¹ The resuming function of VsX clauses in MG is, apart from the reactive function (s. below), the only one which has been identified as such in the linguistic literature: Valioui (1994) calls it *retopicalization*.

(6-73) MG (Ciao, p. 46)

... ola orea ke kala, mexri pu i Ketula diapistose oti den tha kaθisi proto trapezi
all nice and good, till REL the K. noted that not will sit first table
pista, ala sto trapezi N° 46 ...Pikrathike i Ketula. Pije na skasi apo to kako tis...
stage but at-the table Nr. 46 was-embittered the K. went to burst of the bad her
‘... everything was OK, till Kaite understood that she would not sit at a table next
to the stage, but at the table Nr. 46... Our Kaite was embittered. She almost burst
in anger...’

(6-74) MG (Matesis, p. 82)

Stas Aθinas tus efere ena fortiyo, ke liyo protu bi stin poli, »Afiste mas edo«,
in-the A. them brought a truck and little before enters in-the town leave us here
tu lei. ... Tus afise o fortigatzis s' ena VUNALAKI erimo tote...
him says them:CLIT left the truck-driver in one little-hill desolate then
‘They came to Athens in a truck, and little before it entered the town, she said
»You can leave us out here«. ... The truck driver left them out on a little hill which
was at that time still uninhabited...’

In both cases, it is the reaction of a ratified topical discourse referent to an action performed by another participant that is described with a VsX clause: Kaite reacts to the fact that her host does not give her the table she wants, the truck driver to the order given to him by a passenger. Let us call this kind of reactive VsX clauses *interactive*. What they achieve in discourse is an explicit marking of the fact that action A is a reaction to action B, due to the two fundamental properties of VsX: topical continuity and the discontinuity of the perspective. The speaker signals that the set of topical referents (the host and Kaite, the passenger and the truck driver) remains the same, but that the camera moves from one referent to the other.

Now, in interactive contexts, at least one more strategy is available, that with preverbal subject topics; thus, sentence (6-74) could be encoded as *o fortigatzis tus afise...* (‘truck driver left them out...’). The effect would be slightly different: it is not the reactivity that is emphasized, but rather the topic shift, the preverbal position being reserved for non-ratified topics. Whereas in the case a of VsX clause the message is ‘look, we have the same set of discourse referents, and I turn your attention from one to the other’, a subject-verb clause reads as ‘turn your attention to one or another discourse referent; it is of no relevance whether they remain the same through this portion of discourse or not’. The former strategy concentrates on interactivity, the latter on the multitude of referents.

Interactive VsX clauses very often contain verbs of emotion, as illustrated by (6-73), or verbs which are somehow ‘given’ (6-74), the former because human reactions are often emotional, the latter because interactivity often consists in reacting to demands, or in performing the same or similar actions.

Interactive contexts are, however, not the only ones in which reactive VsX clauses are found in MG. They are often employed when the speaker wishes to emphasize that action A is a consequence of action B, without being explicitly interactive:

(6-75) MG (Kapandai, p. 13)

(o Günek) ... epano sta meθisia tu, tin esfakse. ðen perasan tris meres ...

(the G.) on in-the drunkenness his, her he-slaughtered not passed three days

ke XAΘIKE o Günek. ...ðen vreθike puθena, mazi tu omos xaθikan ki

and was-lost the G. not was-found anywhere together him though were-lost also

o aravonjastikos tis kopelas ... ki i mana tu. PJASANE i Turki ton patera tis

the fiancé of-the girl and the mother his took the Turks the father of-the

kopelas, ma ekinos ðen martirise...

girl but he not testified

‘[The Turkish governor of a Greek province, Günek] slaughtered the girl in his drunkenness. After only three days, Günek disappeared. ... No-one could find him, but at the same time the girl’s fiancé and his mother ... disappeared, too. The Turks imprisoned the girl’s father, but he would not testify...’

Günek’s disappearance and the imprisonment of the girl’s father by the Turks are presented as consequences of the previous actions – the former of Günek’s slaughtering a girl, the latter of the suspicions raised by Günek’s disappearance. This kind of reactive VsX clauses will be called *consecutive* henceforth. What is of interesting is that consecutive VsX clauses may be used even when the cause-consequence relationship is not given, or is only to be construed *because* VsX constructions is used. It suffices that the speaker wants to emphasize that action A temporally follows action B, as in the following example:

(6-76) MG (Kapandai, p. 147)

To ’ðenan i ðiki mas me sxinja,... ke to travusan na to anevasun....

it:CLIT they-tied the own our with ropes and it:CLIT drew to it:CLIT raise...

‘[A description of a boar-hunt performed by the Greeks and the Turks in Asia Minor; ‘when the boar was killed,] our people tied it up ... and dragged it away in order to raise it (on a horse...)’

The killing of a boar and the subsequent action of tying it up are not in any direct causal relation, but only follow each other; the consecutive VsX clause *to 'đenan i điki mas me sxinja* simply marks the latter action as (immediately) following the former.

Of course, in consecutive contexts, with or without a causal nuance, subject-verb order may be used as well, if the speaker for some reason decides not to mark the consecutive nature of events narrated.

Now, this is a point where VsX construction becomes applicable in practically all contexts in which events are narrated in their natural sequence: its use is justified by the simple fact that A follows B. With such an extended use, VsX construction becomes a serious rival to the otherwise almost universally applicable subject-verb order: speakers practically have a free choice between VsX and SV, with the former putting some accent on the sequentiality, the latter on the topic shift, the difference being only that of shade, not of color.

What the consequences of this extended discourse role of VsX are, or may be, will be scrutinized in some detail in Sections 6.6. and 6.7. Let us first cast a look at the situation in Alb. and SC and compare it to MG.

The number of reactive VsX clauses in the MG corpus is large: 68 interactive and 123 consecutive clauses, in total 191 (34.3% of all VsX clauses). In contrast, the Alb. corpus contains only one example (0.8%) which can be plausibly described as interactive or consecutive, the SC corpus four (3.4%). For the latter, it must be noted that all four instances stem from one source, Andrić's novel, which is written in a slightly archaizing style (cp. 5.4.). It is precisely the archaic, folklore-like contexts that reactive VsX appears in these two languages; in modern colloquial variants, in modern literature and journalism, it is not a live option for the speakers of Alb. and SC.

6.5.3. Discourse functions of VsX construction: Summary and conclusion

In the final section of this chapter the results will be summarized and the languages compared with respect to the productivity of their VsX constructions. The data are presented in Table (6-77) ('+' reads as 'frequent in the language', '+/-' as 'not especially frequent, but not unusual', '-/+' as 'infrequent, but possible', and '-' as 'nonexistent'):

(6-77) *Discourse functions of VsX construction*

	MG	Alb.	SC
1. polarity focus. confirmation/denial	+	+	+
2. polarity focus. promise /threat	+	+	+
3. polarity focus. questions, wishes, etc.	+	+	+/-
4. polarity focus. Intensifying	?	?	+
5. polarity focus. Concession	+	-/+	-/+
6. polarity focus. Adversativity	+	-/+	-/+
7. focus domain. digression: the copula	+/-	+	+
8. focus domain. paragraph-opening: the copula	+/-	+	+
9. focus domain. paragraph-closing: the copula	-	+	+
10. focus domain. Cleft	+/-	+	-
11. focus domain. digression: other verbs	+	-	-
12. focus domain. paragraph-opening: other verbs	+	-	-
13. focus domain: resumption of narration	+	-	-
14. focus domain: reactive interactive	+	-	-
15. focus domain: reactive consecutive	+	-	-

The whole complicated network of discourse functions may be reduced to a couple of regularities.

In Alb. and SC, the use of VsX construction is, apart from some minor differences, subject to practically identical restrictions:

- (1) VsX with polarity focus is frequently employed in illocution-oriented contexts.
- (2) VsX with broad or narrow focus is frequently employed only when the predicate is the copula, notably in digressions, on the borders of paragraphs, and, in Alb., in clefts.
- (3) VsX with polarity focus is seldom used in text-cohesion oriented contexts.
- (4) VsX with broad or narrow focus is not used when the predicate is a verb other than copula.

In MG, the regularities are somewhat different:

- (1) VsX with polarity focus is frequently employed in both illocution- and text-cohesion-oriented contexts.
- (2) VsX with broad or narrow focus is less frequently employed with the copula.
- (3) VsX with broad or narrow focus is frequently employed in all contexts, especially with verbs other than the copula.

The general impression one gets from these facts is that VsX in Alb. in SC is a moribund construction, surviving only in some semi-petrified predicate-subject combinations and in rather formalized illocutionary contexts, like promises and threats. Since the latter appear with a greater variety of predicates, the polarity focus VsX clauses seem to be much more alive

than the focus domain type, which occurs with some regularity only with the copula combined with pronominal subjects. Some combinations, like COPULA+NSP (*është kjo, bio je to*, ‘was it’) seem to be almost understood as ready-made phrases, so that, in most registers, one relatively rarely encounters the subject-verb order (*kjo është, to je bio*, but see 6.6.). Furthermore, the contexts in which VsX clauses occur in these two languages are very restricted, especially in comparison to MG. All this leads me to the conclusion that postverbal ratified subject topics are a species that is retreating from modern Alb. and modern SC. The fact that the occasional use of predicates other than copula with the broad focus structure in both languages has the status of a learned construction, or of some kind of folklore imitation, is an indication of a previously more prominent status of the construction. I shall return to the diachronic side of the problem briefly.

In MG, VsX is everything but moribund. It covers a wide range of discourse functions, both in its polarity focus variant and when construed with a focus domain. It freely occurs with all predicates and all kinds of subjects. What is even more important, with the reactive use of broad and narrow focus VsX clauses, the construction has reached the level of almost general applicability, to the extent that many native and non-native linguists (most notably Philippaki-Warbuton 1985 and, with some provisos, Horrocks 1994) put forward the claim that the ‘basic word order’ in MG is indeed verb-subject. However, my data cannot confirm this claim: as the statistics presented in 5.5. reveals, subject-verb clauses are significantly more numerous than the verb-subject ones, be it VsX or any other construction, in almost all syntactic contexts and in all registers. If one does not presuppose the existence of a deep structure, where the underlying VsX is transformed into the surface sV, as I do not presuppose, there is no sufficient empirical confirmation for the ‘basic’ status of VsX in MG. The fact that a construction with such a general meaning as the one reached by VsX in its reactive use does not have a much greater frequency is, as I shall try to show below, at least partly reducible to sociolinguistic and diachronic factors.

6.6. Non-structural factors influencing the use of VsX

The last set of features which determine the use of VsX are the sociolinguistic and diachronic ones, which often represent one and the same thing. As I have repeatedly emphasized, as far as the status of VsX is concerned, Alb. and SC side against MG; therefore, the former will be dealt with together, separately from the latter.

6.6.1. Albanian and Serbo-Croat

VsX is in Alb. and SC a construction with a strongly reduced productivity, as indicated by its being restricted to only certain types of subjects and predicates, and to only few discourse functions. Although other lexical fillings and other discourse functions are not excluded, they all sound either very archaic or very folklore-like, and are accordingly extremely rare. Taken together, these two facts point to the construction being on the way of dying out.

The process of the ousting of VsX has taken place in the last two centuries, in the case of Alb. perhaps only in the last century. The texts from the nineteenth and, in Alb., from the early twentieth century, reveal both a greater variety of subjects and predicates employed in VsX, and a greater variety of functions the construction performs, with a concomitant higher overall frequency. I have checked this on a small corpus of oral narrative texts, fairy tales collected by ethnographers and linguists in the 19th century. The collections used are *Mbledhës të hershëm të folklorit shqiptar III*, with texts first published 1908, for Alb., and *Antologija narodnih pripovedaka*, containing texts first published 1852, for SC. Here are the results: out of 1000 clauses in Alb. fairy tales, 42 are VsX clauses, i.e. 4.2%; in SC, the ratio is 1000 : 38, i.e. 3.8%; if only clauses with overt subjects are taken into account, the percentage of VsX clauses amounts to 6.9% (612 : 42) in Alb, and to 5.7% (672 : 38) in SC. These data are compared to those of the modern languages (see (6-5)) in the following table; the first column (VsX vs. SV/VS) refers to the share of VsX clauses within the class of clauses with overt subject, the second to the its share in the whole corpus, including zero-subject clauses:

(6-78) *VsX in the 19th century Alb. and SC texts compared to the modern language*

	VsX vs. SV/VS		VsX vs. Σ	
	old texts	modern texts	old texts	modern texts
Alb.	6.9%	1.5%	4.2%	0.8%
SC	5.7%	1.4%	3.8%	0.8%

It is a telling comparison: in both Alb. and SC, the frequency of VsX clauses has significantly decreased, within one or one and a half century, both with respect to clauses with overt subjects and in general. Of course, I am aware that my modern corpora are much larger and more variegated, whereas the old corpora are restricted to one, narrative, text type, and one register; it cannot therefore be excluded that the ratio of VsX in the nineteenth century Alb. and SC would be somewhat lower if procedural, expository, and conversational texts were taken into account as well. A comparison of old narrative sources with the percentages of

VsX in the narrative modern texts only (Table (6-6)) may mitigate this shortcoming: even in those modern texts which display the highest ratio of VsX (Bishqemi for Alb., Andrić and *Vjesnik* for SC), the percentage within the class of clauses with overt subjects does not exceed 2.5%, which is still significantly less than 6.9% and 5.7% in the old narrative texts.

This picture is further confirmed by the lexical structure of the construction in old texts. Unlike modern VsX clauses, the old ones have a pronominal subject in only about a third of all cases, 35.7% (15 out of 42) in Alb. and 31.5% (12 out of 38) in SC, which is significantly less than 51.7% and 59.5% in modern texts (Table (6-33)). The same holds for the predicates: the copula appears in only about 15% of the cases (14.8% in Alb., 15.6% in SC), whereas in modern sources it covers well more than a third of all instances (35.0% in Alb., 36.2% in SC). All this implies that the construction used to be more productive, i.e. applicable with a greater variety of subjects and predicates.

Finally, most of the discourse functions which are either absent or awkward in the modern languages are very much alive in the 19th century fairy tales. Most important, the reactive function of VsX, which is, due to its generality, partly responsible for the frequency of VsX in MG, is well attested, in contrast to the modern languages, where it practically does not exist. Here are two examples:

(6-79) Alb. (*Mbledhës të hershëm*, p. 240)

... asnjëj nuk iu hap goja, të thotë mirë dhe të nxjerë groshnë të apë,

to-nobody not him-itself opened mouth to say good and to draw groschen to give

po të gjithë mbenë në pushim. Pa udhëtari që asnjëri nuk nxjer pesëshin....

but the all stayed in rest saw passenger that nobody not draws cent

[A passenger suggests to the people he met to give a little money each, so that a dead man lying on the road be buried.] ‘Nobody opened their mouth to say ‘OK’ and to give a couple of dimes, but everybody remained silent. The passenger saw that nobody would give a cent, [so he himself gave 50 dimes].’

(6-80) SC (*Antologija*, p. 270)

Kad sin to čuje, ... otide tamo, a kad dodje pred cara, pita ga car: ...

when son it hears goes there and when comes in-front emperor asks him emperor

‘When the son heard it, ... he went there, and when he arrived to the emperor, the emperor asked him:....’

The Alb. example nicely illustrates how a new paragraph may be opened with a verb other than the copula (the perspective changes from the world in which the suggestion is given and implicitly refused, to the inner world of the passenger), a discourse strategy not existing in

modern Alb. The VsX clause in the SC example is a prototypical reactive clause: the emperor reacts to the appearance of the young man by asking him a question, or alternatively, the narrator only wants to emphasize that the action of asking immediately follows the action of arriving. This function, well attested in modern MG, is completely absent from modern SC.

Even though the definite proof of the claim that Alb. and SC used to be closer to MG with respect to the productivity and frequency of VsX construction would require a much more detailed diachronic exploration, with a larger corpus and more diversified text types from the earlier stages of the two languages, the facts adduced so far at least make this claim plausible. If this is indeed so, what are the reasons for the relatively quick decline of VsX in these two languages? There are probably many factors, but one seems to me to be particularly prominent: the tendency to ‘westernize’ the language, i.e. to form the literary, and subsequently also the colloquial language of the educated and less educated according to the patterns offered by French, English and German, in the Alb. case perhaps also Italian, a tendency which is characteristic of all Balkan languages, most conspicuously with respect to the construction I labeled Inversion (see Chapter 10), but probably also to a certain degree with respect to VsX construction. French, English, etc. have nothing comparable to VsX, so that it is no wonder that the educated avoided it both in speaking and writing for generations. The result is that it is today, apart from some more or less petrified uses, practically only a stylistic ornament in certain types of discourse, subject-verb order being used instead.

In fact, the process of ousting of VsX from the ‘modern’ way of speaking seems to continue even today. As the statistics in (6-6) shows, representatives of journalistic, i.e. official, style, as well as writers creating highly intellectual, urban prose, tend to use VsX very seldom. This practically means that they often use alternative constructions, notably subject-verb order, even in those contexts in which VsX is still common. A good example is Kadare’s novel *Kronikë në gur*, where the combination COPULA+NSP, which is, as indicated in 6.5., almost always expressed by a VsX clause in everyday Alb., is found as VsX only two times, the normal form being NSP+COPULA, i.e. subject-verb, as in the following example:

(6-81) Alb. (Kadare, p. 6)

... **kjo ishte** një gjë e natyrshme, përderisa **ky ishte** një qytet prej guri...

this was a thing the natural, because this was a town from stone

‘[... the town involuntarily caused much pain and wounds;] this was a natural thing, since it was a town made of stone...’

What are the consequences of these non-structural factors for the structure of Alb. and SC sentence? Both languages are perhaps on the way of losing the postverbal structural position

designed for ratified topics. In this fashion, the possibility to encode the difference between non-ratified and ratified topics within a changed discourse frame, already restricted, could completely vanish. The most important corollary to this is the reinterpretation of the preverbal topic slot, at least as far as subject topics are concerned: what once was a position where only non-ratified topical subjects were allowed to land is turning into the default position for all types of topical subjects.

6.6.2. Modern Greek

Unlike Alb. and SC, VsX construction in MG is quite productive and has reached a rather high degree of generality of meaning. This, however, does not imply that it is not subject to sociolinguistic variation. On the contrary, as the statistics presented in (6-6) shows, this kind of variation not only exists, but it runs precisely along the same lines as in Alb. and SC: the more official and ‘modern’ the style, the less VsX; the more traditional and colloquial the style, the more VsX. There are, however, two important differences with respect to Alb. and SC: first, in all text types and registers, VsX is significantly more frequent than in their Alb. and SC counterparts; second, all kinds of VsX represent live options in everyday communication, and not only certain types of polarity focus, as in Alb. and SC. In other words: MG is subject to the same influences of the Western European languages as Alb. and SC, but since the construction is productive in the spoken register, it cannot be suppressed enough in the more official and ‘westernized’ registers, so that its frequency surpasses the one observed in the other two languages.

I cannot help mentioning a conversation with a party neighbour in Athens, who told me about the storm in which she found herself while travelling with a ferry-boat from an Aegean island to Athens. The dramatic situation depicted implied the emphasis on reactivity and temporal sequencing, so that I had the impression that the whole report consisted only of VsX clauses. Here is one I wrote down:

(6-82) *Po po! Foviθike o kosmos, trexane jinekes san treles...*

oh oh feared the world ran women like mad

‘[Then a crash was heard.] Oh no! People got scared, women started running around like mad...’

This kind of oral narration cannot be found in modern Alb. and modern SC. Furthermore, many MG written registers, among them the traditional narrative, still very popular in the Greek culture, consciously use VsX to ‘enliven the narration’ (*na zoirepsun tin afijisi*, as a

Greek native speaker told me), the impression of lively narration having to do with the implication of reactivity often connected with the construction. A very good example of this are the short stories by Kapandai, who, in an attempt to enliven her otherwise rather monotonous narrative style, uses VsX to the point of oversaturation.

Now, all this means that VsX in MG is, despite conscious or subconscious efforts of the intelligentsia, a productive pattern, but still a ‘marked’ one, i.e. one which is felt to be well-suited for lively narratives depicting dramatic events, but too clearly marked for reactivity and sequentiality to be applicable in less tense situations. This perhaps explains the fact that, although relatively nonspecific in meaning, it is not the ‘basic’ word order, judged by both frequency and markedness criteria.

Why VsX in MG is not dying out as in Alb. and SC is a question I cannot answer, at least not definitely. A possible reason is the gap that existed in Greece between the official language and the spoken language for almost two centuries, till the middle of the seventies: the proclaimed standard language, *katharevusa*, was too distinct from the colloquial variant, *dimotiki*, to be able to influence it. This diglossy perhaps contributed to the better preservation of VsX than is the case in Alb. and SC, in which the standard language was always based on the vernacular.¹

The structural consequence of the productivity of VsX is the fact that the postverbal topic slot in MG is a more prominent part of the sentence structure than is the case in Alb. and SC. Accordingly, the preverbal topical slot is more clearly marked for non-ratified topics. MG syntax displays thus a higher degree of discourse-pragmatic orientation in encoding topical subjects than Alb. and SC, in which the preverbal slot has practically lost the feature [non-ratified] when the topic is expressed by the subject.

6.7. VsX: Conclusion

The discussion of VsX construction has brought to light some insights located beyond the narrow array of the construction itself, concerning both the inner structure of the languages under consideration and the general ways in which discourse-pragmatically marked structures function.

¹ The presentation of sociolinguistic factors given in 6.6. is merely a sketch; for more detail, and for more conclusive argumentation (however, with respect to another VS construction, Inversion) see 9.3.2.1.3.

Syntactically, Alb. and MG side together, whereas SC displays a slightly divergent structure; in all other respects, semantically, informationally and sociolinguistically, it is Alb. and SC that are similar, while MG strongly diverges.

First the syntax. The facts presented in 6.3.2. and 6.6. force us to modify the simplified sentence pattern proposed for the Balkan languages in (5-1), both with respect to the positional regularities, notably those of ratified topics, and to the discourse-pragmatic features associated with certain sentence slots, notably those of the preverbal topical slot:

(6-83) *Modified sentence templates of the Balkan languages*

Alb.: [Non-Rat.Top./Subject] [Nar.Foc.] [[Verb] [Rat.Top.] X]_{Focus Domain} [Rat.Top.]

MG: [Non-Rat.Top.] [Nar.Foc.] [[Verb] [Rat.Top.] X]_{Focus Domain} [Rat.Top.]

SC: [Non-Rat.Top./Subject] [Nar.Foc.] [[Verb] [Rat.Top.] X]_{Focus Domain}

The templates reveal a number of things: First, Alb. and MG, but not SC (or only marginally), have two postverbal slots for ratified topics, directly after the verb and clause-finally. Further, Alb. and SC, but not MG, have partly reinterpreted the preverbal slot as the default position of topical subjects, ratified or non-ratified.

What is identical is the central part of the template, the focus domain: in all three languages, it projects to the right, with the element X (i.e., the nonverbal part of the focus domain) carrying the sentence stress; the postverbal ratified topic slot makes the focus domain potentially discontinuous, in so far as non-focal material may intervene between the parts of the focus domain (see Lambrecht 1994 and Matic' 2003 for more on discontinuous focus domains).

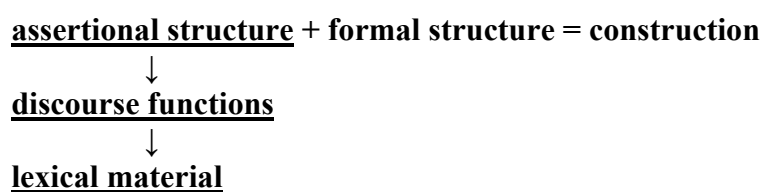
In all other respects apart from syntax, there is a clear line that divides Alb. and SC on one hand and MG on the other, all of them reducible to the fact that the former are in the process of losing VsX construction, while the latter employs it freely and relatively frequently. I have tried to show that the decay of VsX in Alb. and SC is, at least in part, triggered by sociolinguistic factors, and that the long-lasting diglossy in Greece may have contributed to the survival of the construction in MG. Discourse-pragmatically marked constructions are thus shown to be susceptible to nonstructural variation, a fact which will become even more patent in the following chapters, devoted to inversion. Before turning to inversion, however, two issues of more general relevance should be pointed out.

The analysis of VsX has revealed that the assertional structure of constructions strongly correlates with their discourse functions, and with the semantic and informational properties of the elements of the constructions. The combination of features inherent to VsX, the

referential continuity of the topic and the change of the discourse frame, makes it compatible with polarity focus contexts, and with various transitional and reactive contexts (digressions, left and right borders of paragraphs, resumptions, etc.). The prototypical informational and semantic properties of the subjects and the verbs in VsX are derivable from these discourse functions: The subjects tend to be either ‘old’ and ‘given’ or of inherent relevance to the course of communication because they have to be ratified; this is also the reason pronominal subjects are so often found. The frequency of epithet subjects is explained by the inherently transitional nature of the clauses containing them. In polarity focus contexts, verbs are often ‘given’ or ‘derivable’ from the scene, or negated, or ‘weak’ (especially as verbs of existence and appearance). This has to do with the assertional structure of polarity focus: for the descriptive content of focused predicate to be unworthy of asserting, so that the hearer has to look for the assertion in the polarity of the proposition, in its TAM features, or similar, the predicate must be either extremely general in meaning, or already asserted (‘old’). The high correlation between negation and VsX polarity focus is triggered by the illocutionary force of polarity focus, which is often used to deny or correct the assumed beliefs of the interlocutor(s). In the focus domain type of VsX, the most frequent verbs are the copula, modal verbs, and verbs of cognition and emotion. The copula (with a nominal predicative) is so prominent because most digressions and paragraph openings are given in the form of descriptive or identificational statements; verbs of cognition also occur in these functions, especially in digressions, since they aptly express the source of knowledge. The other two verb classes are mostly found in reactive contexts, since changes of emotional states are the typical way of reacting to events.

What I would like to conclude in view of these facts is that there is a kind of hierarchy of factors which determine the identity of a construction. On the highest level, a certain assertional structure is mapped onto a certain formal structure, which is the basis of every construction. The assertional structure determines which discourse functions the construction can perform; the discourse functions, in their turn, determine the preferred lexical and informational filling. Schematically:

(6-84) *Hierarchical structure of constructions*



Constructions, at least those which have a discourse-pragmatic loading, thus appear to be a bundle of lexical, discourse-oriented, and assertional properties united in one formal expression. If a construction works only according to the scheme (6-84), it may be considered fully productive and fully transparent, with the basic meaning, a certain assertional structure, being regularly expressed in certain contexts with certain lexical material. VsX in MG is closest to this status. However, as the description of Alb. and SC, and, to a much lesser extent, of MG data has shown, lexical items and/or discourse properties primarily triggered by the assertional structure may become partly independent from it, i.e. they may become independently associated with the formal structure of the construction, rendering the following picture:

(6-85) *Further development of constructions*

$$\text{assertional structure} + \text{formal structure} = \text{construction}$$

$$\begin{array}{ccc} & \uparrow & \uparrow \\ & \text{discourse functions} & \text{lexical material} \end{array}$$

When the development depicted in (6-85) is completed, the construction is not a productive construction anymore, but at best an idiom, with certain lexical items or certain discourse contexts invariably being expressed in a certain way, regardless of the assertional structure conveyed¹. In none of the three languages has this stage been reached with respect to VsX; the combination COPULA+NSP in Alb. and SC is perhaps close to it, but even it, as (6-81) shows, still allows for alternative ways of expression.

My point is that VsX in Alb., MG and SC, like most most other constructions in natural languages, represents a mixture of (6-84) and (6-85): it is in part a productive pattern allowing for a theoretically infinite number of lexical fillings and discourse functions, although with clear statistical preferences, and in part a petrified or semi-petrified structure, with certain lexical items and certain discourse functions being almost automatically expressed by VsX, which, in turn, excludes all other lexical items and discourse functions. The difference between the languages lies in the point of the scale between (6-84) and (6-85) occupied by a language: MG is much closer to (6-84), Alb. and SC to (6-85), though none of them represents a pure type. This picture will be confirmed in the present study by the investigation of the two other VS constructions, inversion and vS.

The second general issue raised by the scrutiny of VsX in Alb., MG and SC is a typological one. In Section 2.3., an approach to verb-subject order has been presented which

¹ That lexical restrictions are a reliable sign of the loss of productivity of constructions is nicely demonstrated on the example of preverbal objects in the history of English by Koopman and Van Der Wurff (2000).

analyzes it as a narrative device encoding topic continuity (Givón), or temporal sequentiality and foregrounded information (Hopper), whereas subject-verb order is said to have the opposite effect. Givón, Hopper, and their adherents (see 2.3. for references), consider this to be a universal in all languages displaying SV/VS alternation. Myhill (see 2.3.) claims that the kind of VS Givón and Hopper have in mind is indeed only one subtype of VS, incidentally the dominant one in the languages on which they base their theories, notably Semitic, Malay, Old Germanic, etc. He also contends that this, ‘narrative’, type of VS represents a prominent construction only in those languages in which the ‘basic’ word order is verb-subject, whereas it is only a minor, or even non-existent, phenomenon in languages which are basically SV.

How does the Balkan VsX construction, which for a greater part fits the descriptions given by Givón, Hopper and others, fit into the story of the VS/SV alternation as a universal marker of continuity/discontinuity? First, it obviously confirms Myhill’s conclusion: verb-subject order which encodes temporal sequentiality, etc.,¹ is indeed only *one* of the constructions in which VS is used, the others, as we shall see, having nothing to do with the notion of continuity. As far as the other part of Myhill’s conclusion is concerned, namely that the importance of VsX is directly proportional to the overall frequency of VS in a language (which is only the other way of saying that it is because of the frequency of VsX that a language becomes a VS language in the first place), the generality of meaning arrived at by VsX in MG (and probably in the earlier stages of Alb. and SC) may be considered to be a partial confirmation of the claim.

I also hope to have shown that discourse notions like foregrounding, temporal sequentiality, even topic continuity², claimed or at least felt to be unanalyzable primitives in the hard-core functionalism as represented by the above mentioned scholars, are *derived* notions, i.e. that they are by-products of the basic function of the construction, namely to encode ratified topics within changing discourse frames: foregrounding and temporal sequentiality are the features of one of the most prominent uses of VsX, the reactive use, whereas topic continuity is one of the major *manifestations* of the ratified status of topics (not the only one, cp. 6.2.), i.e. not a basic notion in itself but merely a way in which the basic notion of ratifiedness appears in natural discourse. In other words, I hope to have shown that

¹ For the sake of brevity, the abbreviation used in the rest of the present study, VsX, will be used to denote what Givón, Hopper, Myhill, and others call VS encoding topic continuity, foregrounded information, etc.

² As a matter of fact, VsX construction is not necessarily associated with foregrounded information and temporal sequentiality, at least not in the three languages under consideration: many uses of VsX imply backgrounded, additional information, often excluded from the narrative chain.

it is the pairing of a cluster of certain formal features with a certain assertional structure that is basic in discourse-pragmatically marked constructions, with discourse functions, lexical properties, etc. being the logical consequences of the applicability of a certain assertional structure in discourse. This implies that there indeed exists a thing called grammar, i.e. a conventionalized pairing of forms and meanings, and not only a cluster of half-grammaticalized direct responses of the speakers to the needs of the actual discourse, which is the underlying assumption of much of the research done in the functionalist paradigm. It is indeed the case that grammar is repeatedly confronted with fossilization processes, whereby the earlier statistical tendencies are understood as rules, so that the basic form-meaning pairing loses some of its transparency, as shown on the example of VsX in Alb. and SC. This, however, still does not mean that there is no grammar at all, but merely that it is not such a magnificent machine, producing without error only one-to-one form-meaning pairs, as many formal linguists would like it be.

7. Inversion: Introduction

The second type of verb-subject order occurring in the languages of the Balkans is what I, borrowing the terminological convention from Sasse (1995), call **inversion**. Note that for many linguists ‘inversion’ is synonymous to ‘verb-subject order’; in the present study, the former is a hyponym of the latter.

Unlike both VsX construction and vS construction, inversion is a very diversified phenomenon, with two main subtypes, **focal inversion** and **subordinate inversion**, both further comprising a number of subclasses, viz. **quotation inversion**, **wh-inversion** and **fronted focus inversion**, which belong to the focal inversion, and **relative inversion** and **embedded inversion**, which are subclasses of subordinate inversion. For this reason, the chapters on inversion are both longer and more numerous than is the case with the other two verb-subject constructions: each of the subclasses has specific features of its own, which have to be described in some detail.

There are two basic features of inversion: (a) It appears in presuppositional contexts; (b) Inversion clauses always contain a clause-initial element of some kind. The term *inversion* is chosen to clearly demarcate one of the principal properties of the construction, stemming from the feature (b): in many cases (though not all), verb-subject order is triggered by a formal constraint, requiring the verb and the clause-initial element to be immediately adjacent (henceforth **adjacency constraint**), so that the only place left for the subject is the one after the verb. It is in this sense that the subject is ‘inverted’, i.e. somehow moved from its proper position. However, as I shall try to show, the verb-subject order occurring in presuppositional contexts also has a functional load of its own. Even worse, as in the case of VsX, sociolinguistic factors play a certain role in the use of inversion as well. Inversion is thus a mixture of formal, functional, and sociolinguistic factors, which, by the way, makes it a rather difficult object of investigation.

In the following chapters it will become clear that the functional motivation of inversion is more or less identical in all three languages under consideration. The field of variation is primarily that of formal motivation: MG has a rather strong adjacency constraint in most types of inversion; so does Alb., although in a reduced number of contexts; in SC, there is practically no adjacency constraint apart from quotation inversion. Furthermore, SC, but not MG and Alb., has a special rule which practically forbids the inversion of pronominal

subjects. As a result, one would expect MG to have most inversion clauses, SC least. The expectations are only partly fulfilled:

(7-1) *Inversion clauses in the corpus: absolute values*

	original texts	translation	Σ
Alb.	893	189	1082
MG	792	196	988
SC	594	125	719

(7-2) *Inversion clauses in the corpus: percentages*

	Inversion vs. VS	Inversion vs. SV/VS	Inversion vs. Σ
Alb.	1082 vs. 2225 = 48.6%	1082 vs. 7979 = 13.6%	1082 vs. 15813 = 6.8%
MG	988 vs. 2673 = 37.0%	988 vs. 6641 = 14.9%	988 vs. 14463 = 6.8%
SC	719 vs. 2065 = 34.8%	719 vs. 8458 = 8.5%	719 vs. 14176 = 5.1%

In Table (6-5), the frequency of inversion clauses is measured with respect to the sum of all verb-subject clauses in the corpus (the first column), to the sum of all clauses with overt subjects (the second column), and to the total number of clauses in the corpus (the third column). As expected, the percentage of inversion clauses in SC is low in all three categories. The relationship between Alb. and MG is more complex. The higher ratio of the Alb. inversion clauses with respect to the MG ones in the first column has to do with the fact that other types of verb-subject order are more frequent in MG than in Alb. (notably VsX, see 6.2.), so that the percentage of inversion clauses in MG automatically sinks. In the last column, the ratios are identical. This may be traced back to the higher frequency of zero-subject clauses in MG than in Alb. (cp. 5.1. and 5.5.): with the higher ratio of zero-subject clauses, the ratio of inversion clauses sinks. It is thus the second column, the percentage of inversion clauses within the class of clauses with overt subjects, that gives the relevant data: when the subject is expressed, MG has slightly more inversion clauses than Alb.

This notwithstanding, the number of inversion clauses in Alb. is still rather high, although, as indicated, one of the major factors of inversion, the adjacency constraint, seems to be rather weak in subordinate contexts in this language (cp. Sections 9.1. and 9.2.). As will become clear in the course of the following chapters, this frequency is primarily triggered by non-structural reasons and by a slightly unbalanced corpus: first, Alb. often uses embedded questions where MG and SC use abstract nouns (cp. 5.1.), so that the number of wh-inversion clauses is relatively high; second, my Alb. texts contain a great deal of quotations, and, consequently, a much greater number of quotation inversion clauses than the texts in the other two languages.

8. Focal Inversion

The first group of inversion types I am going to deal with is characterized by the presence of a clause-initial narrow focus expression: hence **focal inversion**. The basic formal features of the postfocal rest of the clause are the direct adjacency of the verb and the narrow focus expression (so that the verb precedes the subject and the other elements, if present), and a specific intonational contour. Informationally, the postfocal rest represents a relationally presupposed propositional function with the subject as the topic expression, or one of the topic expressions. Verb-subject order is thus embedded in a presuppositional context, which is the crucial part of my definition of inversion as opposed to vS and VsX constructions. There are three types of focal inversion: quotation inversion, wh-inversion, and fronted focus inversion (ff-inversion).

8.1. Quotation Inversion

The first type of focal inversion I am going to deal with is the obligatory postposition of subjects after verbs of saying, writing and thinking which follow quoted direct speech. I shall call this construction **quotation inversion**¹:

(8-1) Eng: “*What’d it look like?*”, **Henry asked**. (London, p. 11)

Alb: “*Po çfarë ishte?*”, **pyeti Henriku**. (p.11)

asked Henry-the

MG: “*Me ti emjaze?*”, **rotise o Xenri**. (p.16)

asked the Henry

SC: “*A kako je izgledalo?*”, **upita Henri**. (p.14)

asked Henry

(8-2) Alb. (Bishqemi, p. 56)

“*Qartë, qartë*”, **mendova unë**.

clear clear thought I

‘»Now I see«, I thought.’

¹ The term is borrowed from Green (1980) and Birner (1996); others, like Collins and Branigan (1997), call the same construction *quotative inversion*.

(8-3) MG (Elefterotipia, p. 6)

“Fovunde”, *γραfi i Bolton*, “oti i ... Elinokiprii θα tus ekδjoksun...”

they-fear writes the Bolton that the Greek-Cypriots will them oust

‘»They fear«, Bolton writes, »that the Greek Cypriots will oust them...«’

(8-4) SC (Vjesnik, p. 5)

“...sportaši će čuvati ... uspomenu”, *zaključio je Vrdoljak*.

athletes will keep memory concluded AUX:CLIT V.

‘»The athletes will not forget him...«, Vrdoljak concluded.’

This type of inversion is exceptional in many respects: First, in contrast to other types, it is informationally monotonous; second, it seems to be absolutely obligatory, whereas other kinds of inversion generally allow for some exceptions; third, it can bring about some interesting lexical effects absent from other inversion types; fourth, the syntactic status of the fronted element, the quote, is not entirely clear. There are two reasons why it is, in spite of these aberrant features, dealt with first, as a kind of introduction to the whole domain of focal inversion. On one hand, the invariant information structure characteristic of focal inversion in general is particularly conspicuous and easy to capture intuitively in quotation inversion, so that it is in this respect ideal as an introduction; on the other, it is extremely well represented in my corpus, so that it allows for a fine-grained syntactic description of the whole domain of inversion.

First some statistics. The first three columns of Table (8-5) contain data on the number of instances of quotation inversion in the original texts from the corpus (column 1), in the parallel translations of Jack London’s *White Fang* (column 2) and the total number of occurrences in both text types. The last three columns display the ratios of quotation inversion clauses within the sum of VS sentences in the given language (column 4), within the sum of sentences with expressed subjects (column 5), and within the sum of all sentences in the given language (column 6). The relationship is expressed in percents, so that, e.g., column 4 reads as ‘of all VS sentences in the given language, quotative inversions make up 22.1%’, etc.

(8-5) *Quotation inversion in the corpus*

	1. original	2. translat.	3. total	4. QI vs. VS	5. QI vs. SV/VS	6. QI vs. Σ
Alb.	398	93	491	491 vs. 2225 22.1%	491 vs. 7979 6.2%	491 vs. 15813 3.1%
MG	190	72	262	262 vs. 2673 9.8%	262 vs. 6641 3.9%	262 vs. 14463 1.8%
SC	276	95	371	371 vs. 2065 18.0%	371 vs. 8458 4.4%	371 vs. 14176 2.6%

As the statistics shows, the Alb. corpus and, to a smaller extent, the SC one contain significantly more quotation inversions than the MG corpus. This is only partly due to the corpus itself, with MG texts accidentally containing less quotations than the Alb. and the SC ones. In Section 8.1.5. I shall try to identify some non-structural reasons which may be responsible for this disproportion. Before that, however, the syntax, discourse-pragmatics and semantics of quotation inversion will be investigated in some detail.

8.1.1. Syntax and prosody

8.1.1.1. Basic features

Quotation inversion is known from a number of languages, including English (see Green 1980, Birner 1996: 20ff, Collins and Branigan 1997). In contrast to English, however, the inversion in the Balkan languages is not optional: in Alb., MG and SC the subject *must* stand behind the verb which follows a quote¹. Furthermore, apart from some well-defined and independently motivated exceptions, the verb and the quote have to be directly adjacent, i.e. the verb follows the quote immediately, as witnessed by all the examples in this section. The prototypical scheme of quotation inversion in Alb., MG and SC thus looks as follows (the sign ‘~’ marks immediate adjacency):

(8-6) **[quote]~[verb] [subject]**

Apart from this relatively rigid word order rule, quotation inversion is characterized by a specific intonation pattern. The quote, which builds a separate intonational phrase, has its own independent intonational contour, which is of no interest here. In the verb-subject part, the intonational contour is flat, and the sentence accent (always a low pitch, L*)², is carried by the verb, not by the subject: the scheme is **VERB subject**³.

There is, however, one possible deviation from this rule. If the subject is heavy, it can carry the secondary, postnuclear stress (see Section 5.3.): the pattern is **VERB SUBJECT**. Consider the following example:

¹ If the subject is placed in front of the verb following a quote, then the quote and the subject-verb complex form two separate clauses (see 8.1.3).

² Ladd (1996: 220) notes that the type of pitch accent used for the verb-subject complex following a quote in English may depend on the intonation contour of the quote itself. I was not able to detect this kind of interdependency in Alb., MG and SC.

³ See Section 5.3. for the intonational notation used in the present study.

(8-7) Alb. (Koha Ditore, p. 9)

[quote], *THEKSUAN përfaqësueset e sektorit politik-ushtarak*

emphasized representatives-the of-the sector political-military

të Misionit të SHBA-së në KOSOVË.

of-the Mission of-the USA in K.

‘[quote], emphasized the representatives of the political-military sector of the USA Mission in Kosovo.’

The subject *përfaqësueset e sektorit politik-ushtarak të Misionit të SHBA në Kosovë* is so heavy that native speakers invariably pronounce the sentence with a secondary stress on its last element. In other, less drastic instances, the frequency of the secondary stress diminishes. The scheme of the prototypical quotation inversion in the Balkan languages given in (8-6) is thus to be modified as follows:

(8-8) [quote] | ~[VERB] [subject]

[quote] | ~[VERB] [HEAVY SUBJECT]

8.1.1.2. Elements of the construction

In the overwhelming majority of cases, sentences with quotation inversion are extremely simple, containing, apart from the quote, only the verb and the subject, as in (8-1) - (8-7). In Alb., 362, i.e. 73.7% of all quotation inversions, are of this form; in MG, 193, i.e. 73.8%; in SC, 289, i.e. 77.8%. This means that in some three quarters of all instances there is not much more to report about the syntax of the construction than has already been said in 8.1.1.1. and summarized in (8-6) and (8-8).

The remaining one quarter (129 sentences in Alb., 69 in MG, 82 in SC) contain more material. In what follows, I shall try to determine what, if any, rules regulate the position of the sentence elements following the verb, most frequently the relative order of the subject and one or two additional elements.

8.1.1.2.1. General tendencies

There are two general tendencies regulating the relative position of the postverbal elements: the heaviness principle and the principle of cataphoric relevance (Section 5.2.). Since the way they function in quotation inversion has been sufficiently illustrated in 5.2. (see examples (5-

6)-(5-9)), I shall confine myself to describing an instance of a language-specific heaviness rule and to a brief summary.

As has been pointed out in 5.2., the heaviness principle is basically the same in all three languages under consideration, with heavier elements tending to be placed more to the right than the light ones. However, there is one important crosslinguistic difference, that of what counts as a light element. In Alb. and MG, pronominal subjects are on the same level of heaviness as full NPs and PPs; in SC, they count as lighter in a relevant way. This means that, other things (length, complexity) being equal, pronominal subjects will precede all other elements in SC, but not necessarily so in Alb. and MG. Consider first the following SC examples:

(8-9) SC (Antologija, p. 232)

[quote], *rekao je on čovjeku.*

said AUX:CLIT he to-man

‘[quote], said he to the man.’

(8-10) Eng: *How’d it happen?, he asked apathetically.* (London, p. 14)

SC: [quote], *upita on utučeno.* (p.18)

asked he dejectedly

Take (8-9): both *on* and *čovjeku* are equally long and complex; in spite of this, the reverse order (*rekao je čovjeku on*) would be, if not absolutely ungrammatical, then certainly very odd. Example (8-10) shows that pronominal subjects even precede modal adverbials, which, if not under narrow focus, tend to either directly precede or follow the verb (see 5.1. and 8.1.1.2.3.). Thus, pronominal subjects in SC are considerably lighter than lexical items, and this is a rule rather than a tendency: any kind of postposition of pronominal subjects makes the sentence sound awkward. In Alb. and MG, on the contrary, pronominal subjects, when combined with the elements of the same heaviness grade, display the usual kind of free variation (see 8.1.2.2. below):

(8-11) MG (Ciao, p. 5)

[quote], *apandise kofta ekinos.*

answered brusquely he

‘[quote], he answered flatly.’

(8-12) MG (Matesis, p. 69)

[quote], *tis ipa eyo sto mnimosino.*

her:CLIT said I at-the commemorative-meeting

‘[quote], I said to her at the commemorative meeting.’

(8-13) Alb. (Bishqemi, p. 74)
 [quote], *u përgjigj menjëherë ajo*.
 PASS:CLIT answered at-once she
 ‘[quote], answered she at once.’

(8-14) Alb. (Koha Ditore, p. 13)
 [quote], *theksoi ajo në fjalim*.
 emphasized she in speech
 ‘[quote], she emphasized in her speech.’

As the comparison between (8-11) and (8-12) on one hand, and between (8-13) and (8-14) on the other, shows, when the conditions of length and complexity are in equilibrium, pronominal subjects in Alb. and MG behave exactly like full NPs, i.e. they vary freely between pre- and postposition.

Now let us consider the cases in which the principle of cataphoric adjustment and the heaviness principle are not operative, i.e. when the subject and the other elements of the construction are both equally heavy and equally relevant in the following discourse. For the reasons which will become clear shortly, the expressions other than the verb and the subject are divided into two classes, labeled *independent* and *dependent elements* respectively.

8.1.1.2.2. Independent elements

The first class of expressions occurring in quotation inversion is made up of two main syntactic-semantic groups: first, adverbial expressions which set the spatio-temporal parameters of the event (setting adverbials), like *yesterday*, *then*, *at the conference* or *in the yard*, in the form of adverbs or prepositional phrases, and second, expressions denoting the addressee of the speech act, expressed by the dative or by PPs. These expression types are grouped into one class because they do not stand in any kind of close-knit lexical relation to the verb and are not directly selected by it, in contrast to the second class (see below). Statistically, the distribution of these expressions is as follows (**XVS** reads as ‘before the verb’, **VXS** as ‘between the verb and the subject’, and **VSX** as ‘after the subject’):

(8-15) *Quotation inversion: independent elements*

	Alb.			MG			SC		
	XVS	VXS	VSX	XVS	VXS	VSX	XVS	VXS	VSX
setting	0	15	12	0	7	5	0	9	11
addressee	0	4	3	0	5	3	0	3	2
Σ	0	19	15	0	12	8	0	12	13

As expected, independent elements are not allowed to appear before the verb in any of the three languages, conforming thus to the adjacency constraint. As far as two other positions are concerned, there seems to be no clear statistical preference, apart, perhaps, from an insignificant prevalence of the postsubjectal independent elements in SC in contrast to an equally weak prevalence of the presubjectal ones in Alb. and MG. In other words, the impression is that this is, basically, a field of free variation. The subject and independent elements may stand in any order, provided the two principles discussed above are not operative.

(8-16) SC (Vjesnik, p. 5)

[quote], *poručio je u svojoj propovijedi Kuharić.*

concluded AUX:CLIT in his sermon Kuharić

‘[quote], concluded cardinal Kuharić in his sermon.’

(8-17) SC (Kiš, p. 132)

[quote], *tvrdio je Paresijan u svojoj izjavi.*

claimed AUX:CLIT Paresian in his statement

‘[quote], claimed Paresian in his statement.’

8.1.1.2.3. Dependent elements

The second class of elements appearing in quotation inversion is defined as the class of *dependent elements* on two criteria: they either bear a close semantic tie to the verb, serving as what Behrens and Sasse (2003) call attributes to the predicate, or they are narrowly selected by the verb. That the first criterion holds for modal adverbials and depictive predicates, which constitute the first group within the class, is self-evident. The ontological basis for the third group, which I have for lack of a better term labeled *objects*, is the second criterion, that of selection. This group comprises direct objects proper, i.e. expressions in accusative, then genitive and dative objects selected by the verb, and finally, selected PPs (prepositional objects). Here is the statistics:

(8-18) *Quotation inversion: dependent elements*

	Alb.			MG			SC		
	XVS	VXS	VSX	XVS	VXS	VSX	XVS	VXS	VSX
modal adv.	0	31	6	0	23	4	18	6	20
depict. pred.	0	16	8	0	14	1	1	0	2
object	0	23	3	0	4	1	0	1	3
Σ	0	71	17	0	41	6	19	7	25

The most striking feature of the distribution of dependent elements is certainly the violation of the immediate adjacency rule holding between the quote and the verb in SC: modal adverbials and depictive predicates seem to occur freely between them. Furthermore, there are some clearly divergent statistical preferences. While dependent elements in Alb. and MG tend to precede the subject rather than to follow it, with the ratio ranging between two to one and eight to one, the tendency in SC is rather the other way around: when they are not in the preverbal position, dependent elements follow the subject in more than four fifths of the cases. The first approximation concerning the ordering tendencies of dependent elements is, then, that their preferred position is that between the verb and the subject in Alb. and MG, whereas in SC they are most commonly found either before the verb or after the subject.

The situation in Alb. and MG is quite straightforward: when the subject and the dependent element are equally heavy, and there is no special relation of cataphoric relevance, all kinds of dependent elements are preferably positioned between the verb and the subject (71, i.e. 80.7%, cases in Alb., 41, i.e. 87.2%, in MG). The preverbal position is absolutely ungrammatical, whereas the position after the subject is rather rare. The typical scheme is thus **[quote]~[verb] [dependent element] [subject]**. This scheme holds for all kinds of dependent elements (adverbials, depictive predicates and objects, as defined above).

(8-19) Alb. (Bishqemi, p. 79)

[quote], *ia preu **shkurt** komandanti.*

him-it:CLIT cut short commander-the

‘[quote], the commander interrupted him brusquely.’

(8-20) MG (Ciao, p. 5)

[quote], *epemene **ksafniasmenos** o Liatsos.*

insisted surprised the Liatsos

‘[quote], insisted Liatsos, surprised.’

(8-21) Alb. (Koha Ditore, p. 4)

[quote], *ka dhënë **mesazhin e tij** ministri i Rendit.*

AUX gave message the his minister-the of-the order

‘[quote], the minister of internal affairs gave his message.’

(8-22) Alb. (Bishqemi, p. 100)

[quote], *kërceu **në këmbë** kuzhinieri.*

jumped in feet cook-the

‘[quote], the cook jumped to his feet.’

Even if there are two or more dependent elements, they almost invariably stand in the position between the verb and the subject:

(8-23) MG (Kapandai, p. 118)

[quote], *eleje arya ta lojia tis i kapetanisa.*

said slowly the words her the captain-wife

‘[quote], the captain’s wife was pronouncing her words slowly.’

The ordering tendencies in SC are more complex, both concerning the number of positions typical for dependent elements and the homogeneity of the group of dependent elements itself. Because of the latter fact, I shall separately deal with the behavior of modal adverbials and depictive predicates on one hand and with that of objects on the other.

As is visible from Table (8-18), there are two basic positions for modal adverbials and depictive predicates, before the verb and after the subject, without there being any relevant semantic or statistical (18:20) difference between them. The pattern is: [quote]~[mod.adv/dep.pred]~[verb] [subject] [mod.adv/dep.pred].

(8-24) SC (Kiš, p. 4)

[quote], *kaže Mikša zagonetno.*

says Mikša mysteriously

‘[quote], says Mikša mysteriously.’

(8-25) SC (Vjesnik, p. 5)

[quote], *tužno nam je pričala Sofija Spajić.*

sadly us:CLIT AUX:CLIT told Sofija Spajić

‘[quote], Sofija Spajić told us sadly.’

It is the special status of modal adverbials that is responsible for the fact that the adverbials placed preverbally virtually break the immediate adjacency constraint holding between the quote and the verb: in contrast to independent elements and other dependent elements, they play the role of verb modifiers in the narrow sense of the word (cp. 5.1.). This semantic closeness enables the adverbials to form a constituent with the verb, in the way adjectives and nouns form noun phrases. Adverbs may thus be subject to the rules of internal syntax, i.e. to the rules holding on the constituent level. My contention is now predictable: If adverbials form a constituent with the verb, they behave like all other modifiers in SC, i.e. they precede their head. This claim has two consequences. First, when adverbials stand between the quote and the verb, the immediate adjacency rule is not broken, since adverbials count as a part of the verbal phrase. The quotation inversion pattern for SC can thus be altered as follows: [quote]~[[mod.adv/dep.pred]~[verb]] [subject] [mod.adv/dep.pred]. Second, the fact that

modal adverbials in quotation inversion do not appear preverbally in Alb. and MG is explained by different rules holding on the phrase level in these languages (see 5.1 for more details).

Table (8-18) shows further that there is yet another position for modal adverbials in SC: that between the verb and the subject. Its characteristic distribution in the corpus (all examples stem from Andrić, a writer who incorporates many features of the SC traditional folklore style) speaks clearly of its status in modern SC: it is an archaic feature used for stylistic effects. The following example from Andrić (8-26) has a close parallel in a folk fairy-tale (8-27):

(8-26) SC (Andrić, p. 74)

[quote], *odgovorio je mirno fra-Petar.*

answered AUX:CLIT quietly friar-Peter

‘[quote], answered quietly brother Peter.’

(8-27) SC (Antologija, p. 322)

[quote], *odgovori mu polako pop.*

answered him:CLIT slowly priest

‘[quote], answered the priest slowly.’

Objects in SC display a slightly different syntactic behavior: since they do not constitute a phrase with the verb¹, they are not allowed to intervene between the quote and the verb. The only position remaining (apart from some special cases, see below) is the position after the subject – **[quote]~[verb] [subject] [object]**:

(8-28) SC (Andrić, p. 60)

[quote], *prekinuo je Haim svoje kazivanje.*

interrupted AUX:CLIT H. his story

‘[quote], interrupted his story Haim.’

Consequently, the general scheme for all dependent elements in SC looks approximately like this: **[quote]~[[mod.adv/dep.pred]~[verb]] [subject] [dependent element]**.

8.1.1.2.4. Lexical solidarities

Some dependent elements are especially semantically close to the verb – as parts of idioms, semi-idiomatic expressions, support verb constructions, etc. Cross-linguistically, sentence

¹ Note that I do not subscribe to the opinion that objects form constituents with the verb, at least not in the languages of the Balkans; for an opposite view, see Horrocks (1994).

elements of this kind tend to be adjacent to the verb (Matić 2002). The consequences of this tendency for dependent elements in SC are predictable. When semantically integrated into the verb, modal adverbials, whose place otherwise freely varies between the preverbal and the postsubjectal positions, are restricted to the former:

(8-29) Eng: [quote], *Bill cogitated aloud*. (London, p.13)

SC: [quote], *glasno je mislio Bil*. (p.17)

aloud thought AUX:CLIT Bill

The expression *glasno misliti* ('to think aloud') is an idiom, since it is not semantically compositional (it does not mean 'to think loudly' as opposed to 'silently', but 'to speak for oneself'). The preverbal position of the adverb is the only one allowed if one wants to keep this sense: with the adverb in the postsubjectal position, the meaning is compositional (*mislio je Bil glasno* would mean something like 'Bill thought loudly [he was yelling or screaming]'). Note that this is in accordance with the hypothesis expressed above, according to which preverbal adverbs in SC form a constituent with the verb, which is not the case with the postsubjectal ones. As for objects, even when an object is idiomatically bound to the verb, or stands in any kind of narrow lexical solidarity with it, the preverbal position is excluded. The only remaining position adjacent to the verb is thus the intervening position between the verb and the subject:

(8-30) SC (Vesti, p. 16)

[quote], *skrenuo nam je pažnju Milanović*.

turned us:CLIT AUX:CLIT attention M.

'[quote], Milanović called our attention (to this).'

The expression *skrenuti pažnju* is (semi)idiomatic: it does not mean 'to turn attention', but 'to direct someone's attention to something'. Note that it is only in this semantic context (an idiomatical or semi-idiomatical unity between the subject and the verb), of course apart from the heaviness-induced shifts, that objects appear between the verb and the subject in SC.

The position of lexically close elements, both adverbials and objects, in Alb. and MG is the one directly after the verb, i.e. the position designed for all lexical types of dependent elements in these languages.

There is, however, one highly interesting interpretative tendency in Alb. which points to the relevance of the adjacency criterion for lexically close expressions in this language. Eight out of seventeen cases of light dependent elements in the position after the subject in my corpus are depictive predicates, which stand against sixteen light depictive predicates in the position between the verb and the light subject:

(8-31) Alb. (Bishqemi, p. 84)
[quote], *tha Partini i menduar*.

said P. the pensive

‘[quote], said Partin, pensive.’

(8-32) Alb. (Bishqemi, p. 116)

[quote], *tha i menduar komandanti*.

said the pensive commander-the

‘[quote], said the commander, pensive / pensively.’

English translations point to the fine interpretative difference triggered by different positions of the depictive predicate. When after the subject, depictive predicates tend to be understood as describing exclusively the state of the subject referent at the time delimited by the event, i.e. as small clauses proper: ‘Partin said something, and in doing it, he was pensive’. When in the contact position to the verb they may also be understood as verb modifiers: *i menduar* in (8-32) can be interpreted both as a small clause (‘the commander said, and in doing it, he was pensive’) and as an adverbial modifier (‘the commander said in a pensive manner’).

How is this to be explained? Since depictive predicates describe the temporary state of an argument during the event, they can indirectly modify the predicate itself: when a pensive person speaks, s/he usually does it in a manner characteristic of pensive persons, namely pensively. They can thus be ambiguous between the argument- and the predicate-related reading.

Now, since the immediately postverbal syntactic slot in Alb. marks the close semantic connection of the element in that position and the predicate, it is probable that the reading in which the predicate is modified will be chosen, or at least possible, when a depictive predicate is in that slot. On the other hand, the distant position in relation to the verb, the one after the subject, is thus used every time the speaker wants to explicitly exclude the predicate-related reading. This explains why it is depictive predicates that most frequently occur after the subject in Alb. in the cases in which the principles of heaviness and cataphoric adjustment are not involved.

The postsubjectal position is the regular way of avoiding the close semantic tie between the verb and the depictive predicate. This tendency to disambiguate semantics through syntax is simply not present with other kinds of dependent elements (or it is at least not present with

that regularity), since their semantic relation to the verb is disambiguated on the lexical or morphological level¹.

Finally, let me note that the rule which forbids the postposition of pronominal subjects in SC is stronger than the tendency to keep lexically related elements as close to each other as possible: if a pronominal subject cooccurs with an expression lexically related to the verb, then it is the subject that occupies the immediately postverbal position.

8.1.1.2.5. Elements of the construction: Summary

Now let me summarize this complicated section. For the sake of clarity, I shall represent all the tendencies enumerated above as rules; it should be kept in mind, however, that we are actually dealing with more or less strong statistical *tendencies*, and that exceptions are almost always possible. The only rules in the narrower sense of the word are the obligatory adjacency of the verb and the quote in all three languages, the obligatory adjacency of the verb and the pronominal subject in SC, and the almost obligatory adjacency of lexically close elements to the verb in all three languages.

The most frequent type of quotation inversion consists solely of a verb and a subject. If further expressions are present, the following rules hold in all three languages for all kinds of expressions:

¹ Two notes are in order here. First, the whole story of depictive predicates in Alb. is based solely on the intuitions of two native speakers which I translated into the language of semantics. Since I was not able to devise any kind of test which would substantiate this (admittedly volatile) evidence, I suppose the only way to check it would be to collect a large number of naturally occurring attestations of quotation inversion with depictive predicates, which is beyond the scope of the present study. Second, one may wonder why it is only in Alb. that this effect of ambiguity of depictive predicates is observed (native speakers of MG and SC failed to notice any interpretative differences triggered by different positions of these expressions). The answer probably lies in the extreme frequency of depictive predicates in Alb.: Whereas MG and SC seem to pattern with the AME language type in using depictive predicates only in the argument-related reading, in Alb. they often appear in contexts where one would expect adverbial expressions from the AME point of view, i.e. in the predicate-related reading. It may be the case that the division of labor between depictive predicates and modal adverbials in Alb. is different from that in MG and SC, and in AME in general, in that the differentiation between the argument- and the predicate-related readings is achieved syntactically instead of morphologically, at least in some cases. Of course, this is only a speculation which should be verified against a much greater amount of data than is accessible to me (see Buchholz and Fiedler 1987: 452ff., 460ff. and 550ff. for some background information on the Alb. depictive predicates).

(1) *Cataphoric adjustment*: Subjects which are syntactically relevant in the following discourse tend to be postponed:

Alb, MG, SC: [quote]~[verb] [any element] [subject_i], [X_i...]

(2) *Heaviness principle*: Those expressions which are relevantly heavier than the others tend to be postponed; the heavier the expression, the more obligatory its postponement:

Alb, MG, SC: [quote]~[verb] [lighter element] [heavier element]

This is where the first divergence among the languages under consideration appears: Whereas in Alb. and MG pronominal subjects side with all other types of light subjects, in SC they are counted as significantly lighter:

(3) *Pronominal subjects*:

SC: [quote]~[verb]~[pronominal subject] [any element]

If none of the three conditions mentioned above (cataphoric relevance of the subject, relevant heaviness of one of the elements, pronominal subject) obtains, it is syntactic and semantic features of additional expressions apart from the verb and the subject that influence the order of the elements in quotation inversion.

(4) *Independent elements*: Expressions which are neither narrowly selected by the verb nor stand in any kind of lexical solidarity with it (setting adverbials, addressee expressions) are placed either between the verb and the subject or after the subject in all three languages. There is no difference between these two positions (i.e. they are in free variation):

Alb, MG, SC: [quote]~[verb] [independent element] [subject]

[quote]~[verb] [subject] [independent element]

(5) *Dependent elements*: Expressions which are either narrowly selected by the verb or are lexically close to it (modal adverbials, depictive predicates, objects) are subject to following ordering rules:

modal adverbials, depictive predicates

Alb, MG: [quote]~[verb] [mod.adv/dep.pred] [subject]

SC: [quote]~[[mod.adv/dep.pred]~[verb]] [subject] [mod.adv/dep.pred]

objects:

Alb, MG: [quote]~[verb] [object] [subject]

SC: [quote]~[verb] [subject] [object]

(6) *Lexical solidarity*: The term refers to a close semantic tie occurring between verbs and dependent elements. Expressions standing in this kind of relationship to the verb (marked in the schemes with the index **LS**) are as a rule adjacent to it:

modal adverbials and depictive predicates:

Alb, MG: [quote]~[verb]~[mod.adv/dep.pred]_{LS} [subject]

SC: [quote]~[[mod.adv/dep.pred]_{LS}~[verb]] [subject]

objects:

Alb, MG, SC: [quote]~[verb]~[object]_{LS} [subject]

Points (1) and (2), which reflect some presumably universal word order tendencies¹ and are therefore less interesting for a typological comparison, may be neglected here. The same holds for point (4), which describes the behavior of independent elements, since this seems to be a domain of free variation. Points (3) and (5) – (6), which capture some specific features of the syntax of the Balkan languages, may, with some simplification, be represented in a tabular form. Since Alb. and MG display nearly identical positional tendencies, the data from these two languages are represented in a single row. The abbreviations read as follows: *adv* – modal adverbial/depictive predicate; *obj* – object; _{LS} – lexical solidarity; *XVS* – the element is before the verb; *VXS* – between the verb and the subject; *VSX* – after the subject.

(8-33) *Quotation inversion: elements of the construction*

	Adv	obj	adv_{LS}	obj_{LS}
Alb., MG	VXS	VXS	VXS	VXS
SC	XVS / VSX	VSX	XVS	VXS

The table makes the main difference between Alb. and MG on one hand and SC on the other perspicuous: While the position between the verb and the subject is the preferred location for all kinds of elements in Alb. and MG, it is precisely this position that is generally left empty in SC. In other words: Alb. and MG tend to postpone their subjects in quotation inversion; SC shows a relatively strong preference for the direct adjacency of the verb and the subject. These basic tendencies are confirmed by the data from the following section.

8.1.1.3. Complex predicate forms and complex sentences

This section deals with the position of the subject in quotation inversion with respect to auxiliaries and copular verbs and with respect to matrix verbs in complex sentences.

¹ Cataphoric adjustment and heaviness are relevant determiners of the relative order of elements in quotation inversion even in a rigid word order language like English – see Green (1980: 592) and Collins and Branigan (1997: 5ff).

First the complex predicate forms. The question that is of interest here is whether the subject is positioned between the auxiliary/copula and the participle/predicative or after both of them. This question can be asked only with respect to Alb. and MG, since the past tense auxiliaries, practically the only ones occurring in quotation inversion, are clitics in SC, which makes their place in the clause predictable and irrelevant for the discussion (see 5.1.). In Alb. and MG, the combinations AUX-PTCP and COP-PRED seem to be inseparable in quotation inversion. Subjects are invariably placed only after participles and predicatives, as illustrated by following examples:

(8-34) Eng: [quote], *Judge Scott said*. (London, p. 191)

Alb: [quote], *ka thënë gjyqtari Skott*. (p. 188)

AUX said judge-the S.

MG: [quote], *exi pi o dikastis Skot*. (p. 230)

AUX said the judge S.

(8-35) Alb (Kosovarja, p. 14)

[quote], *qe i pakënaqur z. Avdiu*.

was the dissatisfied Mr. A.

‘[quote], Mr. Avdi was dissatisfied.’

Deviations from this order are not possible. Thus, unlike VsX (cp. 6.3.2.3.), the subject in Alb. and MG always follows both components of the predicate, which in their turn always display the firm order AUX-PTCP and COP-PRED. The pattern for Alb. and MG is thus:

(8-36) **Alb, MG:** [quote]~[aux/cop]~[ptcp/pred] [subject].

As for complex sentences, quotation inversion clauses usually contain complement clauses embedded in a matrix clause with a phasal verb (‘to begin’, ‘to continue’, ‘to stop’), a volitional (‘want’) or a control verb (‘try’). What is common to all these matrix verbs is that they take the complement clauses with the subject coreferent to the subject of the matrix verb itself.¹ The number of examples of this syntactic context in my corpus is not very large (6 in Alb, 4 in SC, 2 in MG), so that the data had to be confirmed through elicitation. The results are as follows: if the subject is not exceedingly heavy, it preferably stands either after the complement clause ((8-37) and (8-38)), or after the verb of the complement clause ((8-39)) in Alb. and MG, and between the matrix verb and the complement clause in SC ((8-40)):

¹ Note that Balkan languages either do not have infinitives at all (MG and Standard Alb.), or use them only to a very limited extent (SC), so that SS matrix verbs take finite complement clauses, and not infinitives, as in most West European languages (see Joseph 1983 and the papers in Rivero and Ralli 2001 on the topic of finiteness in the Balkans; see also Section 9.2).

(8-37) Alb. (Bishqemi, p. 40)

[quote], *deshti të sigurohej së fundi **Partini***.

wanted that be-sure of-the end Partin

‘[quote], Partin wanted to be completely certain about it.’

(8-38) MG (Fakinou, p. 17)

[quote], *arxise na tu lei kleyondas **i Rinula***.

started that him:CLIT tell crying the Rinoula

‘[quote], Rinoula started to tell him, crying.’

(8-39) Alb. (Koha Ditore, p. 3)

‘[quote], *është cituar të ketë thënë **Koshtunica** në një intervistë të botuar të enjten*

AUX quoted that AUX said K. in an interview the published the Thursday
në të përditshmen “El Mundo”.

in the daily-paper E.M.

‘[quote], Koštunica is quoted to have said in an interview published on Thursday
in the daily paper »El Mundo«.’

(8-40) SC (Andrić, p. 64)

[quote], *pokušava **kadija** da brani mladića*.

tries qadi that defend young-man

‘[quote], the qadi tried to defend the young man.’

Of course, if the subject is heavy, it tends to follow the complement clause even in SC:

(8-41) SC (Vesti, p. 2)

[quote], *nastavlja da priča **teško bolesna Nataša Ivanović***.

continues that tells heavily ill N.I.

‘[quote], Nataša Ivanović, seriously ill, continues her story.’

The orderings caused by heaviness aside, the patterns for quotation inversion with complex sentences look as follows:

(8-42) **Alb, MG:** [quote]~[matrix verb] [subordinator+verb] [subject]

SC: [quote]~[matrix verb] [subject] [subordinator+verb]

If we treat auxiliaries, the copula and matrix verbs on a par, as carriers of TAM features (with or without additional meaning), and participles, predicatives and complement clauses as their lexical complements, then we may for syntactic purposes mark the former as finite verbs (‘verb’ in my informal notation) and the latter as complements (‘X’ in my notation)¹. If we

¹ A similar analysis of the relationship between auxiliaries and participles is given by Fillmore (1999: 114). For a comparable analysis of matrix verbs in the Balkan languages see Roussou (2001); for an approach in terms of

apply this notation to the patterns (8-36) and (8-42), we get the unified pattern for complex predicate forms and complex sentences:

- (8-43) **Alb, MG:** [quote]~[aux/cop]~[ptcp/pred] [subject].
 [quote]~[matrix verb] [compementizer+verb] [subject]
 = **[quote]~[verb] [X] [subject]**
- SC:** [quote]~[matrix verb] [subject] [subordinator+verb]
 = **[quote]~[verb] [subject] [X]**

8.1.1.4. Summary

As noted above, apart from the iconic tendencies of cataphoric coherence and the postposition of heavy elements, there seem to be no recognizable rules pertaining to the position of independent elements in quotation inversion. This fact seems to be connected to the looseness of semantic and syntactic ties between these elements and the clause, or better, the verb, and is, if not always in such a drastic form as in quotation inversion, characteristic of the word order systems of Balkan languages. Dependent elements, on the other hand, seem to be subject to somewhat stricter rules (which nevertheless may be overridden by the factors of cataphoric coherence and heaviness), apparently due to their stronger semantic and syntactic ties to the verb. Since the rules they are subject to are at least partly syntactical, i.e. conventional (as opposed to iconic), this is the area one expects some crosslinguistic variation. The expectations are fulfilled: Alb. and MG on one hand and SC on the other display some clearly diverging tendencies concerning the positions of dependent elements, tendencies which pervade the whole system (adverbs, depictive predicates, objects, complex verb forms, complement clauses). Abstracting away from the details, these tendencies may be formalized with two templates representing the prototypical quotation inversion clause containing a light dependent element and a light subject in Alb, MG, and SC ('X' means 'any dependent element'):

- (8-44) **Alb, MG:** **[quote]~[verb] [X] [subject]**
- SC:** **[quote]~[[mod.adv.]~[verb]]~[subject] [X]**

Dependent elements tend to be placed between the verb and the subject in Alb. and MG, whereas in SC there exists a strong tendency to keep the verb and the subject together, which

PRO, Control, etc., see references in Alexiadou and Anagnostopoulou (2000), as well as Dobrovie-Sorin (2001) and Krapova (2001).

results in the placing of modal adverbials before the verb and the rest of dependent elements after the subject. Furthermore, in all three languages there is a tendency to place the elements semantically close to the verb in a position adjacent to it; due to the specific syntax of modal adverbials in SC, they are in this case always before the verb; all other elements in all three languages are placed in the position immediately after the verb, which thus seems to be the general position for the elements in the relationship of lexical solidarity to the verb.

The two templates in (8-44) represent, as I hope to demonstrate in the following sections, the basic syntactic pattern of all kinds of focal inversion in all-ratified contexts (see 8.1.2. and 8.2.2. for this notion), which is the reason why this section on the syntax of quotation inversion turned out to be so long.

8.1.2. Information structure and discourse function

A careful reader will have noticed that I have described none of the syntactic rules concerning quotation inversion in terms of topic, focus, and other phenomena of information structure, although the underlying assumption of the present study is that word order in the Balkan languages is to a great extent determined by discourse-pragmatic considerations. Obviously, I do not think that the small amount of word order variation present in this construction is accountable for in terms of information structure. Quotation inversion conveys an invariable presuppositional structure, namely that of a presupposed propositional function, meaning that the whole material within quotation inversion is presupposed, informationally monotonous. Now, the fact that a clause (minus the focused element, of course) is presupposed as a whole, as will become clear later (Sections 8.2., 8.3., and 9.), does not automatically mean that certain discourse-pragmatic functions cannot be differentiated via word order and intonation. I shall therefore argue that the real reason why information structure phenomena do not play any role in the word order variation within quotation inversion is the restricted role this construction plays in discourse, which, then, allows for only one particular kind of presupposed propositional function, the all-ratified one.

Let me start with the claim that quotation inversion always contains only presupposed material. I am aware of only two types of analysis of this construction in terms of presuppositional structure, incidentally both in the works dealing with Balkan languages. The first, proposed by Ulrich (1985:157ff.) and Popović (1997: 67ff.) with respect to Rumanian and SC, respectively, is firmly rooted in the Praguean FSP tradition: the quote is considered to be the theme, and the verb-subject complex the rheme of quotation inversion sentences. The

second was proposed by Sasse (1995a: 150), with respect to MG: in his view, quotation inversion is ‘a fossilized case of focus fronting, the stretch of direct speech being the fronted focus constituent’¹.

I shall side with Sasse, for both formal and interpretational reasons. The formal reason is the intonational pattern of quotation inversion. Namely, if Ulrich’s and Popović’s analysis is translated into the language of the present study, it would mean that quotation inversion is an instance of vS construction with a topicalized object. Intonational data speak against this classification: recall that in vS construction the main sentence accent (always H*) generally falls on the subject, whereas in quotation inversion it is invariably assigned to the verb, and presumably has the value L*. My tests have shown that it is impossible to put the sole accent on the subject: a sentence like (8-1) cannot be pronounced as [*quote*], *pyeti HENRIKU* etc. in Alb., MG, and SC. The interpretational reasons have to do both with the primary intuition of what is actually being asserted when someone’s speech is quoted and with the distribution of quotation inversion across textual contexts. When a speaker decides to render someone’s words within her/his own utterance, then her/his intention is arguably to use them as the main point of this utterance. The quote is thus focused, although, since it plays the role of the term of the predicate within the speaker’s proposition, it must be existentially presupposed (see 4.2.3.). The information that the quote has been uttered is automatically presupposed via existential presupposition of quotes (if a speech exists, it has been uttered; see 4.2.4. on rich existential presuppositions). The same informational status is given to the participants in the speech act – the speaker and the addressee – as well as to the circumstances under which it took place and the manner in which it was performed, judged by the nature of the use of quotes in discourse (see below). The speaker thus presupposes the propositional function ‘someone said X (to someone, under certain circumstances, in a certain manner)’ and makes her/his assertion by identifying the variable X with the quote. The first approximation of the information structure of the clauses containing quotes can be represented as follows:

(8-45) [**quote**]_{NARROW FOCUS} [**verb, subject, (+/-X)**]_{PRESUPPOSED PROPOSITIONAL FUNCTION}

The interpretation represented in (8-45) suggests that, in quotation inversion, the speaker expects that the hearer is either already familiar with the fact that the material contained in the quote is/was uttered by a certain person, etc., or else that s/he is ready to take it from her/him as uncontroversial (see 4.2.1.). Why should this information then be uttered at all? Recall that natural language utterances generally contain presupposed material as a kind of ordering

¹ Green (1980: 590ff) seems to have a similar configuration in mind when she says that the main point in quotation inversion sentences is the quote, whereas the non-quote parts of the sentence are ‘inconspicuous’.

instruction: the hearer is to relate the asserted quote to a certain set of possible worlds determined by the presupposed material (containing in quotative contexts the one who uttered the quoted material and, optionally, other circumstantial information). When the speaker has the impression that the relationship between the asserted quote and the set of possible worlds to which it is to be related is for some reason not entirely present in the auditor's mind, s/he adds the quotation inversion clause to the quote, as a sort of parenthetical reminder, the instruction reading approximately as 'in the case you have forgotten, this was said by X on the occasion Y...'.¹ Two basic discourse contexts in which quotative inversion is used confirm this analysis:

(a) **Retold long monologues.** When longer monologue passages are quoted, especially in the journalistic style (reports from press conferences are the prototypical example), speakers tend to interrupt the quote from time to time with a quotation inversion clause, doubtlessly in order to remind the reader that what s/he is reading is a quote, i.e. in order to remind her/him to which set of possible worlds s/he is to relate the information contained in the text. Needless to say, the content of quotative inversion clauses thus used is extremely easy to presuppose. The principle is nicely illustrated by the following passage, representing a report from the funeral mess for the Croatian president Tudjman. The journal article begins with the announcement that the sermon was held by cardinal Kuharić; what is rendered as [quote] here is the actual content of the article, the extensive quotation of the sermon itself:

(8-46) SC (Vjesnik, p. 6)

[quote], *dodao je Kuharić*, [quote], *rekao je Kuharić*, [quote], *poručio je*
 added AUX:CLIT K. said AUX:CLIT K. concluded AUX:CLIT
u svojoj propovijedi Kuharić.

in his sermon K.

'[quote], added Kuharić, [quote], said Kuharić, [quote], Kuharić concluded in his sermon.'

By the use of quotation inversion clauses, the reader is occasionally (every ten to twenty clauses) reminded that what s/he is reading is a quotation. The final quotation inversion in the example is the last sentence of the journal article: the whole situation in which the quoted speech takes place (the speaker and the type of speech act) are mentioned, as a kind of recapitulation.

¹ See Collins and Branigan (1997: 11-12) and, especially, Green (1980: 591), whose idea that quotation inversion 'puts the non-quote part of the sentence out of the way of the reader' is very similar to my analysis.

(b) **Retold dialogues with a predictable sequence of speakers.** Quotation inversion is also used when conversational interactions are reproduced. What is less obvious is that it is used *only* when the sequence of speakers is either completely predictable or when it is at least either noncontroversial or irrelevant for the development of conversation, i.e. under the conditions which facilitate presuppositions (see 4.2.1., 4.2.4.)¹. Consider (8-47), where in the preceding text a secret meeting of two lovers has been described:

(8-47) Alb. (Kosovarja, p. 8/50)

“Fol”, më thoshte Faca. “Fol”, i thosha unë atij.

“say” me:CLIT said F. “say” him:CLIT said I him

“Fol, moj Dije”, më thoshte Faca. “Jo, fol ti”, i thosha unë atij.

“say, PTCL D.”, me:CLIT said F. “no, say you” him:CLIT said I him

‘»Say something«, Facë told me. »Say something«, I told him. »Say something, oh Dije!«, Facë told me. »No, you say something.«, I told him.’

In a conversation in which it is explicitly stated that only two parties participate, the sequence of speakers is a noncontroversial matter, even in cases where the content of the conversational turns is somewhat richer than in the above example. This is why quotation inversion is especially frequent in certain journalistic genres, like interviews or reports from the court. In cases where there are more than two participants, quotation inversion is used when the content of the quote makes one of the persons present on the scene a particularly good candidate for the utterer of the turn in the conversation. In (8-48), the scene looks as follows: A noble couple sit at the evening table with their guests; after a conversation about a book, in which both the guests and the hosts partake, there is a moment of silence. This is where (8-48) is uttered (the narrator is the couple’s daughter):

(8-48) MG (Kapandai, p. 22)

“Kira mu ke arxondisa”, jirise sti mana mu o pateras mu, “se parakalo na

“lady my and mistress” turned to-the mother my the father my “you I-ask to timisis tus kalesmenus mu me to trayuđi su.”

honor the guests my with the song your

‘»My lady and mistress«, my father turned to my mother, »may I ask you to honor my guests with your song?«’

It is only the husband who can address the lady of the house with *kira mu ke arxondisa*, so that the content of the quote makes the speaker predictable.

¹ This point has been noted by Green (1980: 592): ‘[T]he usual case in the ordinary narration of dialog is that it is obvious who is going to speak next, and perhaps that that person is going to speak.’

Finally, there are cases where the dialogue situation is not established before the quote at all, so that neither the number of speakers nor the content of the quote can help the hearer predict who uttered it, but quotation inversion is nevertheless used. Consider the following example, which is preceded by a passage describing the life of Mrs. Kanello, in which the narrator explains that she got her last child in the middle of the Second World War:

(8-49) MG (Matesis, p. 33-4)

Otan eðikse pja i kilja tis, “Kira mu, lolaθikes”, tis lei

when showed already the stomach her “lady my, you-got-crazy” her:CLIT says

i mana tis Afroðitis “Eyo”, tis ðikeolojite i kiria Kanelo, “ðen to poliθela”.

the mother of A. “I” her:CLIT justified-herself the Mrs. K. “not it much-wanted”

‘When her stomach started to show, »Madam, you are completely nuts«, Aphrodite’s mother told her, »I didn’t really want it«, Mrs. Kanello tried to defend herself.’

The person who uttered *kira mu, lolaθikes* is hardly to be predicted from the immediately preceding context. The solution to this is the broader context. The whole book from which this passage is taken deals with the fate of a few women in a provincial Greek town during the Second World War. The interaction between these women represents the main narrative line in the novel. Consequently, when a quote is given without further context, the reader is expected to assume that it has been uttered by one of the main characters (the person named ‘Aphrodite’s mother’ is one of them). In this particular passage, the writer does not expect the reader to be able to predict exactly which one of the women is the speaker; it suffices for his purposes that the fact that one of the constant *dramatis personae* uttered a sentence is not controversial for the reader. The use of quotation inversion is in thus justified by the noncontroversial nature of the information presented in it, i.e. by its being easily presupposable.

Let us reconsider the informational status of the subjects of quotation inversion clauses in the light of what has been said in this section. The contexts in which these sentences are used are those of a high referential presence of the referents of their subjects, meaning that they are either continuously topical over a certain stretch of discourse, or that they are major participants on the level of the whole text, even if not directly present in the narrowest surroundings of the quotation inversion clause. A rough statistics I undertook in my corpus confirms this unequivocally. Out of 1124 quotation inversion clauses in all three languages taken together, 1054 (93.8%) have a subject which is immediately present in context, either as a previously introduced participant in a dialogue or as the sole speaker of a monologue. Some

61 (5.4%) cases are covered by the notion of major participant, where presuppositions work in the way demonstrated for (8-49). The remaining 9 (0.8%) cases are instances of pretense in Stalnaker's sense (see 4.2.1.).

What does this strong referential continuity of quotation inversion subjects mean in terms of information structure? I should like to argue that they are ratified topics (see 6.2.) within presupposed propositional functions. Recall that ratified topics represent the kind of topic which is mutually agreed upon by the interlocutors, and that the major indicator of this mutual agreement is referential continuity, in the narrow sense, as the presence in the immediate context, or in somewhat broader sense, as a potential presence in the mind of interlocutors due to a major role played in the broader discourse, or due to frames (see 4.2.4. and 6.2.), or for some other reason. Quotation inversion subjects fulfill this formal criterion, and the typical contexts of their occurrence are those of continuous, ratified topicality. The pattern of the information structure of quotation inversion clauses given in (8-45) is therefore to be altered accordingly (PPF = 'presupposed propositional function):

(8-50) **[quote]_{FOCUS} [[subject]_{RATIFIED TOPIC} + rest of the quotation inversion clause]_{PPF}**

To summarize: Quotation inversion occurs in a small group of contexts, comprising retold monologues and dialogues with relatively predictable speech act participants. Its narrow applicability in discourse is due to its extremely restricted information structure: the quote has to be focused, the rest of the sentence has to be presupposed and ratified, and the subject has to be a ratified topic. I shall call the kind of information structure quotation inversion clauses invariably have *all-ratified presupposed propositional functions*, since all the elements of the proposition (except the focused element, of course) are not only presupposed, but also ratified, i.e. accepted in advance as the basis for making assertions.

8.1.3. Semantics of quotation inversion and the syntactic status of the quote

The analysis proposed above presupposes that the quote is integrated into the clause, playing the role of the object of the verb of saying, thinking, etc. This is not an uncontroversial assumption. Thus Collins and Branigan (1997:10-11) assume that 'the quote simply identifies the content of an empty quotative operator, which occupies Spec-C (...) The quote itself is external to the clause, possibly adjoined to CP.' Now, as far as the examples they discuss are concerned (all of the form '[quote]+verb of saying, writing, thinking + subject'), I do not see any theory-independent reason to assume the extracausal nature of the quote: it is subcategorized by the above mentioned classes of verbs, it can be paraphrased by an

anaphoric pronoun ('she said/thought/wrote **it**'), and the fact that it represents a separate intonational phrase is simply due to its length. The argument adduced by Popović (1997: 68) deserves more attention: he notes that the quote can be followed by verbs which do not subcategorize it (see examples below), and concludes that, since it cannot play the object role in these cases, it must represent an independent clause with respect to the quotation inversion clause. Here are some more or less spectacular instances of this phenomenon (see also (8-19) – (8-22), (8-30), (8-35), (8-37), (8-48)):

(8-51) Alb. (Camaj, p. 102)

“Vërtet”, qeshi Lakmuesi.

“true” laughed Envious-person

‘»You are right«, the Envious one laughed.’

(8-52) Alb. (Bishqemi, p. 56)

“Tanku?!”, hapëm sytë ne.

“tank” opened eyes we

‘»Tank!?!«, we opened our eyes wide (in surprise).’

(8-53) MG (Kapandaj, p. 142)

“Vai vai”, travuse tora eksalos ta jenja tu o muftis.

“woe woe” drew now hysterical the beard his the mufti

‘»Woe, woe«, the mufti, beside himself, was tearing his beard.’

(8-54) SC (NIN, p. 25)

“Odakle sad kredit ...”, zbunjen je Miodrag Sedlarević.

“whence now credit ...” confused is M. S.

‘»Where has the bank-credit come from ...«, Miodrag Sedlarević is confused.’

Both objecthood criteria mentioned above (subcategorization by the verb, substitution by an anaphoric pronoun in accusative) fail here: neither do ‘laugh’, ‘tear one’s beard’, ‘be confused’, etc., select objects of the ontological class ‘speech act content’ nor can the quotes be substituted by ‘it’ (*‘we opened our eyes it’). Is it then necessary to allow for two kinds of quotation inversion, one with the integrated object quote and one consisting of two clauses, or even to generalize the disjunct-clauses analysis to all cases of quotation inversion (and thus admit that Collins and Branigan are right, after all)?

I do not think so, although my solution is only a tentative one. Roughly, I shall argue that every predicate used in quotation inversion is lexically reinterpreted as a verb of saying, thinking, etc., modified by the ‘proper’ meaning of the predicate itself. The verb *qesh* (‘laugh’) in (8-51) is thus to be read as ‘say laughing’, *travo ta jenja* (‘tear the beard’) in (8-

53) as ‘say (while) tearing one’s beard’, etc. The principle I propose to account for this meaning change is strongly reminiscent of Levin and Rapoport’s (1988) *lexical subordination* and Jackendoff’s (1990) *constructional idioms*¹. Certain syntactic/semantic configurations allow only for certain kinds of lexical fillings; if a lexical filling other than that required by the syntactic/semantic configuration is inserted into a construction, the inserted lexeme is semantically adjusted to the required lexical structure. In this process, the basic meaning of the inserted lexeme appears as a value subordinated to the adjusted, newly developed meaning. In the case of quotation inversion, this means that in the syntactic/semantic configuration consisting of a focused quote followed by the verb-subject complex (in this order), only verbs of saying, thinking, etc., are licensed as lexical fillings. When another verb or phrase is used in this construction, it develops a new meaning, namely that of saying, thinking, etc., while the original meaning appears as an adverbial modification of this new meaning. I shall call this process *lexical reinterpretation*.

Now, in Section 2. I have criticized a number of approaches to VS phenomenon because of their unrestricted use of the mechanism of lexical reinterpretation. As I do not want to fall victim to this kind of critique myself, I shall try to adduce some evidence for my interpretation of non-prototypical verbs and phrases in quotation inversion. First some soft evidence. In all the examples adduced above, the non-prototypical verbs are paraphrasable with the construction ‘say + gerund’ without any change of meaning – thus (8-52) can be construed as *thamë ne duke hapur sytë* (said we PTCL opening eyes-the), (8-53) as *ipe o muftis travondas ta jenja tu* (said the mufti tearing the beard his). This variant, [say + gerund], is actually sometimes used instead of a simple non-prototypical verb in all three languages. The following translations of the same text make this point particularly conspicuous:

(8-55) Eng: “It’s darn good coffee”, Henry said enticingly. (London, p.15)

MG: “*Ine poli kalos kafes*”, **ton proetrepse o Xenri**. (p.20)

him:CLIT encouraged the Henry

SC: “*A jako je dobra kafa*”, **reče Henri mameći ga**. (p.18)

said Henry enticing him

Alb: “*Ama, ç’ kafe e shijshme!*”, **e ngacmonte Henriku**. (p.15)

him:CLIT enticed Henry-the

¹ For a critique of these and similar approaches (with respect to the English resultative construction) see Levin and Rappaport Hovav (1995: 72-78).

(8-56) Eng: “Ain’t a bit scairt [sic] of you”, Henry laughed. (London, p.17)

MG: “*Siya, mi s’ akusi*”, *jelase o Xenri*. (p.23)

laughed the Henry

SC: “*Nimalo te se ne plaši*”, *nasmija se Henri*. (p.21)

laughed REFL Henry

Alb: “*As që ia bën syri tërr*”, *tha duke qeshur Henriku*. (p.18)

said PTCL laughing Henry-the

In (8-55), it is in SC that we find the avoidance of lexical reinterpretation achieved through [say+gerund] (as in the English original); in (8-56), it is in the Alb. translation.

The evidence based on paraphrases of this kind is not really overwhelming, so I tried an experiment. I presented a number of native speakers with pairs of sentences, one of them consisting of a quote followed by a quotation inversion containing a non-prototypical verb, the other of a quote and a sentence with SV order. Example (8-52), for instance, was used in its given form, “*Tanku?!*”, *hapëm sytë ne* (‘tank?!’, opened eyes-the we) and with a SV clause, “*Tanku?!*” *Ne hapëm sytë*. (‘tank?!’. we opened eyes-the)¹. The subjects were asked to determine the temporal relationship between the speech act in which the quote came into being and the action conveyed by the non-prototypical verb.

The answers matched my own native speaker intuitions concerning SC: With quotation inversion, the speech act and the action conveyed by the following clause are interpreted as simultaneous; when the quote is followed by a SV clause, both the simultaneous and the non-simultaneous interpretations are possible. Thus, ‘*»Tanku?!«*, *hapëm sytë ne*.’ can only mean that we uttered ‘tank’ and opened our eyes wide at the same moment. In contrast, ‘*»Tanku?!«* *Ne hapëm sytë*.’ can also mean that we first uttered ‘tank’ and opened our eyes wide only afterwards, or even before this.

Even better, a native speaker of MG claimed that in (8-56) [“*Siya, mi s’ akusi*”, *jelase o Xenri*. vs. *O Xenri jelase*.], there is also a difference in the attribution of the speech act to a particular person. With quotation inversion, only Henry can be the utterer of the quote; with the SV clause, the preferred interpretation (though not the only one) is that it is someone else

¹ The sentences were read, not shown to the subjects in their written form, in order to avoid a possible interpretational impact of the orthographic convention to write SV clauses after the quote with the initial capital letter. (Incidentally, the soundness of this convention is indirectly confirmed by my test.)

who utters *Siya, mi s' akusi*, while Henry reacts to this utterance with a laughter. The same kind of judgment was then elicited from the native speakers of Alb. and SC, too¹.

Both results speak in favor of the lexical reinterpretation analysis. SV clauses with non-prototypical verbs display the freedom of interpretation typical for intersentential contexts, where it is solely pragmatic principles that determine the temporal and referential issues. Thus, Collins and Branigan's and Popović's analyses, which operate with two disjoint clauses, seem to work perfectly here.

On the other hand, quotation inversion clauses with non-prototypical verbs are interpretationally restricted with respect to time and person reference: speech act and the action expressed by the verb must take place simultaneously, the performer of the action conveyed by the non-prototypical verb and the utterer of the quote have to be the same person.

This is not the kind of semantic behavior one would expect from two separate clauses, but it is rather characteristic of intrasentential contexts, where semantic roles and temporal reference are established syntactically and morphologically. Furthermore, the two interpretational restrictions closely match the features of the lexical interpretation I proposed: if 'laugh', for instance, is constructionally reinterpreted as 'say laughing', then it is precisely the temporal coextension of 'say' and 'laugh' and the identity of the 'sayer' and 'laugher' that is to be expected.

In conclusion: The quote in quotation inversion is syntactically the object of the verb which immediately follows it independently of the semantic and syntactic class the verb itself belongs to.

¹ As a matter of fact, I have one attested example which perfectly illustrates the ability of SV clauses following quotes to convey the action which is both non-simultaneous to the speech act and performed by a person other than the speaker: Bill and Henry are having a conversation, and Bill says "... that's what I wisht [sic]". The sentence that follows this quote is:

Eng.: *Henry grunted and crawled into bed.* (London, p. 8)

Alb.: *Henriku diç mërmëriti nëpër dhëmbë dhe ra të flinte.* (p. 8)

Henry-the something murmured over teeth and fell to sleep

MG: *O Xenri ɣrilise ke sirθike kato apo ta skepasmata.* (p. 13)

the Henry grunted and drew-himself under from the blankets

SC: *Henri progundja, pa se uvuče u postelju.* (p. 10)

Henry grunted and REFL drew-in in bed

It is clear from the context that the sentence describes Henry's reaction to Bill's utterance. With quotation inversion, this interpretation would be impossible – *mërmëriti Henriku...*, *ɣrilise o Xenri...* and *progundja Henri...* would mean that it was Henry who uttered the quote, and that he did it grunting.

8.1.4. Alternatives to quotation inversion

The quote is the narrowly focused object of the predicate of quotation inversion clauses; the verb itself, the subject, and the optional elements of the construction represent an all-ratified presupposed propositional function, within which the subject obligatorily plays the role of the ratified topic. The informational pattern given in (8-50) can thus be further simplified:

(8-57) $O_{\text{NARROW FOCUS}} [V(X) - S_{\text{RATIFIED TOPIC}}]_{\text{ALL-RATIFIED PRESUPPOSED PROPOSITIONAL FUNCTION}}$

This is the pattern which, *mutatis mutandis*, describes the information structure of all kinds of clause-initial narrow foci, with one exception:

In all other types, topics (primarily subjects, though other elements are not excluded) need not be of the ratified type, in which case they are marked by specific word order permutations, frequently by putting the topic expression in front of the narrow focus expression (recall that the narrow focus expression and the verb have to be adjacent, so that an intermediate position is not licensed; for details, see below 8.2. and 8.3.).

This possibility does not exist in quotation inversion – a sentence like **John “Go away” said* sounds equally bad in Alb., MG and SC as it does in English. This restriction is presumably connected with the specific structure of the narrow focus expression, the quote. First, the quote is often very long, so that a preposed topic expression would be at a great distance from the verb.

Second, and more important, it represents a different kind of language material from the rest of the clause, being an instance of ‘mention’ whereas the verb of saying, etc., and its subject are instances of ‘use’ (see Sperber and Wilson 1981 and Carston 1998, 1999 on the use-mention distinction). This ontological difference between the quote and the rest of the sentence may, in some languages, lead to the rule of keeping the ‘use’ part and the ‘mention’ part of the sentence apart^{1,2}.

¹ In Alb, MG and SC, as in English, this tendency has the status of a rule only for the verb-subject complex (the ‘use’ part); the quote may be discontinuous, as witnessed by sentences of the following type:

Alb. *“Fatalizmi ...”, ... tha Markoci, “erdhi nga thatësia e verës...”* (Camaj, p. 210)
“fatalism” said Markoc “came from dryness-the the of-summer-the”
‘»The fatalism...«, said Markoc, “came from the dryness of the summer...”

² Emphasis should be put on the phrase ‘in some languages’: there are languages where there is no such a restriction at all, so that both the ‘mention’ part (the quote) and the ‘use’ part (the verb-subject complex) may be freely intertwined, to the effect that sentences like the above mentioned **John “Go away” said* are perfectly

Be it how it may, the fact is that quotation inversion is more pragmatically restricted than most other kinds of narrow focus fronting. The question is what Alb., MG and SC use when the subject topic is of the non-ratified type – new, resumed, contrastive, etc. Green (1980: 591), speaking of English quotation inversion, notes that ‘new characters, change of scene, or other relevant facts may be introduced into a narrative in quotation frames, and here it seems unnatural to use a preposed quote’. What is done instead in Alb., MG and SC (as in English) is to put the verb-subject complex, in SV order, in front of the quote. These languages thus exploit the other possibility of their word order systems to express narrow focus, placing the narrow focus expression at the right edge of the focus domain, clause-finally. In other words, when the utterer of the quoted passage is treated as a non-ratified topic, the subject and the verb (in this order) precede the quote; the subject thus occupies the slot for non-ratified topics, whereas the verb and the quote build a focus domain with a narrow focus interpretation on the quote. The syntactic pattern is:

[subject]_{NON-RATIFIED TOPIC} [verb [quote]_{NARROW FOCUS}]_{FOCUS DOMAIN},

or, in terms of presuppositional structure only:

[[subject]_{NON-RATIFIED TOPIC} verb]_{PRESUPPOSED PROPOSITIONAL FUNCTION} [quote]_{NARROW FOCUS}.

Example (8-58), being the first sentence in an article, makes both the syntax and the pragmatics of this type perspicuous:

(8-58) MG (Avji, p. 4)

Stin epistoli ... pros ton proedro tis Vulis, o proedros tu ΔS tu SEVE

in-the letter to the president of-the parliament the chairman of EC of APNG

k. Kiriakopulos anaferi: “Kirie proedre...”

Mr. K. refers “mister president”

‘In a letter to the President of the Parliament, the Chairman of the Executive Committee of the Association of the Publishers of Northern Greece, Mr. Kyriakopoulos, says: »Dear Mr. President...«’

Both the subject and the setting adverbial convey information which, although topical, is not yet ratified as such by the interlocutors, since the sentence represents the introduction to the article, which is reflected by the S-V-quote order. Quotation inversion could be used in such a

acceptable. I am aware of two such cases, Latin and Ancient Greek. Compare the following Ancient Greek sentence (Plato, Phaedo, 60a):

kai ho Sōkratēs ... “ō Kritōn”, ephē, “apagetō tis autēn...”

and the Socrates “o Crito” said “may-take-away someone her”

And Socrates said: “Crito, someone should take her away from here...”

context only with a strong presuppositional pretense (4.2.1.). The same structure is also used in dialogues, either as the opening sequence, or when the reactive nature of communication (speaker A reacts to the utterance of speaker B) is to be highlighted. The latter case is illustrated by the following example. The character called Leader says to the character called Novelist that he is ideologically an enlightener:

(8-59) Alb. (Camaj, p. 56)

I acaruem prej fjalës ‘vëzhgues’, Novelisti u përgjigj: “Unë jam mësue...”

the stung from word ‘enlightener’, Novelist-the PTCL answered: ‘I AUX learned’

‘Upset by the word ‘enlightener’, the Novelist answered: »I was taught ...«’

Novelist reacts to the previous utterance of his interlocutor. That his speech is to be understood as a reaction is made explicit not only through the depictive predicate *i acaruem...*, but also through S-V-quote order. Quotative inversion could be employed in such a context, but the contrastive character of the topic ‘Novelist’ and the reactive nuance would be lost.

In both quotation inversion and S-V-quote order, it is the quote that is focused, whereas the rest is presupposed, with ratified or non-ratified topics. If something else – the utterer, the manner of speaking, the spatio-temporal coordinates of the speech act – is to be focused, the only strategy available in the Balkan languages is to use two separate clauses, one for the quote and one for the description of the speech act itself. One very popular mechanism to put the utterer in focus is to resume the quote with a pronominal object and embed it as a topic in a vS sentence:

(8-60) SC (Vesti, p. 11)

“Izbegličke logore... je koristio NATO.” To tvrdi Roni BRAUMAN...

“refugee camps AUX used NATO.” this claims R.B.

‘»The refugee camps ... were used by NATO.’ This is claimed by Ronny BRAUMAN...’

The quotation inversion construction is thus restricted to the contexts in which the speaker is a ratified topic, with the speech act itself under the scope of presupposition. If any other kind of information structure is required, the quote cannot remain preposed – it is either relegated to the right periphery of the focus domain or treated as a separate clause, depending on the needs of the discourse.

8.1.5. Discourse-pragmatic differences between Alb., MG, and SC

This chapter is devoted to that kind of crosslinguistic variation which is not to be explained structurally, but is rather to be attributed to language-specific discourse-pragmatic conventionalizations. Not surprisingly, the number of differences of this sort in a construction which is so extremely restricted semantically and discourse-pragmatically is not very large. Actually, I was able to pin down only one, which is observable already on the level of statistics, as represented in Table (8-5): Whereas the proportion of quotation inversion clauses within the group of VS clauses in Alb. and SC amounts to approximately 20%, it is less than 10% in MG. As already indicated, this discrepancy is partly due to the composition of the corpus itself: my MG texts contain somewhat less quoted material than the Alb. and SC ones. Apart from this, however, there are two discourse conventions in MG which slightly deviate from Alb. and SC and seem to be partly responsible for the difference in the statistics:

(a) Retold dialogues and monologues in Alb. and SC are more often than not interrupted by quotation inversion clauses reminding the reader of the speaker and the circumstances of the speech act reported. In MG, although the former strategy is of course possible and frequently used, one often encounters longer stretches of quoted direct speech without any explicit indication that it is a quote. For instance, in a Greek newspaper, a journalist reporting from a press conference is more likely to indicate that the speech quoted was uttered by a certain person at a press conference only at the beginning of the article and to continue quoting without reminding the reader of the speaker, than in an Albanian or a Croatian/Serbian newspaper. In the latter, it is rather to be expected that an interruption of the quote with a quotation inversion clause will occur every ten to twenty clauses.

(b) MG, as indicated in 5.1. and 5.5., tends to leave its subjects unexpressed much more often than Alb. and SC. This tendency is confirmed by quotation inversion clauses: in many cases, where Alb. and SC would have a quotation inversion clause, MG has only a verb of saying, etc., with an unexpressed subject. Thus, in the parallel translations of Jack London's "White Fang", in quotative contexts, there are 24 cases of unexpressed subjects in MG, one in Alb. and none in SC. Or, the other way around, there are 18 cases in which Alb. and SC have a quotation inversion, and MG only a verb of saying with an unexpressed subject. This is especially frequent with pronominal subjects. There are only 11 (4.1%) quotation inversion clauses with pronominal subjects in MG in my whole corpus, as opposed to 62 (12.6%) in Alb. and 33 (8.9%) in SC. Here is a characteristic example:

(8-61) Eng: *How many cartridges did you say you had left?*, **he asked**. (London, p. 8)

MG: *Poses ipes oti su apomenun?*, **rotise**. (p.12)

asked

SC: *Koliko ti je još ostalo metaka?*, **upita on**. (p.10)

asked he

Alb: *Sa fishekë të kanë mbetur?*, **pyeti ai**. (p.8)

asked he

These two discourse conventions, together with a slightly unbalanced corpus, account for the relatively low percentage of quotation inversion clauses in MG when compared to Alb. and SC. In all other respects, the discourse pragmatics of the construction is identical in all three languages.

8.1.6. Conclusion

Quotation inversion is a type of focal inversion, albeit a lexically and syntactically very specific one. The most important insight it offers lies, notwithstanding the spectacular nature of the semantic effects described in 8.1.3., in syntax: since the construction has a grammaticalized all-ratified presuppositional structure, with subjects (and other, optional, elements) as ratified topics, it is representative of the syntactic behavior of the presupposed, all-ratified material in focal inversion in general.

When the cataphoric adjustment and the heaviness principles are not at work, the placement of independent elements is a matter of free variation, whereas dependent elements show slightly diverging tendencies, with Alb. and MG tending to posit them between the verb and the subject, and SC to keep the verb-subject complex together. Furthermore, SC has, in contrast to Alb. and MG, a strong tendency to place pronominal subjects as close to the verb as possible. In all three languages there is a tendency for the elements that are semantically close to the verb to be immediately adjacent to it, though the slots for these elements are not identical in all three. Finally, all three languages display a strong immediate adjacency constraint holding between the focused preposed element, the quote, and the verb. Keeping this structure in mind, we may now turn to the informationally more complex types of focal inversion, wh-inversion and fronted focus inversion proper.

8.2. Wh-inversion

The second type of focal inversion, which I am going to call **wh-inversion**, occurs in constituent questions after clause-initial question words carrying grammatical relations other than subject ('whom', 'to whom', 'where', 'why', 'how', etc.)¹:

(8-62) Eng: *Matt, how much is a good sled-dog worth?* (London, p. 141)

Alb: *Mett, **sa** kushton një qen i mirë slite?* (p. 138)

M. how-much costs a dog the good of-sled

MG: *Mat, **poso** aksizi ena kalo skili ja elkiθro?* (p. 171)

M. how-much is-worth a good dog for sled

SC: *Mate, **koliko** vrijedi jedan dobar pas za prezanje u saonice?* (p. 158)

M. how-much is-worth one good dog for harnessing in sled

As in other types of focal inversion, the postposition of the subject has to do with the immediate adjacency constraint holding between the fronted element (question word) and the verb, and with the informational status of the subject itself. Compared to quotation inversion, however, where the constraint is absolute in all three languages, wh-inversion represents a domain of a much greater variation, both within and across languages, as will become patent in the discussion that follows. Furthermore, in contrast to quotation inversion, since wh-inversion clauses do not have a monotonous information structure, their word order and intonation are at least partly determined by the presuppositional structure.

Now let us take a look at the statistics. Table (8-63), organized according to the same principle as Table (8-5), contains the relevant data:

(8-63) *Wh-inversion in the corpus*

	1. original	2. translat.	3. total	4. whI vs. VS	5. whI vs. SV/VS	6. whI vs. Σ
Alb.	149	31	180	180 vs. 2225 8.1%	180 vs. 7979 2.3%	180 vs. 15813 1.1%
MG	95	18	113	113 vs. 2673 4.2%	113 vs. 6641 1.7%	113 vs. 14463 0.8%
SC	38	11	49	49 vs. 2065 2.4%	49 vs. 8458 0.6%	49 vs. 14176 0.3%

¹ The anglocentric label *wh-* is used for the simple reason that I could not find a better name: *question inversion* would be misleading because of the possible confusion with yes-no questions (see 6.5.1.1.), and the employment of the initial letters of the question words in Alb, MG and SC would result in the impossible abbreviation *k-/c-/ç-/t-/p-/š-/g-inversion*, which I decided to spare both the readers and myself.

The small percentage of *wh*-inversion clauses in SC is straightforwardly explained by the optionality of inversion in this language (8.2.2.); Alb. shows a disproportionately high presence of the construction due to a number of semantic and discourse-pragmatic factors to be investigated in 8.2.4. It is interesting to note that, in comparison to Alb., the percentage of *wh*-inversion clauses in MG is relatively low within the group of VS clauses (column 4), relatively high within the group of clauses with expressed subjects (column 5), and sinks once again against the background of all the clauses in the corpus. The last fact has to do with extensive subject-dropping in MG, the first with the high frequency of some other VS types in this language (cp. Section 6.1.). The situation in the parallel translations is revealing: out of 36 sentences with *wh*-words in the English original, Alb. has 16 *wh*-inversions, MG 12, and SC only 5 (the rest of the original *wh*-clauses being covered by dropped subjects, the clitic copula and non-inversion in SC, prefocal topics in Alb. and MG, and diverging translations). On the other hand, Alb. has the strongest tendency to translate non-clausal elements (mostly participial complements) with dependent *wh*-clauses and to add clauses not present in the original, though such cases of addition are also, if seldom, found in MG and SC.

All this leads to a picture in which Alb. has a lot of *wh*-inversions, MG a lot of constituent questions with zero subjects, and SC a lot of non-inverted (i.e. SV) clauses following a question word. The prototypical state of affairs is illustrated by the following example:

(8-64) Eng: *How'd it happen?* (London, p. 14)

Alb: *Si ndodhi kjo?* (p. 15)

how happened that

MG: *Pos ejine?* (p. 20)

how happened

SC: *Kako se to dogodilo?* (p. 18)

how REFL:CLIT that happened

8.2.1. The pragmatic role of question words

Syntax and pragmatics of question words have been more than excessively scrutinized in the past forty years, primarily within the generative framework, so that even the Balkan languages have been discussed in this context (see Turano 1994, Tsimpli 1998 and Bošković 2002 and the literature cited therein for Alb., MG and SC, respectively). A line of research which treats question words as a subclass of focus expressions and the question word fronting as a kind of

focus fronting, originating with Horvath (1985), É. Kiss (1987) and Brody (1990)¹, has been particularly successful, having achieved the status of an axiom in the past decade. Although I generally agree with this theoretical tenet (see below), there are some important differences between my treatment of the notion of focus and the one current in generative descriptions. For me, focus is a discourse-pragmatic notion which may be explicitly coded in syntax and phonology (or morphology, for that matter; see 4.4.); for generative linguists, it is a syntactic feature which has to be checked by a particular functional head. Practically, this means that my task is to demonstrate the plausibility of the equation of question words with focus expressions and then to look for the ways the pragmatic configurations with question words are encoded, whereas generative linguists are mostly interested in the derivational history of *wh*-sentences, taking the focus feature of question words for granted, or even divorcing it from its discourse-pragmatic roots.

What kind of focus expressions are question words, and what kind of evidence do we have for their focushood? First the latter question. In great many languages, narrow focus expressions and question words consistently show up with identical or near-identical syntactic and prosodic features (see Raymond and Homer 1996 for a survey). In a number of languages, including Hungarian, Turkish, Bengali, Korean, Basque, etc., the identity of the two expression classes is absolute, both syntactically and prosodically (Horvath 1985, É. Kiss 1987, Kim 1988, Ladd 1996: 171, Lambrecht and Michaelis 1998: 511ff). In another group, represented by Romanian (Ladd 1996: 171ff), the prosody is the same, though the syntax only partially overlaps. In yet another group (Alb, MG and SC, together with English, German, Russian, etc.), the syntax overlaps partially, as in the previous group, whereas the prosody is different, with focus expressions carrying the sentence accent and question words being inherently deaccented (Ladd 1996: 170ff, Lambrecht and Michaelis 1998). On closer inspection, there are even more differences between ‘normal’ focus expressions and question words, at least in the two latter groups of languages: all question words must be preposed, but only some focus expressions may be so; a clause can contain only one focus expression, but more than one question words are always allowed (*Who read what?*); etc.

The best way to account for these differences is to try to answer the first question raised above, concerning the kind of focus question words convey. At this point (as in many other points in this chapter), I shall follow Lambrecht and Michaelis (1998)². First, question words

¹ A similar conclusion was reached within a functionalist-typological paradigm by Kim (1988).

² For a different kind of explanation, based on the assumption that both foci and question words are operator-like elements, the latter being individual, the former quantificational, see Tsimpli (1998).

of necessity convey narrow focus, mostly of the identificational or contrastive type (see 4.4.2. for these notions): in order to ask a question about an element of a situation, one has to presuppose the existence of the situation itself (therefore narrow focus), and it is more often the case than not that the number of alternatives expected to be eliminated by the focus is limited by the context and the common ground of the interlocutors. Thus, in order to ask *Who read this book?*, I have to presuppose that someone read it (otherwise the question makes no sense), and it is pragmatically typical that I have in mind a list of candidates who might have performed the unusual act of reading a book.

Mutatis mutandis, this holds true also for lexically expressed foci in declaratives. The problem is that at the first blush it is difficult to conceive of questions as (pragmatic) assertions. If one understands assertion as the creation of a new proposition in the hearer's mind conducted by the speaker, then constituent questions do not carry assertions, since they contain only presupposed material and an element without intension, the question word. If a non-intensional element is the carrier of an assertion, then this assertion is empty, i.e. nonexistent. Now, I have defined communication in a slightly different fashion (Section 4.1.): the essence of assertion is a change in the state of knowledge of the interlocutors brought about by uttering a proposition, through the search-and-eliminate principle, as described in Sections 4.1. and 4.4. In declaratives, this is achieved by searching all the possible worlds according to the instructions contained in the presupposed material and by eliminating alternatives to the focused portion of the proposition, leaving thus only the proposition in which presupposed material is brought into relation with the denotatum of the expressed focus. In constituent questions, the presupposed material (i.e. the proposition minus the questioned element) also delimits the set of possible worlds the speaker refers to, but the non-intensional nature of the focused element instructs the hearer that the speaker is not able to eliminate any of the possible worlds evoked through the presupposition. What the hearer can infer from this is that the speaker wants to know which possible worlds may be eliminated, or, the other way around, what is the identity of the focused element in the situation described by the uttered proposition. The rest is the work of pragmatic principles (and intonation, which in questions generally signals the incompleteness of the expressed proposition). As Lambrecht and Michaelis (1998: 513) put it, '... utterance of a wh-question pragmatically asserts the desire of the speaker to know the identity of the referent inquired about via the wh-expression'. Thus, question words do function as narrow focus expressions, albeit of a rather specific sort.

Now let us see what is to be done with the formal differences between focus expressions and question words enumerated above. Why is it the case that all (non-echo) question words, but not all narrow focus expressions are fronted? As far as languages of the Balkan type are concerned, the answer is straightforward (and not particularly new – see Tsimpli 1995, É. Kiss 1998b). The sentence-initial focus slot houses narrow focus expressions which are closer to the contrastive than to the additional end of the scale given in 4.4.2; all question words convey this type of focus (due to the real world conditions under which they are usually uttered), but not all lexical focus expressions do. Consequently, it is all question words, but only some lexically expressed foci that are fronted (cp. 8.3.2.). Furthermore, question words are in a way ‘grammatical’ words (as opposed to ‘lexically full’), and it is a well-known fact that such words tend to have idiosyncratic syntactic properties, for instance a firmly grammaticalized sentence position.

The second problem is the deaccentuation of question words in some languages, including Alb., MG and SC. There are at least two solutions to this. First, the one based on economy: since question words are inherently focused, there is no need to mark them prosodically as foci. The second, proposed by Lambrecht and Michaelis (1998: 513ff), is based on the function of sentence accents: according to L&M, sentence accents instruct the hearer that an additional effort is needed in order to process the accented clause element, either because it conveys a non-ratified topic or because it conveys focus. Since in the case of question words the speaker does not commit her/himself to the identity of the referent of the expression, the effort required in matching lexical descriptions with specific referents is simply not present. This can, in some languages, lead to the grammaticalized deaccentuation of question words.

It is for the third problem, multiple question words in one clause, that I cannot offer a straightforward account. A number of explanations has been proposed (for a short assessment, see Erteshik-Shir 1997: 181ff). The only one I find intuitively appealing is the explanation proposed by Erteshik-Shir (1997: 180ff, 186ff), based on the observation that multiple questions allow only for paired-list readings (*pace* Bošković 2002: 357ff). Roughly, she claims that in a sentence like *Who bought what?*, the first question word is a topic expression, defining a contextually determined restrictive set. For each individual in this set, a new constituent question is asked. Thus, if *who* defines the set {Peter, Mary}, then the question *Who bought what?* is a shorthand for *What did Peter buy?* and *What did Mary buy?*. The only problem with this explanation is that it is unclear through what mechanism nonreferential and nonintensional expressions receive the purported topical interpretation. Be it how it may,

since multiple questions are a marginal phenomenon (both statistically and pragmatically) in the languages under consideration, I shall leave the question of their interpretation open.

8.2.2. Syntax, prosody, and information structure

In the chapter on quotation inversion, the syntax and the information structure were dealt with in separate sections, being for a greater part independent from each other due to the uniform discourse-pragmatic function of quotation inversion clauses. Since the presuppositional structure of constituent questions is not uniform, and since much of the word order and intonational variation depends on different presuppositional structures a wh-clause can be ascribed, both areas have to be described simultaneously.

8.2.2.1. General rules

The major difference between quotation inversion and wh-inversion lies in the fact that the former is obligatory in all three languages, whereas the latter exists as a syntactic rule only in Alb. and MG. This, however, does not imply that there is no wh-inversion in SC, but rather that it is a pragmatically triggered (as opposed to purely syntactic) phenomenon in this language. To make this statement somewhat clearer: In Alb. and MG there exists a fairly strict syntactic rule of immediate adjacency between the question word and the verb; in SC, there is no such rule, as witnessed by (8-64). The fact that wh-inversion nevertheless occurs also in SC, as witnessed by (8-62), has to do with some discourse-pragmatically driven word order permutations.

Yet another difference is that the immediate adjacency rule for wh-inversion in Alb. and MG seems to be somewhat weaker than in the case of quotation inversion. Independent elements, primarily sentence adverbials and setting adverbials, may sometimes intervene between the fronted element and the verb:

(8-65) Alb. (Koha Ditore, p. 7)

... *ndëgjoni për çka vërtet do të angazhohet LDK-ja ...*

listen for what really FUT:CLIT SUBJ:CLIT engage-itself LDK-the

‘...listen what the LDK is really going to work on...’

(8-66) MG (Elefterotipia, p. 12)

... *eθese to erotima jati eos tora ðen exi efarmosti o nomos.*

put the question why till now not AUX applied the law

‘... (he) asked why the law has not been put into effect by now.’

Although theoretically possible, this type of ordering is extremely infrequent (I have found only three examples), alternative orderings with postposed adverbials being generally used instead. Since cases like this represent extremely rare exceptions to the rule, I feel justified in ignoring them in what follows.

The same holds for the examples of non-inversion in Alb. and MG. Namely, when the question word is not an object of the verb, i.e. when it does not carry any of the major grammatical relations but is, say, a locational or a modal adverbial or a PP instead, the inversion does not have to take place (see Horrocks 1983 and Mackridge 1985: 236 for MG). Thus, in MG, the sentence **ti o Janis γrafi* (‘what the Ioannes writes’) is ungrammatical, whereas *pote o Janis eyrapse* (‘when the Ioannes wrote’) is acceptable. The adjacency constraint holding between the question word and the subject thus seems to be absolute only when the question word is an object (or a subject, but this case is of no interest here). If it carries any other syntactic function, there appears to be no such constraint. On the other hand, the cases of non-inversion with non-object question words are extremely rare. I do not have a single example in my Alb. and MG corpora (each containing some 15000 clauses). Furthermore, my informants unanimously judge variants with inversion better, less ‘marked’, more natural, etc. than those without it: *pote eyrapse o Janis, kur shkroi Djoni* (both: ‘when wrote John’) are better and much more frequent than *pote o Janis eyrapse, kur Djoni shkroi* (both: ‘when John wrote’).

On the basis of this evidence, I should like to reformulate the adjacency constraint for question words in Alb. and MG: When the question word is the object (or the subject), it must be immediately followed by the verb; when it has any other syntactic role, it is preferably followed by the verb, though intruding elements are possible. Since non-adjacency is highly marked and infrequent, I shall, for the lack of a better solution, ignore it in what follows.

The question word itself has to be fronted¹, standing in the sentence-initial position proper, the only elements allowed to precede it being certain kinds of topics – more often in Alb. and MG than in SC (see 8.2.3. and 8.2.4.). The question word never carries the sentence accent, which is therefore always on one of the following elements. However, in contrast to quotation inversion, the position of the stress is pragmatically determined and therefore not predictable (see 8.2.2.). The preliminary scheme of wh-inversion thus looks as follows:

¹ I leave echo-questions aside, primarily because they are in all three languages much more often formed with a fronted question word than with this element *in situ* (see Bošković 2002: 355).

(8-67) **Alb., MG: ([topic]) [wh_{deaccented}]~[verb] [subject]**

SC: ([topic]) [wh_{deaccented}] [subject] [verb] / [verb] [subject]

As in the case of quotation inversion, wh-inversion clauses are usually very short, consisting of the question word, the verb and the subject. In my corpus, there are 127 (70.6%) clauses of the form [wh][v][s] in Alb., 75 (66.3%) in MG, and 41 (83.7%) in SC. The two templates given in (8-67) cover this kind of constituent question entirely.

As far as the position of other elements of the construction is considered, wh-inversion shares the two processing principles – cataphoric adjustment and heaviness principle – with quotation inversion. Non-verbal elements which are syntactically relevant in the following discourse or significantly heavier than other elements tend to come last. These tendencies are observable in all kinds of constituent questions.

Since all other rules and tendencies depend on the information status of the elements following the question word, they will be dealt with according to this criterion. Before that, a clarification is in order.

How can the information structure of constituent questions vary if the question word is narrowly focused, with the rest of the sentence denoting a presupposed propositional function? Recall that topics may be ratified and non-ratified, depending on the speaker's assumptions on the hearer's capability and readiness at the given moment to take a certain topic for granted. The more prominent the topic referent in the context/common ground, the greater the probability that the topic will be treated as ratified (cp. 6.2.)¹. Now, in quotation inversion, topical subjects (and other elements) are, due to specific pragmatic and syntactic restrictions holding for this construction, obligatorily ratified, so that quotation inversion clauses (minus quote) are invariably all-ratified. Constituent questions, being free of these restrictions, can contain both ratified and non-ratified elements within the presupposed material following the question word. The variation of the information structure in this clause type thus concerns the ratified versus non-ratified status of the elements of the clause.

8.2.2.2. All-ratified constituent questions

Let us first consider the cases which contain only ratified topical elements, as in quotation inversion. Since the line between wh-inversion and non-inversion in SC is drawn according to

¹ I borrow both the terminology and the analytical apparatus from Lambrecht and Michaelis (1998). Ladd (1996), whose analysis in many respects resembles mine, speaks of broad focus of the question where I speak of all-ratified questions and of narrow focus where I speak of non-ratified elements.

the number of constituents in the question, I shall both in this and in the following sections present constituent questions with three elements, the question word, the verb and the subject (henceforth *short questions*) separately from those which contain more material (henceforth *long questions*). Consider the following examples of short questions¹:

(8-68) Alb. (Koha Ditore, p. 2)

... *si do të PËRFUNDONTE ky kontest gjyqësor, ...*

how FUT:CLIT SUBJ:CLIT ended this procedure judicial

‘[He tried to stop the building of a church within the University campus by bringing the whole issue to the court. He knew] how this trial was going to END, [but he wanted us to have documents of the misuse of the court.]’

(8-69) MG (Fakinou, p. 22)

... *ti tha tis EÐINE afto to epangelma?*

what FUT:CLIT her:CLIT gave that the profession

[Ioanna’s mother is opposed to her wish to become a painter. ‘What kind of future will Ioanna have as a painter? And if something goes wrong,] what can that profession GIVE her?’

(8-70) SC (Kiš, p. 141)

Šta RADI Novski?

what does N.

[Novski was sentenced to a long penalty in the Gulag. ‘In those days, the following anecdote went round in Moscow:] “What does Novski DO?” [“He drinks tea with red currant jam and plays »The Internationale«.]’

As the context reveals, we are dealing with a highly ratified topical material in all three cases: to take someone to the court means to accept the relevance of the existence of a trial, which is going to have some outcome, and ‘the profession’ and ‘Novski’ are activated discourse topics. In all three languages, all-rated short questions have the syntax of quotation inversion: they are accented on the verb, presumably with a low tone, L* (e.g. *si do të PËRFUNDONTE ky kontest*)², the subject follows the verb, the verb and the fronted element are immediately adjacent.

¹ In all examples in this and the following sections, the sentence accent will be marked (with small caps) both in the original sentences and in the English translations. Note that the place of accent in Balkan languages does not always correspond to that in English – a problem that will be addressed in 8.2.2.5.

² The phonological interpretation of the sentence accentuation of constituent questions given here is based on Ladd 1996 (see especially pp. 212ff. for the analysis of MG questions).

However, in contrast to Alb. and MG, SC allows for non-inversion in this context:

(8-71) SC (Vijesnik, p. 5)

Kuda ovaj tramvaj VOZI?

where-to this tram drive

[‘Some of them get on the tram, the others, uncertain, turn around and ask loudly:]

Where is this tram GOING?’

Albeit relatively rare (only five examples in my corpus), the pattern without inversion in this context reveals that there is a significant difference between SC and Alb./MG. Whereas the latter have a relatively strong rule of adjacency between the question word and the verb, in SC this rule does not exist, so that inversion has to be attributed to some other factor. Even better, there is a sentence type in SC where inversion almost never takes place, namely that with pronominal subjects, as witnessed by (8-64). The tendency for pronominal subjects to be preposed is strong, so that wh-inversion in SC is generally not allowed when the subject is pronominal. The basic pattern is thus¹:

(8-72) **all-ratified short questions**

Alb, MG: [wh]~[VERB] [subject]

SC: [wh] [VERB] [subject] or [wh] [subject] [VERB]

[wh] [pronominal subject] [VERB]

In long all-ratified questions, the differences between Alb. and MG on one hand and SC on the other become more evident:

(8-73) Alb. (Bishqemi, p. 88)

Si po i RUANTE për ne pikat e fundit të ujit komandant Ylli!

how PTCL:CLIT them:CLIT saved for us drops the last-of-the water commander Y.

[Commander Yll saved the last drops of water for the children. ‘Thirst kills faster than hunger. I didn’t know that, and now I shall have to pay dearly for that lesson.] How (nice that) commander Yll saved the last drops of WATER for us!

(8-74) MG (Matesis, p. 32)

Ti ØES eđo esi?

what want here you

[Mrs. Kanello, collaborating with the Resistance during the Second World War, leaves messages for her comrades in a public toilette for gentlemen. ‘Once, an old man got confused and asked her:] ‘What are you doing here? [Are you a man?’]

¹ Heavy subjects may carry a secondary, postnuclear sentence accent, in the same way as quotation inversion heavy subjects do (see 8.1.1.).

(8-75) SC (Vesti, p. 31)

Kako je deset skupih zvezda IZGUBILO od šestorice “Djetića”?

how AUX:CLIT ten expensive stars lost from six Montenegrinian-chaps

[A press conference after a basketball match between the Greek club ‘Panathenaikos’ and the Montenegrinian ‘Budućnost’. ‘An unusual question was asked by a young colleague:] How come that ten expensive basketball stars lose a match to six »Montenegrinian chaps«?’

Alb. and MG display the same syntactic behavior as in quotation inversion and in short all-ratified questions; SC generally does not have inversion in this configuration. First Alb. and MG: as witnessed by the examples above, there is a tendency for the elements other than the verb and the subject to stand between the two, although this is not obligatory, just like in quotation inversion. SC, as indicated, does not invert in long all-ratified questions, although wh-inversion sometimes occurs (three examples in my corpus), but only when the additional element in the question is a setting adverbial or a sentence adverbial. Here is an example:

(8-76) SC (Kiš, p. 174)

... šta je zapravo REKAO Gorki, ...

what AUX:CLIT really said G.

[There were rumors that Gorky said something not very laudable about the poet Darmolatov. ‘It is difficult to state] what Gorky actually SAID, [but he seems to have alluded to Darmolatov’s indomitable temper.’]

The intonation pattern is, in contrast to the syntax, the same in all three languages: the sentence accent is carried by the verb. The basic patterns are:

(8-77) **all-ratified long questions:**

Alb., MG: [wh]~[VERB] [X] [subject]

SC: [wh] [subject] [VERB] [X]

To conclude: When all the elements following the question word are ratified, i.e. when the hearer is considered to be able to take them as presupposed without further effort, the syntax and the prosody of wh-inversion in Alb. and MG are identical to those of quotation inversion. The situation in SC is somewhat more complex. In short questions, the inversion usually applies, with the exception of pronominal subjects, and its syntax and prosody are the same as in quotation inversion. In long questions, the inversion usually does not apply, although the prosody remains identical to that of the prototypical focal inversion, with the verb carrying the sentence accent. In both short and long questions, the opposite of usual may occur, i.e. non-inversion in the short and inversion in the long questions, but this is exceptional.

8.2.2.3. Constituent questions with non-ratified elements

To use a non-ratified element in a question means to give the hearer approximately the following instruction: ‘contrary to your expectation, and even though you did not have this in mind, I would like to know the identity of an element in *this* situation rather than in some other’. Let me illustrate this with an example borrowed from Lambrecht and Michaelis (1998). In the context where I say ‘I went to the mall with Audrey yesterday.’, you may ask *What did AUDREY buy?*. The message you are sending me in uttering this is that you count on the fact that I have expected the question about buying (therefore no accent on *buy*), but that I might have been uncertain whether it was my buying or Audrey’s buying you would be interested in. By accenting *Audrey*, you establish the topic you consider to be not absolutely uncontroversial (since it could have been me you were interested in), i.e. you *ratify* it. Note that both the ratified and the non-ratified elements are of necessity presupposed. In order to ask *What did AUDREY buy?*, you have to count on my readiness to take for granted that Audrey did buy something; the difference lies merely in the *grade* of certainty as to my readiness to take an element of the proposition for granted¹.

In short questions, either the verb or the subject can be non-ratified. Let us first deal with the non-ratified subjects. Consider the following sentence pair:

(8-78) Alb. (Kosovarja, p. 6)

... *ku do të jetë babai i TYRE.*

where FUT:CLIT SUBJ:CLIT is father-the the their

[Albanian prisoners are coming back home, having been released from the Serbian prisons. But not all of them are back: ‘One brother’s children cannot

¹ Different accentuation patterns in constituent questions have been a puzzle for linguists for quite a long time (for an overview, see Ladd 1996: 170ff. and Lambrecht and Michaelis 1998: 482ff), one of the most prominent theories being that the accented word is the second focus of the question (Culicover and Rochemont 1983, Ladd 1996, and if I understand her right, Tsimpli 1998). Under the assumptions on focus embraced in this study (Section 4.4.), the presence of a second focus in a clausal unit is impossible, so that this account is to be ruled out already on purely theoretical grounds, notwithstanding its counterintuitive nature. As L&M note, the second-focus theory is based on what they (1998: 478) call *assumption of iconicity*, according to which, roughly, everything that is accented has to be focus. As amply demonstrated by Lambrecht (1994), this is an erroneous belief, since the role of the sentence accent is much more diversified than the one-to-one focus-to-accent theory is able to predict.

sleep, overwhelmed by happiness, whereas the children of the other brother, who is still missing, wonder] where THEIR father could be.’

(8-79) Alb. (Bishqemi, p. 69)

... *ku NDODHEJ pafla e kërkuar.*

where finds-itself tile-the the required

[Two boys are playing a game of finding numerically marked fields on a large table. ‘One of us would say what the coordinates of the field are: »A4!«; – »Blue«, the other one would say, having found the color] and where the required field was LOCATED.’

In (8-78), the subject referent is obviously contrastive (‘their father’ as opposed to ‘other fathers’), and therefore non-ratified. The result is the movement of the sentence accent: it is placed on the non-ratified element, i.e. on the subject¹. The similarity in syntax ([wh]~[v][s]) and the difference in prosody between non-ratified subject clauses and all-ratified clauses is obvious when a short question with a very similar lexical filling, but with no non-ratified elements, like (8-79), is compared: here, the sentence accent falls on the verb. Even in SC, which otherwise allows for some word order variation even in short questions, this pattern – wh-inversion with the accent on the subject – seems to be the only pattern occurring in this context. The scheme is therefore:

(8-80) **short questions with non-ratified subjects**

Alb., MG, SC: [wh]~[verb] [SUBJECT]

It was difficult to find an instance of non-ratified verbs in short questions in written texts, so that I had to elicit them (the MG example stands for both MG and Alb., since both languages behave identically). The context is as follows: A teacher speaks to a colleague about the journey to Greece to which he took his class. In the course of the conversation, it turns out that they have not visited Athens, Thebes, Thessaloniki, etc. Enervated, the colleague asks:

(8-81) MG (elicited)

Pjes polis iðAN ta peðja?

what cities saw the children

(8-82) SC (elicited)

Koje gradove su deca VIDELA?

what cities AUX:CLIT children saw

‘What cities DID the children see?’

¹ I have the impression that the tone is generally still low (L*; see also Ladd 1996: 212ff. on MG), but in the pronunciation of some of my informants it seemed rather high.

The verb ‘see’ stands in contrast to ‘not see’ and is therefore non-ratified. Alb. and MG on one hand and SC on the other behave differently: In Alb. and MG, the syntax and the prosody are the same as in all-ratified short questions (inversion, the accent on the verb), with the possible rise in the intensity and pitch of the sentence stress. In SC, the inversion is virtually impossible, the SV order with a somewhat raised pitch on the verb being the only structure licensed:

(8-83) **short questions with non-ratified verbs**

Alb., MG: [wh]~[VERB] [subject]

SC: [wh] [subject] [VERB]

When the verb or the subject are non-ratified in long questions, a comparable situation obtains. Here is an example for a non-ratified subject:

(8-84) MG (Matesis, p. 50)

... ti ipe sto rađjofono o kirios TSORTSIL.

what said in-the radio the mister Ch.

[The women in a Greek provincial town during the Second World War meet every evening to spend some time together. ‘All they talked about was] what Mr. CHURCHILL had said in the radio.’

In all three languages, non-ratified subjects are placed last in the sentence, after both the verb and the additional element(s), carrying the sentence accent:

(8-85) **long questions with non-ratified subjects**

Alb., MG, SC: [wh]~[verb] [X] [SUBJECT]

When it is the verb that is not ratified, SC again behaves differently (the context is the same as for (8-81) and (8-82)):

(8-86) MG (elicited)

Pjes polis iđAN ta peđja stin ekđromi?

what cities saw the children in-the excursion

(8-87) SC (elicited)

Koje gradove su deca na izletu VIDE LA?

what cities AUX:CLIT chidden on excursion saw

‘What cities DID the children see on the excursion?’

In Alb. and MG, the situation is identical as in short questions with non-ratified verbs. In SC, the verb tends to be placed at the very end of the sentence, with the subject and the other elements preceding it:

(8-88) **long questions with non-ratified verbs**

Alb., MG: [wh]~[VERB] [subject] [X]

SC: [wh] [subject] [X] [VERB]

The last possible combination in long questions, the one in which an element other than subject or verb is non-ratified, is illustrated in the following examples:

(8-89) Alb. (Bishqemi, p. 78)

... si mund të na vinin në ndihmë njerëzit tanë në thellësi KAQ të madhe?

how can SUBJ:CLIT us come in help people-the our in depth so the big

[A group of people is stuck in a cave deep under the earth. ‘We cannot call our people to help us because the radio doesn’t work. And even if it worked,] how could our people come to help us at such a DEPTH?’

(8-90) MG (Elefterotipia, p. 12)

Se pia posotita tha ftasi i sinepia tus istera apo merika XRONIA.

to what percent FUT:CLIT come the adherence-to-principles their after some years

[The Greek Socialist Party proposed to keep the state ownership of the telephone company at the level of 34%, explaining this with their adherence to principles. ‘One may wonder:] What percentage will their adherence to principles reach in a couple of YEARS?’

(8-91) SC (Vesti, p. 29)

... kako će košarkaši objasniti ovaj poraz UPRAVI kluba.

how FUT:CLIT basketball-players explain this defeat to-management of-club

[‘It will be easy to find some excuse for the fans, but the real problem is] how the basketball players will explain this defeat to the club MANAGERS.’

In all three cases, the last sentence element has to be ratified by the interlocutors: in the Alb. example, the proposition ‘our people come to help’ is already present, what has to be added to it is ‘at such a depth’; in MG, it is clear that the adherence to principles reaches certain percentages, but what has to be ratified is that this situation will also obtain in a couple of years; in SC, it is the contrast between ‘the fans’ and ‘the club managers’ that is responsible for the non-ratified status of the latter. All three languages use the same strategy, placing the non-ratified element at the end of the sentence and assigning it the primary sentence accent.

The difference lies in the presence of wh-inversion: in Alb. and MG, it is obligatory in this context, as well; in SC, it is not used at all, SV order occurring instead¹.

(8-92) **long questions with a non-ratified X**

Alb., MG: [wh]~[verb] [subject] [X]

SC: [wh] [subject] [verb] [X]

8.2.2.4. Complex predicate forms and complex sentences

This section should shed some light on the relationship between the subject and the verb in constituent questions with complex predicate forms ([auxiliary + participle], [copula + predicative]) and in complex sentences ([matrix verb + complement clause]).

Since in SC the auxiliary and the copula are mostly clitic, it is only Alb. and MG data that are of interest with respect to complex predicate forms. As in quotation inversion, the AUX-PTCP and the COP-PRED complexes are in principle not separable (i.e. the subject never intervenes between them) and they occur in the firm order – first AUX/COP, then PTCP/PRED:

(8-93) MG (Ciao, p. 6)

Ti exi zitisi aloste o Janis?

what AUX asked-for otherwise the I.

[Ioannis prayed for the health of his children.] ‘What else did Ioannis ask for?’

There is one exception to this, however. When the question has the form ‘how + predicative + copula’ (as in *How old are you?*), the questioned predicative may be adjacent to the question word or separated from it by the copula, or by the copula *and* the subject, with the result which resembles what is called stranding in generative grammar. All three stages are shown in the following examples:

(8-94) Alb. (Koha Ditore, p. 2)

... sa demokratike dhe properëndimore janë ato ndryshime.

how democratic and pro-western are these changes

[‘We will have to see in what direction they move] and how democratic and pro-western these changes are.’

¹ There are cases where the non-ratified element, the subject or X, does not occur at the very right edge of the clause, but is followed by another element instead. The few instances I collected are all due to the heaviness principle (the element following the non-ratified element is relevantly heavier than it).

(8-95) Alb. (Koha Ditore, p. kult/8)

... *sa është i madh* produksioni i artit shqiptar në përgjithësi ...

how is the big production-the of-the art Albanian in whole

[Many great works of art are not mentioned here, which is only understandable in view of the fact] how big the Albanian art production is in general...?

(8-96) Alb. (Koha Ditore, p. 10)

Po *sa do të jetë* pushteti i ri në Serbi *realist dhe demokrat*?

PTCL how FUT:CLIT SUBJ:CLIT is power the new in Serbia realistic and democratic

[The future relationships on the Balkans depend on the new Serbian government.]

‘But how realistic and democratic will the new Serbian government be?’

All three types are synonymous: what is questioned is the degree of the presence of a certain property with the referent of the subject. The variation in the first two cases is, as far as I can see, free, and it is triggered by the special status of the predicative, which is syntactically bound both to the copula and to the question word. If the first syntactic relationship prevails, it stands after the copula (8-95). If the latter prevails, it is after the question word and *before* the copula (8-94). The third type has apparently to do with ratification: if the denotation of the predicative is still to be ratified, as is the case in (8-96), then the tendency to put non-ratified elements at the right edge of the clause may (but need not) override its syntactic dependencies and lead to the discontinuous picture we have in (8-96).

(8-97) Alb., MG: [wh]~[aux/cop]~[ptcp/pred] [subject]

[wh~pred]~[cop] [subject]

[wh]~[cop] [subject] [PRED]

Complex sentences are of interest in all three languages. As with quotation inversion, in all-ratified contexts, Alb. and MG tend to place their subjects after both the matrix verb and the verb of the complement clause (this is the pattern found in 10 out of 11 instances of complex predicate forms in Alb., and in 8 out of 8 such instances in MG):

(8-98) Eng: *What d’ye mean by that?* (London, p. 6)

MG: *Ti pai na pi afto?* (p. 11)

what goes that says that (‘What does that go to say?’)

If the subject is non-ratified, however, it may appear between the matrix verb and the complement clause¹:

¹ I have only one (Alb.) example, but my informants (MG and Alb.) reassure me that this pattern is common enough.

(8-99) Alb. (Bishqemi, p. 33)

Ç' duan TANËT të ecin fshehurazi nën dhe?!

what want ours-the that they-go secretly under earth

[The people we met on our way through the cave cannot be Albanians.] 'Why would OUR people want to secretly go under the earth? [It is certainly spies!]

In SC, all depends on the length of the question. In short questions, with both ratified and non-ratified contexts, the situation is the same as in quotation inversion: the subject tends to stand between the matrix verb and the complement clause (although the position after the verb in the complement clause is not excluded). If it is non-ratified, it carries the sentence accent.

(8-100) SC (Vijesnik, p. 13)

Što je mogao taj politički program sadržavati?

what AUX:CLIT could that political program contain

'What could that political program contain?'

In long questions, the subject precedes both the matrix verb and the complement clause, i.e., no inversion occurs, although, as in the case of simple sentences, the position after the matrix verb is not excluded when the only elements apart from the verb and the subject are setting adverbials. The typical situation is, however, the one without wh-inversion:

(8-101) SC (Vesti, p. 5)

Kome bi nova intervencija mogla da ide u prilog?

to-whom AUX:CLIT new intervention could that goes in use

'[There will be no new military intervention against Yugoslavia.] Who would profit from a new intervention?'

The scheme for complex sentences is thus, even with significant simplification, somewhat more complex than the one given for quotation inversion:

(8-102) Alb, MG: [wh]~[matrix verb] [subordinator+verb] [subject] [+/-X]

[wh]~[matrix verb] [SUBJECT] [subordinator+verb] [+/-X]

SC: [wh] [matrix verb] [subject] [subordinator+verb]

[wh] [subject] [matrix verb] [subordinator+verb] [X]

Note that in wh-inversion the behavior of complex predicate forms and complex sentences in Alb. and MG is not identical, as is the case with quotation inversion (cp. (8-43)). The obvious reason for this lies in the fact that the syntactic tie between AUX/COP and PTCP/PRED is narrower than that between matrix verbs and complement clauses, so that non-ratified subjects cannot intervene between the former, but can freely do so between the latter elements. The deeper

reason is pragmatic in nature: in contrast to wh-inversion, there are no non-ratified subjects in quotation inversion, so that the difference simply does not surface.

8.2.2.5. Ratified and non-ratified, short and long: A summary

This is the right place to summarize the results of the previous sections. As with quotation inversion, the syntax and the prosody of Alb. and MG wh-inversion clauses may be subsumed under one sentence template.

As has been repeatedly indicated above, the positioning principles holding for the elements other than subject and verb established for quotation inversion obtain also in wh-inversion, *if* the superordinate topic-ratifying principle is not at work. Thus, in all-ratified constituent questions, one is likely to encounter the cataphoric adjustment principle, the heaviness principle, the free variation of independent elements and the tendency for dependent elements to be placed between the verb and the subject in Alb. and MG, and either before the verb or after the subject in SC. Since all these tendencies are effusively discussed in Section 8.1.1., I shall take them for granted here and ignore them in the presentation which follows.

Let us first consider the list of sentence templates given in the previous chapters. Short and long questions are subsumed in one scheme (the additional element in long questions being represented by a parenthesized X), except where the length of the question determines the sentence form, i.e. in all-ratified contexts in SC.

(8-103) *Sentence templates for wh-inversion:*

Alb., MG:

all-ratified:

[wh]~[VERB] ([X]) [subject]

non-ratified subject:

[wh]~[verb] ([X]) [SUBJECT]

non-ratified verb:

[wh]~[VERB] [subject] ([X])

non-ratified X:

[wh]~[verb] [subject] [X]

SC:

all-ratified:

[wh] [VERB] [subject]

[wh] [pronominal subject] [VERB]

[wh] [subject] [VERB] [X]

non-ratified subject:

[wh] [verb] ([X]) [SUBJECT]

non-ratified verb:

[wh] [subject] ([X]) [VERB]

non-ratified X:

[wh] [subject] [verb] [X]

The most obvious syntactic difference between the three languages is certainly the complete lack of adjacency constraint in SC, represented by the lack of a tilde in the SC sentence templates. This basic difference straightforwardly explains the occasional lack of inversion in this language, but is unable to account both for the fact that in certain contexts the inversion in SC does occur, and for the similarities between the Alb., MG and SC prosodic patterns. I shall therefore contend that we need three superordinate word order and prosodic rules holding in all three languages to account for the similarities, and the simple opposition [+adjacent] versus [−adjacent] to account for the differences. Here is the complete set of rules:

(8-104) In all-ratified contexts, the sentence accent is on the verb.

(8-105) In all-ratified contexts, the sentence accent must not be carried by the last sentence element.

(8-106) If an element is to be ratified, it must stand at the end of the sentence and carry the sentence accent.

(8-107) In Alb. and MG, the question word and the verb have to be adjacent; there is no such constraint in SC.

(8-108) In SC, pronominal subjects practically never license inversion.

First the simpler case, that of constituent questions with a non-ratified element. The sentence form of these questions is easily derivable from rules (8-106) and (8-107). In all three languages, non-ratified elements occupy the final sentence position and carry the sentence accent, independently of the syntactic role they have. Rule (8-106) explains the inversion in SC questions with non-ratified subjects, which, due to their informational status, have to follow the verb, as well as the lack of inversion in other contexts. For Alb. and MG data both rules are needed: Rule (8-107) accounts for the obligatory inversion in all contexts, rule (8-106) for the position of the sentence accent with non-ratified subjects and X elements. In constituent questions with non-ratified verbs in Alb. and MG, the rules are in conflict: According to (8-106), the verb should be in the sentence-final position, whereas (8-107) predicts its adjacency to the fronted question word. It is the latter that prevails, the non-ratified status of the verb being marked by the position of the sentence accent only.

Three rules are needed for all-ratified questions: (8-107), which predicts the obligatoriness of inversion in Alb. and MG, and (8-104) and (8-105), which describe the verb as the carrier of the sentence accent and prevent it from appearing in the last sentence position. The Alb. and MG data are thus straightforwardly accounted for. In SC, the difference between short and long questions becomes somewhat clearer: Rule (8-105) demands the carrier of the

sentence accent, the verb, to be in the non-last position. In short questions, only the subject can cover the last position, so that wh-inversion has to occur. In long questions, the ‘covering’ function is overtaken by X, so that wh-inversion does not occur.

Finally, rule (8-108) accounts for the specific behavior of pronominal subjects in SC: even in all-ratified short questions, or in the questions with non-ratified subjects, pronominal subjects do not allow for inversion. In the case of all-ratified short questions, rule (8-108) thus runs contrary to the rule (8-105), according to which the verb must not occupy the last sentence position in all-ratified contexts. In the case of questions with non-ratified subjects, it runs contrary to the rule (8-106), which demands the non-ratified element to be placed in the last sentence position. In the vast majority of cases, it is the rule (8-108) that has the priority: the inversion does not take place with pronominal subjects.

There are three points worth noting in this context. The syntactic and prosodic behavior of non-ratified elements strongly resembles that of narrow focus expressions, a fact which lead some people to conclude that these elements are in fact foci of constituent questions. In the present framework, which discards the one-to-one focus-to-accent assumption, this fact represents an indication that the basic function of word order and intonation permutations, at least in the languages under consideration, may be to mark the difference between the informationally integrated and the non-integrated readings. In the majority of cases, where these permutations pertain to the assertive part of the clause, a certain position of accent and a certain sentential position indicate the informationally non-integrated reading, separating the narrow focus from the rest of the clause; in those cases in which the sentence material is unequivocally presupposed (constituent questions being the prototypical instance of this type), they separate the non-ratified element from the ratified rest of the clause.

The second point pertains to all-ratified questions. If non-ratified elements, with their non-integrated reading, resemble narrow focus expressions, then one would expect the all-ratified material to resemble broad foci, having the same kind of integrated reading. This is, obviously, not the case: whereas broad foci in Alb., MG and SC expand postverbally with the sentence accent serving as the right boundary (see Sections 4.5. and 5.1.), all-ratified material is marked by a non-final accented verb. According to Ladd (1996: 168ff.), the ‘neutral’ (= all-ratified) accentuation and, one may add, word order, of wh-questions seems to be a major typological variable, with some languages (English) having the same sentence accentuation pattern for questions and declaratives, and some (Hungarian) displaying divergent patterns. Alb., MG, and SC belong to the latter group, where integrated readings are marked by non-identical syntactic and prosodic devices in presupposed and in assertive contexts.

The third point I should like to emphasize is the interplay of the discourse-pragmatically driven word order rules and the purely formal syntactic constraints. The former group of rules seems to be identical in all three languages under consideration, whereas the latter is absent in one of them, with the resulting (relatively strong) typological divergence. This is a kind of situation which one encounters permanently in comparing word order patterns across languages, and, more specifically, which is going to reappear in the present study, so that it is worth remembering.

8.2.3. Alternatives to wh-inversion

As far as SC is concerned, the basic alternative to wh-inversion is non-inversion. Alb. and MG are syntactically more constrained, so that it is interesting to see what happens when the syntactic constraints run contrary to the discourse-pragmatic needs of the speakers.

In 8.2., it was demonstrated that both ratified and non-ratified topics, including the contrastive ones, may be expressed in wh-inversion, by being accented and placed on the right periphery of the clause. There are, however, two subtypes of non-ratified topics which do not, or only seldom do, fit into the wh-inversion pattern. I shall call the first type *newly introduced non-ratified topics* (NITop). It is used when a new discourse referent is introduced into the universe of discourse directly, as a sentence topic (i.e., without introductory clauses asserting its existence or some similar presentational device). In many languages, there are special syntactic devices specialized for NITops, the most prominent being *as-for* constructions (*As for Mary, she is ill*), hanging topics (*Elephants, the nose is long*) and left-dislocation (*These Romans, they're crazy*; cp. Prince 1998 for the typological aspects of these constructions).

Now, in Alb. and MG, NITops tend to occur in the leftmost sentence slot. This is obviously incompatible with wh-inversion, where they would have to be placed at least after the verb. There are basically two ways of solving this. The first is to treat them as simple non-ratified topics (i.e. accent them and place them at the very end of the wh-clause). The second, which is of interest here, is to place them in front of the question word. Consider the following example, overheard on a bus in Athens, 07.11.2000. An elderly lady is talking to a boy about his family; after the father has been exuberantly dealt with, the following question is asked:

(8-116) *Ke i MANA su ti KANI?*

and the mother your what does

‘And how is your MOTHER?’

‘Your mother’ is a typical NITop, being introduced into the universe of discourse without any introductory mechanisms, and is therefore placed *in front of* the question word. The sentence has no audible pauses (it does not have the form *ke i mana su | ti kani*), so I should like to exclude the possibility that the preposed NITop expression is in any way dislocated or extraclassical. There are two intonation peaks, one on the NITop expression (*i mana su*), and one on the verb (*kani*), the first being a secondary, topic-ratifying, accent, the second the default primary accent on the verb in all-ratified questions. Note that the canonical wh-inversion form would also be perfectly felicitous in this context: *Ke ti kani i MANA su?* occurs probably just as frequently as (8-116).

The same holds for the second type, *new paragraph topics* (NPTop). When a discourse referent which has not been the main topic of the previous paragraph is raised to the topic status in the new one, the paragraph boundary being marked by the explicit change of the topic referent, the position before the question word seems to be relatively frequent. Consider the following example: After a detailed report on the procedural troubles the management of a football club had to organize a meeting, it is said that at that meeting, when finally held, the old trainer was replaced by a new one, named Bak Sokol. The new paragraph begins with the following sentence:

(8-117) Alb. (Koha Ditore, p. 18)

SOKOLI mbetet të shihet se **si** do t'i ZGJIDHË problemet ...

S. remains that be-seen that how FUT:CLIT SUBJ:CLIT them:CLIT solves problems-the
 ‘It remains to be seen how Sokol is going to cope with the problems...’

The whole paragraph that follows deals with Sokol’s future duties as a trainer. The intonational pattern and the syntax are the same as those described for NITops above, and the alternative with the canonical wh-inversion is also present (*Mbetet të shihet se si do t'i zgjidhë problemet SOKOLI*). What is interesting here is the fact that the preposition of NITops/NPTops obviously also takes place in embedded syntactic contexts, the preposed topic expression being placed before the embedded question. Example (8-117) illustrates the more frequent variant of this, where the topic expression is located at the very left edge of the sentence, before the matrix verb. There is, however, a somewhat less frequent variant with the topic expression between the matrix verb and the question word, i.e. *Mbetet të shihet SOKOLI se si do të zgjidhë problemet*.

Although SC has the possibility of non-inversion in all contexts, prefocal NITops and NPTops are sometimes, if rather rarely, found also in this language:

(8-118) SC (from Popović 1997: 242)

Novi i stari svijet kad se razumiju?

new and old world when REFL:CLIT understand

‘When do the new and the old word understand each other?’

Thus, wh-inversion may be obviated by placing topic expressions in the prefocal position. Although the examples adduced in this section illustrate only the prefocal position of those topics which are the subjects of constituent questions, all other grammatical relations are allowed in this position, too. In contrast to this generosity as to the syntactic status of prefocal elements, the range of discourse-pragmatic roles they may carry is rather limited, with only NITops and NPtops appearing in this position. There is no evidence that prefocal topic expressions are in any relevant way extraclassical, so that sentence templates given in (8-103) should all be expanded with an optional slot for NITops and NPtops in front of the question word slot: **([NI/NPTop]) [wh]~[VERB] ([X]) [subject]**, etc.

Note that this option, the prefocal position of certain topic types, is absent from quotation inversion, although both the quote and the question words represent, in my interpretation, fronted narrow focus expressions. As already indicated in 8.1.4., the restrictiveness of quotative contexts probably has to do with the length of the preposed element, the quote, and with its specific ontological status.

8.2.4. Discourse-pragmatic differences between Alb., MG, and SC

The statistical differences between the three languages, are, as indicated in connection with Table (8-63), partly triggered by diverging syntactic constraints (SC has relatively few wh-inversions because its grammar does not allow for inversion in a number of contexts), partly by very general discourse-pragmatic tendencies already dealt with elsewhere (MG has less wh-inversions than Alb. and SC *inter alia* because MG subjects are much more frequently left unexpressed than the Alb. and SC ones, see 5.1. and 8.1.5.).

One further point worth noting is the tendency in Alb. to use embedded clauses rather than abstract nouns for coding states of affairs in argument positions, described in some detail in Section 5.1. The consequences of this tendency for the frequency of wh-inversion mostly concerns the complement clauses of cognitive and perception verbs: with these verbs, Alb. simply uses more embedded questions than MG and SC, so that the number of wh-inversion clauses proportionally rises:

(8-119) Eng: ... *the elder saw **the opportunity***. (London, p. 34)

Alb: ... *plaku e pa se ç' rast i mirë ishte ky*. (p. 34)

what opportunity the good was that

MG: ... *o γerondas idë **tin efkeria***. (p. 43)

the opportunity

SC: *Stariji vodja ... vidje **tu zgodnu priliku***. (p. 39)

that fine opportunity

There is also a minor difference between Alb. and MG on one hand and SC on the other with respect to the use of prefocal NITops and NPTops. Namely, they seem to be generally much more frequent in Alb. and MG than in SC, both in written and spoken language – an intuition which I can only partially prove, namely for the written language: In my whole corpus, there are 37 examples of prefocal topics in Alb., 45 in MG, and none in SC. Although I do not subscribe to functionalist theories of language economy (what you have on one side, you lose on the other), I consider it possible that this has to do with the non-obligatoriness of wh-inversion in SC, i.e. with the fact that much of the work done by prefocal topics in Alb. and MG may be done by non-inverted postfocal subjects in SC.

8.2.5. Conclusion

Like quotation inversion, wh-inversion is a subtype of focus inversion in which the verb-subject complex carries the sentence accent (here, in contrast to quotation inversion, because of the inherent unaccentability of question words). Also like quotation inversion, the main interest of wh-inversion lies in its syntax and prosody: it represents a very good model for the description of the ratifying mechanisms within the presupposed clausal material and of their interaction with purely formal restrictions and tendencies, the main points being the principle according to which non-ratified elements tend to occur clause-finally and carry the nuclear stress, and the principle of accenting the non-final verb in all-ratified presupposed contexts. Similarities and differences between these and the mechanisms of focus domain formation noted in 8.2.2.5 will be further elaborated in 10.1.2.. Unlike quotation inversion, wh-inversion is not obligatory in all three languages, with SC lacking the adjacency constraint between the verb and the fronted focused expression. Also unlike quotation inversion, wh-inversion has a simple alternative, the construction with prefocal topic expressions, reserved only for a subclass of non-ratified topics, NI and NPTops. With some minor divergences, this is also the

basic pattern of fronted focus inversion, the third and the last type of focal inversion in the Balkan languages.

8.3. Fronted focus inversion (ff-inversion)

When a narrow focus expression is fronted, the verb has to follow it immediately, and the subject follows the verb. This is a short definition of the third type of focal inversion, **fronted focus inversion** (henceforth **ff-inversion**):

(8-120) Eng: *NEVER, in all his fighting, had this thing happened.* (London, p. 136)

Alb: *KURRË ndonjëherë nuk i kishte ndodhur diçka e tillë.* (p. 133)

never once not:CLIT him:CLIT AUX happened something the such

MG: *POTÉ ðen ixë ksanasimvi tetjo prayma stis maxes tu.* (p. 166)

never not:CLIT AUX again-happened such thing in-the battles his

SC: *NIKAD se njemu, u svem njegovom borenju, nije dogodilo tako nešto.* (p. 153)

never REFL:CLIT him in all his fighting NEG:AUX happened so something

Ff-inversion works in the same fashion as wh-inversion, with respect both to the language-specific formal syntactic constraints and to the crosslinguistic discourse-pragmatically driven rules, the only point of divergence being the intonation (8.3.1.). Since narrow focus expressions more often appear clause-finally, at the right edge of the focus domain, the particular point of interest with respect to ff-inversion are the conditions under which these expressions are fronted. Unlike the other two types of focal inversion, ff-inversion is in some cases difficult to identify, and is subject to massive discourse-pragmatic crosslinguistic variation (8.3.2. and 8.3.3.).

This kind of focal inversion is, apart from MG, not extremely frequent in my corpus, as witnessed by Table (8-121), organized according to the same principles as Table (8-5):

(8-121) *Fronted focus inversion in the corpus*

	1. original	2. translat.	3. total	4. FFI vs. VS	5. FFI vs. SV/VS	6. FFI vs. Σ
Alb.	58	7	65	65 vs. 2225 2.9%	65 vs. 7979 0.8%	65 vs. 15813 0.4%
MG	140	7	147	147 vs. 2673 5.5%	147 vs. 6641 2.2%	147 vs. 14463 1.0%
SC	40	1	41	41 vs. 2065 2.0%	41 vs. 8458 0.5%	41 vs. 14176 0.3%

Focus fronting is in general more frequent in the spoken than in the written language, so that the numbers for Alb. and SC are not surprising, especially since in SC, similar to wh-inversion, the condition of adjacency between the focus expression and the verb does not exist. Furthermore, the typical discourse contexts in which focus fronting is used favor the use of zero subjects, in which case, of course, no inversion occurs, as in the following example:

(8-122) Alb. (Camaj, p. 167)

Vetëm për të shikue ATO vijnë në Qytecë.

only for to seeing those:ACC they-come to small-town

‘[It is well-known that foreigners like these landscapes.] It is only in order to see THEM that they come to the Small Town.’

What is surprising is that MG displays a relatively high ratio of this construction even in a corpus consisting only of written texts. This indicates that MG, in contrast to Alb. and SC, has a relatively unrestricted, productive pattern of focus fronting. This numerically based syllogism will be confirmed in the following chapters.

8.3.1. Syntax, prosody, and information structure

Informationally and syntactically, ff-inversion is more or less identical to wh-inversion. Rules (8-104) – (8-107) and the sentence templates given in (8-103) also obtain for ff-inversion, only that the slot [wh] has to be replaced by [ff], fronted focus. Even some details not represented in the templates and rules coincide. Thus, as in wh-inversion, the adjacency constraint in Alb. and MG seems to be (almost) absolute with respect to major grammatical relations (subjects, objects, or similar, may in no way intervene between the ff and the verb), whereas expressions with looser syntactic ties to the predicate, notably setting adverbials and sentence adverbials, sometimes occur in the slot between the ff and the verb¹, if only rarely (one example in Alb., three in MG):

(8-123) MG (Kapandai, p. 124)

KAΘE fora etsi katelijan oles tu i prosπαθjes.

every time thus ended all his the attempts

¹ The adjacency constraint has been one of the important topics in the generative literature on focus fronting in MG: Agouraki (1990) and Tsimpli (1990, 1995, 1998) claim that it is absolute and structural, not allowing for any kind of intervening elements, Giannakidou (1997), and after her Alexiadou (2000) and Alexiadou and Anagnostopoulou (2000), note that adverbials like *xtes* ‘yesterday’ freely occur between the focused phrase and the verb. My corpus data confirm Giannakidou’s intuitions.

‘All his attempts ended up like this EVERY time.’

The similarity between wh-inversion and ff-inversion is one of the reasons I have included fronted predicatives into the group of fronted narrow focus expressions. Consider the following example:

(8-124) Alb. (Kosovarja, p. 5)

Të KOTA ishin përpyjekjet e mjekëve për ta ndihmuar.

the futile were attempts-the of-the physicians for the-him helping

‘The attempts of the doctors to help him were FUTILE.’

With its form, [PREDICATIVE] [copula] [subject], this sentence is virtually identical to (8-94) (represented as a sentence template in (8-97)), [wh~PREDICATIVE] [copula] [subject]. This is the first indication that the predicative in such cases is a fronted narrow focus expression. The second indication is the identity of fronted predicatives and other fronted narrow focus expressions with respect to discourse functions and semantic features (see 8.3.2.). The third one is the fact that this form, [PRED] [cop] [sub], stands in contrast to the forms [cop] [PRED] [sub] and [cop] [sub] [PRED], which, as I hope to have demonstrated in 6.3.2.3., convey presuppositional structures clearly distinct from the one with the fronted accented predicative. Thus, I consider accented predicatives in the clause-initial position to be instances of narrow focus expressions, for both formal and functional reasons.

There are also two differences between wh-inversion and ff-inversion. The first one is formal: unlike question words, which may not be accented, fronted focus phrases *must* carry the nuclear stress, as witnessed by (8-120) through (8-124). This, again, has consequences for the intonation pattern of the rest of the clause. If the presupposed material following the ff-expression is all-ratified, it remains completely unaccented, as in (8-120) and (8-123)¹. If there are non-ratified elements, they carry only the secondary, ratifying accent (the primary accent being on the ff-expression), but are otherwise subject to the rules (8-106) and (8-107). Sentences (8-125) and (8-126) illustrate the position and the intonation of non-ratified subjects and additional elements, respectively:

(8-125) MG (Elefterotipia, p. 10)

Ke KANENA ðen epjase o ISANGELEAS!

and no-one not caught the public prosecutor

‘[This is one of the greatest robberies of the 20th century. And nobody went to prison!] And NOT A PERSON was prosecuted by the public prosecutor!’

¹ Of course, if the subject or any other element is long, it may carry a secondary, postnuclear accent, in keeping with the principles described in 8.1.1.1.

(8-126) Alb. (Koha Ditore, p. 8)

Se VËRTET ka ndodhur kjo në VOKSH, këtë dëshmon z. Ali Nitaj ...

that really AUX happened this in V., this testifies Mr. A.N.

[Cases were reported of disrupted meetings of the political party LDK] ‘Mr. Ali Nitaj confirms that this REALLY happened in VOKSH...’

The patterns for wh-inversion given in (8-103) thus have to be altered so as to describe the prosody of ff-inversion:

(8-127) *Sentence templates for ff-inversion:*

Alb., MG:

all-ratified:

[FF]~[verb] ([X]) [subject]

non-ratified subject:

[FF]~[verb] ([X]) [SUBJECT]

non-ratified verb:

[FF]~[VERB] [subject] ([X])

non-ratified X:

[FF]~[verb] [subject] [X]

SC:

all-ratified:

[FF] [verb] [subject]

[FF] [subject] [verb] [X]

non-ratified subject:

[FF] [verb] ([X]) [SUBJECT]

non-ratified verb:

[FF] [subject] ([X]) [VERB]

non-ratified X:

[FF] [subject] [verb] [X]

Many of the templates represented in (8-127) are based solely on elicitation/introspection, since, in contrast to wh-questions, many of the possible informational/syntactic types are not represented in the corpus at all – thus I have no authentic examples of long ff-clauses with non-ratified subjects, and there are no non-ratified verbs at all. As a matter of fact, ff-clauses are, even more often than quotation inversion and wh-inversion clauses, short, consisting only of the focus expression, the verb and the subject – 56 (86.2%) in Alb., 112 (76.2%) in MG, and 36 (87.8%) in SC.

This brings us to the second difference between ff-inversion and wh-inversion, which is not formal, but rather discourse-pragmatic in nature: ff-inversion clauses only very rarely contain non-ratified material (even on a very generous interpretation, I have found only 13 examples in all three languages), resembling at this point the quotation inversion rather than the wh-inversion. The reason for the scarcity of non-ratified topics in ff-clauses lies partly in

thier discourse function and partly in the fact that ratification in clauses containing narrow foci is usually done with alternative constructions. These two phenomena – discourse functions of ff-inversion and alternative constructions – are the topics of the two following sections.

8.3.2. Meaning of fronted narrow focus expressions

At least since Mathesius (e.g. 1939), it has been a common place that the kind of word order in which the ‘new’ information precedes the ‘old’ information is somehow special, although synonymous with the more common ‘old-new’ order. The basic idea is that ‘new-old’ sentences differ from ‘old-new’ sentences not in meaning or information structure, but merely in the degree of emotional involvement of the speaker, the special, ‘new-old’, order being somehow more passionate¹.

The development of the theory of focus in the last twenty years has brought with it the rephrasing of the problem. In many languages (Alb., MG, and SC, *inter alia*), narrow focus expressions generally occur at the end of the clause (see 4.5.). However, they may also appear on the left periphery, preferably at the very beginning of the clause.

The question is whether there is some semantic or discourse-pragmatic difference between these two types of focus expressions. The solution which seems to have achieved a rather broad acceptance is the one proposed by É. Kiss (1998b, see also Szabolcsi 1983). Roughly, she claims that there are two kinds of narrow foci, the information focus, which simply conveys new information, and the identificational focus, which is an operator-like entity performing exhaustive identification within a contextually given set (for more details, see 4.4.2.). The former type is said to be expressed by clause-final, the latter by clause-initial narrow focus expressions.

Now, as indicated in 4.4.2., although I do not consider the ontological dichotomy implied in É.Kiss’s system justified, I do agree that there is a pragmatically and contextually determined continuum between foci with a strongly limited number of alternatives and those with a practically unconstrained set of alternatives. I also agree that different languages may choose a certain point on this continuum and code the narrow foci to the left of this point differently than those to the right of it. The first thing to do, then, is to determine which point

¹ This is the reason the terms like ‘subjective’ or ‘emotional word order’ were commonly used. Mathesius’ term was ‘pathetic’, whereas Bolinger (1954b), in an important study on Spanish word order, spoke of ‘passionate accent/word order’

on the continuum is chosen¹. But even this is not enough: as noted, e.g., by Tsimpli (1995, 1998), Popović (1997), and Bošković (2002), there is a considerable overlapping between the clause-initial and the clause-final ('normal') narrow focus expressions in MG and SC (the same holds for Alb. as well). That is, even if we agree that, say, contrastive foci are fronted and the other focus types not, we have to admit that many contrastive foci are *not* fronted, i.e. that they are clause-final. The favorite solution to this is to say that focus fronting in these languages is optional (Tsimpli 1995, Bošković 2002). Of course, the possibility of free variation is not to be discarded *a priori*, but one should first try to find out if there are any pragmatic or semantic differences between the two apparently synonymous forms.

Both questions – at what point of the focus continuum does the narrow focus fronting begin, and why are narrow foci to the left of this point sometimes fronted and sometimes not – beg for an empirical investigation. What I shall therefore do in this section is to enumerate the semantic properties of phrases which are most often found in the ff-position in my corpus (8.3.2.1) and the typical discourse contexts in which ff-clauses occur in texts (8.3.2.2).

8.3.2.1. Semantic and informational features of fronted narrow foci

Let us first take a look at the expressions which occupy the ff-position with a frequency greater than average. The data are organized according to mixed criteria – semantic (groups 1.- 4.), formal and pragmatic (5.), and purely pragmatic (6.-7.). It is important to understand that the classes enumerated below represent only the prototypical, most frequent semantic/pragmatic types of fronted focus expressions, but not the only expression types which may and do appear in the ff-position. Here is the list:

(1) Quantified arguments of the verb (objects, PPs, etc.). Both the simply and the universally/existentially quantified arguments of the verb occur in the ff-position in all three languages, though they seem to be used in MG more frequently than in Alb. and SC, both in my corpus (24 instances in MG vs. 5 in Alb. and 7 in SC) and in the spoken language.

¹ There have been, as far as I know, two attempts to identify this point for MG, both rather rudimentary. In Tsimpli (1995, 1998), it is simply claimed that fronted foci are 'contrastive' or 'identificational', without further explanation. É. Kiss (1998b) tries to establish a typology of 'identificational' foci along the parameter of the presence of features [exhaustive] and [contrastive]. In this system, MG belongs to the group of languages whose 'identificational' (i.e. fronted) foci are [+exhaustive, +contrastive] (in contrast to e.g. Hungarian, where they are only [+exhaustive]). This is a more substantial claim, which I am going to try to test in this section.

(1.a) *Universal and existential quantification* (all, none). For an existentially quantified ff-object in MG ('nobody'), see (8-125); (8-128) is an example of a universally quantified ff-object in SC:

(8-128) SC (Andrić, p. 70)

SVE je znao i SVE predvidjao ... ovaj Haim iz Smirne.

all AUX:CLIT known and all foreseen this H. from Smyrna.

'This Haim from Smyrna knew EVERYTHING and foresaw EVERYTHING.'

A very interesting phenomenon occurs in MG (as far as I can see, it is absent from Alb. and SC): fronted focus expressions *without* a universal or existential quantifier sometimes receive a universally or existentially quantified reading:

(8-129) MG (Kapandai, p. 140)

Mia zoi tin kratisan makria ta matia tu. [= *mia zoi olokliri*, 'one life whole']

one life her:CLIT held at-distance the eyes his

'His eyes kept her at the distance one [whole] life (=during all of his life)'

(8-130) MG (Kapandai, p. 142)

LEKSI ðen katalavene o Filipos. [= *kamia leksi*, 'none word']

word not understood the Philippos

'Philippos was not able to understand a single word.'

Giannakidou (2000: 244), referring to the type represented by (8-130), suggests a semantic explanation: it is said that bare singulars serve as minimizers and belong to the group of affective polarity items. Examples of the type (8-129), with an indefinite article and universally quantified interpretation, point in another, namely pragmatic, direction.

(1.b) *Simple quantification* (much, many, little).

(8-131) Alb. (Bishqemi, p. 79)

PAK interes kishte sqarimi i kësaj pike.

little interest had clarification-the of-the this point

'The clarification of this point was of NEGLIGIBLE importance.'

(2) Quantifying adverbials, depictive predicates, and predicatives.

Universally/existentially quantified adverbials and adverbials of duration (2.a. and 2.b.) are placed in the ff-position with approximately the same frequency in all three languages. The other two classes (adverbials/predicatives of simple quantification and metaphorically quantifying expressions, 2.c. and 2.d.) seem to be frequently fronted only in MG.

(2.a) *Universal and existential quantification* (always, never). The adverbials meaning 'always', 'never' and the like are relatively often fronted in all three languages. For a

universally quantifying ff-adverbial in MG ('every time'), see (8-123); for an existentially quantifying ff-adverbial in all three languages ('never'), see (8-120).

(2.b) *Adverbials of duration* (four months, for months). Adverbials which express determinate or indeterminate duration are probably the most frequent lexical filling of the ff-slot in all three languages, with 11 instances in Alb., 27 in MG, and 8 in SC.

(8-132) SC (Kiš, p. 76)

DVA meseca se šetao Taube ... ulicama Moskve ...

two months REFL:CLIT walked T. through-streets of-Moscow

'Taube walked through the streets of Moscow for [whole] two months...'

(2.c) *Simple quantification* (very, much, often; alone; big, small, few, many). This kind of fronted focus expressions is in general use only in MG.

(8-133) MG (Matesis, p. 53)

POLI timiθike tote m' afto i mana mu.

much was-honored then with that the mother my

'My mother felt VERY honored with it at that time.'

(8-134) MG (Elefterotipia, p. 9)

POLAPLASIES ine i sinepies se varos tis politikis kata ton narkotikon.

multifarious are the consequences in burden of-the policy against the narcotics

'The consequences at the expense of the policy against the narcotics are MULTIFARIOUS.'

(2.d) *Comparative and metaphorical quantification* (like a wolf, like a thunder; in waves). These expressions seem to regularly occupy the ff-position only in MG, whereas in Alb. and SC they generally occur clause-finally.

(8-135) MG (Kapandai, p. 40)

San ASTRAPI taksiðepse to neo.

as lightning traveled the news

'[The Turks imprisoned the greatest captain of the region, captain Mesovuniotis.]

The news traveled (as quickly) as a LIGHTNING, [and our Petros learned that he was being taken to Ioannina.]'

(8-136) MG (Kapandai, p. 29)

KIMATA kimata efevje o kosmos apo to Moria.

waves waves fled the people from the Moreas

'The people were fleeing from Moreas in WAVES.'

(3) Adverbials and predicatives of expressive qualification. This class is not easy to define. It comprises:

(3.a) *Modal adverbials and predicatives which convey an implicit or explicit moral or esthetic judgment of the speaker* (adverbials like ‘terribly’ [as in ‘to make a terrible mistake’], ‘beautifully’ [as in ‘to sing beautifully’], predicatives like ‘black’>‘terrible, dark’ [as in ‘his soul was “black”], ‘a miracle’ [‘our success was a miracle’], ‘an abyss’ [‘man’s soul is an abyss’]);

(3.b) *Adverbials which designate a high grade of the presence of a property*, which is often expressed by reduplication in MG and Alb. (MG *kato kato* ‘very much below’, *arya arya* ‘very slowly’); and finally,

(3.c) *Adverbs meaning ‘in vain’ and their predicative counterparts meaning ‘futile’.*

The common property of all these expressions is the strong note of subjective judgment on the part of the speaker, be it with respect to a moral or esthetic canon (the first group), to the grade of the presence of a property (the second group), or to the probable outcome of the state of affairs described (‘in vain’). More often than not, this judgment is negative. All three languages make use of the ff-slot when an adverbial or a predicative of this kind is focused, but SC seems to use it with some regularity and with a non-pathetic flavor only in the case of ‘in vain/futile’, whereas in Alb. and MG all the classes enumerated above are freely fronted. An example of the fronted ‘futile’ in Alb. is given in (8-124); here are some representatives of the other two classes¹:

(8-137) MG (Fakinou, p. 53)

AVISOS i psixi tu anθropu.

abyss the soul of-the man

‘Man’s soul is an ABYSS.’

(8-138) Alb. (Kosovarja, p. 8/50)

Dhe gabuam, TMERRËSISHT gabuam unë dhe Faca...

and we-made-mistake, terribly we-made mistake I and F.

‘And we were wrong, oh how TERRIBLY wrong we were, me and Facë...’

¹ Note that the meaning of the predicatives/adverbials of this class makes it especially easy for them to appear in proverbial or quasi-proverbial statements, which is a context in which the copula is frequently left out in Balkan languages, (8-137) being a nice example of this phenomenon.

(8-139) Alb. (Bishqemi, p. 52)

I ÇMUESHËM është rubini...

the precious is ruby

‘Rubies are PRECIOUS’

(8-140) MG (Kapandai, p. 116)

Sto KATO kato eđo itan to Aivali...

in-the below below here was the A.

‘Aivali was at the VERY lowest part (of the island)...’

(4) Sentence adverbials denoting polarity. Words corresponding to English ‘really’, ‘it is true that’, etc., as well their negative counterparts, often appear in the ff-position in all three languages. An example with the Alb. particle-like word *vërtet* ‘true’ is given in (8-126). Other similar words and expressions I have found in this position: Alb. *me siguri* (‘with certainty’, ‘certainly’), *pa dyshim* (‘without doubt’, ‘doubtless’), MG *praymati* (‘and really’), *aliθja* (‘true’), SC *zaista/stvarno* (‘really’), *fakat* (colloquial: ‘it is true that’), etc. There is also a construction in which fronted negative particles carrying the sentence accent are followed by a factive subordinator (Alb. *se/që*, MG *oti/pos*, SC *da*) and a non-accented presupposed quasi-subordinated clause:

(8-141) MG (Matesis, p. 72)

OXI pos psonizo eyo eki, ala ...

not that buy I there, but

‘NOT that I do my shopping there, but...’

(5) Pronominal arguments, adverbials, and predicatives. When discourse-active, pronominally expressed notions are narrowly focused, then the focus fronting is the preferred alternative in all three languages, perhaps most of all in SC, where focused pronominal words are rarely found in other focus positions. Focused pronominal arguments are, of course, always in their full, non-clitic form. The adverbials found most frequently are ‘so’ (*kështu, etsi, tako*), ‘in other way’ (*aljos, drugačije*), ‘for that reason’ (*ndaj, j’afto*), ‘here, there, now, then’, etc. The typical predicatives are ‘such’ and ‘so big/many’ (*i tillë, tetjos/tosos, takav/toliki*). Here are some examples:

(8-142) Alb. (Bishqemi, p. 121)

NDAJ them unë se ... është i vërtetë.

therefore say I that is the true

‘THAT’s why I say that it ... is true.’

(8-143) SC (Vjesnik, p. 4)

TAKAV bijaše Franjo Tudjman!

such was F.T.

‘THAT’s how Franjo Tudjman was!’

(6) **Focus expressions modified with *only***. Expressions modified by focus particles meaning ‘only’ (*vetëm, mono, samo*), which play an important role in É. Kiss’s definition of exhaustive identification (1998b), are, as expected, frequently placed in the clause initial ff-position in all three languages (8 instances in Alb, 11 in MG, 9 in SC). An illustration of a fronted *vetëm*-phrase is given in (8-122).

(7) **Focus expressions modified with *just/precisely***. Phrases modified with focus particles denoting exact identification (*pikërisht, akriivos, upravo*) can occupy the ff-position in all three languages; the phrase itself is often pronominal.

(8-144) Alb. (Bishqemi, p. 66)

Pikërisht ATË shkëlqim rrezatonin sytë e tij kur u takuam në fillim.

precisely that shine beamed eyes-the the his when PASS:CLIT we-met in beginning

‘It was precisely that SHINE that his eyes had when we met for the first time.’

The expression types enumerated above cover between 70 and 80 % of all the instances of ff-phrases I collected. What is interesting, however, is the fact that some expression types do not appear in all languages, or not with the same frequency and the same stylistic value. In order to facilitate comparison, the data are summarized in Table (8-145); ‘++’ means ‘when focused, most commonly found in the ff-slot’; ‘+’ means ‘when focused, not unusual in the ff-slot, without strong stylistic effects’; ‘-’ reads as ‘rarely occurring in the ff-slot, or only with a strong pathetic, poetic, etc. stylistic effect’.

(8-145) *Fronted focus phrases in Alb., MG, and SC*

	MG	Alb.	SC
1.a. <i>all/nothing</i> -phrases	++	+	+
1.b. <i>much/little</i> -phrases	+	+	-
2.a. <i>always/never</i>	+	+	+
2.b. adverbials of duration	++	++	++
2.c. <i>very/big/small</i>	++	-	-
2.d. <i>like a wolf</i>	+	-	-
3.a. moral/esthetic judgment	++	+	-
3.b. high grade of a property	++	+	-
3.c. <i>in vain/futile</i>	++	++	++
4. sentence adverbials of polarity	+	+	+
5. pronominal expressions	++	++	++
6. <i>only</i> -phrases	++	++	++
7. <i>just/precisely</i> -phrases	+	+	+

The typical semantic filling of the ff-slot looks as follows. The common semantic core of this slot is made up of *only*-phrases (6.), focused pronominal expressions (5.), adverbials of duration (2.c.) and words meaning ‘in vain’ (3.c.). Universally or existentially quantified and quantifying expressions (1.a., 2.a.), polarity items (4.) and *just/precisely*-phrases are relatively common in all three languages. This is where the differences begin. In SC, none of the other expression types is used, at least not without a strong ‘bookish’ flavor. In Alb., simply quantified verbal arguments (1.b.), adverbials and predicatives of moral/esthetic judgment (3.b.) and adverbs denoting high grade of a property (3.c.) are not uncommon. Finally, in MG, almost all expression types enumerated appear in the ff-slot very or relatively often. Thus, there is a scale along which the three Balkan languages may be ordered, with MG using the most diversified types of expressions in the ff-slot, Alb. displaying some lexical restrictions, and SC having an ff-slot which is relatively strongly restricted to four or five lexical types.

8.3.2.2. Discourse functions of fronted narrow foci

In this section, I shall try to identify contexts in which clauses with a filled ff-slot are employed. Before presenting the list, I shall describe the basic common function of the expressions in the ff-slot and the basic divergences between the languages under consideration, in order to make the principles underlying the discourse functions transparent.

If one compares a clause like (8-128) with its counterpart without focus fronting (*Ovaj Haim iz Smirne je znao SVE*), the first impression one gets is, as Mathesius noted almost a century ago, that the version without ff-inversion is somehow emotionally flat, a mere statement of a fact, whereas the actually attested ff-clause (*SVE je znao ovaj Haim...*) conveys a shade of surprise, unexpectedness. Consider the context in which it is uttered: A prisoner named Haim tells to another prisoner, friar Peter, some facts which an imprisoned person is not expected to know. Friar Peter, who is the narrator, is surprised by this, and his comment is *SVE je znao ovaj Haim...* (‘EVERYTHING knew this Haim’). What is implied is that the speaker had expected Haim not to know everything, but that the information Haim entrusted him with convinced him that he, in fact, did.

Such clauses function on the basis of contrast. The expression in the ff-slot evokes the set of alternatives, in this case, the set {everything, not-everything}. The alternative which is not explicitly expressed – {not-everything} – is interpreted as that what the speaker had expected, the expressed alternative – {everything} – as that what he got. Fronted foci are thus indeed

contrastive. And not only that: This focus type expresses a particular kind of contrast, that between the expected state of affairs and the state of affairs actually obtaining. Let us call this type of contrast *contrary-to-expectation contrast (CTE)*.

What kind of contrast is CTE? The expression classes usually appearing in this context may help us here. This is where the languages start to diverge.

In all three languages, universally/existentially quantified or quantifying phrases (groups 1.a. and 2.a. above), words and phrases meaning ‘in vain/futile’ (3.c.), sentence adverbials of polarity (4.), pronominal expressions (5.), *only*-phrases (6.) and *just/precisely*-phrases (8.) occur in the ff-position more or less frequently. What all these expression types have in common is that they function on the binary basis. For the first four groups, the binarity is easily derivable from the semantics: either ‘all’ or ‘some = not-all’, either ‘some = not-none’ or ‘none’, either ‘with success’ or ‘in vain’, either ‘yes’ or ‘no’. For the latter three types, it is rather their discourse-pragmatic function from which the binarity is derived, as will be demonstrated in greater detail below: pronominal expressions, *only*- and *just/precisely*-phrases are used in the ff-slot then, when the alternative which they exclude is meant as a single contrastive entity.

The phrases appearing in the ff-slot which are not in the list of the prototypical expressions given above often denote entities or properties naturally occurring in pairs, like ‘up’ and ‘down’. The first type of contrary-to-expectation fronted focus, common to Alb., MG, and SC, is thus the *binary contrary-to-expectation focus (binary CTEFoc)*.

The second type, with one exception, does not occur in SC, or only in archaic, poetic, and similar registers. It does occur in Alb., albeit with some restrictions, and represents a rather frequent type in MG. This kind of CTE focus is based on the notion of graduality. Many denotata in natural languages allow for gradation: a state of affairs can last long, very long, extremely long, etc.; a person can be tall, somewhat taller than expected, extremely tall, etc. In creating expectations, the speaker can start from a certain grade of the presence of a property in the state of affairs to be conveyed. For instance, in (8-140), the writer expects, or expects the hearer to expect, that the town of Aivali is somewhere in the lower part of the island. Now, fronted foci function contrary to expectation: Aivali is not only somewhere in the lower part, it is located in the VERY lowest part of island. The contrastive set thus contains {a low part, the lowest part}, whereby the first member of the set represents what is expected, the latter what is asserted. And again, the variant with a clause-final focus expression (*To Aivali itan sto KATO kato edo...*) would mean the same, but without the contrary-to-expectation nuance. This type is represented by simply quantified/quantifying expressions (1.b, 2.c), by

adverbials of duration (2.b), metaphorically quantifying expressions (2.d), and by expressions of moral/esthetic judgment and of high grade of a property (3.a, 3.b). Let us call this kind of CTE focus *gradual CTE focus*. Note that there is some variation here: although it practically does not have the category of gradual CTE focus, SC frequently uses duration adverbials in the ff-slot; speakers of Alb. do not seem to place adverbs like ‘very’ and predicatives like ‘big’ in the ff-slot, and generally seem to use the gradual CTE foci less often than the speakers of MG.

Now let us see how the hypothesis of the contrary-to-expectation nature of fronted focus phrases and of the double nature of this contrariety is reflected in discourse. Here is a list of discourse functions of the clauses with fronted narrow foci:

(1) Corrections. In all three languages, fronted narrow focus expressions are used in order to correct explicitly expressed false assumptions on the part of the hearer:

(8-146) MG (Ciao, p. 29)

ðen ine zoi afto pu zume, to karnavali tu Rio se ful sezon ine afto.

not is life that what we-live the carnival of-the R. in full season is it

‘What we live is not a life, it is the Carnival in Rio at the height of the season.’

The corrective narrow focus operates over the binary set of alternatives – the false assumption of the hearer (what we live is a life), and the correction of the assumption (what we live is the Carnival in Rio): the contrastive set has the form {life, Carnival in Rio}, and the function of the focus is to exclude the former and assert the latter alternative. Although pronominal, quantified and quantifying expressions represent the most frequent lexical filling in corrective contexts, this discourse type seems to be less restricted than all others. Actually, most of the examples from my corpus in which the narrow focus expression does not belong to one of the eight classes enumerated in the previous chapter belong to corrective statements. The lack of lexical restrictions is easily understandable in view of the fact that corrections represent one of the few naturally occurring contexts in which the number of alternatives over which the focus quantifies is necessarily reduced to two, the right and the wrong assumption. Whereas in most other contexts the binarity condition for fronted narrow foci is fulfilled by the lexical class of the focus expression, here it is the pragmatics of the corrective situation itself that does the job. The second important property of fronted foci apart from the binarity of the alternatives, the contrary-to-expectation effect, is also of necessity given in a corrective context.

Of course, the same principle obtains if the false assumption is left unexpressed, as in the following example: After some time of absence, Mrs. Kanello appeared one day with a baby.

Her curious neighbors try to find out whose child it is by asking indirect questions (and assuming that it must be hers). Guessing what the whole thing is about, she says:

(8-147) MG (Matesis, p. 57)

Tis AðERFIS mu ine to moro.

of-the sister my is the baby

‘It is my SISTER’S baby.’

The contrastive set is {Mrs. Kanello’s, her sister’s}, the former being the assumed expectation, the latter the correction of that expectation¹.

(2) Comments. In all three languages, when a person gives a comment on a situation, if this comment is supposed to express the person’s surprise by any aspect of the situation, focus-fronted expressions are likely to be used. Thus, in the SC example (8-128), commented upon above, the preceding text describes the action of a prisoner talking about things a prisoner is not supposed to know. The sentence (8-128) itself is pronounced as a comment on this state of affairs: the speaker is surprised that the prisoner Haim knows EVERYTHING, the focused denotatum forming the binary contrastive set {not-everything, everything}. This is the typical situation of a comment: first the description of a state of affairs, then a sentence in which the speaker states how s/he judges the state of affairs, one attribute of which runs contrary to her/his expectations. The coding of this attribute as a ff-phrase makes the unexpectedness explicit.

In all three languages, quantified and quantifying phrases, as well as pronouns, are often used in this context, as well as *only*-phrases and *just/precisely*-phrases, all of them triggering the binary interpretation of the context (expected, not expected). Another nice example of this kind, with a *just/precisely*-phrase in the ff-slot, is (8-144). Smerald has a strange shine in his eyes. The author’s comment upon this fact is that it is precisely that shine that he had when they first met. What is presupposed is that Smerald’s eyes were shiny when he and the author met, the fact the readers are familiar with, since it is described some twenty pages before. What is asserted is that the shine at that situation was absolutely identical to the shine appearing in the present situation, not, as the reader might have supposed, some other shine. The contrastive set is {a different shine, exactly the same shine}.

Never in SC, sometimes in Alb., and frequently in MG, gradual CTE foci also occur in comments:

¹ The corrective strategy with fronted narrow foci is also operative, if not even more frequent, with zero subjects: *ðen itan Iðrea, Kranidjotisa itan* (MG, Kapandai, p.117) (not was Hydriot, Kranidiot was) – ‘She was not from Hydra, she was from Kranidi.’

(8-148) MG (Matesis, p. 72)

ΘΑΥΜΑ ΘΕΟΥ ine i epitixies pu ixame tote ston andra.

miracle of-god are the successes which we-had then in-the man

‘[Women at that time did not have all the make up, etc., that women have today.]

It is UNBELIEVABLE that we had so much success with men at that time. (= Our success with men at that time is a miracle.)’

There are many grades of being miraculous, and after the description of the possibilities women had during the war to make themselves attractive, the reader does not really expect to hear that the degree of being miraculous was extremely high as far as the success with men is concerned. But it is precisely this that the writer asserts, and her assumption that the reader has expected a lower degree of the property in question than asserted is explicitly marked with a narrow focus phrase ‘miracle of god’ in the ff-slot.

(3) Questions and answers. In Alb. and MG, but, contrary to expectation, not in SC, yes/no questions with a narrowly focused element are often asked with this element in the ff-slot. Consider the following example: Mr. Scott is trying to separate two dogs that are fighting; a person approaches him, and Mr. Scott asks:

(8-149) Eng: *YOUR dog?* (London, p.140)

Alb: *JUAJI është qeni?* (p.138)

your is dog-the

MG: *ðIKOS su ine o skilos?* (p.170)

own your is the dog

SC: *Je li ovo VAŠ pas?* (p.157)

is QU:CLIT this your dog

Mr. Scott presupposes that the dog belongs to someone (‘the dog is of-X’) and uses the question as a directive speech act in order to convey to the hearer his wish to know the identity of X (see 6.5.1.1. and 8.2.1. for the same analysis applied to yes-no questions with polarity focus and to constituent questions). In contrast to constituent questions, in this kind of question the speaker has to narrow down the number of alternatives to the minimum. Namely, what he has to presuppose in order to ask a question like (8-149) is not only ‘the dog is of-X’, but also ‘the dog is yours or the dog is not-yours’, since his intention is not merely to find out to whom the dog belongs, but rather *whether* the dog belongs to the person who has just approached the fighting scene or not, because only this is relevant in the given context: Mr. Scott wants to know if the person may somehow influence the dog.

If this analysis is correct, then we are again dealing with the binary CTE focus: the set of alternatives has the form {X, not-X}. This is confirmed by the preferred lexical items appearing in this question type: pronouns and words whose denotata naturally occur in pairs. Why is it, then, that SC does not use this strategy, given the fact that its ff-slot does accept binary CTE foci? The reasons are purely formal in nature, and have nothing to do with the semantics of focus. In modern SC, yes/no questions are formed with the Wackernagel clitic particle *li* preceded by the non-clitic form of the copula/auxiliary *jesam* (am), *jesi* (are), *je* (is), etc., even in those tenses which are not formed with auxiliaries (in this case the form *je* is generalized for all persons). In other words, the ff-slot is already occupied by the complex AUX+*li*, and the narrow focus expression cannot be placed there, to the effect that in questions only the clause-final narrow focus construction is used, as witnessed by the SC translation of (8-149)¹.

Once the binary alternatives are established by the question, they remain as the only relevant possible worlds in the common ground. Hence it does not come as a great surprise that questions of this kind are frequently answered with the clause-initial, fronted narrow focus phrase, i.e. with a binary CTE focus. The following exchange was heard at the entrance to the Institute of Byzantine Studies in Athens (04.11.2000). A guest asks the porter whether professor X is in his office; having received the positive answer, he proceeds:

(8-150) guest: *KATO ine to yrafio?*

down is the office

porter: *Ne, KATO ine to yrafio.*

yes, down is the office

‘»Is his office downstairs?« »Yes, his office is downstairs«’

Remarkable for the porter’s repetitive effort², this sentence nicely illustrates the principle of asking and answering with fronted narrow foci. The same construction exists in Alb., but, of course, not in SC.

¹ In the 19th century literary language, as well as in some contemporary eastern Serbian dialects, the possibility to front the narrow focus phrase and place *li* directly after it is a living alternative: the sentence *Pijan li je...* (‘drunk *li* is’), taken from Popović (1997: 272), is a nice example of this now obsolete structure.

² Which is, by the way, in no way exceptional in a culture in which ‘effusiveness is a highly valued ... component of behavior between people’ (Mackridge 1985: 338), as also witnessed by the following example:

Peđi điko su? Ne, điko mu to peđi. (Matesis, p. 67)

child own your yes own my the child

(4) Conclusions. Paragraphs are sometimes concluded with summarizing sentences meaning something like ‘this is how it was’, or ‘that is what he said’, or similar. In Alb., MG, and SC, the pronominal focused expressions referring to the content of the whole previous paragraph are regularly placed in the ff-slot, so that the summarizing sentence receives an exclusive binary interpretation: ‘THIS is how it was, and not, as you may have thought before, some other way’, with the contrastive set {some other way, this way}. The following example illustrates this:

(8-151) SC (Andrić, p. 66)

TAKO je izgledala Ćamil efendijina istorija.

so AUX:CLIT looked of-Ć. efendi history

‘[After a long description of Ćamil’s life:] THAT is how Ćamil-efendi’s life story looked like. (=THAT was Ćamil-efendi’s life story.)’

(5) Openings. The opening of a new paragraph with a fronted narrow focus phrase is a strategy which is, with one exception, non-existent in modern SC, but represents a productive narrative device in Alb. and, in particular, MG. The state of affairs explicitly stated or easily derivable from the previous paragraph is presented as presupposed, and an attribute modifying it is focused, introducing the new temporal, modal, or some other frame. Nice examples of this principle are (8-133) and (8-134). Thus, the context for (8-133) (*POLI timiθike tote m’ afto i mana mu*) is as follows: Mrs. Phani plans an excursion to the seaside and wants to invite the whole neighborhood. She first goes to the narrator’s mother to inform her of her plans. Now comes the sentence *POLI timiθike ...* (My mother felt VERY honored with it at that time), followed by an explanation (the narrator’s mother was despised by many because of an affair she had with an Italian soldier). The new paragraph is opened with a ff-sentence: the writer presupposes that the mother was honored (it is a non-controversial fact that people feel honored when invited as first), and it is asserted that the grade of honor was high, by focusing the adverbial *poli*, ‘very’. The topic of honor is thus introduced by the presupposed material, and further elaborated upon in the paragraph thus opened.

It is primarily the expressions belonging to the class of gradable lexems (groups 1.b, 2.b.c.d., 3.a.b.c.) that occur in this context, so that it is no wonder that SC, which generally does not allow for gradable CTE foci in the ff-slot, does not use this strategy.

There is one exception, though. As noted above, of all gradable CTE focus expressions, only adverbials of duration occur in all three languages, i.e. also in SC (see e.g. (8-132)). Consequently, when adverbials of duration are involved, SC also uses the ff-strategy for paragraph openings. Here are some examples from Alb and SC:

(8-152) Alb. (Kosovarja, p. 8/50)

PEŠË vjet më gënjeu zemra mua, dhe Facën tim.

five years me:CLIT lied heart-the my and Facë-the my

[The narrator had a long and passionate affair with certain Facë. Then her family inform her that she has to marry someone else, with whom they had arranged the marriage. She is desperate.] ‘My heart has been cheating on me for whole FIVE years, me and my Facë. [Young people should not rely only on what the heart says, but also on reason...]’

(8-153) SC (Kiš, p. 91)

OSAM ga je godina pratila Segidulinova senka, kao zla kob ...

eight him:CLIT AUX:CLIT years followed of-S. shadow, as evil fate

‘[Two prisoners of the Gulag, Segidulin and Korshunidze, had a fight, which the latter lost; Korshunidze is sent to other concentration camps, where he often tries to commit suicide, feeling disgraced by the defeat.] ‘Segidulin’s shadow persecuted him for whole EIGHT years, like a fate... [He received insulting letters from Segidulin, in which he was continuously reminded of his disgrace.]’

In both cases, the new theme is introduced by the presupposed material more or less easily derivable from the context (my heart lied to me, Segidulin’s shadow persecuted him); it is the unexpectedly high grade of duration of these presupposed states of affairs that is focused.

Paragraph openings with a fronted narrow focus phrase are a highly conventionalized narrative device based on the hearer’s assumed expectations concerning the grade of the presence of a property in a presupposed state of affairs. In quite a number of cases, like (8-153), one has the feeling of pretense (see 4.2.1.): what we feel should be asserted, because newsworthy and ‘new’, is treated as presupposed. Thus, in (8-153), it is not self-evident that Segidulin plays an active role in Korshunidze’s feeling of shame, so that the proposition ‘Segidulin’s shadow persecuted him...’ should rather be asserted than presupposed.

However, all three languages have a conventionalized discourse strategy according to which new paragraphs may be opened with an *ff*-clause if a gradable term is present in the proposition and if the grade of the presence of this term’s denotatum is higher/lower than the world knowledge allows us to expect (the difference between the languages being that SC allows only for adverbials of duration in this use, whereas in Alb. and MG all kinds of gradable terms occur). This conventionalized discourse strategy sometimes overrides the demands of optimizing information processing by asserting less self-evident states of affairs. What we get is the impression of compressed information, with the marking of the contrary-

to-expectation focus taking the overhand over the marking of easily presupposable proposition chunks.

In only one of the three languages – as could be expected, MG – the possibility to presuppose hardly presupposable propositional functions with gradable CTE foci in paragraph-opening contexts has led to a further conventionalization. Article titles and opening sentences in MG journals are often construed as ff-inversion clauses. In my corpus, there are 27 such sentences, most of them from the daily newspaper *Eleftherotipia* (23). Here are some examples of article titles:

(8-154) MG (*Eleftherotipia*, p. 6)

Tin ipoyrafi mas se lefko XARTI θeli o Denktas!

the signature our on white paper wants the D.

‘Denktaş wants us to give him a carte BLANCHE!’

(8-155) MG (*Eleftherotipia*, p. 20)

Tin ARSI ton kiroseon kata tis Afstrias apofasise i Evropaiki Enosi.

the raising of-the sanctions against the A. decided the E.U.

‘European Union decided to RAISE the sanctions against Austria.’

Newspaper article titles give a minimal number of context clues for the formation of presuppositions. Thus, the reader starting to read the article whose title (8-154) is has no indication whatsoever from the context that the president of the Turkish part of Cyprus, Denktaş, wants something; the behavior one would expect from the journalist writing the article would be to first assert this fact, and then to add the information on what it is Denktaş wants, or at least to make one assertion comprising all this: either something like [Denktaş wants something]_{asserted} [he wants [carte blanche]_{asserted}], or at least [Denktaş wants a carte blanche]_{asserted}. The latter variant would be successfully conveyed by vS-construction, in which the verb and the subject build a focus domain, i.e. represent the asserted part of the proposition conveyed, and which very frequently, precisely because of its assertional structure, occurs in newspaper article titles (see 11.5.2.2.1.).

Why is it, then, that I contend that we are dealing with ff-inversion in the clause type exemplified by (8-154) and (8-155) rather than with vS-construction, i.e. that the assertional structure of (8-154) is [[Denktas wants]_{presupposed} [a carte blanche]_{asserted}] rather than any of the above? I have two structural and one discourse-pragmatic reason: First and foremost, the intonation of these sentences is not that of the vS-construction, with the nuclear stress on the subject, but is rather identical to the tune described for ff-inversion, with a high tone nuclear stress on the preposed element and the unstressed postfocal rest. Second, and less conclusive,

the position of assertional elements other than the verb and the subject in the MG vS-construction is typically not the one before the verb but rather between the verb and the subject. Third, this kind of journal article titles as a rule refers to the states of affairs which are potentially highly present in the consciousness of the members of the linguistic community – (8-154) is a title of an article dealing with the topic which was of utmost importance for Greece at the time the journal was published, namely the negotiations between Greece and the Turkish part of Cyprus, and (8-155) refers to a long-lasting debate on the question of the sanctions against Austria. What the journalists using ff-inversion titles do is to presuppose the states of affairs which are potentially expected as article topics by their readership and assert an unexpected aspect of these states of affairs. Thus, the journalist writing the title (8-155) exploits the fact that the readers of a serious daily newspaper may expect information on the outcome of the meeting of EU prime ministers concerning the sanctions against Austria and presupposes the propositional function ‘EU decided X’; then he asserts that X is raising of the sanctions, and by using the fronted focus expression signals that among the two alternatives, raising vs. keeping, the former was – unexpectedly – chosen. Precisely this is the flavor all journal titles of this kind have: a new development in a socially highly relevant state of affairs is described as unexpected, surprising choice among the alternatives.

Note again that this discourse strategy of title giving is limited only to journalistic style and only to MG; in other registers and in the other two languages it is not attested.

(6) Exclamations. In Alb. and MG, but not in SC, exclamatory clauses conveying surprise often contain fronted narrow focus expressions, which mostly belong to the class of quantified/quantifying expressions (groups 1. and 2. above) or to the class of expressive qualification (group 3.). Syntactically, it is almost only predicatives and adverbials that are fronted in these contexts¹.

(8-156) Alb. (Bishqemi, p. 118)

Kaq SHQETËSUES qenka ky ndryshku!

so damaging is:ADMIR this rust

‘[After a scientist had explained to the children why rust is dangerous, one of the children exclaimed:] »Boy, is rust DANGEROUS!«’

The English translation of (8-156) tries to emphasize the functional (though not the syntactic) similarity of the Balkan exclamatory clauses with the English exclamatory inversion

¹ Of course, other types are also allowed, as witnessed by the MG example (8-125), which contains an existentially quantified object phrase in the ff-position, but the prototypical exclamatory clause has a focused predicative/adverbial.

(McCawley 1973, Michaelis and Lambrecht 1996, Fillmore 1999): what the speaker wants to convey is the fact that s/he did not expect that the property expressed by the focused phrase is so good or bad, or present to so high a degree (the effect of surprise being additionally emphasized by the admirative morphology in the Alb. example). Thus, the focused phrase in exclamatory clauses works as a gradual contrary-to-expectation device.

A final note on examples (8-129) and (8-130), where non-quantified items in the ff-position are interpreted as universally and existentially quantified, respectively. As indicated, Giannakidou (2000) argues for a semantic interpretation of this phenomenon. I think that my two types of CTE contrastive focality represent a less *ad hoc* explanation. In (8-130), the bare singular *leksi* ('word') in the ff-slot of a negative sentence is understood as 'no word at all', which is easily accounted for by the principle of pragmatic enrichment (Sperber and Wilson 1986): the hearer finds a word not belonging to the class of words which trigger binary interpretation in the slot which requires such an interpretation; in order to solve the conflict, s/he adds the most plausible meaning to the word so as to *make* it fit into the syntactic slot it occupies. This meaning happens to be the existential quantification in the given context. *Mutatis mutandis*, the same holds for (8-129): in order to fit it into the ff-slot, the hearer interprets the non-gradual expression 'a life' as a gradual notion ('a part of life', 'whole life'), and the maximal duration, i.e. universal quantification, is chosen among the grades which constitute the contrastive set.

Now let me summarize the results of this section in the form of a table. The signs used read as those in Table (8-145):

(8-157) *Discourse functions of ff-clauses in Alb., MG, and SC*

	MG	Alb.	SC
1.corrections [binary CTE]	++	+	+
2.a.comments [binary CTE]	++	+	+
2.b.comments [gradual CTE]	+	+	–
3.questions/answers [binary CTE]	++	++	–
4.conclusions [binary CTE]	+	+	+
5.a.openings [gradual CTE]	++	+	–
5.b.openings [adverbials of duration]	++	++	++
5.c.openings [article titles]	++	–	–
6.exclamations [gradual CTE]	+	+	–

Note that in SC none of the discourse functions demanding gradual CTE foci is present (apart from adverbials of duration), and that the differences between Alb. and MG are those of frequency rather than of ontology: though in both languages ff-phrases cover the same range

of meanings, in certain contexts, MG seems to use these phrases with a frequency which by far surpasses the one they have in Alb.

8.3.2.3. Summary

Fronted narrow focus expressions in Alb., MG, and SC encode the narrow scope of assertion marked by a type of contrast which runs contrary to the assumed expectations of the hearer. The contrast conveyed by these focus expressions is either binary, meaning that the number of alternatives to the one asserted is narrowed down to one, or gradual, meaning that the alternatives are ordered on a scale comprising different grades of the presence of a property. Whereas Alb. and, especially, MG, make use of both types of CTE foci, in SC only the former, the binary CTE focus, is present. The only exception to this are adverbials of duration, which, although gradable terms, appear in paragraph openings even in SC.

These basic pragmatic features of the CTE focus determine both the kind of lexical items prototypically paired with it and the discourse functions which the clauses containing it fulfill. The contrary-to-expectation feature is responsible for the quantified/quantifying and ‘expressive’ expressions employed in the ff-slot, the binarity/gradability feature for the lexical items intrinsically connected with binarity or those which are inherently gradable. What is especially interesting is that, apart from corrective contexts and from *only*- and *just/precisely*-phrases, which are lexically unconstrained, all other discourse functions display clear preferences or even restrictions as to the type of lexical items filling the ff-slot. In other words, if the discourse context itself does not have an inherent binary/gradual interpretation, it is the lexical semantics of one of the elements of the proposition itself that licenses focus fronting: if it allows for a binary/gradual reading, then the proposition will, under appropriate discourse conditions, be coded as a ff-clause. In some cases, as in paragraph openings, especially when adverbials of duration are involved, the focus-fronting strategy is so conventionalized that it may override the needs of optimizing the flow of information.

All this means that both the Praguean notion of ‘emotional’ word order and É. Kiss’s idea of ‘contrastive’ foci are essentially correct, though rough and imprecise. Clause-initial narrow foci are more ‘emotional’ and ‘subjective’ than their clause-final counterparts, since they explicitly mark that the state of affairs conveyed runs contrary to expectations in one of its aspects. They are contrastive, since they operate over contrastive sets. It is even true that they are exhaustive, since the number of alternatives included is always so limited that the exclusion of one of them means the exhaustive reading of the asserted alternative. What I have

tried to show in this chapter is that such general characterizations, though correct, are not enough if one tries to determine both the differences between languages and the variation within languages.

8.3.3. Alternatives to ff-inversion

As indicated in 8.3.1., ff-inversion clauses only rarely contain non-ratified material, mostly because of the fact that when a topic is to be ratified, alternative constructions are used. There are several such constructions.

The first appears only in SC, since it is only in this language that there is no adjacency constraint. The focus expression remains fronted, but the inversion does not take place. Surprisingly, this kind of alternative is used almost only with pronominal subjects (which, as in all other types of focal inversion, resist postverbal position), whereas full NPs normally occur in one of the following two alternative constructions.

The second alternative construction is identical to the prefocal NITops and NPTops in wh-inversion: a newly introduced topic or a new paragraph topic is placed in front of the focus phrase. The construction is relatively rare in Alb. and SC and relatively frequent in colloquial MG. Thus, the MG sentence (8-150) (*KATO ine to yrafio?*) can be construed as

(8-158) *To yRAFIO KATO ine?*

the office down is

if the person asking this question introduces the discourse referent ‘office’ for the first time or feels the need to ratify its topical status for some other reason.

The alternative strategy that is most common in all three languages is the simple use of clause-final focus expressions, i.e. of the typical Balkan focus domain with narrow focus construal. Recall that it has often been noted that clause-final narrow focus expressions can also be ‘contrastive’ (8.3.2.). In my terms, this means that they can mark CTE narrow foci. As CTE is only one of the readings of the clause-final narrow foci, I should like to propose an analysis in terms of markedness. In the pair [clause-initial narrow focus] [clause-final narrow focus], it is the former that represents the marked member of the opposition, whereas the latter is the unmarked member, comparable to the opposition between the lexemes [she-wolf] [wolf].

In the previous chapter, it has been shown under which conditions the marked member of the opposition is used. What is of interest here is when it is *not* used, i.e. when the clause-final focus is used instead, all other things being equal.

First, in ratification contexts. If the topic of the sentence is to be ratified, and if the construction with prefocal topics described above is not used (and it is generally not used in Alb. and SC), then the speakers are likely to use the canonical construction with the fronted non-ratified topic and the clause-final focus expression:

(8-159) SC (from Popović 1997: 98)

KRITIKA je značila POHVALU i PREPORUKU.

Critique AUX:CLIT meant praise and recommendation

‘[A description of the strange situation in the literary life of the 19th century.] The critique meant praise and recommendation.’

Although the focus expression *pohvalu i preporuku* carries the essential properties of CTE foci in SC – the contrary-to-expectation interpretation and binarity (the set being {reproach, praise}) – it is not fronted. The reason for this is the need to ratify the NITop *kritika* (which could be paraphrased by ‘as far as the critique is concerned...’), which is achieved by placing it clause-initially. The focus expression, although fulfilling the requirements for the ff-slot, is encoded as an unmarked (in the above defined sense of the term) clause-final focus. The result is non-inversion.

Second, the two principles frequently invoked in the whole section on focal inversion, namely *the cataphoric adjustment principle* and *the heaviness principle* (cp. Section 5.2.) may override the need to mark CTE-interpretation overtly and lead to the clause-final focus. If the focus expression is in any way syntactically or pragmatically relevant in the following discourse, it will be placed clause-finally rather than clause initially. The same tendency is at work if the focus expression is very long. Sentence (8-160) illustrates the heaviness principle: in contrast to many focus-fronted (short) predicatives modified with the pronominal adverb *aq* (‘so’), the long predicative *aq e komplikuar ...* is clause-final, although it is used in the typical ff-context, that of paragraph opening:

(8-160) Alb. (Koha Ditore, p. 4)

*Çështja e kthimit të refugjatëve kosovarë nuk është **aq e komplikuar siç është e** problem of-the return of-the refugees of-Kosovo not is so complicated as is the **prezentuar në opinion e Maqedonisë.***

presented in opinion of-the M.

‘The question of the return of the Kosovo refugees is not as complicated as it is presented in the public opinion of Macedonia.’

Thus, there are three possibilities to avoid focus inversion, mostly in order to ratify a topical subject. The first, fronted focus + non-inversion, occurs only in SC and is basically restricted

to pronominal subjects. The second, the use of prefocal topical expressions, is relatively frequent only in MG. It is the third alternative that is in general use: instead of focus fronting, the unmarked construction with a narrowly construed focus domain is used. This alternative is employed not only when the subject is to be ratified as the sentence topic, but also when the focus phrase is either very heavy or somehow cataphorically bound.

This is, then, (a part of) the explanation for the optionality of ‘focus movement’ (as clause-initial foci are frequently formalized in the generative literature). Since clause-final narrow foci represent the unmarked member of the opposition clause-initial vs. clause-final focus, they can take over the task of clause-initial focus expressions every time there is a need to posit the topical subject in the clause-initial position for ratifying reasons, or when the length or the cataphoric relevance of the focus expression itself favor the clause-final position for processing reasons.

8.3.4. Conclusion

The type of verb-subject order I labeled ff-inversion stands somewhere between quotation inversion and wh-inversion. Syntactically, it resembles wh-inversion, in that the adjacency constraint exists in Alb. and MG, but not in SC, in that the presupposed postfocal material may, though infrequently, contain non-ratified elements, and in that it allows for an alternative construction with prefocal topical material. Pragmatically, it is more like quotation inversion. It generally occurs only in contexts which facilitate all-ratified presupposed propositions, whereas ratification is most frequently done with an alternative construction containing a narrowly construed focus domain. A specific formal feature of ff-inversion is the nuclear accent on the fronted phrase and the concomitant lack of accent on the rest of the clause.

The most interesting feature of this construction, however, are the conditions under which it is used. I have tried to show that the principal criterion is the nature of assertion conveyed – it has to be a CTE assertion, with focus operating over binary or graded alternatives in Alb. and MG, and only over the binary ones in SC. Different paths of discourse conventionalizations, i.e. non-structural parameters, also play a role: MG has generalized the use of ff-inversion to many discourse contexts in which it only seldom occurs in Alb. and, even less, in SC. The reason why focus fronting does not seem to be obligatory is its marked status with respect to the clause-final narrow focus, so that the latter can, due to its unmarked nature,

overtake the function of the former whenever higher-order principles, like ratification of topics or the heaviness principle, demand it.¹

¹ Everything that has been said in 8.3. obtains also in those cases in which the subject itself represents the narrow focus expression. Thus, it can be fronted if it fulfills the conditions for narrow focus fronting (understandably, no ff-inversion occurs in this case), as in the following Alb. sentence:

Vetëm njerëzit PRIMITIVË mund të kenë pasur të tilla këmbë. (Bishqemi, p. 20)

only people-the primitive can SUB:CLIT have had the such feet

‘Only PRIMITIVE people could have had such feet!’

Narrowly focused subjects with CTE-reading can also be placed clause-finally, under the same conditions as other such focus expressions; in the following Alb. sentence, the principle of cataphoric adjustment is at work:

Me çajin e mbusha gojën UNË, e jo AI. (Bishqemi, p. 67)

with tea it:CLIT filled mouth-the I, and not he

‘I swallowed that tea, not he.’

More on subject foci in Section 11.5.1.

9. Subordinate inversion

Subordinate inversion is a verb-subject construction occurring in existentially presupposed subordinate clauses with a subordinating word in the initial clause position. The basic formal features of this inversion type are the sporadic occurrence of obligatory adjacency between the subordinator and the verb, so that the verb has to precede the subject and the rest of the clause, and, as in the case of focal inversion, a specific intonational contour. In contrast to focal inversion, the principal discourse-pragmatic feature of the verb-subject complex is not the relational presupposedness (not all subordinate clauses have to be topical), but rather the presence of the existential presupposition: subordinate clauses in which inversion occurs are existentially presupposed, since the propositions they convey function as terms on the semantic level. As in focal inversion, verb-subject order in these subordinate clauses is embedded in a presuppositional context, fulfilling thus the chief criterion for treating a construction as inversion.

There are two types of subordinate inversion, the one appearing in relative clauses and the other in embedded complement and adverbial clauses. These two types are conveniently labeled relative and embedded inversion.

9.1. Relative inversion

When the relative word introducing a relative clause is an object, an adverbial, or has any other syntactic function other than subject, the verb and the subject of the relative clause are often found in verb-subject order. I shall call one subtype of this clause type **relative inversion**:

(9-1) Eng: *all the evil that had been wrought him at the hands of men.* (London, p. 150)

Alb: *gjithë ato të këqia që i kishin sjellë duart e njerëzve.* (p. 148)

all those the evils REL him:CLIT have brought hands-the the of-the-men

MG: *ola ta martiria pu tu ixan epivali ta xerja ton theon.* (p. 182)

all the torments REL him:CLIT have inflicted the hands of-the gods

SC: *sve zlo koje su mu nanele ljudske ruke.* (p. 169)

all evil which AUX:CLIT him:CLIT brought human hands

9.1.1. Balkan relative clauses

Before taking up the issue of the verb-subject order occurring in relative clauses, a couple of notes on the structure of these clauses in Alb., MG, and SC are in order.

As in many other languages, relative elements in Alb., MG and SC have to be clause-initial, the only exception being relative pronouns in genitive, which may occur after their head noun (which then has to be clause-initial).

There are two types of relative elements in all three languages. The first one is a relative pronoun typical for most Indo-European languages, with all the morphological features of a pronoun, including gender, number, and case marking. Its form is *i cilë* (m.), *e cilë* (f.) in Alb., *o opios* (m.), *i opia* (f.), *to opio* (n.) in MG, *koji* (m.), *koja* (f.), *koje* (n.) in SC.

(9-2) Eng.: *the hand that was now bandaged* (London, p. 148)

Alb: *dorën, e cila tani ish e lidhur* (p. 145)

hand-the:FEM which:FEM:NOM now was the bandaged

MG: *tu xerju, to opio tora itan ðemeno* (p. 180)

of-the hand:NEUT which:NEUT:NOM now was bandaged

SC: *ruku, koja je sad bila u zavoju* (p. 167)

hand:FEM which:FEM:NOM AUX now was in bandage

In the second type, the relative element is a particle without gender, number, and case marking, partly resembling the English pronoun *that* when used as a relative. In Alb., its form is *që*, in MG *pu*, and in SC *što*. The number of grammatical roles the relativized element marked with this particle can play is quite limited in Alb. and SC (subject, direct and indirect object), whereas in MG objects of prepositions and genitive possessors are also allowed¹. Here is an example of *që*, *pu* and *što* in the subject role:

(9-3) Eng: *two men who were not yet dead* (London, p. 4)

Alb: *dy njerëz, që nuk kishin hequr dorë ende nga jeta* (p. 4)

two men REL not have drawn hands yet from life-the

MG: *i ðio andres pu ðen ixan akomi peθani* (p. 8)

the two men REL not have yet died

SC: *ona dva čovjeka što još ne bjehu mrtvi* (p. 6)

those two men REL yet not were dead

¹ The differences in the range of grammatical roles played by the relative particle notwithstanding, there is a common tendency in all three languages to use the relative pronoun instead of the particle when the relativized element is anything else apart from subject or direct object.

When the relativized element expressed by the relative particle is not the subject or the direct object, it has to be additionally marked by a resumptive pronoun in all three languages, formally identical to the clitic forms of the personal/demonstrative pronoun. Here is an example from MG:

(9-4) MG (Kapandai, p. 20)

makri trapezi pu epano tu ixan idi aploθi ta fajita

long table REL upon it were already spread the dishes

‘a long table upon which the dishes have already been spread out’

In case the relativized element is the direct object, it *may* be resumed with a clitic pronoun. In SC, the relevant criterion seems to be ontological in nature: if the relativized element is a state of affairs, no resumption occurs, if it is an entity, the resumption is obligatory. The conditions under which the resumptive pronoun is used in Alb. and MG are not entirely clear: there seems to exist a consensus that restrictive relative clauses do not take a resumptive pronoun, whereas the appositive (non-restrictive) ones do (see Stavrou 1984, Mackridge 1985:225ff., Haberland and van der Auwera 1990: 142ff., Alexiadou and Anagnostopoulou 2000: 194ff., and the references therein). However, as Stavrou, Mackridge (p. 225), and others observe, there are cases in which the resumptive pronoun does occur in restrictive relative clauses, which is confirmed by the data from my corpus (five examples in MG, nine in Alb.). Thus, at the present state of knowledge, we may only say that in Alb. and MG resumptive pronouns with relativized direct objects tend to occur in appositive relative clauses and not to occur in the restrictive ones, but this has a status of a tendency, not of a rule.

The two types of relative elements described above do not have the same sociolinguistic status in the three languages under consideration. Not untypical for the general attitude towards Balkanisms in Albanian and Greek linguistic communities, relative particles (*që* and *pu*) are considered to be colloquial (although they do occur, especially in the last couple of decades, even in highly elaborate styles), whereas relative pronouns (*i cilë* and *o opios*) are more formal. In SC, the reverse holds: If referring to an entity, the relative particle *što* is poetic, archaic, etc., and as such practically non-existent in the colloquial language; on the other hand it is universally applicable if it refers to a state of affairs. The relative pronoun *koji* represents the normal choice of the speakers of contemporary SC, save for the reference to states of affairs, where, as indicated, the relative particle is normally used.

9.1.2. Statistics

In Section 5.5. it has been noted that verb-subject order is relatively frequent in relative clauses in comparison to other clause categories: If all the relatives are taken into account (i.e. also those in which the relative element is the subject of the relative clause, so that the verb-subject order is *per definitionem* excluded), the number of verb-subject clauses amounts to 13.5% in Alb., 17.0% in MG, and 13.1% in SC.

Let us now look at the statistics in some more detail, starting with the grammatical role of the relative element itself. In all three languages, the relative element is the subject of the relative clause (as in *the man **who** came in*) in the majority of the cases – more precisely, out of 1432 relative clauses in my Alb. corpus, 1024, i.e. 71.5%, are of this form, out of 1356 MG relative clauses, the number of relative subjects is 859, i.e. 63.3%, while in SC the ratio is 1866 versus 1061, i.e. 56.9%. In what follows, I shall ignore this type of relative clause, since the obligatory subject-initial position makes it irrelevant for the question of verb-subject order.

What remains are 408 clauses in Alb., 497 in MG, and 805 in SC, in which the relative element is not the subject, i.e. in which the verb-subject order *may* occur; in 5.5.5. I called this clause type **oblique relative clause**. The grammatical roles of relative elements are distributed as follows¹:

(9-5) *Relative elements in oblique relative clauses*

	object	setting adv.	modal adv.	genitive	Σ
Alb.	184 – 45.1%	168 – 41.2%	48 – 11.7%	8 – 2.0%	408 – 100%
MG	303 – 61.0%	151 – 30.4%	34 – 6.8%	9 – 1.8%	497 – 100%
SC	355 – 44.2%	336 – 41.7%	88 – 10.9%	26 – 3.2%	805 – 100%

Apart from the fact that in the MG corpus the relative elements tend to be objects somewhat more often than in the Alb. and the SC ones, the distribution of grammatical roles seems to be relatively consistent across languages.

More interesting is the ratio of zero subjects, verb-subject and subject-verb orders within the class of oblique relative clauses. I repeat the results from Section 5.5.:

¹ Direct, indirect and prepositional objects are subsumed under the label ‘object’ (i.e. *the man whom I saw, the man to whom I spoke, the man about whom I spoke*); by ‘setting adverbial’, both locative and temporal expressions, both adverbs proper and prepositional phrases, are meant (*the place where I lived, the day when I was born, the house in which you saw him*); relative modal adverbials are the words meaning ‘as’ (*so, as I did it*).

(9-6) *Ratio of zero-subjects, SV and VS orders in oblique relative clauses*

	zero subject	subject-verb	verb-subject	Σ
Alb.	125 – 30.7%	89 – 21.8%	194 – 47.5%	408 – 100%
MG	221 – 44.5%	45 – 9.0%	231 – 46.5%	497 – 100%
SC	331 – 41.1%	229 – 28.5%	245 – 30.4%	805 – 100%

Almost half of all oblique relative clauses in Alb. and MG, and almost a third in SC, have verb-subject order. The predominance of this order becomes even clearer if only the clauses with overt subjects are taken into account: the ratio of VS and SV in Alb. is 69.6% vs. 31.4%, in MG 83.7% vs. 16.3%, and in SC 51.2% vs. 48.3%¹.

These ratios reveal two things. First, in contrast to most other formally definable clause types, like complement, adverbial, main declarative, etc., clauses, in oblique relative clauses verb-subject order is the dominant word order in all three languages, though somewhat less in SC than in Alb. and MG. Second, in spite of the prevalence of VS order, in all three languages there is a possibility to place the subject between the relative expression and the verb (in SC, not surprisingly, more than in Alb. and MG). This means either that there is no adjacency constraint between the fronted (relative) element and the verb at all, so that the subject inversion is to be explained by some other principle, or that this constraint is in some relevant way not absolute. I shall investigate this issue in the following section.

9.1.3. Adjacency constraint and restrictive vs. appositive relative clauses

The situation in SC is clear: as in the case of question words and fronted focus expressions, this language does not have any kind of adjacency constraint holding between the relative element and the verb, so that the frequent use of verb-subject order is, similar to wh-inversion and ff-inversion, to be explained by a superordinate discourse-pragmatic principle. This is in accordance with the fact that the predominance of VS order in SC relative clauses is significantly less pronounced than in Alb. and MG.

Surprisingly, Alb., which has a rather strong adjacency constraint with question words and fronted focus expressions, seems to side with SC rather than with MG with respect to relative elements: according both to my corpus data and my informants, there is no adjacency constraint holding between the relative element and the verb. The high incidence of verb-

¹ A comparable statistical investigation of oblique relative clauses exists only for MG, with the results very similar to mine: the ratio given in Seifert 1984: 88 (quoted after Haberland and van der Auwera 1990: 154) is 76.1% VS versus 23.9% SV clauses.

subject order in relative clauses in Alb. is thus to be explained, as in SC, by some discourse-pragmatic principle. The statistical difference between the two languages with respect to the VS : SV ratio (51.2% : 48.3% in SC versus 69.6% : 31.4% in Alb.) is reducible to the independently functioning rule which practically forbids the postposition of pronominal subjects in SC (cp. 5.2. and 8.1.1.2.1; more on this in Section 9.1.5.).

What is problematic is MG: one has the impression that there is something like adjacency constraint, but the data show that it is not absolute. In what follows I shall argue that the opposition [+adjacent] vs. [-adjacent] is a mirror image of the opposition [-resumptive pronoun] vs. [+resumptive pronoun], i.e., that in restrictive relative clauses there is a strong tendency to keep the relative element and the verb together (and not to use a resumptive pronoun), whereas this tendency is absent from appositive relative clauses. Comparable to resumptive pronouns, there are some exceptions in both directions which I am not able to account for.

As in wh- and ff-inversion, independent elements (setting adverbials, sentence adverbials) with loose syntactic ties to the verb may sometimes intervene between the relative element and the verb *both* in restrictive and appositive relative clauses. Here is an example of a restrictive relative clause:

(9-7) MG (Kapandai, p. 24)

*...tus kanones pu **stin arxi** mu epevale o pnevmatikos...*

the penitence-measures REL at-the beginning me:CLIT imposed the confessor

‘... the penitential punishments which my confessor inflicted upon me at the beginning...’

This is, as with wh- and ff-inversion, a minor phenomenon, both statistically (15 examples) and syntactically (independent elements have a greater positioning freedom than the elements which are more closely tied to the verb), so that I shall take no further notice of it¹.

More interesting are the cases in which a major grammatical relation – the subject – appears between the relative element and the verb, so that no relative inversion occurs. As shown in Table (9-6), there are 45 instances of non-inversion in my corpus. In 37 (82.2%) cases, the relative element is a setting adverbial (‘where’, ‘when’, ‘in which’, etc.), often in

¹ The same holds true for the cases where the intervening element is the head of the phrase of which the relative pronoun in genitive is the modifier:

*enas andras **pu; o moxthos tu; ixe teljosi** (MG, London, p. 8)*

a man REL_i the toil his_i has ended (‘a man whose toils were over’)

the form *pu*, or as a prepositional phrase with *o opios* (*me ton opion* ‘with which’, *sto opio* ‘in which’, etc.):

(9-8) MG (Ciao, p. 22)

tora, pu to simvoleo tis me ti Sony teljoni
 now REL the contract her with the S. ends
 ‘now, when her contract with Sony expires’

(9-9) MG (Ciao, p. 72)

to metopo tu, apo to opio ta malja exun arxisi eðo ke xronja na areonun.
 the forehead his from the which the hair have begun here and years to thin-out
 ‘his forehead, from which the hair started to disappear years ago.’

In four instances, the relative element is in genitive (see above); the remaining four instances (8.9%) contain a relative element playing the role of direct object:

(9-10) MG (Kapandai, p. 18)

t’ ayapimena mu t’ adelfja, pu o Θεος na tus exi kala
 the beloved my the brothers REL the God SUB:CLIT them:CLIT have well
 ‘my beloved brothers, who I hope God will save’ (= ‘my beloved brothers, may God save them’)

This distribution is not accidental: when the relative element has a narrow syntactic tie to the verb, which is the case with objects, the non-inversion becomes increasingly difficult, whereas independent elements like setting adverbials tend to allow for it more readily (for complete statistical data, see the last column in Table (9-11)). But this is not the whole story: inversion vs. non-inversion (adjacency vs. non-adjacency) has also something to do with the function of the relative clause, with restrictive relatives having the feature [+adjacent], the appositive ones [–adjacent]. This claim can be substantiated by some statistics. The data presented in Table (9-5) show that out of 497 oblique relative clauses in my corpus, 303 have relativized objects and 151 relativized setting adverbials ($\Sigma = 454$; genitive relatives and modal adverbials are left aside). Out of this number, in 201 cases the subject is not expressed at all; 253 clauses remain, 148 with relativized objects, 105 with setting adverbials. This is the basis for my statistics, presented in Table (9-11) (*VS* means ‘verb and relative element are adjacent, the inversion takes place’, *SV* the opposite):

(9-11) *Adjacency vs. non-adjacency in restrictive and appositive relative clauses in MG*

		restrictive	appositive	Σ
relativized object	VS	105	39	144
	SV	1	3	4
	Σ	106	42	148
relativized setting adverbial	VS	36	32	68
	SV	5	32	37
	Σ	41	64	105
Σ	VS	141	71	212
	SV	6	35	41
	Σ	147	106	253

The last row illustrates my point: in restrictive relative clauses non-inversion occurs only in a handful of cases (six instances out of 147, i.e. 4.1%), whereas in appositives the non-adjacency between the relative element and the verb is in no way exceptional (35 instances out of 106, 33.0%). The appositives with non-inversion are represented in all non-inverted examples adduced so far, i.e. (9-8) – (9-10). Here are two examples of non-inversion in restrictives:

(9-12) MG (Elefterotipia, p. 14)

... *ipeyrapse afto pu emis arniθikame*

signed that REL we refused

‘... (he) signed what we refused (to sign)’

(9-13) MG (Kapandai, p. 18)

... *ti mera pu o Mixail Ralis kirikse to sikomo*

the day REL the M.R. announced the uprising

‘... on that day on which Michael Ralles announced the uprising’

Obviously, non-adjacency is possible in restrictive relatives, although it is extremely rare. An interesting result was obtained when native speakers of MG were asked to judge the possibilities of inversion in restrictives like (9-12) and (9-13) in contrast to appositives like (9-8) and (9-9). They claimed that the inverted version of the former (*pu arniθikame emis, pu kirikse to sikomo o M.R.*) would not be semantically/pragmatically different from the attested non-inverted version. In the case of appositives (*pu teljoni to simvoleo, apo to opio exun arxisi na areonun ta malja*), they claimed that they felt some kind of difference in ‘emphasis’.

The conclusion I am prone to draw from these facts is the following: In MG restrictive relative clauses, the relative element and the verb are preferably adjacent, non-adjacency being a rather rare exception. When non-adjacency (and non-inversion) occurs, it does not carry any kind of specific meaning which would be absent from the inverted, adjacent cases.

Non-adjacency and the concomitant non-inversion in restrictive relative clauses thus appear to be marginal variants without a functional load – perhaps a result of a diachronic process (the end phase or the initial phase of a process¹), or of a dialectal mixture, or something else. Be it how it may, the fact remains that in restrictives the adjacency between the relative element and the verb has practically the status of a rule.

MG appositive relative clauses, on the other hand, simply have no adjacency constraint, and consequently no obligatory inversion. The fact that verb-subject order does occur very frequently, but brings about a different ‘emphasis’, points to a discourse-pragmatic source of this frequency, which I hope to identify in the following section.

9.1.4. Information structure, restrictive relative clauses, and relative inversion

Inversion, as I use the term, differs from both vS and VsX constructions in that VS order is embedded in a presuppositional context, i.e. in those portions of the propositions which are of necessity interpreted as existentially (and/or relationally) presupposed.

Restrictive relative clauses differ from the appositives precisely in this feature. In order to use a relative clause as a part of a description of a term, the speaker must treat its content as noncontroversial, i.e. s/he must assume that the auditor is ready to take the content of the relative clause for granted. If I say *A/The woman I saw yesterday had a nice hat*, I expect my interlocutor to take it from me that there is a person corresponding to descriptions ‘X is a woman’ and ‘I saw X yesterday’, so that s/he can go on to process the actual assertion of my utterance, namely that X had a nice hat. Restrictive relative clauses are thus, just like all other linguistic material used for encoding terms (as opposed to predicates), existentially presupposed (cp. 4.2.3.). Appositive relative clauses, on the other hand, function *de facto* as independent main clauses, meaning that they do not constitute a (part of a) term in the proposition, but rather an independent proposition. Thus, if I say *Yesterday I saw your brother, who refused to say ‘hello’*, I do not create the term ‘your brother, the one who refused ...’ in order to bind it to the predicate ‘meet’, but I rather assert two independent

¹ Haberland and van der Auwera (1990: 152ff.), following Hesselning (1927), explain the adjacency between *pu* and the verb by the fact that *pu* is an ‘incomplete clitic’: ‘... there is an *Endstation* Clitic for *pu*, which it has not yet reached’.

propositions, ‘I met your brother’ and ‘he refused to say *hello*’. This, again, means that appositive relative clauses are not instances of the presuppositional, but of assertive context¹.

What is of interest in this chapter, devoted to the relative *inversion*, i.e. to the verb-subject construction occurring in presuppositional contexts and/or under adjacency constraint, are thus only restrictive relative clauses. The verb-subject order occurring in appositive clauses belongs to the other two verb-subject constructions postulated for Alb., MG and SC, the vS and the VsX construction, and has nothing to do with the relative inversion. The fact that appositive relatives also display a ratio of verb-subject order higher than average has to do with the type of topic these clauses have, namely indirect topic, which is the major trigger of vS construction. For this reason, appositive oblique relative clauses will be dealt with in the chapters devoted to this construction.

Obviously, restrictive oblique relative clauses represent the same kind of context as the clauses described in the section on focal inversion – they contain a fronted element, which may be obligatorily adjacent to the verb, and the rest of the clause represents a presupposed propositional function. However, as I mentioned in the introductory note to Section 9, there is an important difference concerning the nature of the fronted element: In all kinds of focal inversion, this element invariably denotes the assertive part of the proposition, the narrow focus, whereas in relative inversion the relative pronoun/particle denotes the primary topic of an entirely presupposed clause. Thus, whereas in *Where does he go?* ‘where’ carries the narrow focus (‘he’ being a ratified topic), in *the place where he went* ‘where’ is the primary (indirect) topic of the clause, and ‘he’ the secondary, ratified topic. This difference should be kept in mind despite the virtual syntactic identity between the two inversion types.

¹ Surprisingly, this apparently self-evident fact is not accepted by everyone. Thus, Levinson (1983: 183) claims that restrictive relatives are affected by the negation of the main verb, and thus not presupposed, whereas appositives are presupposed and immune to the negation of the main verb. The former claim is simply false (*The woman I met yesterday did not have a nice hat* still presupposes the existence of a woman I saw), the latter is true, but interpreted wrongly. The negation of the main verb does not affect the appositive clause for the simple reason that the two clauses (the main and the appositive), in contrast to, say, clauses containing a factive verb and its complement, convey independent propositions. In saying *I liked the new performance of ‘King Lear’, which is rather long* I convey two propositions independent from each other, with the relative clause functioning almost as a coordinate clause, so that it is only logical that the negation of one of the clauses does not affect the truth value of the other, as witnessed by *I did not like the new performance of ‘King Lear’, which is very long*.

9.1.5. Syntax, prosody, and information structure of relative inversion

9.1.5.1. General rules

Now that relative inversion is defined as verb-subject order appearing in restrictive relative clauses, in a presuppositional context, let us concentrate on the syntax of the construction and its relationship to the information structure. Here are the numerical data on relative inversion (MG numbers are altered with respect to Table (9-11) by adding the few cases of relativized modal adverbials and genitives).

(9-14) *Restrictive oblique relative clauses with relative inversion*

	object	setting adv.	modal adv.	genitive	Σ
Alb.	52 – 47.3%	38 – 34.5%	17 – 15.5%	3 – 2.7%	110 – 100%
MG	105 – 68.6%	36 – 23.5%	10 – 6.5%	2 – 1.4%	153 – 100%
SC	63 – 57.8%	22 – 20.2%	23 – 21.1%	1 – 0.9%	109 – 100%

The semantic and syntactic functions of the relativized element are approximately equally distributed across languages, and the number of restrictive clauses themselves is similar, the slight deviation in MG being due to the composition of the corpus (one author, Kapandai, uses restrictive relatives disproportionately often – some 52 instances come from this source alone).

The propositions conveyed by restrictive relative clauses are, as indicated, existentially presupposed. Now, presupposed material may be informationally unified, if it is all-ratified, or it can contain an informationally separated element, a non-ratified topic, as discussed in the chapters on focal inversion. It has also been argued that the difference between all-ratified clauses and the clauses containing different kinds of non-ratified elements is responsible, at least partly, both for the occurrence of inversion and for its non-occurrence (where possible), and for the minor word order variation in inversion contexts (cp. especially 8.1.2. and 8.2.2.). This holds true also for relative inversion: together with the length of the clause, and with the presence or absence of the adjacency constraint between the verb and the relative element in a language, it is the ratification status of the clause elements that determines the word order in restrictive oblique relative clauses.

The syntax and prosody of these clauses find their fullest parallel in the syntax and prosody of wh-inversion: the fronted element – here the relative pronoun/particle – is never accented, so that the nuclear stress has to fall on one of the elements in the verb-subject-X complex. In Section 8.2.2.5., the word order system of wh-inversion (and of focal inversion in general) was presented in the form of rules ((8-104) through (8-108)). Since these are also the rules

that govern the behavior of sentence elements in relative inversion, I repeat them here with necessary adjustments as (8-104) – (8-108):

- (9-15) In all-ratified contexts, the sentence accent is on the verb.
- (9-16) In all-ratified contexts, the sentence accent must not be carried by the last sentence element.
- (9-17) If an element is to be ratified, it must stand at the end of the sentence and carry the sentence accent.
- (9-18) In MG, the relative element and the verb have to be adjacent; there is no such constraint in Alb. and SC.
- (9-19) In SC, pronominal subjects are practically never inverted.

The major difference between relative inversion and wh-inversion is the lack of adjacency constraint in the former in Alb. (its presence in MG has been, exceptions notwithstanding, demonstrated in 9.1.3), so that this language now sides with SC, not with MG, as is the case in all types of focal inversion. Let us now look at the details.

9.1.5.2. All-ratified restrictive relative clauses

In all-ratified restrictive relative clauses, i.e. in those, in which the whole presupposed proposition conveyed is currently lit up in the consciousness of the hearer (or the speaker considers that s/he and the hearer have already agreed on its presupposed status for some other reason), MG has the inversion in practically all cases, irrespective of the length of the clause, conforming thus to rules (8-104), (8-105) and (8-107), the last with all the provisos mentioned above:

- (9-20) MG (Ciao, p. 36)

I kalesmeni tu ...dionde pros timin tu ikoðespoti ... me tus famous firmes
the guests his dress for honor of-the host with the famous labels
pu ANDIPROSOPEVI o epixirimatias.

REL represents the businessman

‘[Mr. Lakis organizes the best parties on Mykonos.] To honor their host, his guests ... always wear the famous labels which the businessman represents.’

- (9-21) MG (Matesis, p. 63)

Mono s' afto xrisimevi i eklisia. Ke ja andiðoro pu sas ELEI kaθe kiriaki
only for that is-useful the church and for host REL you:CLIT gives every Sunday
o papas.
the priest.

‘[Hide in the church and wait till the rain is over.] That’s the only thing the Church is good for. And for the host which the priest gives you every Sunday.’

Example (9-20) illustrates the syntax of relative inversion in short all-ratified clauses, with the proposition ‘the businessman (=Mr. Lakis) represents (some) famous labels’ presupposed and ratified: the inversion is obligatory, the nuclear accent is on the verb. In (9-21), a long all-ratified clause is given: the proposition ‘the priest gives you host every Sunday’ is both presupposed and ratified in the context of the discourse topic ‘the use of the Church’. The inversion is obligatory, the nuclear accent is on the verb, the element X is (as usual in MG) placed between the verb and the subject¹.

In Alb. and SC, on the other hand, the inversion in all-ratified contexts occurs primarily in short clauses, due to rules (8-104) and (8-105), according to which the verb carries the nuclear stress which must not be placed on the last element of the clause:

(9-22) Alb. (Bishqemi, p. 59)

Tingujt e veçantë që LËSHON lakuriqi i natës janë të pakapshëm nga ne.

voices the special REL emit bat-the the of-night are the inaudible from us

[Bats communicate through sounds.] ‘These special sounds which bats emit are inaudible to us.’

(9-23) SC (Kiš, p. 10)

Mikša je ... obećao da će se ... osvetiti za uvredu koju su mu NANELI

M. AUX promised that will REFL avenge for insult which AUX:CLIT him:CLIT brought *talmudisti*.

Talmudists

[Since being dismissed by his boss, the Jewish merchant reb-Mendel, Mikša cannot find a job anywhere.] ‘Mikša ... promised that he will avenge the insult which he received from the Talmudists.’

These examples (both patently presupposed and ratified) show that the syntax and prosody are the same as in MG.

Rule (8-108) is the cause of the major difference between Alb. and SC in this context: whereas in Alb. pronominal subjects behave like all other subjects, they practically never invert in SC, even when in this way rules (8-104) and (8-105) are broken. Consider the following Alb. example:

¹ Here as elsewhere in this chapter, the variations caused by the heaviness principle, the principle of cataphoric adjustment, and similar, will be simply ignored, since they have been sufficiently described in the section on quotation inversion (especially 8.1.1).

(9-24) Alb. (Kosovarja, p. 9)

Çfarë përmban arti që e KRIJONI ju?

what contains art-the REL it create you

[An interview with an artist] ‘What does the art which you create contain?’

The only acceptable SC translation of this sentence is *koju vi KREIRATE* (‘which you create’), the variant with relative inversion (*koju KREIRATE vi*) sounding rather odd. The ban on the inversion of pronominal subjects in SC is thus stronger than the rule prescribing the non-final position of the sentence accent in all-ratified contexts.

Rule (8-107) shows that neither Alb. nor SC have an adjacency constraint. Consequently, relative inversion in long all-ratified restrictive relative clauses is not obligatory, since the last position in the clause may be ‘covered’ by an element other than subject, as in the following Alb. example:

(9-25) Alb. (Koha Ditore, p. 5)

... nga koha që ju VIZITUAT Tiranën ...

from time-the REL you visited T.

‘...from the time you visited Tirana ...’

However, relative inversion does occur in all-ratified contexts in long clauses, both in Alb. and SC. This happens mostly when the additional element (‘X’) is a setting adverbial, usually somehow deictically or anaphorically bound (although non-inversion is more ‘normal’ and more frequent in such cases):

(9-26) Alb. (Koha Ditore, p. 13)

Këshilli i Apeleve vlerësoi se funksioni i lartë që KISHTE atëbotë Kambanda

council the of-appeals judged that function-the the high which had then K.

bën që të përjashtohet çdo mundësi për të zbutur dënimin.

does that is-excluded every possibility for to mitigate sentence-the

[Ian Kambanda, the former prime minister of Rwanda, is the first to receive a life-sentence from the International Tribunal.] ‘The appellate court considered the function which Kambanda had at that time to be the sufficient reason to exclude every possibility of mitigating the sentence.’

(9-27) SC (Kiš, p. 106)

Ova aluzija ... na šešir što ga je u to vreme NOSIO Novski

this allusion on hat REL it:CLIT AUX:CLIT in that time carried N.

nije ...lišena političkog konteksta

is-not deprived of-political context

[Boris Novsky appears in Paris with, as one biographer says, “the most beautiful hat that could be seen in Paris at that time”.] ‘This allusion ... to the hat which Novsky wore at that time is not without political connotations.’

Suffice it for now to state the fact; an explanation of this apparently unnecessary inversion (i.e., unnecessary within the system proposed here) in all-ratified long clauses will be offered in Section 9.3.2.1.

9.1.5.3. Restrictive relative clauses with a non-ratified element

When a restrictive relative clause contains a non-ratified topical element, MG behaves in exactly the same way as in *wh*-inversion and *ff*-inversion. First, because of rule (8-107), the verb comes directly after the fronted element, in this case the relative pronoun/particle; when it is itself non-ratified, it carries the sentence stress (in contrast to all-ratified cases, where this accent seems to be L*, it is clearly H* here); when any other element (subject or X) is non-ratified, it comes at the end of the clause and carries the sentence accent (H*); the relative clause in (9-28) illustrates the syntax and prosody of a non-ratified verb (*ðen proekipse*, ‘has not appeared’, stands in implicit contrast to ‘has appeared, has been proven’), whereas (9-29) contains a non-ratified subject (‘Sophia Alimberti’ is both newly introduced and implicitly contrastive with respect to ‘Iannis Parios’):

(9-28) MG (Elefterotipia, p. 9)

Ine ðikeolojimenō to xasma, θa skefti kanis, oso anaferomaste se aθlites

is justified the gap, will think someone, as-far-as we-refer to athletes

ja tus opius ðen PROEKIPSE i xrisi mempton usion.

for the which not appeared the use of-harmful substances

[There is a large gap between drug addicts and athletes, the former being ostracized, the latter respectful members of the society, although many athletes use doping substances which are no better than other narcotics.] ‘One could consider the existence of this gap justified, if we have in mind those athletes for which the use of harmful substances has not been PROVEN’. (=“in connection with which the use of harmful substances has not been mentioned”). [But this does not mean anything, since there are doping substances which cannot be traced.]

(9-29) MG (Ciao, p. 6)

Apenandi apo to paljo tu spiti, ekino sto opio meni tora i Sofia ALIMBERTI,

opposite from the old his house, that in-the which stays now the S.A.

o Janis Parios ... kitakse ti nea tu katikia.

the I.P. saw the new his residence

[Iannis Parios loves the island of Paros; ‘There, his new house awaits him as a new page in his life.] Opposite his old house – the one in which Sophia ALIMBERTI now lives – Iannis Parios envisaged his new residence.’

Alb. and SC ratify their non-ratified topics in the same fashion as SC does in wh-inversion and ff-inversion (as predicted by rules (8-106) and (8-107)): what is felt to be non-ratified is positioned clause-finally and receives the sentence stress. Inversion thus occurs regularly only when it is the subject that is non-ratified, as in (9-30) below. Non-ratified verbs and non-ratified X-elements either do not allow for inversion, which is the case with non-ratified verbs, since they have to be clause-final, or at least disfavor it, which is the case with non-ratified additional elements. Again, rule (8-108) prevents pronominal subjects in SC from appearing postverbally, though this position is more acceptable here than in all-ratified contexts. The following Alb. example illustrates the position of a non-ratified subject (the International Tribunal in the Hague is contrasted to the Yugoslav courts of justice):

(9-30) Alb. (Koha Ditore, p. 15)

Mirëpo, megjithatë, precizoi Landeyl, një aktakuzë e tillë nuk do të kishte kurfarë efekte ligjore për aktakuzën të cilën e ka ngritur kundër Milosheviqit ... Tribunali

international the of-the-Hague

[Jim Landeyl of the Hague Tribunal announced that the Tribunal does not oppose to the wish of the new government in Belgrade to raise an accusation against Milošević before a Yugoslav court of justice.] ‘»However«, Landeyl explained, »such an accusation has no legal effect on the accusation which was raised against Milošević by the International Tribunal in the HAGUE.«’

9.1.5.4. Summary

In contrast to focal inversion, where Alb. and MG display an identical syntactic behavior, different from SC, in relative inversion we have an opposition between Alb. and SC on one hand and MG on the other, due to lack of adjacency constraint in the former and its presence in the latter. The behavior of the sentence elements in relative inversion is predicted by the

rules (8-104) – (8-108). Here is a set of sentence templates which represent the syntax and prosody of relative inversion:

(9-31) *Sentence templates for relative inversion:*

MG:	Alb., SC:
all-ratified:	all-ratified:
[rel]~[<u>VERB</u>] ([X]) [subject]	[rel] [<u>VERB</u>] [subject]
	(SC: [rel] [pronominal subject] [<u>VERB</u>])
	[rel] [subject] [<u>VERB</u>] [X]
non-ratified subject:	non-ratified subject:
[rel]~[verb] ([X]) [<u>SUBJECT</u>]	[rel] [verb] ([X]) [<u>SUBJECT</u>]
non-ratified verb:	non-ratified verb:
[rel]~[<u>VERB</u>] ([X]) [subject] ([X])	[rel] [subject] ([X]) [<u>VERB</u>]
non-ratified X:	non-ratified X:
[rel]~[verb] [subject] [<u>X</u>]	[rel] [subject] [verb] [<u>X</u>]

9.1.6. Semantic and informational properties of subjects and verbs

Unlike other types of focal inversion, relative inversion displays some lexical properties of the verb and the subject which are best described as textual and/or lexical boundness.

First the easier case, that of textual boundness. In Alb., in nine out of 110 oblique restrictive relative clauses (= 8.2%), the verb has already been mentioned in the previous text; in MG the relationship is 19:153 (12.4%), in SC 14:109 (12.8%). This is in all three languages a percentage much higher than usual for ‘old’ verbs. Here is an example:

(9-32) SC (Andrić, p. 84)

Bajazit je tražio od pape da zadrži Džema kod sebe pod istim uslovima

B. AUX asked from pope to keep:PERF C. by himself under same conditions

pod kojim su ga držali vitezovi sa Roda ...

under which AUX:CLIT him:CLIT kept:IMPF knights from Rhodes

‘Bajazet asked the Pope to keep Cem, under the same conditions under which he was kept by the Knights of Rhodes.’

Under the lexical boundness of the verb and/or the subject I subsume two phenomena.

(1) The verb of the restrictive relative clause is a part of the existential presupposition of the head of the relative clause. Recall that some entities have rich existential

presuppositions (cp. 4.2.4.). If I say *house*, the state of affairs ‘be built’ is easily presupposed, since being built is a part of the rich existential presupposition of ‘house’. Thus, in the relative complex *the house I built*, the denotatum of *built* is a part of the existential presupposition of ‘house’. Consider the following examples:

(9-33) Alb. (Bishqemi, p. 75)

fjalët e mira që tha komandant Ylli

words the good REL said commander Y.

‘kind words which the commander Yll said’

(9-34) MG (Ciao, p. 9)

sta parti pu e xhiti o shtetshpunës

at-the parties REL has given the designer

‘at the parties given by the designer’

The state of affairs denoted by ‘say’ is a part of the existential presupposition of ‘word’, and in order for a party to exist it has to be ‘given’ by someone. This kind of lexical relationship between the verb and the head of the relative clause – I shall call it *lexical solidarity between the relative element and the verb* – is rather frequent in my corpus: 40 instances out of 110 in Alb. (36.4%), 77 out of 153 in MG (50.3%), and 49 out of 109 in SC (44.9%).

(2) The verb of the restrictive relative clause is a part of the rich existential presupposition of the subject, like in *the valley through which the river flows*, *the room in which the tap leaks*, or of the simple existential presupposition, as in *the yard in which a large tree stands*. Here is an example:

(9-35) Alb. (Bishqemi, p. 29)

atje ku u dëgjua krisja e shkëmbit

there where PASS heard creaking-the the of-the-stone

‘there where the creaking of the stone was heard’

‘Creaking’ exists only if heard, so the state of affairs denoted by ‘hear’ is a part of the existential presupposition of ‘creaking’. This kind of lexical relationship – *lexical solidarity between the verb and the subject* – is also not infrequent in restrictive relatives: 30 instances out of 110 in Alb (27.3%), 15 out of 153 in MG (9.8%), 32 out of 109 in SC (29.4%).

Taken together, these three categories (‘old’ verbs, lexical solidarity between the verb and the relative element or the subject) make up the majority of oblique restrictive relative clauses in my corpus – 79 out of 110 in Alb. (71.8%), 111 out of 153 in MG (72.5%), 95 out of 109 in SC (87.1%).

In the section on sources of presuppositions (4.2.4.) it has been said that certain contextual clues facilitate (but do not automatically trigger) presupposing: physical context, textual context and lexical frames have been singled out. Statistically, this facilitating relation is expressed by a relatively high coincidence of these context clues and presuppositional constructions. Of course, as I have repeated a number of times, one can presuppose even when none of the facilitating clues are there, and one does not have to presuppose even when they are all there. The coincidence between the context clues and the presuppositions is thus only statistically significant, but not absolute.

This is the explanation of the strong prevalence of textually or lexically bound elements in restrictive relative clauses: since restrictives are of necessity existentially presupposed, speakers tend to use the lexical material which is easily presupposable – ‘old’, textually present lexical items or those which are either automatically presupposed when the head of the relative is mentioned (lexical solidarity between the relative and the verb), or are easily presupposable as a whole (lexical solidarity between the verb and the subject). Since contextual clues are only statistically connected with presuppositions, restrictive relative clauses only *tend* to contain these types of lexical material, but need not do so.

9.1.7. Alternatives to relative inversion

In all kinds of focal inversion, there is at least one construction which may be used instead of inversion when the demands of discourse pragmatics are in some kind of conflict with the inverted position of the subject. Relative inversion is different, in that it has no real alternatives.

Namely, none of the three languages allows for the pre-relative position of any element of the relative clause (or, if you prefer, for the extraction out of the relative clauses), so that the mechanism of ‘pre-positions’ (frequently used as the mechanism of prefocal positions in Alb. and MG instead of wh-inversion, and in MG instead of ff-inversion) is not possible: a sentence like **ekino i Sofia Alimberti sto opio meni* (‘that S.A. in which lives’) instead of the attested form *ekino sto opio meni i S.A.* (‘that in which lives S.A.’, cp. (9-29)) is simply not possible. The other type of alternative noted in the section on focal inversion, that of using clause-final instead of clause-initial focus (frequent in quotation inversion and ff-inversion in all three languages) is impossible since the clause-initial element here is not a focus phrase, but a subordinator with a grammatically fixed position.

Thus, restrictive relative clauses must be formed in the fashion described in 9.1.5. even when one of the topical elements is of the kind which preferably stands at the absolute beginning of the clause, like NITops or NPTops.

9.1.8. Non-structural differences between Alb., MG, and SC

As noted in 9.1.5., in the commentary to Table (9-14), the number of inverted oblique restrictive relative clauses in Alb., MG, and SC is approximately the same. This, however, is not to say that there are no non-structural differences between the three languages, but merely that they are not reflected in the statistics of verb-subject clauses. Namely, since Alb. and SC frequently use the non-inversion strategy, which is practically banned from MG, the percentage of all oblique restrictive relative clauses is not the same as the percentage of the inverted ones. Thus, in order to get the real impression on the frequency relationships, one has to count subject-verb and zero-subject clauses, too:

(9-36) *Inverted and non-inverted oblique restrictive relative clauses*

	VS	SV	zero S	Σ
Alb.	110 – 29.0%	131 – 34.6%	138 – 36.4%	379 – 100%
MG	153 – 56.0%	6 – 2.2%	114 – 41.8%	273 – 100%
SC	109 – 31.1%	144 – 41.0%	98 – 27.9%	351 – 100%

Alb. and SC are statistically very similar, but MG displays three relevant divergences. First, the almost complete lack of subject-verb clauses, which is explainable in structural terms (9.1.3.). Second, the slightly higher percentage of zero-subject clauses, which is due to the general tendency for MG to use the zero strategy more often than the other two languages (cp. 5.1. and 5.5.).

What is of special interest here is the third divergent feature, that of absolute frequency: 273 restrictive relative clauses in MG correspond to 350-400 in Alb. and SC. At least partly, this is due to the fact that MG tends to use restrictive adjectives and participles instead of restrictive clauses more often than Alb. and SC. For instance, in the parallel translations of Jack London's *White Fang*, English attributive participles/adjectives are never translated by a restrictive relative clause in MG, whereas the results for Alb. and SC are 16 and 9, respectively. Here is an example:

(9-37) Eng: *In spite of the menacing hand...* (London, p. 150)

Alb: *Me gjithë rrezikun që fsheh dora në vetëvete...* (p. 147)

with all danger-the REL hid hand-the in itself

MG: *Para to apilitiko xeri...* (p. 182)

despite the threatening hand

SC: *I pored te ruke što je prijetila...* (p. 169)

and despite that hand REL AUX threatened

This tendency to use clausal instead of nominal attributes resembles the tendency observable in Alb. to use clausal complements instead of abstract nouns (cp. 5.1. and 8.2.4.). SC, in which the latter tendency is not observable, shares the former feature with Alb. Written MG is in both cases more ‘learned’, with the prevalence of abstract nouns and nominal attributes.

9.1.9. Conclusion

Relative inversion is indicative of two important issues concerning inversion in general. First, it demonstrates how important it is to determine whether a particular syntactic and/or pragmatic environment of verb-subject order is an instance of a presuppositional or of an assertional context. When the former is the case, as with restrictive relatives, we are likely to be dealing with inversion; when the latter is the case, as with the appositive ones, it is vS-construction or VsX-construction. Second, in a fashion only partly similar to ff-inversion, relative inversion shows the interplay of lexical semantics, textual parameters and presuppositional behavior. It is important to notice that at least two kinds of lexical fillings appearing in the presuppositional context in relative inversion, namely lexical solidarities between the relative and the verb and between the verb and the subject, are also characteristic of vS-construction when they are used in assertional contexts.

9.2. Embedded inversion

Verb-subject order is frequently found in adverbial and complement clauses introduced by a subordinator; when this order is found in a presuppositional context, with or without an adjacency constraint holding between the subordinator and the verb (see below for details), I shall call it **embedded inversion** (note that the term ‘embedded’ is used as a shorthand for

complement and adverbial clauses, admittedly a rather idiosyncratic feature of my terminology):

(9-38) Eng: *We'd have six dogs ..., if it wasn't for her.* (London, p.18)

Alb: *Të mos qe kjo shtazë, tani do të kishim gjashtë qen ...* (p. 18)

SUBJ:CLIT not:CLIT was that beast now FUT:CLIT had six dogs

MG: *An ðen itane tuti, tha ixame eksi skilia ...* (p. 24)

if not:CLIT was that-one FUT:CLIT had six dogs

SC: *Mi bismo sada imali šest pasa ... da nije bilo nje.* (p. 22)

we would now have six dogs if not-AUX was she

9.2.1. Balkan embedded clauses

As with relative clauses, before the analysis of embedded inversion is presented, a note on the structure of embedded clauses in the Balkan languages is in order.

Adverbial clauses (temporal, conditional, concessive, etc.) are formed in a way similar to the one known from the better-known modern European languages: a subordinator defining the semantic and syntactic relationship of the embedded clause to the matrix clause is followed by the embedded clause itself, with only very restricted possibilities of extraction, i.e. of positing the material from the embedded clause before the subordinator.

It is the complement clauses that are partly different from the AME type. There are two major points of difference. First, Alb. and MG have three, SC two elements serving to introduce complement clauses: in Alb., *se*, *që* and *që të/të* are found, in MG *oti/pos*, *pu* and *na*, in SC *da* and *što*, all roughly translatable into English as *that*. At least some of them are ambiguous as to their syntactic status.

The Alb. *se*, the MG *oti/pos* (the difference between *oti* and *pos* is merely that of register) and the SC *da* are indisputably subordinators. They are, for instance, used to introduce reported speech – *He said that he was ill* is rendered as: (Alb.) *Tha se ishte i sëmurë*, (MG) *Ipe oti/pos ine arostos* and (SC) *Reče da je bolestan* (for all three: ‘he-said that he-is ill’). The elements *që*, *pu* and *što* are formally identical to relative particles, and they are at least in some cases also semantically and syntactically very similar to them. In some cases, however, the relative interpretation is rather difficult, so that, at least for these cases, it is probably best to treat them as subordinators. The prototypical subordinating use of *që*, *pu* and *što* is that after verbs of emotion – *I am sorry that he is ill* is rendered as: (Alb.) *Më dhimbet që është i sëmurë*, (MG) *Lipame pu ine arostos*, (SC) *Žalim što je bolestan* (all three: ‘I-am-sorry that is

ill’). The subordinators *se*, *oti/pos* and *da* are generally used when the speaker does not want to take stand on the factivity of the embedded clause; *që*, *pu* and *što* mark the clause as explicitly factive¹.

The Alb. and MG forms *që të/të* and *na* are the somewhat more problematic, both syntactically and semantically. On one hand, *të* and *na* look like subordinators, since they may introduce embedded clauses, determining the syntactic and semantic relationship with the matrix clause. On the other hand, they display some characteristics of affixes, more precisely modal prefixes (or subjunctive markers, as they are sometimes called). They are practically inseparable from the verb, they may be used after question words and occur also in main and relative clauses. The following examples illustrate this ambiguity:

(9-39) (Alb.) *dua të shkruaj.* (MG) *thelo na yrafo.*

I-want *të* I-write I-want *na* I-write = ‘I want to write.’

(9-40) (Alb.) *ai trim që të njohë këtë ...*

that hero who *të* he-knows this = ‘the hero who (might) know this’

(adapted from Buchholz and Fiedler 1987: 135)

(9-41) (MG) *den kseri pjon na pari mazi tu*

not he-knows whom *na* he-takes with him = ‘he doesn’t know who to take with him’ (from Philippaki-Warburton 1992: 273)

The question of the syntactic identity of these particles (especially that of the MG *na*) has been hotly disputed for more than a decade (cp., e.g., Agouraki 1991, Philippaki-Warburton 1992, Turano 1993, Roussou 2000, and Alexiadou and Anagnostopoulou 2000 for an overview). Reference grammars – rightfully, I think – do not seem to be impressed by the subordinator/modal particle dispute, and generally describe *të* and *na* as subjunctive markers (Buchholz and Fiedler 1987: 133ff., Mackridge 1985: 247ff., Holton et al. 1997: 203ff., 450ff.). This is also the position taken in this study: *të* and *na* are treated as subjunctive markers, and the embedded clauses introduced by them (like (9-39)) as subjunctive subordinate clauses without a subordinator. Note that there is a difference between Alb. and MG in this respect: whereas in Alb. the subjunctive form with *të* may in certain cases be preceded by the subordinator *që*, this is not the case in MG, where *na*-verb forms may not (at least in complement clauses) be combined with subordinators (more on this in 9.2.4.). Note also that SC has no subjunctive, and consequently no subjunctive marker.

¹ The division of labor between the two sets of subordinators given here is simplified to the point of being partly wrong; I refer the interested reader to Svalberg (1992) for a functional description, and to Giannakidou (2000) for an overview of the relevant formal semantic literature.

The second specific feature of the complementation in the Balkans is the almost complete lack of non-finite verb forms. Namely, standard variants of Alb. and MG have no infinitive, so that all complement clauses, including those following control verbs like ‘try’, phasal verbs like ‘begin’, modal verbs like ‘can’, etc., are construed with (subjunctive) finite verb forms. Thus, the sentence *He began to cry* is expressed as *Filloi të qajë* in Alb. and as *Arxise na klei* in MG, both of them literally meaning something like ‘he-began (that) he-cries’. Standard SC is more ambiguous in this respect: it does have infinitive forms which are theoretically able to perform the functions they usually perform in AME, but these forms are freely interchangeable with a construction resembling the Alb. and MG subjunctive embedded clauses, namely *da* + finite verb. A SC translation of the sentence *He began to cry* is thus either *Počeo je plakati* (‘he-began AUX cry:INFIN’) or *Počeo je da plače* (‘he-began AUX that he-cries’). The distribution of the two alternative forms is partly dependent on the register and partly on the geography: the higher the register and the more one moves to the north-west, the greater the probability that the infinitive will be used.¹

9.2.2. Embedded clauses and presuppositions

The word *inversion* in the term *embedded inversion* presupposes that the construction in question occurs in presuppositional contexts. The present chapter is therefore dedicated to the analysis of the presuppositional behavior of embedded clauses.

From the moment presuppositions entered the logico-linguistic scene, embedded clauses have been, together with definite descriptions, considered to be the major presupposition carrier. Thus Frege (in his seminal 1892 paper *Über Sinn und Bedeutung*) considered temporal clauses to be of necessity *vorausgesetzt*, i.e. presupposed, since a sentence like ‘After Schleswig-Holstein had separated from Denmark, Prussia and Austria quarreled ...’² could not be interpreted without the speaker presupposing that the separation in question really had taken place.

In the Seventies and early Eighties, the linguistic market was virtually flooded with different types of presupposed embedded clauses – factive, implicative, phasal, etc. verb complements, counterfactual conditionals, concessives, etc. (cf. Karttunen 1973, Gazdar

¹ The spectacular lack of nonfinite forms in the Balkan languages has spawned a rich literature on the subject – apart from the standard works on Balkan linguistics, see Joseph (1983) for a diachronic overview, and the papers in Rivero and Ralli (2001) for an attempt to deal with the topic within the generative framework.

² The slightly modified English translation is taken from Levinson (1983: 169).

1979, Soames 1982) – the main criterion for the presupposedness being the preservation of the truth value under the negation of the matrix clause. On this criterion, some embedded clauses, like complements of verbs of saying, are not presupposed. Unfortunately, it soon turned out that the neat presupposed vs. non-presupposed distinction created along these lines is not tenable: for instance, complements of factive verbs may, under certain conditions, lose their presupposed status, and the other way around, complements of verbs of saying may sometimes acquire it (see Levinson 1983: 186ff. for details and further references). Since it was obviously not possible to capture the presuppositional behavior of embedded clauses with a simple formal/semantic classification of clause types, they gradually lost their prominent status in the debate, so that they, for instance, do not figure at all in some more recent accounts of presuppositions (e.g. van der Sandt 1992, Horn 1996, Mohanan and Mohanan 1999).

As has been abundantly expounded in Section 4.2. (see especially 4.2.2.), I do not think that the question of truth value is relevant for defining presuppositions: for a proposition to be (existentially) presupposed, it suffices that the speaker assumes that the hearer entertains a propositional thought, even without commitment to its being true. Hence I contend that every term of a proposition carries at least one, namely existential presupposition (and may, but need not, be relationally presupposed).

The logical consequence of this attitude would be to claim that *all* embedded clauses are necessarily existentially presupposed, regardless of their resistance to negation (or of their behavior with regard to other standard acid tests of presuppositionality), since they regularly function as terms of their matrix clauses. However, although I do subscribe to the second part of this claim (the negation test and other similar devices are an insufficient diagnostics for presuppositions), I do not think that every embedded clause has to be treated as existentially presupposed, even though my reasons are quite different from those based on the truth value.

Let me first clarify the difference between my pragmatic account of the existential presupposedness of some embedded clauses and the truth-conditional account endorsed by semantic presuppositionalists (the term is borrowed from Carston 1998). One of the classical examples for the cancellability of presuppositions is Levinson's sentence *Sue died before she finished her thesis* (1983: 187). According to the truth-conditional account of presuppositions, even though *before*-clauses are usually presupposed, this one is not, because the matrix clause shows that it is not true that the state of affairs 'she finished her thesis' obtains. In my account, according to which presuppositions are simply non-asserted propositions present or evoked in the mind of the interlocutors, *before she finished her thesis* is a perfect example of

an existentially presupposed proposition: the speaker invites the hearer to form the mental representation of the state of affairs ‘the finishing of Sue’s thesis’ and uses this mental representation as a temporal modification of the predicate ‘die’. That ‘she finished her thesis’ is obviously false is of as little relevance for the presupposed status of this clause as the non-existence of the present king of France is for the existential (and relational) presupposedness of the phrase *the present king of France* in the sentence *The present king of France is bold* (see Horn 1996): in both cases, the speaker posits a discourse referent (‘Sue’s finishing of her thesis’, ‘the present king of France’), expecting the hearer to create/recall the mental representation of this discourse referent and use it as a context in the creation of discourse.

If this is so, what embedded clauses are not existentially presupposed? There are three classes I am able to think of. First, those which are only formally embedded, but represent semantically independent propositions (as in the case of appositive relative clauses, cf. 9.1.4). A nice case in point are the temporal clauses called *cum-inversum clauses* in traditional Latin grammars: the main clause describes the circumstances holding at the time of the event, the temporal clause the event itself, as in the following example:

(9-42) Eng: *Matt was pegging up, when there was an outcry ... without.* (London, p. 158)

Alb: *Metti po numëronte, kur befas jashtë u dëgjuan ca britma ...* (p. 156)

M. PROGR counted when suddenly outside PASS heard some cries

MG: *O Mat metruse, otan akustike kravji...* (p. 191)

the M. counted when was-heard cry

SC: *Mat je brojao, kad se spolja začu nečiji krik...* (p. 178)

M. AUX counted when REFL outside was-heard someone’s cry

The clause *when there was an outcry* (as well as its translational equivalents) does not function as a temporal modification to the predicate *peg up*, and is consequently not its term, but encodes instead an independent proposition which temporally partly overlaps with the proposition conveyed by what is formally the main clause, despite the fact that *when there was ...* is formally identical to the ‘normal’ temporal clauses which are terms *qua* presupposed (as in the case of *before she finished her thesis* mentioned above).

This mismatch between form and function (subordinate clause, independent proposition) is not found only with *cum-inversum* clauses. Subordinators corresponding to English *while* are frequently used as adversative coordinators; causal clauses are often used as independent

explanations of the speech act itself, partly parallel to independent main clauses introduced with particles corresponding to *namely*¹, etc.

The second group of existentially non-presupposed embedded clauses are the complements of modal and similar verbs, like in *I can say, I should say*, etc. (recall that in the Balkan languages these complement clauses are expressed with finite verbs: ‘I can [that] I say’, ‘I should [that] I say’). The reason for their non-presupposedness is again semantic: modal and similar verbs are rather instances of predicate-modifying operators than of (semantic) predicates, and are only accidentally encoded as verbs in most European languages. The complement clauses of these verbs represent thus semantically full-fledged independent propositions, not terms of a predicate, and are accordingly not to be treated as automatically existentially presupposed.

The third group are the complements of verbs of propositional attitude (*think, believe*, etc.) and of saying, like in *She said that she was ill, I think that you are wrong*. As is patent in the case of propositional-attitude verbs, the reasons for my assumption that these embedded clauses are not necessarily existentially presupposed are the same as in the case of modal verbs: at least in some cases, these verbs are semantically equivalent to predicate-modifying operators, not to autonomous predicates, so that their complements are not predicate terms, but rather independent propositions. It is perhaps less self-evident that this sometimes also holds in the case of the verbs of saying. I should like to argue that, at least in some cases, these verbs are used as predicate modifiers referring to the attitude of the speaker towards the proposition expressed in the embedded clause: in saying *She said she was ill*, I explicitly add the modal property ‘claimed, but not necessarily factual’ to the proposition ‘she was ill’.

To sum up: under the definition of presupposition embraced in the present study, embedded clauses are existentially presupposed in all cases in which the propositions they encode semantically figure as terms of the predicate of the matrix clause, irrespective of their truth value. When the embedded status of a clause is not reflected in the semantics, as in the cases enumerated above, it is not automatically existentially presupposed.

¹ In some languages, the causal clauses which are truly subordinated, i.e. which are terms on the semantic level, are formally differentiated from those which are only formally subordinated. In colloquial German, for instance, only the former have the verb-final word order characteristic of embedded clauses. Even more interesting in the present context is the system existing in spoken SC: in the truly subordinated clauses, the clitics obligatorily follow the subordinator *jer* ‘because’ (*jer je video duha* [because AUX:CLIT seen ghost] ‘because he saw a ghost’); in pseudo-subordinate clauses, it is the first word/phrase after *jer* that is followed by the clitics (*jer video je duha* [because seen AUX:CLIT ghost] ‘namely, he saw a ghost’).

9.2.3. Statistics

As the data furnished by statistics show (5.5.3.), embedded clauses are one of the syntactic contexts in which VS order occurs more frequently than average. However, this preferential status is characteristic only of Alb. and MG, with SC showing no clear preferences with respect to embedded clauses. To repeat some of the figures: In Alb., the average value for VS clauses within the group of clauses with overt subjects is 27.9%; among the embedded clauses with overt subjects, the percentage is 39.1%; in MG, the average is 40.2%; among embedded clauses, those with VS order make up 54.8%; in SC, unlike the previous two languages, the ratio is 24.4% vs. 25.0% (cp. 5.5.3.).

The first step in defining embedded inversion is to see how often existentially presupposed embedded clauses occur in the corpus. When all embedded clauses are taken together, the percentages look as follows: out of 4437 embedded clauses in the Alb. corpus, 2604, i.e. 58.7% may with some plausibility be considered existentially presupposed; in MG, the ratio is similar: 4425 vs. 2624, i.e. 59.3%; in SC, 3531 vs. 1935, i.e. 54.8%.

In order to see how often the inversion occurs, as opposed to its non-occurrence, it is best to exclude the clauses without overt subject, i.e. zeroS clauses. The percentages sink slightly: in Alb., within the group of clauses with overt subjects (1340), 553, i.e. 41.3% are existentially presupposed; in MG, the ratio is 1260 vs. 543, i.e. 43.1%; in SC, 1491 vs. 689, i.e. 46.2%.

Now we have the basis for the investigation of embedded inversion: 553 existentially presupposed clauses with an overt subject in Alb., 543 in MG, and 689 in SC. Table (9-43) shows the frequencies of SV and VS order within this class:

(9-43) *VS and SV orders in existentially presupposed embedded clauses*

	SV	VS	Σ
Alb.	317 – 57.3%	236 – 42.7%	553 – 100%
MG	230 – 42.4%	313 – 57.6%	543 – 100%
SC	540 – 78.4%	149 – 21.6%	689 – 100%

As could be expected on the basis of the general distribution of SV and VS in embedded clauses in the three languages (see above), embedded inversion occurs most frequently in MG, closely followed by Alb., while SC displays only the average presence of this construction, meaning that the frequency of embedded inversion (= VS order in presupposed embedded clauses) is roughly equal to the overall frequency of VS order both in embedded

clauses and in the corpus as a whole. Similar ratios also occur when the clauses from the parallel translations of London's *White Fang* are counted: out of 92 presupposed embedded clauses in Alb., 52 (56.5%) are SV, 40 (43.5%) VS; in MG, the ratio is: SV = 51 (41.1%) vs. VS = 73 (58.9%); in SC, SV = 29 (67.5%) vs. VS = 14 (32.5%).

Embedded inversion thus occurs with the greatest regularity in MG, and, somewhat less often, in Alb., while it is relatively weakly represented in SC. In most types of inversion presented so far, this numerical relationship (MG, Alb. > SC) is triggered by the existence or non-existence of the adjacency constraint holding between the fronted element and the verb, SC usually being the language without such a constraint. Though the picture is essentially correct in the case of embedded inversion, too, the situation is much more complicated, since we are not dealing with one kind of fronted element, but with a host of different subordinators and related words, each of which has a potential to display a specific syntactic behavior. For this reason, the following chapter will be dedicated to the question of adjacency¹.

9.2.4. Adjacency constraint

SC is, as usual, easiest to describe: none of the subordinators in this language has to be adjacent to the verb in the embedded clause.

In Alb. and MG, the situation is less unequivocal. When a subordinator takes the indicative, no obligatory adjacency to the verb is required, in both languages. This practically means that adverbial subordinators, like Alb. *pasi* and MG *afu* ('after'), and complementizers, like Alb. *se* and MG *oti/pos* ('that'), which always or almost always take the indicative, do not display the adjacency constraint. Here are some examples of non-adjacency, with the subject intervening between the subordinator and the verb:

¹ A note on distribution of different VS constructions in embedded clauses: embedded inversion, which occurs in presuppositional contexts, and VsX and vS constructions, which are assertional, normally make up between one third and one half of all VS embedded clauses; other inversion types (notably wh-inversion) amount to some 10-20%. This is illustrated by the following table:

	embedded inversion	other inversion types	vS/VsX	Σ
Alb.	236 – 45.6%	97 – 18.8%	184 – 35.6%	517 – 100%
MG	313 – 45.9%	53 – 7.8%	316 – 46.3%	682 – 100%
SC	149 – 40.6%	40 – 10.9%	178 – 48.5%	367 – 100%

(9-44) MG (Fakinou, p. 108)

...otan to kefali tis eftase sto ipsos tu patomatos...

when the head her reached at-the height of-the floor

‘...when her head reached the level of the (attic) floor...’

(9-45) Alb. (Kadare, *Kronikë*, p. 6)

... përderisa ky ishte një qytet prej guri...

because this was a town from stone

‘... because this was a town of stone...’

(9-46) Alb. (Camaj, p. 199)

... se ajo zbutej kur ky i përulej asaj ...

that she calmed-down when he to-her:CLIT bent-himself to-her

‘[He knew] that she always calmed down when he submitted to her...’

Indicative embedded clauses of this type are responsible for the majority of SV embedded clauses in Alb. and MG¹.

The real problem are the subordinators taking the subjunctive². As indicated above, in Alb. and MG, complement clauses may be introduced by the subjunctive markers *të* and *na* without subordinator, so that these particles serve also as introductory markers of the clause, as illustrated by (9-39). Now, the relevant fact here is that *të* and *na* have almost reached the affixal status, since they may be separated from the verb of the embedded clause only by clitics and by the negative particle.

In other words, *të* and *na* are instances of clause-initial elements with an adjacency constraint surpassing in its rigidity all the instances of this constraint we have encountered by now: the variants of (9-39) with the subject (or any other element) intervening between *të/na* and the verb are impossible: **dua të ti shkruash*, **thelo na esi yrafis* (both: ‘I-want *të/na* you write’). In sum: *të* and *na* introducing complement clauses have to be directly adjacent to the verb of the embedded clause; the pattern is [**Ø-subordinator**] [**të/na~subjunctive verb form**] (see 9.2.6. for details).

¹ The claim that there is no adjacency constraint with indicative-taking subordinators is intended to depict the present state of affairs in Alb. and MG. When the diachronic dimension is taken into account, at least for MG, there seems to have existed once a difference between the newly developed subordinators, like *pos*, and the old ones, like *oti* (both meaning ‘that’), the former being (almost) obligatorily adjacent to the verb, the latter not (see Mackridge 1993 and Horrocks 1997: 59ff., 208ff.; more on this in Section 10).

² There is one excellent description of subjunctive clauses in MG, Mackridge 1985: 276ff. (see also the relevant sections in Holton et al. 1997). I am not aware of a comparable survey of the Alb. usage (see, however, Buchholz and Fiedler 1987 and, from the generative viewpoint, Turano 1993).

A number of adverbial subordinators may or must take the subjunctive, formed with *të/na* and an adjacent finite verb with the subjunctive morphology, the basic structure for both languages thus being **[subordinator] [të/na~subjunctive verb form]**. The slot between the subordinator and the *të/na*-verb complex is free for all kinds of constituents, including subjects, i.e. there is no obligatory adjacency between the subordinator and the verb which would force the subject to appear after the verb:

(9-47) Eng: ... *so that...they sent surges of fear through the toiling dogs...* (London, p. 10)

MG: ... *toso ... oste ta zemena skilja na rijisun apo fovo...* (p. 15)

so that the harnessed dogs *na* shiver from fear

(9-48) Eng: ...*by the time Henry had helped him ...* (London, p. 21)

Alb: ...*ndërsa Henriku të ndihmojë ...* (p. 20)

until Henry-the *të* help:SUBJUNCTIVE

The full pattern is thus **[subordinator] [X] [të/na~subjunctive verb form]**. There are, however, two structures diverging from this scheme which are relevant for the question of adjacency.

Temporal and conditional clauses referring to the future or to the events that have not yet taken place at the time of the main clause event are regularly formed with the subjunctive in both languages. Alb. has only the structures corresponding to the scheme given above, with a free slot for subjects and other constituents between the subordinator and the *të*-verb complex, as in (9-48).

Unlike Alb., MG has yet another possibility, using the subjunctive form of the finite verb *without* the particle *na* (bare subjunctive); in this case, the verb has to follow the temporal subordinator immediately:

(9-49) MG (Kapandai, p. 29)

...*prin klisi o xronos...*

before closes:SUBJUNCTIVE the year

‘... before the year was over...’

The sentence with the subject between *prin* and the subjunctive form *klisi* would not be perfectly grammatical: *?prin o xronos klisi*. Thus, with temporal subordinators taking the bare subjunctive, the adjacency constraint holding between the subordinator and the verb is fairly strong: the scheme is **[subordinator]~[bare subjunctive verb]**. Apart from *prin*, this

syntactic behavior occurs with *protu* ('before'), *mexri*, *ospu*, *oso* ('until'), *molis* ('as soon as'), *otan* ('when'), *an* ('if'), etc.¹

The second exception to the scheme [subordinator] [X] [*të/na*-subjunctive verb form] are the clauses in which the subordinator and the particles *të/na* are merged into a word-like entity. The clearest case in point are purpose clauses. In Alb., they are formed with *që+të*-subjunctive; in MG, with *ja+na*-subjunctive: 'in order (for him) to come here' is expressed with *që të vijë këtu, ja na 'rthi eđo* (for both: *që/ja të/na*-come:SUBJUNCTIVE here). In MG, the subordinator (diachronically a preposition) *ja* and *na* are not separable at all. Consequently, the subjects are always inverted. In Alb., at least the negative particle *mos* is allowed to sometimes intervene between *që* and *të*. As for nominal constituents, I was not able to find a single example of an element between *që* and *të*-verb in my corpus, and my two Alb. informants, although none of them discarded the variants with an intervening subject as ungrammatical, unaminously proclaimed the adjacent variants (*që~të*-verb-subject) to be more natural and better. Here are two typical examples:

(9-50) Alb. (Bishqemi, p. 125)

... *që të mos përsëritej më historia e hidhur* ...

që të not repeat-itself more history-the the bitter

'... in order for that terrible thing not to happen again...'

(9-51) MG (Elefterotipia, p. 4)

... *ja na epanaliturjisi i Sxoli*...

ja na again-functions:SUBJUNCTIVE the school

'... in order for the School to reopen...'

The variant **ja i Sxoli na epanaliturjisi* is, as indicated, completely ungrammatical, whereas *që historia e hidhur të mos përsëritej*, though not absolutely excluded, sounds 'weird' (as one of my informants said). The pattern for the purpose clauses can thus be loosely represented as **[subordinator~*të/na*~subjunctive verb form]**.

In Alb. this pattern is, along with purpose clauses, present in the type of conditional clauses formed with *po* (*të*); in MG, some further combinations of the subordinator and *na* seem to be well advanced on their way to wordiness, though they have not yet achieved the level of the *ja na* combination. Thus, the manner clause subordinator *san* is practically never separated from

¹ Interestingly enough, bare subjunctive does not cause adjacency effects in another class of clauses in which it occurs with some regularity, in the complements of verbs of fearing introduced by *min/mipos*, as witnessed by the following example (from Mackridge 1985: 300): *fovate mipos o Kostis đen pai* [he-fears that the K. not goes] 'he is afraid Kostis might not go'.

na and from the subsequent verb (the complex *san na* meaning ‘as if’), although the subjects (and other elements) placed between *san* and the *na*-verb complex are not so readily characterized as ungrammatical as in the case of *ja na*. A clause like *san o Petros na ine enoxos* (as-if the Peter *na* is guilty) is virtually impossible to find in natural texts and sounds a bit strange, but it is not ungrammatical (see Holton et al. 1997: 464, Lascaratou 1998: 161). Exactly the same holds for the subjunctive clauses nominalized with the article *to*: *to* and the complex *na*-verb are virtually inseparable. The pattern is [**subordinator~të/na~subjunctive verb form**], although, as indicated, not so strictly as with *ja na*.

To sum up: SC displays no adjacency constraint. As far as Alb. and MG are concerned, in all indicative and most adverbial subjunctive embedded clauses, the subordinator and the verb do not have to be immediately adjacent. Adjacency constraint occurs only in MG temporal clauses with bare subjunctive, in MG manner clauses with *san na*, in MG nominalized subjunctive clauses, in Alb. and MG purpose clauses, and in Alb. and MG complement clauses introduced by the subjunctive markers *të* and *na* without subordinator. A summary is given in Table (9-52):

(9-52) *Adjacency constraint in Alb. and MG embedded clauses:*

	SC	Alb.	MG
indicative embedded clauses	–	–	–
subjunctive adverbial clauses	/	–	–
manner clauses with <i>san na</i>	/	/	–/+
conditional clauses with <i>po të</i>	/	+	/
nominalized subjunctive clauses	/	/	+
clauses with bare subjunctive	/	/	+
purpose clauses	–	–/+	+
subjunctive temporal clauses (<i>të/na</i>)	/	+	+

9.2.5. Syntax, prosody, and information structure of embedded inversion

What we have at hand is a presuppositional context (some embedded clauses) and a sporadic adjacency constraint: an ideal surroundings for an inversion. Embedded inversion is thus to be defined as the verb-subject order appearing in existentially presupposed embedded clauses, with or without obligatory adjacency between the introductory element (subordinator or subjunctive marker) and the verb. In this section, the syntactic and intonational properties of the construction and its informational value will be explored. The material used are the existentially presupposed VS clauses (as opposed to both non-presupposed embedded clauses and to presupposed SV clauses): 236 in Alb., 313 in MG, and 149 in SC.

Like other inversion types, embedded inversion complies to rules (8-104) through (8-108), repeated in Section 9 as (8-104) – (8-108). The rules defining the interplay between syntax/intonation and information structure ((8-104), (8-105) and (8-106)) are operative in all inversion types and in all three languages: in all-ratified contexts, the nuclear sentence stress is placed on the non-final verb; when an element needs to be ratified, it is placed in the final position and carries the sentence stress. The rule defining the behavior of pronominal subjects in SC (8-108), which avoid inversion, is also a constant. What varies from one to another inversion type and across languages is rule (8-107), describing adjacency phenomena. With relative inversion, there is a clear-cut distinction between Alb. and SC on one hand, and MG on the other, with the former having no adjacency constraint, the latter employing it almost with no exceptions. In embedded inversion, as shown in 9.2.4., the state of affairs is more complex: SC, Alb. and MG indicative clauses, most Alb. and some MG subjunctive clauses, do not display the adjacency constraint; Alb. and MG purpose and complement subjunctive clauses, as well as some other clause types in MG, strictly demand that the introductory element and the verb be adjacent. Since this situation does not allow for dividing *languages* along the adjacent/non-adjacent line, as in other inversion types, the description that follows will be organized not according to languages, but according to the presence of adjacency.

9.2.5.1. Clauses with the feature [–adjacent]

This group comprises all SC embedded clauses and many clause types in Alb. and MG. Since subordinator need not be directly adjacent to the verb, the relevant rules governing the word order and the intonation are those with discourse-pragmatic motivation ((8-104), (8-105) and (8-106)), as well as the rule disallowing inversion for pronominal subjects in SC (8-108).

In all-ratified clauses in all three languages, the inversion occurs with some regularity only in short clauses, i.e. those consisting only of subordinator, verb, and subject; the sentence stress is on the verb:

(9-53) Alb. (Bishqemi, p. 75)

Pasi u LARGUAN ata të dy, komandanti u dha edhe

after PASS went-away these the two, commander-the to-them gave also

të tjerëve detyra.

to-the to-others duties

‘[The commander ordered Retun and Vetiolë to go to check the radio-transmitter.]

When the two were GONE, the commander also gave orders to the others.’

(9-54) SC (Vesti, p. 32)

Ukoliko se PRIHVATI predlog Komisije, sledeće godine bi odmah posle

if REFL accepts suggestion of-Committee next year would immediately after
prve faze sledile eliminacione borbe.

first phase follow elimination matches

‘[The European Football Committee suggested a radical change in the organization of the competitions.] If the Committee's suggestion is ACCEPTED, the elimination matches would immediately follow the first competition phase next year.’

In both cases, the state of affairs invoked by the embedded proposition (‘their leaving’, ‘the acceptance of the suggestion of the Commission’) is existentially presupposed by the speaker: the hearer is expected to create the mental representations of these situations without their being asserted, without a commitment to their truth value (as is especially patent in (9-54), where the acceptance of the suggestion is presented as a mere possibility). They are also all-ratified: the fact that the commander ordered the two members of the crew to go makes their leaving immediately present in the interlocutors’ minds, as well as the existence of a suggestion makes it plausible to assume that its acceptance is envisaged. The formal reflex of this informational configuration in short embedded clauses is the same as in all other types of inversion: the accented verb and the inverted subject, so that the verb may carry the nuclear stress in a non-final position.

In long embedded all-ratified clauses, i.e. those consisting of more than verb and subject, the inversion need not occur: the additional element(s) cover the last position and allow the verb to carry the non-final sentence stress:

(9-55) MG (Elefterotipia, p. 4)

Apo to apotelesma ton ekloyon sto Veliyraði tha eksartiði to an i Evropei tha

from the result of-the elections in-the B. FUT depend the if the Europeans FUT

EPANAÐIATIPOSUN tin politiki tus ja tin arsi ton kiroseon se varos tis Jugoslavias.

reformulate the policy their for the raising the sanctions in detriment of-the Y.

‘[In view of the elections to be held next week, the EU wishes to help the people of Yugoslavia get out of the permanent crisis.] Whether the Europeans will reformulate their policy with respect to the raising of the sanctions against YUGOSLAVIA depends on the results of the elections.’

The element X (the direct object *tin politiki tus...*) covers the final position, so that the subject need not be inverted in order to enable the verb to carry the non-final sentence stress.

When a presupposed embedded clause contains a non-ratified element, this element tends to be placed last, and always carries the nuclear stress. When this element is the subject, inversion occurs:

(9-56) SC (Andrić, p. 21)

Čim se razidje JEDAN krug, on ... prilazi drugom krugu.

as-soon-as REFL disperses one circle he approaches other circle

‘[Every morning, in front of his room a small circle around a certain Zaim gathers, and he speaks of his life.] When ONE circle disperses, he ... approaches the other.’

(9-57) Alb. (Koha Ditore, p. 6)

Djemtë e Gjakovës nuk shkuan në luftë pse i thirri RAMUSHI,

sons-the of-the Gj. not went in war because them called R.

por se i thirri KOSOVA, atdheu.

but because them called K., fatherland

‘Sons of Gjakovë did not go to war because RAMUSH called them, but because KOSOVO, the fatherland, called them.’

In (9-56), the fact that a circle disperses is existentially presupposed, but the term ‘one circle’ is in cataphoric contrast to ‘another circle’, and therefore non-ratified. Consequently, it is placed clause-finally and carries the nuclear stress. The existential presupposition and the contrast at work in (9-57) are self-evident: the proposition ‘Ramush called them’ is presupposed (the speaker acts as if the hearers might have expected that this was the reason the people went to war), with one of its elements, ‘Ramush’, treated as non-ratified, so that it can be contrasted with the correct assumption, namely that it was the fatherland that called them.

Inversion normally does not occur when an element other than subject is non-ratified:

(9-58) MG (Fakinou, p. 62)

... an o Andonis epestrefe NORITERA, tha paraksenevotan siçura pu ðen tha tin

if the A. came-back earlier FUT was-surprised surely that not FUT her

evriske eki.

found there

‘[If Andonis does not return home in two days, the groceries won’t be fresh anymore]; only, if Andonis returned home EARLIER, he would surely be surprised not to find her there.’

That Andonis comes back home at some time is presupposed; his coming home *earlier* is in contrast to the expected time of his arrival, and therefore non-ratified. No inversion occurs, the non-ratified element is accented and postposed.

In SC, pronominal subjects practically never invert, rule (8-108) being stronger than all other rules (see 9.1.5.2).

The syntax and intonation of presupposed embedded clauses whose subordinators have the feature [-adjacent] is essentially the same in all three languages (with the exception of pronominal subjects in SC, which do not invert at all), and essentially identical to all other inversion types in which the adjacency constraint does not exist.

9.2.5.2. Clauses with the feature [+adjacent]

Adjacency constraint is a feature of a number of subordinators/subordinator-like words in MG, and of a couple of such words in Alb. The consequences of its existence for the inversion are predictable: inversion is obligatory, independently both of the ratification status of particular elements and of the length of the embedded clause. Only the intonation varies along the parameters of information structure: in all-ratified clauses, the sentence accent is on the verb; in the clauses with non-ratified elements, the accent is on the non-ratified element.

First two examples of all-ratified embedded clauses with adjacency constraint:

(9-59) MG (Matesis, p. 43)

... *molis KIMIΘI to peði se θelo.*

as-soon-as falls-asleep the child you I-want

‘[My mother put my little brother to bed and told me] ‘As soon as the child falls asleep, I shall need you.’

(9-60) MG (Elefterotipia, p. 15)

J’ afto ke θa apoðexti na TEΘI sto Sindayma i sxetiki ðjataksi.

for that also FUT accept *na* is-put in-the Constitution the relevant decree

‘[The opposition has prepared a decree against corruption and wants to see this decree as a part of the Constitution. The prime minister Simitis wants to neutralize the impression that he is doing nothing against corruption.] This is why he is going to agree that the relevant decree be incorporated into the Constitution.’

The state of affairs ‘the child falls asleep’ is to be expected from the fact that the child has just been put to bed; ‘incorporation of the decree in the Constitution’ is directly activated in the previous text; both propositions are all-ratified, so that their verbs carry the sentence accent. As far as the relative position of the postverbal elements with the same, ratified, information

value is concerned (typically the subject and some other element), it is governed by the principles described for quotation inversion (8.1.1.2, 8.1.1.4); thus, in (9-59), the subject is placed after the PP rather than in front of it either because of the heaviness principle or because of the tendency observable in MG (and Alb.) to place the non-subjectal elements with a close semantic tie to the verb first.

Because of the adjacency constraint, the word order pattern with inverted subject cannot change when the clause contains non-ratified elements, the only ratifying mechanism being the position of the nuclear stress. In (9-61), it is the subject that is non-ratified, in (9-62), the temporal adverbial:

(9-61) MG (Matesis, p. 80)

...pos epitrepis na se filane anθropi pu trone VATRAXIA!

how you-allow *na* you:CLIT kiss people REL eat frogs

‘[There was a rumor in the village that the Italians eat frogs; the priest asks a woman having an affair with an Italian:] ... how can you bear to be kissed by people who eat FROGS!’

(9-62) Alb. (Kosovarja, p. 2/44)

Është e para herë në gjithë jetën time të më ketë bërë dikush për vete

is the first time in all life-the my *të* me:CLIT has done someone for himself

brenda një çasti të VETËM.

within one moment the single

‘It was the first time in all my life that someone had enchanted me within one single moment.’

The consequences of the existence of the adjacency constraint are thus no different from those observed in other inversion types in which this constraint is present, the most notable being the obligatoriness of inversion in all contexts.

9.2.5.3. Summary

The sentence patterns for embedded inversion are essentially the same as those for relative inversion and *wh*-inversion, only that the [+/-adjacent] line does not run among the languages, but rather within them, with the exception of SC, which has only non-adjacent subordinators.

(9-63) *Sentence templates for embedded inversion:*

[+adjacent]:

some subjunctive clauses in Alb. & MG

all-ratified:

[sub]~[VERB] ([X]) [subject]

non-ratified subject:

[sub]~[verb] ([X]) [SUBJECT]

non-ratified verb:

[sub]~[VERB] [subject] ([X])

non-ratified X:

[sub]~[verb] [subject] [X]

[–adjacent]: SC, indicative and

some subjunctive clauses in Alb. & MG

all-ratified:

[sub] [VERB] [subject]

(SC: **[sub] [pronominal subject] [VERB]**)

[sub] [pronominal subject] [VERB] [X]

non-ratified subject:

[sub] [verb] ([X]) [SUBJECT]

non-ratified verb:

[sub] [subject] ([X]) [VERB]

non-ratified X:

[sub] [subject] [verb] [X]

9.2.6. Semantic and informational properties of subjects and verbs

Textual and/or lexical boundness of the verb and the subject observable in relative inversion is also characteristic of embedded inversion (see 9.1.6. for the definition of the terms).

First the textual boundness. In Alb., the percentage of ‘old’ verbs is 13.6% (32 out of 236 presupposed embedded clauses with inversion); in MG, it is 10.5% (33 out of 313), in SC, 6.1% (9 out of 149) – in all three cases, much more than average. A beautiful example of an old verb is given in (9-57) above.

Yet another category of bound verbs appears in some of the presupposed embedded clauses: predicates denoting states of affairs inherently retrievable from the stock of encyclopedic knowledge, like ‘be needed’, ‘be present’, various predicate denoting order (‘be first’, ‘be best’), etc. They, however, do not seem to be particularly frequent: There are two such predicates in Alb, five in MG, two in SC. Here is an example:

(9-64) MG (Elefterotipa, p. 10)

*Aftonoito ine **oti xriazonde taxitita ke apofasistikotita.***

self-evident is that are-needed speed and determination

‘It is self-evident that speed and determination are needed.’

The instances of lexical boundness are much more numerous.

(1) Lexical solidarity between the verb and the primary topic (the verb belongs to the existential presupposition of the primary topic) is, not surprisingly, less frequent here than in relative inversion: seven instances in Alb., ten in MG, eight in SC. Here is an example, where the verb *konstrukton* ‘make, construct’ is a part of the existential presupposition of the primary topic ‘caves’:

(9-65) Alb. (Bishqemi, p. 12)

Di se të gjithë këto shpella ...i ka konstruktuar një mjeshtër magjik

I-know that the all these caves them has constructed a master magical

‘I know that all these caves were constructed by a great magician.’

(2) Much more frequent than in relative inversion are the cases in which the verb and the subject are in the relationship of lexical solidarity, the verb being a part of the rich or of the simple existential presupposition of the subject referent: 64.4% in Alb. (152 out of 236 clauses), 54.6% in MG (171 out of 313), 65.7% in SC (98 out of 149). In the following example, ‘being carried on’ (*voditi se*) is the mode of existence of the fight (*borba*):

(9-66) SC (Andrić, p. 103)

... dok se u ćeliji vodila borba u mraku ...

while REFL in the cell was-led fight in darkness

‘... while in the cell the fight was carried on in the darkness...’

These lexical types cover most of the clauses with embedded inversion. The reason for this prevalence is the same as in the case of relative inversion (see 9.1.6., where the explanation is more elaborated): it is easier to presuppose propositions that have either already been mentioned, or are in a way derivable from the encyclopedic knowledge, or from the lexical frames evoked by one element of the proposition.

9.2.7. Alternatives to embedded inversion

For many types of embedded clauses – all embedded clauses in SC, indicative and most subjunctive clauses in Alb. and MG – the simple alternative to inversion is non-inversion, which is possible because of the lack of adjacency constraint holding between the subordinator and the verb.

However, there is yet another possibility. In contrast to relative clauses, which in the Balkan languages do not license extraposition, embedded clauses do allow for placing subjects and other elements in front of the subordinators under certain well-defined

conditions. The principle at work is similar to that observed in connection with wh-inversion (8.2.3.): the subject (or some other element) is placed before the fronted element. Extraposition may occur both in clauses with an adjacency constraint and in those without it.

The naturalness and frequency of extraction vary both across languages and across clause types. Modern SC practically does not use this syntactic device, or only quite exceptionally.¹ In Alb., it is possible with both indicative and subjunctive clauses with the feature [–adjacent] on one hand, and with those subjunctive clauses in which the verb has to be adjacent to the subordinator, but it is significantly more frequent with the former (in my corpus, the relationship is seven to one), a fact I shall try to explain below; (9-67) illustrates the extraction from an indicative clause:

(9-67) Alb. (Kosovarja, p. 10)

E unë kur ia qëlloja përgjigjen, ai heshte.

and I when him-it guessed answer-the he was-silent

‘[He used to ask me questions.] When I had the answer right, he was silent.’

In MG, extraction is significantly more frequent than in Alb.: I have some 23 examples of [–adjacent] clauses and 35 [+adjacent] clauses with subject placed in front of the subordinator (note that the frequency of subject extraposition with the two clause types stands in inverse proportion to that observed in Alb.).

(9-68) MG (Elefterotipia, p. 22)

Ine ađianoito i Xios, pu ine jemati apo petres, na isaji petra apo tin Turkia.

is unthinkable the Ch. REL is full of stones, na import stone from the T.

‘It is inconceivable that Chios, which is full of stone, imports stone from Turkey.’

The conditions under which the subject extraposition is used are similar to those described for wh-inversion, i.e. mostly then when the subject is a non-ratified topic belonging to one of the two topic types defined in 8.2.3., NITops (newly introduced topics) and NPTops (new paragraph topics). As in wh-inversion, extraposition is not obligatory even in those cases, the canonical ratification mechanism (placing the accented subject in the clause-final position) being at least as common. Thus, the NITop *i Xios* (‘Chios’) in (9-68) could as well be at the end of the clause, carrying the nuclear stress: *ine ađianoito na isaji petra apo Turkia i XIOS, pu ine JEMATI apo petres* (for a similar analysis, see Mackridge 1985: 238).

¹ Extraction from the complements of modal and similar verbs (which, being non-presupposed, do not belong in this chapter) was still in use in the 19th century literary language: *Ovaj sud mislim da je pravedan* (this court I think that is just) ‘I think that this court is just’ (from Popović 1997: 150, where the phenomenon is described in some detail).

Alb. and MG display an interesting typological difference concerning extraposition (both of subjects and of other elements) from subjunctive complement clauses, which is responsible for diverging frequencies of extraposition in this clause type.

In 9.2.1., I have claimed that bare subjunctive markers, *të* and *na*, introduce subjunctive complements without a subordinator. This is true for MG in all contexts, but not so for Alb. Namely, in Alb., if an element need be placed in front of the verb for pragmatic reasons, the subjunctive marker *të* may not introduce the clause alone. In this case, the subordinator *që* appears. The contrast between clauses with and without fronted elements is patent in the following two examples¹:

(9-69) Alb. (Koha Ditore, p. 14)

... *nuk ka kuptim të krijohet diçka që është shkatërruar...*

not has sense *të* it-is-created something REL is destroyed

‘... it is pointless to re-create something that is destroyed...’

(9-70) Alb. (Koha Ditore, p. 14)

... *nuk ka arsye që Mali i Zi të shkëputet nga federata...*

not has reason *që* Mountain the Black *të* it-separates from federation-the

‘... there is no reason for Montenegro to secede from the federation...’

Sentences (9-71) and (9-72) (adapted from Dobrovie-Sorin 2001: 48) show the difference between Alb. and MG:

(9-71) (Alb.) *Dua të shkoj nesër.* (MG) *thelo na fiyo avrio.*

I-want *të* I-leave tomorrow I-want *na* I-leave tomorrow

(9-72) (Alb.) *Dua që Agimi të shkojë.* (MG) *thelo o Janis na fiji.*

I-want *që* A. *të* he-leaves I-want the I. *na* he-leaves

Thus, in contexts in which extraposition occurs in MG, Alb. invariably introduces a subordinator, creating thus a structure similar to that described for the majority of subjunctive adverbial clauses in the Balkan languages in 9.2.4., namely [**subordinator**] [**X**] [**të~subjunctive verb form**]². A corollary to this is that extraposition may occur in Alb.

¹ The appearance of *që* in subjunctive complement clauses has become an object of interest only in the last ten years, mostly in generative circles, and mostly in the context of comparison with Rumanian, which displays a similar variation between *ca să* and *să* (see, e.g., Rivero 1994, Turano 1993, 1995, Roussou 2000, 2001, Dobrovie-Sorin 2001).

² Of course, preposed elements other than subject may trigger the appearance of *që* as well, as the following two examples show:

subjunctive clauses only in conditionals introduced with *po të*, which accounts for the extremely low frequency of this phenomenon in this language¹.

9.2.8. Conclusion

The basic principles according to which embedded inversion works are the same as those holding in other inversion types, summarized in rules (9-14) through (8-108). In particular, it is similar to relative inversion, since both occur within clauses denoting existentially presupposed propositions, i.e. terms, both have a non-focused initial element and display the same lexical preferences.

What is different is rule (8-106), defining adjacency conditions, which are more complex and variegated here than in any other inversion type described, and the possibilities of obviating inversion, which are in embedded clauses reminiscent of those occurring in the context of wh-inversion.

9.3. Generalization patterns

In Section 5.1. a number of linguists have been quoted who claim that the basic word order (meaning, presumably, the most frequent word order pattern) in Alb. and MG subordinate clauses is VSO, in contrast to other clause types. This claim has been partly confirmed by my statistical data (Sections 5.5.3. and 9.2.3), and partly explained by the notions of relative and embedded inversions. This section represents an attempt to give a complete explanation.

(a) ... *duhet të bëhet një organizim më i mirë.*

it-should *të* be-made an organization more the good

‘... a better organization should be created’ (Koha Ditore, p. 11)

(b) ... *duhet që nëpërmjet politikës fiskale të favorizohen prodhimet...*

it-should *që* by means of policy fiscal *të* be-favored productions-the

‘... by the means of fiscal policy, the production... should be revived’ (Koha Ditore, p. 11)

¹ Some more remarks on *që të*: First, another context in which *që* occurs is that of the disjoint reference between the main and the complement clause subject (the so-called obviation effect): *do që të vijë* (he-wants *që të* he-comes) ‘He, wants him_j to come’ vs. *do të vijë* (he-wants *të* he-comes) ‘He wants to come’ (see Roussou 2001: 92). Second, it is interesting to note that *që* does not have to occur when the fronted element is placed in front of the matrix verb, as the following example shows: *atë duhet të kenë në dorë të gjithë studentet* (that it-should *të* they-have in hand the all students-the) ‘all the students should have this book’ (Koha Ditore, p. 2).

When the term ‘subordinate clause’ is used, it is meant to comprise both relative clauses and complement and adverbial clauses.

The basic idea is, to expound my results in advance, that a diachronic process is (or was) observable in Alb. and MG, in which verb-subject order is getting divorced from its pragmatic motivation (rules (8-104), (8-105) and (8-106)) and is turning into a simple mechanical syntactic rule without a specific meaning. This process has been retarded or completely stopped in the last decades, at least partly for normative reasons. On the other hand, SC seems to be in the process of generalizing subject-verb order at the expense of pragmatic transparency.

9.3.1. Formal and functional explanations of frequency phenomena

Let us first enumerate the factors causing the high frequency of verb-subject order in existentially presupposed subordinate clauses in Alb. and MG. In these two languages, there are clauses in which the introductory word has to be directly adjacent to the verb, thus triggering inversion irrespective of pragmatic and other conditions. In those clauses in which this is not the case, rules (8-104), (8-105) and (8-106) license the postverbal position of the subject for discourse-pragmatic reasons under certain conditions; these conditions are most commonly met when the clause is short; most subordinate clauses are short (see statistics in 5.5. and 9.1.2.). SV order may thus occur only in those presupposed subordinate clauses whose introductory elements have the feature [-adjacent], which consist of more than two elements and do not fall under the scope of the above mentioned rules (see Tables (9-36) and (9-43) for numerical values).

There is even more to it: in those clause types in which the adjacency constraint is present irrespective of their presuppositional status, like subjunctive complements or purpose clauses, VS order is obligatory even then when the clause is not existentially presupposed, e.g. after modal and similar verbs. Furthermore, as indicated in 9.1.4., appositive, i.e. non-presupposed, oblique relative clauses tend to have VS order for purely discourse-pragmatic reasons.

Taken together, these factors may be considered a sufficient formal-functional explanation for the (relative) predominance of VS order in subordinate clauses, both the presupposed and the non-presupposed ones. The fact that MG systematically displays the higher percentage of VS clauses than Alb. is easily explained by the greater number of introductory words which have to be adjacent to the verb in the former.

In contrast, SC subordinate clauses, as demonstrated in 5.5.3. and 9.2.3, do not seem to display any clear preference for VS order. There are two reasons for this: First, SC subordinators and relative particles/pronouns never display obligatory adjacency to the verb, so that inversion occurs only when it is necessary for discourse-pragmatic reasons (all-ratified short clauses, non-ratified subjects). Second, pronominal subjects, a relatively frequent subject type in subordinate clauses, are practically excluded from inversion (rule (8-108)). Subordinate clauses in SC thus do not represent a preferential syntactic context for the occurrence of VS order.

9.3.2. Generalization patterns

9.3.2.1. VS structure

9.3.2.1.1. Superfluous inversion

In all three languages there are some cases in which inversion rules ((8-104)-(8-108)) do not predict the occurrence of VS order (i.e. there is neither an adjacency constraint nor the need to ‘cover’ the final position or to ratify a non-ratified subject referent), but the inversion nevertheless occurs. Since the inversion in short clauses is always accounted for by my rules, it is only long clauses, i.e. those with at least one element more than the verb and the subject, that occur in this class. This type was briefly mentioned in the discussion of all-ratified restrictive relative clauses in Alb. and SC (9.1.5.2.), with six such instances in Alb. and three in SC (the MG relative particle/pronoun is obligatorily adjacent to the verb); among the presupposed embedded (complement and adverbial) clauses, I have 79 examples in MG, 37 in Alb., and 19 in SC. Note that the number of such clauses, at least as far as MG embedded clauses are concerned, is not negligible: 25.2% of all inverted embedded clauses in this language belong to the group where inversion is not motivated by my inversion rules (the numbers for the other two languages are not so dramatic: 15.6% in Alb., 9.4% in SC). Here are some examples:

(9-73) Alb. (Kosovarja, p. 6/48)

Nuk ia vlen që të vuajë familja ime vetëm për teke të MIAT.

not him-it is-worth that të suffer family my only for pleasures the my

‘It is not worth it that my family suffers only because of my pleasures.’

(9-74) MG (Kapandai, p. 142)

... emathe pos lejete i xina KIZ.

learned that is-called the goose *kiz*

‘[He asked someone who spoke Turkish] and thus found out that goose was called *kiz*.’

(9-75) SC (Kiš, p. 148)

... *da se vrate ... u krilo judejstva, ako im OSLOBODE dušu zakoni hrišćanski ...*

to REFL return in lap of-Judaism if them:CLIT free soul laws Christian

‘[I told them that they could perhaps] return to the Jewish faith, if the Christian laws release their souls...’

In none of the three clauses is the subordinator (*që, pos, ako*) obligatorily adjacent to the verb, so that nothing prevents the subject from intervening between them. In the Alb. and MG examples, the element X (*për teke të miat, kiz*) is non-ratified and, according to rule (8-106), placed clause-finally and stressed; in the SC example, the proposition conveyed by the clause is all-ratified (the possibility of the Christian laws releasing the souls of the Jews baptized by force being the topic of the whole passage), so that the verb is stressed and non-final, according to rules (8-104) and (8-105). However, there is no rule in my system which could account for the postverbal position of the subject in any of the three instances: in the case of non-ratified elements other than subject and verb, it suffices that they are clause-final; in the case of long all-ratified clauses, the non-final position of the verb is ensured by the additional element, so that the subject need not be postposed.

How are cases of this kind to be accounted for? I shall suggest a multicausal explanation, with different types of sentence elements triggering the apparently unnecessary inversion for different reasons. For the sake of brevity, the phenomenon will be called **superfluous inversion**. In the course of the presentation, it will become clear that only the last class, (iv), carries this name with some justification.

(i) **Modal adverbials**. As has been elaborated upon in 8.1.1.2.3., modal adverbials may under certain conditions form a constituent with the verb; in this case, the verb and the adverbial count as one unit for syntactic and informational purposes. Consequently, subordinate clauses in which the verb and the adverbial are present and unified into a constituent are to be treated as short, not as long, so that inversion naturally follows from my rules. The positional regularity of constituent-forming adverbials mentioned in 5.1. is observable also in this case: they stand immediately after the verb in Alb. and MG, and immediately before the verb in SC. The number of instances is not very large: eight in Alb. (18.6% of all superfluous inversions, 3.4% of all inversion clauses), eight in MG (10.1%, i.e. 2.6%), and four in SC (18.2%, i.e. 2.7%).

(9-76) MG (Matesis, p. 42)

Emis omos vjazomaste na teljonun yrlyora i sistasis.

we however hurry *na* end quickly the introductions

‘We, however, are in a hurry to quickly finish the introductory rituals.’

(ii) **Lexical solidarities.** The same principle as with modal adverbials (multi-word syntactic units) is found in the clauses in which the verb and the additional element display a close semantic tie, as idioms or in some other way. As indicated in 8.1.1.2.4., elements in the relation of lexical solidarity are in all three languages placed immediately after the verb. There are six instances in Alb. (14.0% of all superfluous inversions, 2.6% of all inversion clauses), seven in MG (8.9%, i.e. 2.2%), four in SC (18.2%, i.e. 2.7%).

(9-77) SC (Kiš, p. 56)

... kratko vreme kad se ima u vidu transformacija koju je doživeo.

short time when REFL has in sight transformation which he underwent

‘... a short time, when the transformation he went through is concerned (“when the transformation is had in sight”)’

(iii) **Deictically/anaphorically bound setting adverbials.** In all-ratified contexts, when the additional element in a long clause is a setting adverbial which is somehow deictically or anaphorically bound (‘here’, ‘there’, ‘at that time’, ‘in his room’, etc.), it sometimes stands in front of the verb, so that the subject has to be placed postverbally, in order to satisfy rules (8-104) and (8-105). There are 14 such cases in Alb. (32.6% of all superfluous inversions, 5.9% of all inversion clauses), 23 in MG (29.1%, i.e. 7.3%), ten in SC (45.5%, i.e. 6.7%). Here is an example of an all-ratified long temporal clause:

(9-78) Alb. (Koha Ditore, p. 12)

... do të dërgojnë nëndetësen ... para-se atje të ZHYTEN tre zhytës me aparaturën.

FUT *të* they-send submarine before-to there *të* dive three divers with apparatus

‘[Norwegian divers should dive in and try to find out what happened with the Russian war ship. However,] they will send a submarine ... before three divers with the apparatus dive in there.’

The word order pattern one would expect according to my rules – *para-se tre zhytës* (S) *të ZHYTEN* (V) *atje* (X) – is also possible and quite natural. That inversion nevertheless occurs, and so often, has to do with the general tendency of setting adverbials, especially when somehow deictically or anaphorically bound, to occupy the preverbal position independently of inversion rules. When they are placed preverbally (for the reasons which have nothing to do with the inversion), the inversion has to take place, at least in all-ratified contexts, because

the subject is then the only remaining element of the clause capable of ‘covering’ the clause-final position, so that the accented verb may stand non-finally, complying to rules (8-104) and (8-105)¹. Let me note in passing that this is, with one exception, the only kind of superfluous inversion I found in restrictive oblique relative clauses (see examples (9-26) and (9-27) in Section 9.1.5.2.).

(iv) **No functional explanation.** There is no functional explanation for inversion in 15 cases in the Alb. corpus (34.8% of all superfluous inversions, 6.4% of all inversion clauses), 41 cases in the MG one (51.9%, i.e. 13.1%), and four in SC (18.2%, i.e. 2.7%).

Let us first do away with SC: All four instances (one all-ratified relative clause and three all-ratified adverbial clauses) come from the same source, one of the stories from Kiš’s collection which, in the best postmodern tradition, pretends to be a translation of a transcript of an interrogation conducted by the Spanish Inquisition. The language is deliberately archaizing, one of the archaic features being the superfluous inversion. Thus, (9-75), taken from this source, sounds very old-fashioned with inversion (*ako im oslobode dušu zakoni ...*, ‘if to-them release soul laws ...’); the unison judgment of all speakers of SC I interviewed (including myself) is that the SV version would be the ‘normal’ one in the contemporary language (*ako im zakoni oslobode dušu*, ‘if to-them laws release soul’). Superfluous inversion not belonging to one of the three types enumerated above is an archaism in modern SC, a fact I shall return to at the end of this section.

Alb. and MG (especially MG) are quite different, both statistically and with respect to the place superfluous inversions occupy in the language system. Consider the examples (9-73) and (9-74): the attested sentences are of the form [subordinator][verb][subject][non-ratifiedX] (*që të vuajë familja ime vetëm për teke të miat* ‘that të suffer family my only for pleasures the my’, *pos lejete i xina kiz* ‘that is-called the goose “kiz”’). When presented with the variants with SV order (*që familja ime të vuajë ...*, *pos i xina lejete kiz*), native speakers of Alb. and MG do not, like the speakers of SC, claim that they are ‘better’, or ‘more normal’, but simply that they are possible as well. This ‘normality’ of superfluous inversion is confirmed by the statistics (see above): a clause type occurring with the frequency of 5–15% within the class of inversion clauses is not likely to be merely an archaizing ornament.

I should like to argue that the frequency of verb-subject order, triggered by the formal and pragmatic factors depicted in 9.3.1., began at some point to be reinterpreted as a default

¹ The following quotation from Jacennik and Dryer (1992), pertaining to Polish, nicely captures the principle I stipulate: “While in some cases the choice between SVX and XVS may reflect the discourse properties of the subject, it is likely that it will often reflect the discourse properties of the nonsubject”.

ordering pattern for subordinate clauses in Alb. and MG, so that the subject began to be placed postverbally also in those cases in which there was neither a formal nor a functional justification for it. From the data presented in Tables (9-36) and (9-43), which show that SV order occurs in subordinate clauses with the frequency of 30–50%, it is clear that this process has not reached its end (and probably never will – see below), which is the reason I prefer the use of the term *generalization* to the commonly used *grammaticalization*. Its consequences are nevertheless apparent.

9.3.2.1.2. Excursus: Some paths of generalization

Although the present study does not have diachronic pretensions, a word on the possible paths of generalization is in order. Note that a line is drawn between *motivation* of a change (in this case, the frequency of VS order in subordinate clauses for pragmatic reasons) and the *path*, which is conceived as the way the change comes into being in a language. The number of paths does not have to be one, as diachronic studies often assume: a unitary motivation of a change may open a number of different ways for the change to take place.

Apart from vague claims to the effect that VS is an areal feature of the Balkan *Sprachbund* (e.g. Haebler 1957), I am aware of only one proposal concerning the generalization of VS order in subordinate clauses, which, not surprisingly, deals with Greek. Combining the data on the accentuation change of subordinators (Trypanis 1960) and Mackridge's account of the behavior of clitics in medieval Greek texts (Mackridge 1993), Horrocks (1990, 1997: 59sq., 208sq.) tries to account for the apparent dominance of VS order in subordinate clauses by assuming that this order is a compromise between two opposed tendencies having to do with clitics, namely to keep the old Wackernagel law intact by placing the clitics in the second position, i.e. directly after the subordinator, on one hand, and to keep the clitic as close as possible to its governing verb on the other. The optimal way to do both things simultaneously is to use a verb-initial order: [subordinator] [clitic] [verb] [rest of the clause]¹.

There is doubtless something to this hypothesis, though I do not think it is able to account for the whole phenomenon of generalization. Therefore I propose two more possible paths, without pretending to exhaustiveness. First, the simple cliticization of at least one subordinator: Ancient Greek *hína*, as Trypanis (1960) has shown, became *hiná* till the early sixth century AD, and subsequently lost its lexical accent (and the subordinator status proper, see above), so that it now behaves like any other MG clitic, being obligatorily directly

¹ An almost identical solution for the grammaticalization of verb-initial order in Insular Celtic was proposed almost forty years ago by Watkins (1963).

proclitic to the verb complex, which always has the subjunctive morphology¹. A similar analysis is probable, but not provable (because of the meager historical evidence) for the Alb. particle *të*. This pattern – [**na/(të)**]~[**verb complex**] – might have then spread over to all the cases in which the verb complex had the subjunctive morphology without *na*, the pattern being thus [**subordinator/na/(të)**]~[**bare subjunctive verb**]. This is a possible scenario of how the adjacency constraint came into being in the first place, through the loss of phonological independence of one subordinator taking the bare subjunctive and the subsequent generalization of this pattern.

The second hypothetical path must be of a somewhat more recent date, being possible only from the time *na/të+verb* became reinterpreted as a syntactically unanalyzable mood form. As indicated above and in Section 9.2.4., the pattern [**subordinator**]~[**bare subjunctive verb**] holds in all those cases in which a subordinator is used without *na/të* and with a verb with the subjunctive morphology. Now, most of the subordinators which allow for this option may also be used with *na/të+subjunctive* and with the indicative. Thus, sentence (9-59) (*molis kimiθi to peði* – ‘as-soon-as falls-asleep:SUBJUNCTIVE the child’) is an instance of the adjacent pattern with bare subjunctive. The same subordinator can, however, be used with the indicative, when the time reference is to the past:

(9-79) MG (Matesis, p. 62)

Molis katevikan ti skala i ikoðespotes...

as-soon-as went-down:INDICATIVE the stairs the hosts

‘As soon as the host went down the stairs...’

It is probable that the pattern with the obligatory adjacency between the subordinator and the verb, valid only for bare subjunctive clauses, is easily transmitted to the other uses of the same subordinator, with *na+subjunctive*, and with the indicative. In this particular case, with the MG temporal subordinator *molis* (and with other temporal subordinators), the adjacency has never become obligatory with the indicative, but the tendency is recognizable, so that clause (9-79) is an instance of superfluous inversion, probably triggered by the analogy to the use of *molis* with bare subjunctive.

Some other paths are conceivable, too: for instance, the creation of new subordinators out of the combination [subordinator]+[*na/të* complex], a process already completed in the case of *ja na*, introducing purpose clauses in MG, and almost completed with the *që të* complex introducing these clauses in Alb., and with *san na* ‘as if’ in MG (see 9.2.4.). The process

¹ Hesseling (1927) proposes the same development for the relative particle/subordinator *pu*; if this is true, the clitic status has not yet been reached (cp. Haberland and van der Auwera 1990).

seems to be well under way with other such combinations as well, like *oste na* ‘so that’ in MG and with the other uses of *që të* in Alb.

9.3.2.1.3. An interrupted process

As is clear from my statistics, the process of generalizing the VS order independently of discourse-pragmatics is far from completed. In fact, it is today much farther from being completed than it was some fifty years ago: at the time people like Thumb (1910) and even Tzartanos (1963) were writing on MG, and Haebler (1957) on Alb. word order, the system seems to have been almost rounded up¹, so that it is no wonder that they claimed VS to be the ‘basic’ word order in subordinate clauses. The process of generalization seems to have been retarded, or even reversed since then: even in the narrowest field of inversion (presupposed clauses), the share of SV clauses ranges between 30 and 50%. I shall argue that there is one general and one particular reason to this.

The general reason is what Blake (2001) calls ‘global trends in language’, meaning by that nothing else but the enormous influence of the western European languages (AME, Average Modern European) on the rest of the world. In the particular case of the two Balkan languages, this influence was twofold: when the modern Albanian and Greek nations came into being, the creation of the national standard based on western standards, mostly French, Italian, and English, was considered a major task of the local intelligentsia. In Greece, this process commenced in the 19th century, but, because of the very specific sociolinguistic situation in the country, it had little impact on both the spoken and the language of literature until recently (see Section 6.6. and below). In Albania, the standardization based on the Tosk dialect effectively started only after the Second World War, which is approximately the time the generalization of VS order is observed to have been retarded: preverbal subjects in French, Italian, and Russian subordinate clauses were probably reason enough to reduce the occurrence of VS order to inversion proper, i.e. either to contexts in which it is triggered by the discourse-pragmatic factor or to those in which the adjacency constraint had already become grammaticalized to the point. The written standard has exercised some influence on the spoken language of the educated, so that VS order in subordinate clauses in modern Alb. is basically reduced to the above mentioned contexts. The second wave of AME influence

¹ For fairness’ sake, it must be said that both Tzartanos and Haebler worked mostly with the folkloristic material collected in the early twentieth century, so that their conclusions reflect rather the situation of Thumb’s time than of their own.

came with the massive influence of English in the post-war period, which was, however, until recently much stronger in Greece than in Albania.

The particular reason pertains to MG only. Standardization, which started in the 19th century, never had a decisive influence on the spoken and the literary languages because the proclaimed standard, *katharevousa*, was almost closer to Ancient Greek than to the modern language, and was as such neither able to influence the spoken variety nor accepted as standard by most men of letters, who wrote in the colloquial variant called *dhimotiki*. Only of late, practically since the early Seventies, have the two variants merged into the modern standard, with the demotic basis and many elements of *katharevousa*. The result is nicely summarized in the following quotation:

In ... subordinate clauses the most common neutral order is for the verb immediately to follow the conjunction. Nevertheless, in written styles ..., the subject is often placed before the verb. It is possible that in writing, Greek speakers have been influenced by *katharevousa*, which tries to approximate to the SVO order of French, and that this influence extends even to those kinds of subordinate clauses in which this order was not frequently found in natural speech. Moreover, today the written order seems to have influenced the spoken... (Mackridge 1985: 237)

The way the 'global trends in language' found their way into MG was a bit more cumbersome than in the case of Alb., but they did eventually find their way, so that MG is today on a level similar to Alb. as far as VS order in subordinate clauses is concerned. The main difference lies in the fact that in MG, the adjacency constraint was more widespread and more strict when the retardation/reversal of the VS-generalizing process started than was the case in Alb.

The picture we get is thus one of an almost cyclical diachronic process: first only the discourse-pragmatically triggered VS order, i.e. the one which is determined by rules (8-104) through (8-106); then a generalization of this order, resulting on one hand in the creation of adjacency constraint with a number of words, and on the other in the gradual separation of VS order from its pragmatic roots; finally, the generalization process is interrupted (or only retarded), so that only those instances of VS order occur which are justified by the pragmatic principles, but also those which have been grammaticalized during the generalization process due to the appearance of the adjacency constraint. The superfluous inversion clauses described in point (iv) of Section 9.3.2.1.1. are thus to be regarded either as remnants of the older stage, or simply as the results of the fact that the generalization process has not been stopped, but merely retarded. It is perhaps not unimportant that most of the examples of this kind occur in novels and other literary and semi-literary products (12 in Alb., 36 in MG),

whereas the more formal styles, like the journalistic one, seem to avoid them altogether (3 in Alb., 5 in MG).

A note on SC: the fact that the few examples of superfluous inversion I found have an unmistakably archaic flavor suggests that a generalization process similar to the one observed in Alb. and MG may have once also existed in SC. If this is the case, this hypothetical process must have been stopped and reversed (see below) quite a long time ago. However, since there are no investigations on the diachronic development of SC word order, this must remain a highly hypothetical issue.

9.3.2.2. SV structure in SC

The factors of frequency and, probably, the influence from the West that was both longer and more profound than in Alb. and MG (at least in Croatian and northern Serbian areas, the donor languages being mostly German, Latin, and later French), make the SV pattern in SC much more prominent than in Alb. and MG: notably, the former never developed anything similar to the adjacency constraints of the latter two.

Consequently, a generalization pattern opposite to the one described in 9.3.2.1. may be observed in this language: there are cases where my rules predict inversion, but SV order occurs instead. Their frequency in my corpus is not particularly impressive – I counted 19 instances – but they are possible and quite natural and, so my impression, found also in spoken language along with their regular VS counterparts. All the examples are short, consisting only of a verb and its subject, and all-ratified, so that inversion should occur, in accordance with rules (8-104) and (8-105). But it does not:

(9-80) Eng: *And every time **the hand descended**, the ears flattened down and a cavernous growl surged in his throat.* (London, p. 150)

Alb: *E sa herë **ULEJ dora** ...* (p. 148)

and as-many times descended hand-the

MG: *Ke kaθe fora **pu KATEVENE to xeri** ...* (p. 183)

and every time REL descended the hand

SC: *A svaki put **kad bi se ruka SPUŠTALA** ...* (p. 170)

and every time when would:CLIT REFL:CLIT hand descended

Alb. and MG behave properly; SC does not invert, although the clause is both short and all-ratified. The passage in which this sentence occurs describes a master fondling his dog; that his hand descends to reach the dog is in no way controversial in this context, so that it would

be a pure stipulation to characterize the verb as non-ratified and therefore postposed in SC. The only remaining explanation is that of a generalization process: speakers of SC tend to extend the field of SV order over the limits imposed to it by discourse-pragmatics. My impression is that the analogy with pronominal subjects, which, as defined by rule (8-108), almost never invert, may be the decisive factor in this and similar cases.

The process does not seem to be particularly productive and is actually only a minor phenomenon in the language, which still has a full-fledged inversion based on the rules repeatedly quoted throughout this chapter, but it is worth mentioning, since it represents a potential for the gradual disintegration of the principles on which inversion is based.

10. Inversion: Conclusion and summary

The principal feature of the verb-subject order labeled *inversion* is its use as (a part of) the presupposed clause material: The easiest way to characterize it would thus be to say that inversion is a verb-subject construction appearing in presuppositional contexts. Unfortunately, the situation is not that simple. The presuppositional contexts in which inversion occurs are invariably marked with a clause-initial element, marking narrow focus in focal inversion, the subordinated status of the clause in subordinate inversion. This element tends to be directly adjacent to the verb, the degree of strictness of the adjacency constraint varying both across inversion types and across languages. Inversion thus seems to be triggered by a mixture of discourse-pragmatic and purely formal factors. To make things even worse, at least in the case of relative and subordinate inversions, all three languages seem to be involved in diachronic processes of generalizing certain word order patterns, making thus the precise definition of the formal factors operative in these two inversion types extremely difficult.

Inversion is therefore best described as the verb-subject order occurring in presuppositional contexts which are marked as such by a clause-initial element. It is partly triggered by the need to mark different ratification statuses of the proposition conveyed, partly by the restriction imposed by the syntax of the given language that the initial element and the verb be immediately adjacent, and partly by the tendency to turn the ordering pattern predominant for pragmatic reasons into a formal restriction. Discourse-pragmatics, formal restrictions, and deviations from the both are therefore the key words of the discussion that follows.

10.1. Discourse pragmatics

10.1.1. Presuppositional contexts

The discourse-pragmatic side of my analysis crucially depends on the analysis of the inversion contexts as presuppositional.

This is not very problematic as far as focal inversion is concerned: if the interpretation of the initial element in quotative, interrogative and fronted-focus sentences as narrow focus expression is accepted, then the rest of the sentence has to convey a propositional function which is not only existentially, but also relationally presupposed, i.e. the speaker does not only have to treat it as non-controversial and therefore not worthy to be asserted, but also as the point of relevance with respect to which the assertion is to be assessed (see 4.2.3.).

The presupposed status in some subordinate contexts may be more controversial. Recall that my contention is that restrictive, but not appositive, relative clauses are existentially presupposed, as well as those embedded clauses which encode terms (as opposed to predicates) on the semantic level. Obviously, there are two possible bones of contention here: the definition of existential presupposition and the criteria for telling apart the clauses that encode terms from those that do not.

The former issue has been expounded at some length in Section 4.2.: the principal features of my notion of existential presupposition are the treatment of a proposition as non-controversial and therefore superfluous to assert, its status as a term on the semantic level, and the complete independence from the truth value of the proposition. These features justify, or at least make possible, the treatment of restrictive relatives and term-encoding embedded clauses as existentially presupposed.

The latter problem, the criteria for distinguishing terms from independent propositions, is not easy to solve in all particular cases. For relative clauses, both formal semantic representations (e.g. Quine 1960: 110ff., Partee 1975: 230ff., Haberland and van der Auwera 1990: 134ff.), treating restrictives as attributes to the noun and appositives as epenthetical predications, and simple intuitions speak for my analysis. It is rather the embedded clauses that are difficult. First, some of them have been treated as presupposed from the very beginning of the presupposition research (9.2.2.), but on a basis different from that endorsed by me, namely because of the interaction of their truth values with negation. Those who are used to the traditional approach may have difficulties in intuitively grasping the somewhat different notion of presupposition used here, but this is something I shall have to live with.

Second, and more important, I claim that different *uses* of the *same* clause may cause it to be a term in the one, but an independent proposition in the other context (9.2.2.). We are thus effectively not in the position to simply distribute features [+/-presupposed] according to clause types (which would sound like: all temporal clauses are [+presupposed], all causal clauses are [-presupposed], etc.; recall that this is also the problem of all semantic presuppositionalists, 9.2.2.). The issue of ‘termness’ vs. ‘non-termness’, and of the concomitant presuppositional status, is dependent on both the intentions of the interlocutors and the semantic structure of complex propositions, an issue which surpasses the scope of this study. Thus, I am afraid I am not able to do any better than to enumerate those uses of certain embedded clauses in which the propositions conveyed are independent predications (see the list in 9.2.2.) and therefore not existentially presupposed. Such lists, notoriously, witness to the inability of the analyzing linguist to find superordinate regularities and to his/her reliance

on intuitions. I am aware that this is not a particularly rigorous methodological stance, but I consider it better to offer a semi-intuitive account than none.

An indirect support for the existentially presupposed status of subordinate clauses comes from the lexical semantics of the components of relative and embedded inversions. As demonstrated in 4.2.4., certain combinations of lexemes are easier to existentially presuppose than the others, in particular those verb-noun combinations in which the verb belongs to the (rich) existential presupposition of the denotatum of the noun, or those which often and typically occur together. It has also been shown that contextually ‘given’ material is, being non-controversial, rather presupposed than asserted. It is precisely these three configurations that are found in subordinate inverted clauses significantly more often than average¹. Now, these lexical preferences are only preferences, not rules (neither are presuppositions automatically triggered by certain lexemes nor are certain combinations of lexemes obligatorily presupposed), so that this is only an indication of the presuppositional structure of certain subordinate clauses, not a proof. Nevertheless, the reader should keep in mind the prototypical lexical filling of subordinate inversion clauses, since its similarity to vS construction described in the following sections allows for some generalizations beyond the notions of presupposition and assertion.

10.1.2. Ratification

Another important notion in my account of inversion is that of ratification, elaborated upon with respect to nominally encoded topics in 6.2.

My analysis of word order rules regulating inversion is based on the understanding that it is not only nominally encoded (parts of) propositions that may function as topics, but also those propositions or propositional functions which are encoded as clauses. In other words: it is not only subjects, objects, etc., that can be topical, and consequently ratified or non-ratified in their topical status, but also whole clauses, or parts of clauses. This naturally follows from the fact that propositions may function as discourse referents, just like entities (for a strong claim to this effect, see Dryer 1996; cp. also Lambrecht 1994, Lambrecht and Michaelis 1998).

If a proposition/propositional function is relationally and/or existentially presupposed, there are two principal informational statuses it can be ascribed. It can be ratified as a whole,

¹ The fact that no such tendency is observable in focal inversion has to do both with specific conditions on its use and with the relational, not merely existential, presupposition connected with it.

if the speaker assumes that the hearer expects the whole state of affairs conveyed to appear in a presuppositional context (because of its immediate presence in the physical or textual context, or because of the general plausibility of it occurring in the slot in which it occurs, or for some other reason). Thus, if I, while being on holiday, hear from my insurance agent that my house burnt down, and if, on my return, the friend who was supposed to look after it does not know how to confront me with this fact, and I notice his being in trouble, I may, in order to ease his pain (if I am a nice person), say *I KNOW that the house burnt down*, with the whole complement clause deaccented, i.e. ratified. I assume that my interlocutor has the whole proposition ‘the house burnt down’ in the foreground of his consciousness at the moment, and feel therefore justified in treating it not only as presupposed, but also as all-ratified, in order to assert that I am aware of its denotatum.

The other possibility is that the presupposed proposition/propositional function contains non-ratified elements, if the speaker has a reason to believe that the one of the participants in the state of affairs is for any reason either not expected to occur in the given presuppositional environment, or in any way contrasted with any other element of the discourse. Thus, if I, meeting the above mentioned friend, hear from him that my dog died in the fire, I may say *I KNEW that the HOUSE burnt down*. I still presuppose the proposition ‘the house burnt down’ and assert my awareness of this fact, but ascribe to the hearer different expectations than in the above case: I act as if he were not expecting the appearance of the referent ‘the house’ in the presupposed presuppositional function ‘X burnt down’, thus indirectly inducing a contrast with the formerly asserted ‘the dog died = burnt down’.

The principle of ratification within presupposed propositions/propositional functions is, in my interpretation, the actual *raison d’ être* of inversion, not the presupposed status of the propositions themselves, this being sufficiently marked by other means, notably by the initial focal or subordinating elements in the clauses conveying them. In particular, I claim that the syntax and intonation of presupposed clauses reflect the difference between all-ratified propositions and those containing a non-ratified element. Inversion is only an incidental by-product of this fundamental distinction.

More precisely, the rules which define the mapping of discourse pragmatics to syntax/intonation, (8-104) – (8-108) = (8-104) – (8-106), demand merely that (a) in all-ratified clauses, the nuclear stress is placed on the non-final verb and (b) in the clauses with a non-ratified element, this element carries the sentence stress and is placed clause-finally. When a clause is all-ratified (case a), the inversion is necessary only when the subject is the only clause element which can cover the final position, so that the verb-centered nuclear stress may be in

a non-final position, i.e. only in short clauses: the above example with the all-ratified proposition ‘the house burnt down’ (*I KNOW that the house burnt down*) would in all three languages under consideration be expressed with inversion, because the clause is short: ... *se u DOGJ shtëpia* (Alb.), ... *pos KAIKE to spiti* (MG), ... *da je IZGORELA kuća* (SC) [all three: ‘that burnt-down the house’]. If yet another element is present, the inversion is less probable, since it is unnecessary. When a clause contains a non-ratified element (case b), it is only when this element happens to be the subject that the inversion must occur, so that the rule of placing the non-ratified element in the final position is satisfied. The second example, with a contrastive and therefore non-ratified subject referent (*I KNEW that the HOUSE burnt down*), would again be inverted, this time with a different intonation pattern: ... *se u dogj SHTËPIA* (Alb.), ... *pos kaike to SPITI* (MG), ... *da je izgorela KUĆA* (SC). If any element other than subject is non-ratified, the inversion need not occur.

The syntactic and intonational marking of ratification statuses taking place within larger chunks of presupposed material raises three interesting issues. First, it shows that fine differentiations in information packaging, marked formally in order to make the speaker’s intentions transparent (cp. Section 4.6.), are possible, indeed obligatory, not only in assertional, but also in nonassertional contexts. This is an indirect support for my claim that presupposed and asserted propositions are ontologically identical, the only difference being the order in which they are processed (4.2.1.).

The relative weight of this indirect support is, to my despair, seriously diminished by the second issue: if one compares the formal strategies of informationally marking the asserted material with those used in the presuppositional contexts, the picture that emerges displays only a partial parallelism between the two. Namely, non-ratified elements behave in a manner similar to expressions encoding narrow foci: they both tend to be placed clause-finally (though see 8.3. for a specialized divergent strategy of encoding narrow foci) and carry the nuclear stress (see also Section 4.5.). All-ratified propositions should, if the parallelism were perfect, be formally identical to broad foci, since in both cases propositions or proposition chunks are assigned a unitary informational value. This is, however, not the case, at least not completely. Although in both cases the verb occupies a non-final position, there is an intonational difference: with broad foci, the nuclear accent is on the last postverbal element belonging to the focus domain (see Section 4.5.2.); with all-ratified presupposed clauses, the nuclear accent is on the verb. Since I do not have a functional explanation for this phenomenon, I simply state it here as a descriptive fact.

The third issue, however, mitigates the disturbing impact of the second. As already mentioned in 8.2.2.5., Ladd (1996:168ff.) has shown that the accentuation in ‘neutral’ (= all-ratified) wh-questions is a typological variable: some languages (in Europe, mostly located in the East and, we may add, Southeast) accent the verb, some accent postverbal nominal elements (mostly in the West). Lambrecht and Michaelis (1998: 508), albeit with some modifications, demonstrate that the second part of this claim, the one pertaining to the Western European languages, may be extended so as to hold for all nonassertional contexts: in these languages, if a (part of a) clause is all-ratified and presupposed, the accent falls on the postverbal material. I hope that my description of presuppositional contexts in Alb., MG, and SC has made it clear that the first part of Ladd’s generalization, predicting the accentuation of the verb in the East, obtains not only in wh-questions, but all kinds of presuppositional contexts: at least in Alb., MG, and SC, it is the non-final verb, not the postverbal material, that carries the sentence stress when an all-ratified proposition is embedded in a syntactic context marked as presuppositional. Schematically, the accentuation patterns in all-ratified clauses may be represented as follows:

(9-81) *Accentuation patterns in assertional and presuppositional contexts*

	assertional context	presuppositional context
West	postverbal material	postverbal material
Southeast	postverbal material	non-final verb

If anything, this table shows that the mapping of discourse-pragmatics and syntax/intonation is a conventionalized matter, varying from language to language, or from linguistic area to linguistic area. For reasons which are not clear to me, the languages of the Balkans tend to mark the integrated reading in presuppositional contexts differently from the integrated reading embedded in an assertional context.

10.2. Formal restrictions

If information packaging were the only motivation for inversion, we would have a unified system functioning in an equal fashion in all three languages under consideration, the rules defining the mechanism of ratification being identical for all. Of course, the reality is not so neatly leveled: Alb., MG and SC do behave differently in many cases despite identical discourse-pragmatic prerequisites. I have put the burden of explaining divergences on the difference in formal restrictions holding in each of the languages, the principal ones

concerning the syntactic relationship between the initial element, a narrow focus expression or a subordinator, and the verb on one hand, and the behavior of pronominal subjects on the other.

Pronominal subjects, whose syntax is described by rule (8-108) (=8-108)), are subject to specific restrictions only in SC (in Alb. and MG, they are subject to the same rules as lexical NPs), to the effect that they are practically excluded from appearing postverbally, even when one of the discourse-pragmatically motivated rules demands it. The fact is of some interest, since: (a) it is an instance of a formal restriction overweighing the needs of discourse; and (b) it is a construction-specific restriction. A word on (b): As indicated in 6.4.1., pronominal subjects in SC do appear postverbally, even rather frequently, if certain conditions are met. It is only in the nonassertional, i.e. inversion, contexts, that this position is very untypical – an indication of inversion being distinct from the other two VS constructions.

The relationship between the initial element and the verb is a major factor of both the intralinguistic and crosslinguistic variation in the Balkans. Within one language, it may be the case that some elements have to be immediately adjacent to the verb, that others allow for setting adverbials and similar peripheral expressions to intervene, while some are completely void of any kind of adjacency constraint. The same variegated picture appears when the languages themselves are compared. The relevant facts are summarized in the following table:

(9-82) *Adjacency constraint in Alb., MG, and SC*

	MG	Alb.	SC
quote	+	+	+
interrogative word	+	+	–
fronted focus	+	+	–
relative word	+ (–)	–	–
subordinator	+/-	-/+	–

The greater the number of initial elements in a language which display the adjacency constraint, the more obligatory the inversion is. From the above table it is clear that in MG most of the elements flagging the presupposed status of the material that follows are more or less obligatorily adjacent to the verb; Alb. is somewhere in the middle, whereas in SC it is only direct quotes that display this constraint. Consequently, inversion is almost obligatory in MG, irrespective of the ratification mechanism, whereas in SC it appears *only* when the system of ratification imposes it on the speakers. Thus, in many contexts in MG and Alb., ratification is expressed only through intonation, the word order variation being partly divorced from its pragmatic origins.

10.3. Syntactic pattern, deviations, and configurationality

Inversion is partly a discourse-pragmatic phenomenon, partly merely a syntactic feature, the proportion of these two parts varying both within and across languages. With some simplification, two patterns obtaining for all three languages, may be established, one with adjacency constraint and one without it. The pattern for SC pronominal subjects is separately represented (*ie* reads as ‘initial element’):

(9-83) *Patterns of inversion in Alb., MG, and SC*

(a) [+adjacent]

(b) [–adjacent]

all-ratified:

[ie]~[VERB] ([X]) [subject]

all-ratified:

[ie] [VERB] [subject]

(SC: **[ie] [pronominal subject] [VERB]**)

[ie] [subject] [VERB] [X]

non-ratified subject:

[ie]~[verb] ([X]) [SUBJECT]

non-ratified subject:

[ie] [verb] ([X]) [SUBJECT]

non-ratified verb:

[ie]~[VERB] ([X]) [subject] ([X])

non-ratified verb:

[ie] [subject] ([X]) [VERB]

non-ratified X:

[ie]~[verb] [subject] [X]

non-ratified X:

[ie] [subject] [verb] [X]

As the summary in table (9-82) shows, the pattern (a) is most common in MG, the pattern (b) in SC, Alb. being somewhere in between¹.

In relative and embedded inversions, MG and, to a smaller extent, Alb., display a tendency to generalize the pattern (a), either by investing further initial elements (mostly subordinators) with the adjacency restriction or simply by analogically using VS order in the clauses in which it is neither required by the syntax nor by the discourse pragmatics. A somewhat weaker tendency in the opposite direction is observable in SC, where SV order is sometimes

¹ There are some minor intonational deviations from the patterns, notably in quotation inversion and ff-inversion: in the former, the initial element has its own intonational structure, in the latter, it carries the nuclear stress, so that ratification accents (on the verb or on the subject) are instantiated as the secondary stress.

found where the pattern (b) predicts inversion¹. In both cases, the variation between the form predicted by the patterns and the generalized VS or SV order is partly dependent on the register speakers use.

Among other things, this shows that the Balkan languages are not configurational in the sense in which the term is usually used. The fact that their sentence structures are not predictable from the supposed syntactic hierarchies has been acknowledged long ago; attempts to include them in the group of discourse-configurational languages, on a par with Hungarian and other languages in which the sentence structure is (almost) wholly dependent from the underlying discourse configurations, fall short of accounting for the type of variation triggered by the interplay of such different factors as discourse pragmatics, syntax, and register.

¹ If my observations on the possible developments in SC are correct, then Bennett's (1987) and Haiman's (2002) claim that SC (and Slovene) are currently developing a verb-second structure comparable to that in Germanic languages, starting with one of the prototypical inversion contexts, namely constituent questions, is simply wrong.

11. vS Construction

The construction labeled vS in the present study covers the phenomenological field of what has been called *theticity* in the past thirty years or so (cp. Chapter 2), or at least a part of it. In order to follow the principle known as ‘first things first’, however, I shall postpone the discussion of theticity to the end of this chapter and dedicate this section to the description of vS construction only. In other words, I shall first try to come to an empirically based conclusion on vS construction and only then to compare my results to the approaches to the theticity problem expounded in Chapter 2.

The label vS is, as in the case of VsX (cp. 6.1.), meant to capture the prototypical features of the construction: deaccented verbs, accented subjects, the lack of postsubjectal material. As will become patent briefly, however, ‘vS’ should not be understood as a complete description of the formal properties of the construction: there are many instances in which there is some material between the verb and the subject, even some with postsubjectal elements, and the presence of the preverbal material is rather a rule than an exception. The reader should therefore keep in mind that the label is merely a convenient shorthand. Here are some examples:

- (11-1) Eng: *Twenty feet away towered a huge dead spruce.* (London, p. 26)
Alb: *Nja njëzet hapa larg slitës ishte një pishë e madhe e THATË.* (p. 25)
one twenty steps away of-sled-the was a pine-tree the big and dry
MG: *Kamja ðekarja metra pjo pera ipsonotan ena pelorio, KSERO elato.* (p. 33)
some ten meters more away rose a giant dry fir
SC: *Na dvadeset stopa od njega dizala se ... ogromna OSUŠENA jela.* (p. 30)
on twenty feet from him rose REFL giant dried fir
- (11-2) SC (Andrić, p. 57)
Lepu i nesrećnu udovicu prosili su mnogi GRČI.
beautiful and unfortunate widow wooed AUX many Greeks
‘The beautiful and unfortunate widow was wooed by many Greeks.’
- (11-3) MG (Fakinou, p. 61)
Prota payosan ta POĐJA tis ki istera arxisan na muðjazun ta ðaxtila ton XERJON tis.
first froze the feet her and later started to be-numb the finger of-the hands her
‘Her FEET froze first, and later the FINGERS on the hands started to get numb.’

(11-4) Alb. (Bishqemi, p. 43)

... *komandanti shtypi një buton... Lëvizti një derë.*

commander-the pressed a button moved a door

‘... the commander pressed a button... A DOOR moved.’

I shall try to show that vS should be defined through two partly independent phenomena. First, it represents a focus domain consisting of a verb and a subject, with a relatively limited possibility of including further elements into the focus domain. Recall that focus domain in the Balkan languages is defined as a construction marking the maximal scope of assertion, in which the verb represents the left border, the non-verbal element carrying the nuclear stress the right border of the domain (Section 4.5.2.). Consequently, in the case of vS, the unaccented verb is the left, the accented subject the right border of the domain, functioning in absolutely the same way in which in verb-object, verb-adverbial, etc., focus domains work, the only difference being that the non-verbal element in this particular case plays the grammatical role of the subject. Recall further that the term *focus domain* refers to a the syntactic reflex of a focus structure underspecified as to the feature broad/narrow focus. In vS construction this means that two focus construals are possible: the broad focus construal with the denotatum of the verb-subject complex in the scope of assertion, and the narrow focus construal in which only the denotatum of the subject carries the assertion. The focus of a vS clause can thus be [verb-subject], as in (11-1), (11-2) and (11-4), or only [subject], as I suppose is the case in (11-3).

The second part of my explanation of vS concerns the topics to which the assertional structure conveyed by these clauses is related. Not surprisingly, verb-subject focus domains encode assertions about non-subject topics. These topics may be direct, if they are semantically construed as entities and encoded as direct terms of the predicate, i.e. as objects, so that the relevance relationship to the predicate is explicitly marked (cp. (11-2)). They may also be indirect, i.e. construed semantically as non-entities, or playing a non-central role in relation to the predicate (cp. (11-1)), or both (cp. (11-3)), or even completely absent from the expressed proposition, so that they have to be construed in their entirety, like (11-4) (cp. 4.3.3. and 4.5.2.). In the following sections I shall argue that most semantic, informational, and discourse properties often ascribed to ‘thetic’ utterances are easily derivable from the semantics and pragmatics of indirect topics, and, to a smaller degree, direct non-subject, i.e. object, topics.

Before taking up the description of the construction, I should like to emphasize three points of some theoretical and practical relevance:

(1) In my system, focus domain is a syntactic, i.e. conventionalized, construction, but it is not necessarily coextensive with any kind of constituent. In other words, I do not claim that verb and subject form a constituent in vS construction (contra Lambrecht; cp. 2.6.2.).

(2) My indirect topics, expressed overtly or merely construed pragmatically, have nothing to do with the Davidsonian event argument, often stipulated for ‘thetic’ utterances (cp. 2.5.3. and 2.6.3.). Indirect topics are placed on the level of information processing, i.e. of pragmatic inference, and are in no way anchored in syntax.

(3) The broad/narrow focus distinction within a focus domain, which pervades the whole system, is also merely a matter of pragmatic inference, i.e. it is not syntactically or intonationally encoded. In spite of this, in the presentation of the statistics, the syntax and the semantics of the construction, I shall occasionally pay attention to the broad/narrow distinction, because speakers under certain conditions tend to use slightly different elements of the construction when they intend to utter a statement with a broad or with a narrow focus. Of course, this particularly holds for the description of the discourse functions of vS.

11.1. Statistics

The construction labeled vS in this study is extremely well attested in my corpus. Actually, it is by far the most frequent type of verb-subject order in all three languages, with only slight crosslinguistic variation. The relevant data are given in Tables (6-4) and (6-5), the former giving the absolute numbers for original texts and the translation of Jack London’s *White Fang*, the latter the share of vS clauses within the class of verb-subject clauses, within the class of clauses with overt subjects (SV/VS), and the percentage within the whole corpus (Σ):

(11-5) *vS clauses in the corpus: absolute values*

	original texts	Translation	total
Alb.	806	217	1023
MG	865	263	1128
SC	1035	195	1230

(11-6) *vS clauses in the corpus: percentages*

	vS vs. VS	vS vs. SV/VS	vS vs. Σ
Alb.	1023 vs. 2225 = 46.0%	1023 vs. 7979 = 12.8%	1023 vs. 15813 = 6.5%
MG	1128 vs. 2673 = 42.2%	1128 vs. 6641 = 17.0%	1128 vs. 14463 = 7.8%
SC	1230 vs. 2065 = 59.5%	1230 vs. 8458 = 14.5%	1230 vs. 14176 = 8.7%

Unlike the previous two constructions, where it was SC that was numerically the weakest language in the sample, the reverse holds here. However, on a closer inspection, we may observe that the statistical difference between the three languages is not so large as it may seem. The data in the first column of (6-5) show that SC has a significantly higher ratio of vS within the VS group than the other two languages. This is, however, merely a consequence of the fact that Alb. and MG have much more VsX and inversion clauses than SC, so that the ratio of vS within the class of VS clauses in these two languages automatically sinks. The figures in the third column, which represent the ratio of vS within the whole corpus, again show a slightly higher ratio of vS in SC. This has some relevance only in the case of Alb., which, as the other data show (see below), indeed seems to employ vS less frequently than the other two languages. In the case of MG, the lower ratio of vS in the third column is merely a by-product of the much greater number of zero-subject clauses (cp. 5.5.), as becomes clear when the data from the second column are taken into account, where the ratios of vS within the class of clauses with overt subjects are compared. The results show that, when the subject is expressed, MG has somewhat, though not significantly, more vS clauses than the other two languages.

Absolute frequency calculations of the kind performed above obviously do not result in any clear-cut differentiation between the three languages, except for the fact that Alb. seems to consistently display an insignificantly lower ratio of vS than the other two languages. The comparison of the frequency relationships within the sources which make up the corpora basically confirms this picture, and offers an explanation for the slight differences observable:

(11-7) *vS in Alb., MG, and SC sources*

	number	vS vs. VS	VS vs. SV/VS	vS vs. Σ
Alb.	1023	46.0%	12.8%	6.5%
Bishqemi	245	42.5%	15.5%	6.5%
Kadare	35	39.5%	7.0%	5.1%
Koha Dit.	252	46.4%	13.9%	8.4%
Kosovarja	274	47.7%	12.4%	6.6%
translation	217	51.8%	11.4%	5.9%

	number	vS vs. VS	VS vs. SV/VS	vS vs. Σ
MG	1128	42.2%	17.0%	7.8%
Kapandai	263	32.1%	21.1%	8.8%
Matesis	167	41.6%	19.1%	6.7%
Fakinou	67	42.9%	8.9%	4.2%
Elefterotip.	250	52.2%	17.2%	11.4%
Ciao	118	47.5%	16.7%	8.8%
translation	263	51.5%	18.1%	7.5%

	number	vS vs. VS	vS vs. SV/VS	vS vs. Σ
SC	1230	59.5%	14.5%	8.7%
Andrić	252	60.1%	13.8%	7.2%
Kiš	268	55.6%	12.4%	7.3%
Vjesnik	125	60.1%	21.8%	14.3%
Vesti	390	62.8%	19.3%	13.6%
translation	195	57.9%	10.4%	6.0%

Consider first the parallel translations of one text (Jack London's *White Fang*) given in the last rows of the tables: the ratios obtained for the corpus as a whole obtain in the column concerning the share of vS within the class of verb-subject clauses, with the same explanation. In other two columns, the relationships are somewhat different: it is MG that has the highest number of vS, followed by SC and Alb. In the case of SC, the ratio of vS lower than the mean value stems, as far as I can see, from the rather slavish and literal translation of the English text, resulting in the use of subject-verb order even there, where vS would be a more natural solution; in the case of Alb., the difference with respect to the mean value is simply insignificant.

The variation within the languages themselves displays different features in the three languages. In Alb., the percentages are almost the same in all sources (the slight variation being explainable in terms of varying frequencies of other verb-subject constructions), apart from Kadare, a representative of modern, 'urbane' prose, who consistently uses vS less frequently than is the case in other Alb. texts. In MG, the situation is basically the same, with the modern prose writer, Fakinou, using less vS than the mean value. A new moment is the higher percentage of vS in the journalistic style, especially in the daily *Elefterotipia*. In SC, this is even more pronounced: both daily newspapers, *Vesnik* and *Vijesti*, contain significantly more vS clauses than the other sources. In Kiš's postmodern prose, vS is used only insignificantly less often than the mean value.

Modern prose does not use a lot of descriptions, it introduces its major participants as topics, without special introductory devices, avoids dramatic turns, etc. As will become patent in the course of this chapter, it is precisely these contexts that are the typical discourse environment of vS clauses. It is therefore no wonder that in the writings of Kadare, Fakinou and Kiš vS occurs less often than elsewhere. Kiš is the one to have most vS clauses of the three, because he often mimics the pseudo-objective style of a historian, using many passivized vS clauses without an agent phrase.

The texts published in daily and weekly journals are as a rule short, introduce a lot of new discourse referents which occur only once in the text, tend to use passive-like constructions

without an agent phrase in order to look more objective, etc., and the titles of journal articles are often construed so as to present the event conveyed as surprising and sensational. All these facts contribute to the frequency of vS clauses in the journalistic style. It is therefore surprising that the Alb. journals, *Koha Ditore* and *Kosovarja*, do not contain more vS clauses than they do. The two issues I used are only equal to the mean value for Alb., not higher, which is also the explanation why in this language vS is somewhat less well represented than in the other two. I do not have the impression that this has anything to do with the productivity of the construction or with the absence of some conventionalized discourse strategies present in MG and SC, so that I should like to dismiss this slight statistical deviation as due to pure accident.

On the other side of the scale, the two SC journals, *Vesti* and *Vjesnik*, have a disproportionately high percentage of vS clauses. This has two reasons: first, in the ‘lower’ journalistic style, the construction with a verb of existence and a subject denoting a state of affairs (something like *there was fighting* or similar) seems to be exceedingly popular, so that in *Vesti* this lexical constellation is much more frequent than in all other journals in any of the three languages. Second, the articles in *Vjesnik* are all full of descriptions of places and ritual actions, since the whole issue is devoted to the funeral of the Croatian president Tudjman, rituals and place descriptions being one of the typical contents of vS clauses. The high percentage of vS in these two daily newspapers, triggered by very specific content-related issues, is the reason SC has somewhat more vS than the other two languages.

What I am trying to show is that there is no relevant *typological* difference in the frequency of vS between Alb., MG and SC, i.e. that the minor numerical deviations are triggered by the corpus, not by the structure of the languages or by strongly divergent discourse strategies. Unlike VsX and inversion, vS is a fully productive pattern in all three languages. If one language would have to be singled out as employing vS in more contexts than the others, it would probably be MG, contrary to the statistics, but Alb. and SC make up for it by more regularly expressing propositions as vS clauses in some contexts in which MG does it only sporadically. On the whole, statistical differences between the languages are neither typologically relevant nor particularly revealing.

One more thing at the end. It has been repeatedly emphasized that vS is a focus domain construction, and as such subject to two possible pragmatic construals, broad and narrow focus. Even though these two interpretations are not formally marked, and as such partly merely a matter of interpretation which may vary from one hearer to the other, I considered it

useful to make a rough and of necessity partly impressionistic statistics of the frequency of these two readings within the corpora. The results are summarized in the following table:

(11-8) *Broad and narrow focus readings*

	broad focus	narrow focus	Σ
Alb.	681 – 66.5%	342 – 33.5%	1023 – 100%
MG	708 – 62.8%	420 – 37.2%	1128 – 100%
SC	835 – 67.9%	395 – 32.1%	1230 – 100%

Broad focus interpretation seems to be twice as frequent as the narrow one in all three languages, with only slight deviation from this relationship in MG, where the narrow scope of assertion appears insignificantly more often. This is in accordance with the breadth of the scope of discourse functions performed by broad verb-subject foci, in comparison to which narrow subject foci are significantly more restricted. Even if the criteria according to which I performed this counting are partly intuitive, the stability of the two-to-one ratio across languages confirms that my intuitions are not completely wrong.

11.2. Formal properties and information structure

11.2.1. Basic features: Intonation, length, complex predicates and complex clauses

The first feature of vS construction which clearly demarcates it as distinct from the other two is the position of the nuclear accent. Unlike VsX and inversion, where the subject is generally deaccented (save for non-ratified inversion subjects), in vS it is the primary carrier of the sentence stress, quite in accordance with the rules of focus domain formation holding in the Balkan languages (see examples (11-1)–(11-4)). The verb is deaccented, or, if the speaker for any reason considers it necessary (ratification or similar), it carries the secondary stress, which is audibly weaker than the primary accent on the subject. The same holds for other, optional, elements of the construction: no matter what position in the sentence they occupy (in front of the verb, between the verb and the subject, after the subject), they are not accented. However, as far as I could discern, only the preverbal optional elements may carry the secondary, i.e. ratifying, accent. The full intonational pattern thus looks as follows:

(11-9) *Intonational pattern of vS*

[rest] [verb] [rest] [SUBJECT] [rest]
 [rest] [VERB] [rest] [SUBJECT] [rest]
 [REST] [verb] [rest] [SUBJECT] [rest]

There is one extremely rare exception to this rule: in some forty sentences from my corpus, the element behind the subject carries an accent which is perceptually *not* weaker than the accent on the subject, judged by the pronunciation of my informants. I shall attempt to interpret this rare phenomenon in the course of this section. Suffice it for now to state the fact:

(11-10) *An exceptional intonational pattern of vS*

[rest] [verb] [rest] [SUBJECT] [REST]

Although the shorthand vS may suggest that the clauses in question are generally short, this is a wrong impression, at least as far as my corpus is concerned. Most of the clauses I sifted through contain at least one additional element, most frequently in front of the verb (from 50% to 80%), or between V and S (10% to 30%). Clitic pronouns, which are subject to independent, strictly grammaticalized rules in Alb., MG and SC (cp. 5.1.), occur in some 10% to 20% of clauses. Clause elements after the subject occur relatively rarely (less than 10% in all three languages). Sentences containing only a verb and a subject are thus not particularly statistically prominent, amounting to approximately 20% in my Alb., MG and SC corpora. As illustrated by (11-1) – (11-4), the prototypical vS clause may thus be said to have the following form (# marks clause boundary):

(11-11) *The structure of the prototypical vS clause in the corpus*

[rest] [verb] ([rest]) [SUBJECT] # (+/- clitic)

No matter how convincing my statistics at this point may be, one should be cautious to proclaim this sentence scheme for the absolutely prevalent form of vS clauses in the three languages under consideration. My figures are drawn from a corpus which consists solely of written texts, and the pattern given in (11-11) is indeed the prototypical form of vS clauses in the written language. If the data from the spoken language were taken into consideration, I am pretty sure that the number of preverbal elements would be significantly lower, and the number of clauses containing clitics or consisting only of V and S significantly higher, due, among other things, to the abundantly used ellipses, deictic space and time determination, frequent use of first and second person direct and indirect objects, etc. In other words, the written language tends to use vS in the clauses with a preverbal or a directly postverbal element, with an occasional occurrence of clitics. My impression (which I cannot prove statistically) is that in the spoken language vS clauses are shorter, often containing only a verb and a subject, and with a much more extensive use of clitic pronouns.

The third feature which characterizes all instances of vS and represents a construction-specific property is the position of the subject with respect to auxiliaries and the copula and with respect to matrix verbs in complex sentences. As has been demonstrated in 6.3.2.3.,

8.1.1.3. and 8.2.2.4., in VsX and inversion, the subject is either preferably inserted between the auxiliary/copula and the participle/predicative or between the matrix verb and the subordinate clause (most often in SC), or is at least allowed to occupy this position under certain conditions (Alb. and MG). In vS clauses, this is never the case: in all three languages, the subject always follows the whole complex predicate form, and is always placed only after the verb of the subordinate clause, i.e. after both the matrix and the subordinate verb. Here are some examples of complex predicate forms (the SC example is chosen so as to contain a non-clitic form of the copula):

(11-12) MG (Ciao, p. 5)

Poli mejali ine ke i pisina tu epixirimatia Vangeli MITILINEU.

very big is also the swimming-pool of-the businessman V. M.

‘Also very big is the swimming pool of the businessman Vangelis Mityleneos.

(11-13) SC (Vjesnik, p. 3)

Bili su nazočni mnogi strani NOVINARI.

were AUX:CLIT present many foreign journalists

‘Many foreign journalists were there.’

(11-14) Alb. (Koha Ditore, p. 4)

Janë intesifikuar ANGAZHIMET që të kthehen ... gjashtë mijë refugjatë...

are intensified engagements to SUBJ:CLIT return six thousand refugees

‘The efforts have been intensified to enable six thousand refugees to return to their homes ...’

The variants with the subject inserted between AUX/COP and PTCP/PRED are not attested in my corpus, and seem to be almost unacceptable for the native speakers: *ine i PISINA ... poli mejali* (‘is the swimming-pool big’), *bili su ... NOVINARI nazočni* (‘were ... journalists present’), etc. are judged odd by my informants and myself. Note that the relative order of the predicative and the copula is not fixed (compare (11-12) and (11-13)). I will have more to say on this later.

The same rule holds for complex sentences, as witnessed by the following example:

(11-15) Eng: *At once began to rise the cries that were fiercely sad ...* (London, p. 10)

Alb: *Dhe menjëherë filloi t’ dëgjohej një ulërimë e EGËR, e trishtueshme.* (p. 10)

and at-once began to be-heard a howling the wild the sad

MG: *Amesos arxisan na andixun i KRAVJES, endona θlimenes ...* (p. 15)

at-once began to resound the cries strongly sad

SC: *Odmah se stadoše razlijegati GLASOVI, koji bjehu divlje tužni ...* (p. 13)

at-once REFL started resound voices which were wildly sad

As in the case of complex predicate forms, the variants with the subject inserted between the matrix verb and the subordinate clause are odd: *filloi një ULËRIMË të dëgjohej, arxisan i KRAVJES na andixun*, etc. ('began a/the cries to resound') are not normal sentences in Alb., MG and SC.¹ The pattern of vS construction in the context of complex verb forms and in complex sentences is thus:

(11-16) [**aux/cop**] [**ptcp/pred**] [**SUBJECT**] / [**ptcp/pred**] [**aux/cop**] [**SUBJECT**]
[**matrix verb**] [**subordinator+verb**] [**SUBJECT**]

The three features described in this section pervade the whole system of vS clauses, both with broad and narrow focus construal. Let us now look at the behavior of additional elements apart from the verb and the subject. Since their distribution is for the greater part dependent on the informational status, I shall separately describe the topical expressions and those which belong to the verb-subject focus domain.

11.2.2. Topical elements

11.2.2.1. Clitics

Clitic, i.e. unaccented, pronouns, are, as Lambrecht (1994: 172ff.) rightfully observes, one of the most common topic expression types if it is the ratified topicality that is to be expressed (the other two being, at least in the Balkans, the zero strategy and the postverbal position, as in VsX). Clitics used in vS clauses thus encode direct, if objects, or indirect, if free datives, ratified topics. Consider the following examples, with direct or indirect objects, or with both:

(11-17) SC (Kiš, p. 29)

Leže u vlažnu travu, licem prema zemlji. Budi ga štapski KURIR.

he-lies-down in wet grass face to earth wakes him:CLIT of-headquarters courier

'He lies down in the wet grass, face towards the earth. A courier from the headquarters awakes him.'

¹ At best, the forms with the subject between the matrix verb and the subordinate clause are understood as instances of corrective narrow foci *in situ*, i.e. not as instances of vS construction.

(11-18) Alb. (Kosovarja, p. 8/50)

Më gjetën PRINDËRIT e mi.

me:CLIT found parents-the the my

‘[I hid from everyone, wanting to be alone.] My parents found me.’

(11-19) MG (Kapandai, p. 37)

Ton ixan fëri KALOJERI ap’ ta Jerosolima,

it:CLIT have brought monks from the Jerusalem

... *ke tu ton estelne, ap’ ayapi, i MANA tis kopelas ...*

and to-him:CLIT it:CLIT from love the mother of-the girl

‘[a ring] It was brought from Jerusalem by monks ... and it was sent to him by the mother of the girl, as a token of affection...’

The examples illustrate two things. First, and more banal, they confirm what has been said about the clitics in Alb., MG and SC in 5.1.: they are placed in the Wackernagel position in SC, and directly before the verb in Alb. and MG. Second, they clearly show the kind of ratified topicality conveyed by clitic pronouns: in all the examples, we are dealing with discourse referents of immediate interest in the discourse, either because of contextual givenness or because of inherent topicality (speech act participants), which are not subjects (otherwise they would not be expressed at all). The following sentences illustrate what I called free datives:

(11-20) Alb. (Kosovarja, p. 11)

Por në këtë minutë, i dridhet ZËRI.

but in that minute to-him:CLIT trembles voice-the

‘[The old man spoke in a very determinate fashion.] But at that moment, his voice started to tremble.’

(11-21) MG (Matesis, p. 59)

Ke tus ixe vji to kalo tus ONOMA ðja pandos ...

and to-them:CLIT has went-out the good their name through all

‘[There was a dance hall in the town; only few women have ever went there.] And these women lost their reputation for good.’

(11-22) SC (Vjesnik, p. 5)

Odlazi nam VELIKI čovjek.

goes-away to-us:CLIT great man

‘[The president is dead.] We have lost a great man.’

Free datives encode participants with a semantically undetermined relationship to the predicate: they may be interpreted as possessive expressions, as in 0, as malefactive or benefactive, as in (11-22) and 0, etc. Since the relevance relationship is not explicitly marked, but has to be pragmatically construed, I treat these topics as indirect. It should be noted that MG uses significantly less free datives than SC, and especially Alb., where they are ubiquitous, and generally tends to employ possessive markers within the noun phrase instead. Thus, the MG sentence *rajise i karđia tis* ('cracked the heart her', Kapandai, p. 30), 'her heart broke', would be expressed with a free dative in Alb. and SC, *iu dogj zemra* ('to-her-PASS:CLIT burned-down heart-the') in the former, *prepuklo joj je srce* ('broke to-her:CLIT AUX heart') in the latter¹.

In Alb. and MG, but never in SC, clitics often appear as the second instantiation of the referent expressed lexically in the same clause, in a construction known as clitic doubling (cp. 5.1.). It is nowadays generally agreed that the appearance of clitic doubling has something to do with topicality (see e.g. Kallulli 2001, and Alexiadou and Anagnostopoulou 2000 for an overview of the generative literature on the topic), but the precise conditions of its occurrence and non-occurrence are not entirely clear. As the clarification of this issue is beyond the scope of this study, I can only state that doubling clitics do occur in my corpus of vS clauses, although with a rather low frequency – only 29 examples in MG and 39 in Alb.

As indicated in 11.2.1., clitics appear as topic expressions in roughly 10-20% of all vS clauses, this relatively low ratio being probably in part due to the nature of the corpus. Here is a more precise statistics:

	clitics	doubling clitics
Alb.	186 – 18.2%	39 – 3.8%
MG	122 – 10.8%	29 – 2.6%
SC	149 – 12.1%	/

The difference between ratios is primarily triggered by the difference in the frequency of free datives, which highest in Alb. and lowest in MG.

¹ Free datives have been a very popular theme in linguistics for quite a while; for a general overview, see e.g. Kendall 1980; for an attempt at interpretation (based on Polish material), Siewierska 1990. The use of free datives in the Balkan languages is described in detail in Buchholz and Fiedler (1987: 217ff., 447ff.), Hubbard (1985: 105ff.), Holton et al. (1997: 267ff.), Mackridge (1985: 61ff.), etc.

11.2.2.2. Preverbal topic expressions

Some 50-80% of all vS clauses in my corpus have at least one topical expression in front of the verb. I have divided these expressions into five classes: objects, comprising direct, indirect and prepositional objects; free datives; setting adverbials, including various kinds of expressions denoting space and time, cause and similar phenomena; predicatives; and manner adverbs. Consider first the distribution and the overall frequency of these preverbal expressions in Alb., MG and SC:

(11-23) *Preverbal topic expressions in vS clauses*

	object	free dative	setting adv.	predicative	manner adv.	Σ
Alb.	168 – 16.4%	32 – 3.1%	432 – 42.3%	16 – 1.6%	1 – 0.1%	632 – 63.5%
MG	229 – 20.3%	9 – 0.8%	309 – 27.4%	21 – 1.9%	2 – 0.2%	570 – 50.6%
SC	239 – 19.4%	27 – 2.2%	701 – 57.0%	12 – 1.0%	3 – 0.2%	982 – 79.8%

Three things may be concluded on the basis of this table. First, SC has much more preverbal topic expressions than the other two languages; MG trails far behind. Second, the most frequent type of a preverbal topic expression are setting adverbials, invariably in all three languages, although MG is relatively weak even here. Third, predicatives and manner adverbials only very rarely play the role of a preverbal topic expression.

As indicated in 11.2.1., preverbal topic expressions are the only ones which may, and often do, receive the secondary, ratifying, accent. This is in accordance with the sentence template proposed for the Balkan languages in Chapters 5., 6., and 10., according to which the preverbal, often clause-initial, position is the slot reserved for non-ratified topics. As we shall see, some of these expressions even have to carry a specific type of ratifying accent. Let us now consider preverbal topic expressions in some detail.

11.2.2.2.1. Objects as preverbal topic expressions

Direct, indirect and prepositional objects are lumped together in one group on the basis of their syntactic tie to the verb, being the only type of topic expressions used with vS selected by the verb, and on the basis of their informational status, being the only type of topic expressions used with vS which encode direct topics. As Table (11-23) shows, expressions of this kind are most frequent in MG, somewhat less in SC and especially in Alb. In MG and Alb. they may be doubled by a clitic, but not necessarily so (see 11.2.2.1.); they may be expressed lexically or pronominally. Here are some examples:

(11-24) Alb. (Bishqemi, p. 53)

Fjalën ma rrëmbeu nga goja PARTINI.

word-the to-me-it:CLIT grabbed from mouth-the P.

‘Partin said precisely what I wanted to say.’

(11-25) MG (Ciao, p. 5)

Simasia exi mono i AγAPI.

meaning has only the love

‘Only LOVE counts.’

(11-26) SC (Kiš, p. 99)

... *tome je doprinela TUBERKULOZA.*

to-it AUX:CLIT contributed tuberculosis

‘[He almost ruined his family. Some say that] TUBERCULOSIS had its share in this.’

11.2.2.2.2. Free datives

Free datives of the kind described in the section on clitics are rarely expressed as full noun phrases or as orthotonic pronouns, i.e. so as to be able to occupy the preverbal slot: the fact that they denote some kind of intimate involvement of their referents in the state of affairs makes them especially well-suited for ratified topics, which are regularly encoded as clitics. The few preverbal instances I have conform to the pattern observed with the clitics: Alb. and SC use free datives significantly more often than MG. This picture is nicely illustrated by the following example:

(11-27) Eng: ... *while Cherokee’s wounds increased.* (London, p. 135)

Alb: ... *kurse Çerokit i shtoheshin plagët në trup.* (p. 132)

while to-Cherokee to-him:CLIT were-growing wounds in body

MG: ... *eno pliθenan i plijes tu Tseroki.* (p. 164)

while multiplied the wounds of-the Cherokee

SC: ... *dok su se Čerokiju uvećavale rane.* (p. 151)

while AUX REFL to-Cherokee grew wounds

11.2.2.2.3. Setting adverbials

The most frequent preverbal topic expressions in all three languages are setting adverbials, appearing, in one way or the other, in almost 60% of all vS clauses in SC, and somewhat less frequently in Alb. and MG. Semantically, it is a rather heterogeneous class: the prototypical members are expressions of space and time and various metaphoric extensions thereof, but

other kinds of scene-setting modifiers also occur. There are three types of phrases playing the role of setting adverbials: most common are prepositional phrases, then various deictically bound space and time adverbials, like *then, there, later*, etc., and finally embedded clauses, temporal, causal, or conditional. All of them are illustrated in the following examples:

(11-28) Alb. (Bishqemi, p. 6)

Në qiell filluan të ndizen YJET.

in sky began to shine stars-the

‘In the sky, STARS started to shine.’

(11-29) MG (Fakinou, p. 27)

Eki tin afise to TAKSI.

there her left the taxi

‘The TAXI brought her there.’

(11-30) SC (Vesti, p. 7)

Dok su izlazili iz autobusa, iznenada je naišlo vozilo.

while AUX went-out from bus suddenly AUX came-across vehicle

‘While they were getting off the bus, a VEHICLE suddenly appeared.’

Unlike other preverbal topic expressions, which are relatively rarely accumulated in clusters in a single clause (even when they are, then usually only up to two per clause), setting adverbials seem to possess an almost unlimited capability to multiply. Here is a particularly impressive example, with five setting expressions (*nakon svih, na Mirogoju, toga dana, sa Predsjednikom, još satima*):

(11-31) SC (Vjesnik, p. 4)

Nakon svih, na Mirogoju su toga dana sa Predsjednikom još satima ostali

after all on M. AUX that day with president more hours stayed

mnogobrojni gradjani.

numerous citizens

‘Longer than the others, many citizens stayed at Mirogoj (a graveyard) with the President that day, for many hours.’

11.2.2.2.4. Predicatives and manner adverbs

In a small fraction of instances it is nonreferential expressions that are placed in the preverbal topic position. They are without exceptions accented with a secondary stress, or intonation peak, which has a specific fall-rise contour, presumably L*+H. Here is an example:

(11-32) SC (Andrić, p. 32)

Ko ovde dodje, taj je kriv... Pustio sam ih dosta ..., da.

who here comes he is guilty I-released AUX them a-lot yes

Ali kriv je bio svaki.

but guilty AUX was every.

‘Whoever comes here, he must be guilty... I did release many of them..., it is true.

But every single one of them was guilty. (“Guilty were they all”).’

(11-33) Alb. (Kosovarja, p. 16)

– *Merreni edhe me politikë? – Shqiptar jam edhe unë.*

you-deal also with politics Albanian am also I

– ‘Are you interested in politics, too?’ – ‘I am an Albanian, too.’

In (11-32), the speaker relationally presupposes the propositional function ‘be guilty’ and ascribes a specific type of contrast to the denotatum of ‘guilty’, which I in an earlier paper (Matić 2003) called exclusive contrast. The same holds for (11-33), where the propositional function ‘be an Albanian’ is presupposed and the denotatum of ‘Albanian’ contrastively topicalized, and used as an answer to a question about politics on the basis of a specific world knowledge (according to which being an Albanian is the same as being interested in politics). Exclusive contrast is, to put it briefly, characterized by a generation of contrastive sets and by the exclusion of all but the expressed alternative¹. This topic type is used only with the narrow focus construal, a fact to which I shall return later. Needless to say, nonreferential expressions in topic positions encode indirect topics, since both the entity status and the relevance relationship to the predicate have to be construed by the hearer.

As example (11-32) shows, when a predicative is topicalized, it stands in front of the copula. This is the explanation for the phenomenon observed in 11.2.1., namely that the order of the predicative and the copula standing in front of the subject varies between [pred][cop] and [cop][pred]: when the speaker uses the predicative as a topic expression, it is positioned before the copula; when this is not the case, i.e. when the whole verbal complex comprising the copula and the predicative is within the focus domain, the order is [cop][pred].

¹ For a more detailed analysis of exclusive contrastive topics, see Molnár (1998), Buring (1999) and Matić (2003).

11.2.2.3. Topic expressions between the verb and the subject

In Chapter 6., it was demonstrated that there is more than one topic position in the languages of the Balkans, at least one of which is postverbal. Whereas in 6. it was the postverbal topical subjects that were of interest, in this section the postverbal direct and indirect non-subject topic expressions are dealt with.

As with topical subjects, the non-subject topic expressions occurring after the verb encode ratified topics. One corollary to this is the fact that they never carry a ratification accent, i.e. the secondary stress. The distribution of the topic expressions placed between V and S in vS clauses looks as follows:

(11-34) *Topic expressions placed between the verb and the subject in vS clauses*

	object	free dative	setting adv.	Σ
Alb.	39 – 3.8%	8 – 0.8%	75 – 7.3%	122 – 11.9%
MG	84 – 7.4%	1 – 0.1%	306 – 27.1%	391 – 34.6%
SC	3 – 0.2%	1 – 0.1%	47 – 3.8%	51 – 4.1%

And again, as in the case of preverbal topic expressions, a number of points is observable already on the level of bare statistics. In contrast to preverbal topic expressions, SC has only a minimal number of the postverbal ones; exactly the opposite holds for MG, while Alb. is somewhere in between. Setting adverbials are again the most frequent type in all three languages. There are no topical predicatives and adverbs in the position between the verb and the subject, which is understandable in view of the fact that they invariably encode exclusive contrastive topics, which are by their very nature non-ratified, therefore obligatorily preverbal.

11.2.2.3.1. Objects and free datives¹

Like preverbal topical objects, the postverbal ones may be pronominal and lexical. In Alb. and MG, if they are direct or indirect objects, they seem to be more regularly doubled by a clitic than is the case with the preverbal ones. Here are some examples:

(11-35) Alb. (Koha Ditore, p. 7)

*E hapi **tubimin** z. Fatik LILA*

it:CLIT opened meeting-the Mr. F. L.

[A description of an assembly] The meeting was opened by Mr. Fatik Lila.'

¹ Since the number of free datives is extremely small, they are dealt with together with objects.

(11-36) MG (Elefterotipia, p. 9)

Iđios ti ipothalpi ti miθolojia ton narkotikon i latria tis PROTJAS.

especially it:CLIT supports the mythology of-the narcotics the worship of primacy
[There are many reasons why so many sportsmen believe that it is not harmful to
take doping.] The myth of the (harmless) narcotics supports it (this opinion)
especially strongly.

(11-37) SC (Vjesnik, p. 5)

Naglasili su to u svojoj prigodnoj riječi V. Marković... i M. BLAŽEVIĆ ...

emphasized AUX that in their appropriate word V.M. and M.B.

[The late president was a great sports fan.] ‘This was emphasized in the speeches
prepared on this occasion by V. Marković ... and M. Blažević...’

There is an important difference between the Alb. and MG examples on one hand, and the SC one on the other: the former represent normal, colloquial sentences in their languages; the latter sounds archaic. This is partially in conformance with one of the conclusions of Chapter 6, where I claim that the postverbal topic position in Alb. and SC is dying out. In the light of the behavior of non-subject topic expressions, this claim has to be slightly modified: the postverbal topic position is dying out in SC with respect to all kinds of topics; in Alb., it is only the subject topics that are gradually ousted from this position. MG seems to use the postverbal non-subject topic expressions with the same grade of productivity as those which play the role of the subject.

11.2.2.3.2. Setting adverbials

Different kinds of setting adverbials are, as indicated, the most frequent topic expressions occurring between the subject and the verb. In Alb. and MG they seem to represent quite a common position for ratified space and time topic expressions:

(11-38) Alb. (Koha Ditore, p. 18)

U zgjodh në këtë mbledhje edhe SEKRETARI.

PASS elected in that assembly also secretary-the

‘At that assembly, the secretary (of the club) was also elected.’

(11-39) MG (Kapandai, p. 46)

Stekotan mesa stis skies, akiniti, mia ARKUĐA...

stood among in-the shadows motionless a she-bear

‘Among the shadows, motionless, stood a she-bear...’

In SC, even this kind of expression is somewhat restricted: it is either deictically bound adverbials of time and space that are found postverbally with some regularity, as in (11-40), or prepositional phrases, usually denoting time, when the preverbal position is already filled with another setting adverbial, as in (11-41). Note that the latter is only an optional rhetorical device (cp. (11-31), where five preverbal setting adverbials occur).

(11-40) SC (Kiš, p. 9)

*Odjednom se stvori **tu** reb Mendel.*

suddenly REFL creates there reb M.

‘Suddenly, reb Mendel appeared there out of nowhere.’

(11-41) SC (Vesti, p. 18)

*U galeriji "Kosovka" održana je **pre neko več**e humanitarna aukcija SLIKA.*

in gallery K. held AUX before some evening humanitarian auction of-pictures

‘In the gallery “Kosovka” a humanitarian picture auction was held a couple of evenings ago.’

11.2.2.4. Topic expressions after the subject

As noted in Chapter 6, there is, apart from the postverbal topic position, yet another landing site for ratified topic expressions, the one on the right edge of the clause. Although quite productive in the Alb. and MG inversion and VsX construction, this position is rather weakly represented in vS clauses, both in terms of frequency and of syntactic roles of the topic expressions. Consider the following table:

(11-42) *Topic expressions after the subject in vS clauses*

	object	setting adv.	Σ
Alb.	1 – 0.1%	52 – 5.1%	53 – 5.2%
MG	18 – 1.6%	79 – 7.0%	97 – 8.6%
SC	0 – 0.0%	22 – 1.8%	22 – 1.8%

Not surprisingly, SC, in which the clause-final topic position is almost non-existent, has the weakest score, MG the strongest. Topical objects seem to be almost completely excluded from the right periphery of the clause in Alb. and SC, and are rather infrequent in MG. Setting adverbials are thus the only expressions appearing in this position with some regularity. They almost always denote space or cause (more than 90% in all three languages):

(11-43) Alb. (Bishqemi, p. 101)

*Nxituam edhe unë me Partinin **për ta ndihmuar Retunin.***

hurried also I with Partin-the for the-him:CLIT helping Retun-the

‘Me and PARTIN also hurried in order to help Retun.’

(11-44) MG (Fakinou, p. 15)

θa erθi o askimos vasiljas Psari apo kato...

will come the ugly king Fish from down

‘The ugly king Fish will come from the depths...’

(11-45) SC (Andrić, p. 109)

Zbog nje i njene lepote palo je za pola sata SEDAM mrtvih glava oko njene kuće.

for her and her beauty fell AUX in half hour seven dead heads around her house

‘Because of her and her beauty, in half an hour seven men were killed around her house.’

I was not able to find any informational difference between postverbal and clause-final topic expressions; the heaviness principle (cp. 5.2.) may play a certain role, as in the Alb. example, but not necessarily, as the other two clauses show. The other possible factor is the number of setting expressions: in the SC example, for instance, both the preverbal and the postverbal positions are already filled, so that it simply sounds stylistically better, more polished, to place the third setting adverbial at the end of the clause.

Many speakers and writers seem to be uncertain whether to treat these clause-final expressions as integrated into the clause or as appositions, as witnessed by the frequently occurring pauses in the spoken language, and the frequent use of a comma in written texts. Sentence (11-43) could thus, as confirmed by my informants, be written (and pronounced) as *nxituam edhe ne..., për ta ndihmuar...*, (*We also hurried, in order to help*) without a difference in meaning.

11.2.2.5. Clauses without clausal topic expressions

It is possible, indeed quite common, for a vS clause not to have an overt topic expression at the clause level at all. Although in the written language the use of such clauses seems to be more restricted than in the casual speech, the ratio of vS clauses without topic expressions in my corpus is not negligible: 236, i.e. 23.1% in Alb., 308, i.e. 27.3% in MG, and 214, i.e. 17.4% in SC.

A number of distinctions may be made within this class. First, the presence vs. absence of a zero anaphoric/deictic element: In some cases, the lack of an overt topic expression is simply due to the fact that the speaker has chosen to refer to a previously mentioned or situationally given referent by not expressing it at all. Since in vS the topics are non-subjects, and Alb., MG and SC generally use clitics in order to refer to previously mentioned and

situationally given referents when these are not subjects, this is a relatively rare phenomenon in the spoken language, and virtually non-existent in the written variants; consequently, I have not a single example in my corpus. Here is therefore an example from a SC conversation I witnessed a couple of years ago. The situation is as follows: E. offers Ž. a cigarette; he refuses; she says, in order to persuade him that it is indeed a very fine cigarette:

(11-46) *Uzmi! Doneo CIGAN.*

take brought gypsy

‘Take it! A/the GIPSY brought (it)’ (~‘A/the gypsy gave it to me’).

The strategy seems to have some prominence only in SC. My Alb. and MG informants invariably preferred the variants with clitic pronouns.

All other clauses with no overt topic expressions, i.e. all the ‘topicless’ clauses from my corpus, are ‘really’ void of clausal topic expressions. However, a number of them has indirect topic expressions not on the level of the clause, but on the phrasal level, usually as possessives. Lambrecht’s (1994) famous example *My CAR broke down* is a case in point: As I have tried to show in 4.3.3., ‘my’ is an instruction to the hearer to construe the topical referent ‘I’ and assess the assertion in relation to it. In my Alb., MG and SC corpora, the subjects of clauses without overt topic expressions are possessed in some 25-35% of cases (25.6% in Alb., 36.1% in MG, 27.2% in SC). Here is a MG example:

(11-47) MG (Ciao, p. 32)

Jelai to XILOS tis.

laughs the lip her

‘She is smiling (and trying to hide it).’ (=‘Her lip is laughing.’)

It is possible that the higher ratio of possessed subjects in MG has to do with the rarity of free datives in this language. Sentence (11-47) would be expressed as *Smeje joj se brk* (‘laugh to-her:CLIT the moustashes’) in SC, without a possessive pronoun, but with a clitic free dative, i.e. with a clausal level topic expression.

There are also other ways of indicating the identity of the indirect topic, some of which will be mentioned in the course of this chapter. In some cases, however, it is very difficult to imagine what kind of direct or indirect topic the speaker could have in mind, as in the following clause:

(11-48) Alb. (Bishqemi, p. 5)

Kaluan disa SEKONDA.

passed-by some seconds

[A boy is in a cave; suddenly, he hears a terrible voice.] ‘A couple of seconds passed. [Nothing happened.]’

Whether sentences of this kind are truly topicless, i.e. whether they really do not even evoke a pragmatically construed topic referent, is a question I shall address in Section 11.6.

11.2.2.6. Topical elements: A summary

Let us now take a look at the general picture emerging from the facts enumerated above. Table (11-49) gives the absolute numbers and the percentages of the five classes of topic referents identified, irrespective of their position in the clause:

(11-49) *Types of topic expressions in vS*

	object	free dative	setting adv.	predicative	manner adv.
Alb.	291 – 28.4%	143 – 14.0%	559 – 54.6%	16 – 1.6%	1 – 0.1%
MG	435 – 38.6%	28 – 2.5%	694 – 61.5%	21 – 1.9%	2 – 0.2%
SC	304 – 24.7%	115 – 9.3%	770 – 62.6%	12 – 1.0%	3 – 0.2%

The total number is not given, since the sum is much higher than 100%: a clause may contain up to five different topic expressions. However, when a clause contains more than one topic expression of the same kind (say, three setting adverbials), it was counted as one. Now, some language-specific preferences are clear: Alb. uses more free datives and less setting adverbials than the other two languages; in MG, topical objects and setting adverbials are rather frequent, free datives rather infrequent; SC is characterized by a high number of setting adverbials only. Predicatives and manner adverbials rarely play the role of a topic expression in all languages; Setting adverbials are universally the most common type. An explanation for these tendencies will be offered in the section devoted to discourse functions of vS.

Table (11-50) contains data on the positional tendencies of topic expressions, irrespective of the class they belong to. The principles applied to the previous table hold here as well.

(11-50) *Position of topic expressions in vS clauses*

	clitic	preverbal	postverbal	clause-final	no top. express.
Alb.	186 – 18.2%	632 – 63.5%	122 – 11.9%	53 – 5.2%	236 – 23.1%
MG	122 – 10.8%	570 – 50.6%	391 – 34.6%	97 – 8.6%	308 – 27.3%
SC	149 – 12.1%	982 – 79.8%	51 – 4.1%	22 – 1.8%	214 – 17.4%

Apart from clitics, which follow the positioning rules of their own, two extremes are visible: on one hand, MG, where the preverbal and the two postverbal positions are almost in equilibrium, and on the other, SC, in which the postverbal positions are practically not used at all, with some minor exceptions. Alb. is somewhere in between. This confirms the conclusion about the fate of postverbal topics drawn in Chapter 6 and modified in 11.2.2.3.1. The

postverbal topic slots in SC are in the stadium of disintegration, no matter whether they are filled by subjects, objects, or any kind of indirect topic expressions. In Alb., it is only the subject topics that are regularly preposed, other types of topic expressions being still productively placed postverbally when the conditions are met, although less regularly than in MG. In the latter, both the preverbal and the postverbal topic slots are used without syntactic restrictions, encoding the difference between ratified and non-ratified topics.

Finally, a note on frequency of the clauses without overt clausal topic expressions: SC has some 10% less such clauses than MG, Alb. about 5%. This partly due to the fact that free datives are rarely used in MG, so that indirect topics denoting the interested party are more often encoded as possessives, i.e. on the phrase level. I shall try to show in the course of this chapter that there are also other, discourse-specific, reasons for this discrepancy.

11.2.3. Further elements of the focus domain

The verb and the subject in vS construction form a focus domain. This does not necessarily imply that these are the only two parts of the domain, i.e. that they alone may be under the scope of assertion. Verb-object focus domains, for instance, allow for further assertive modifications: apart from *I played FOOTBALL* with the focus domain [played football], one may say *I played football FIERCELY*, where the focus domain is [played football fiercely], or *I played football in the GARDEN*, with [played football in the garden] as the focus domain.

In natural language (as opposed to the examples construed by linguists), extended focus domains are not particularly frequent, especially those with more than one additional element. This is also true in the case of verb-subject focus domains, probably even more than in the case of their verb-object, verb-adverb, etc, counterparts, as is visible from the following statistics. (Modal adverbs are, due to the specific status they have, counted separately; XvS means ‘additional part of the focus domain in front of the verb’, vXS ‘additional part of the focus domain between the verb and the subject’, etc.)

(11-51) *Further elements of the focus domain in vS*

	modal adverbs	XvS	vXS	vSX	Σ
Alb.	7 – 0.7%	0 – 0.0%	26 – 2.5%	20 – 2.0%	53 – 5.2%
MG	11 – 1.0%	0 – 0.0%	35 – 3.1%	17 – 1.5%	63 – 5.6%
SC	19 – 1.5%	0 – 0.0%	21 – 1.7%	8 – 0.7%	48 – 3.9%

In all three languages, extended verb-subject focus domains are very infrequent, especially in SC. Modal adverbs are rarely used; preverbal elements are excluded due to the grammatical

rule of building focus domains in the Balkan languages, which requires the verb to represent the left border of the domain (cp. Section 4.5.); the elements between the verb and the subject and after the subject are also rather infrequent, with slightly diverging ratios in the three languages. Let us now look at these elements in some more detail.

11.2.3.1. Modal adverbials

Modal adverbs which are not topicalized or narrowly focused, i.e. those which play the role of the verb modifier in the narrower sense of the word, are subject to different positioning rules in Alb. and MG on one hand, and SC on the other (cp. Sections 5.1. and 8.1.2.3.): they are immediately postverbal in the former and immediately preverbal in the latter:

(11-52) Alb. (Bishqemi, p. 9)

*Ndërkohë cërriti **fuqishëm** ZILJA...*
meanwhile struck strongly bell-the
'In the meantime, the BELL loudly rang.'

(11-53) SC (Andrić, p. 55)

*To su **odlučno** zahtevali i MORNARI...*
it AUX resolutely demanded also sailors
'Sailors also resolutely demanded it...'

The Alb. example is an instance of broad focus construal (what is asserted is [strongly rang the bell]), whereas the SC one illustrates the narrow focus construal (the focus domain [resolutely demanded also sailors] is pragmatically construed as [[resolutely demanded]_{Presupposed} [also sailors]_{Focus}]).

The fact that modal adverbs as verb modifiers are not very frequent in any of the three languages does not seem to me to be a construction-specific matter. The overall ratio of modal adverbs in the corpus is not significantly higher than in vS clauses: 1.1% in Alb., 2.1% in MG, 2.1% in SC.

11.2.3.2. Preverbal elements of the focus domain

As already indicated, preverbal elements are *per definitionem* excluded from the focus domain in the modern languages of the Balkans. If a preverbal element is found in a context which points to vS construction, it is as a rule an instance of the fronted narrow focus expression, i.e. an inversion clause, not vS (cp. Section 8.3.2.2., Subsection 5).

11.2.3.3. Elements of the focus domain between the verb and the subject

The most frequent type of additional elements of the focus domain in my corpus are the assertive expressions placed between the verb and the subject. They are mostly objects, direct and indirect, or prepositional phrases.

Most of these expressions stand in the relationship of lexical solidarity with the verb, either as idioms, or as support verb constructions, or as semi-idiomatical verb-object combinations. Lexically bound objects and prepositional phrases are actually the only type attested in Alb., and the most frequent one in MG (29 out of 35) and SC (20 out of 21). Consider the following examples:

(11-54) Alb. (Bishqemi, p. 126)

Dhe në të dyja këto gëzime morëm pjesë edhe NE, anëtarët më të RINJ të ekuipazhit

and in the two that joys took part also we members most the young of-the crew

‘Even WE, the youngest members of the crew, took part in these two celebrations.’

(11-55) MG (Matesis, p. 50)

...imuna koketa ..., j’afto ðen m’afisan se xloro klari i ANDRES...

I-was coquette for that not me left in green branch the men

‘...I was a coquette..., so that MEN never left me in the lurch...’

(11-56) SC (Kiš, p. 141)

... doçi će mu glave čelik i OLOVO.

come will:CLIT him:CLIT of-head steel and lead

‘...steel and LEAD will kill him.’

The elements intervening between the verb and the subject are all semantically closely tied to the verb, as full-fledged idioms (*afino se xloro klari*, ‘to leave on a green branch’ = to leave in the lurch), or at least semi-idiomatical (*merr pjesë*, ‘take part’ = take part, *doçi glave*, ‘come of-head’ = kill). These elements automatically fall under the scope of assertion when the verb is asserted, i.e. they are a part of the focus domain.

The other type of intervening assertive expressions, much less frequent, are the expressions which are not in any relevant way semantically tied to the verb. There are 6 examples of this in MG, one in SC, all of which are objects:

(11-57) MG (Kaθimerini, p. 1)

Metavali ta ðeðomena tis ayoras i “INTERAMERIKAN”.

changes the conditions of-the selling the Interamerican

‘»Interamerican« changes the conditions of purchase.’

(11-58) SC (Kiš, p. 24)

...*tu smisljaju atentate lazni svestenici, pesnici i IZDAJNICI*

there plan assassinations false priests poets and traitors

‘... there, false priests, poets and traitors plan assassinations...’

In all eight examples I have (7 in the corpus + (11-57)), it seems counterintuitive to treat the elements between the verb and the subject as ratified topics (as described in 11.2.2.3.). Sentence (11-57) is a title of an article in a journal, i.e. void of previous context; *ta đeđomena tis ayoras* is both too general (what conditions, what purchase?) and too situation-bound, i.e. specific (the reference is not to any old purchase, but to a specific transaction) to be treated as an inherently relevant issue by the writer of the article. In (11-58), which is a description of a pub in Dublin at the beginning of the 20th century, my feeling is that *atentati* is a part of a complex predicate ‘plan assassinations’ (so that it perhaps belongs to the first group of intervening assertive elements), or, alternatively, it is like *ta đeđomena tis ayoras*, both too general and too specific in order to be inherently topicalizable. In other words, the expressions between the verb and the subject are indeed parts of the focus domain, even without close semantic tie to the verb.

This kind of construction, with semantically non-bound intervening elements, seems to be somewhat rhetorically marked. I find Kiš’s sentence (11-58) difficult to process and stylistically strange, but this may be merely a linguist’s prejudice. My MG informants did not find anything strange with sentences like (11-57). However, they are marked with respect to the discourse contexts in which they are found: all seven examples are journal article titles.

The extended verb-subject focus domains with an additional element between the verb and the subject are triggered either by the close semantic tie between the verb and the additional element, so that they both have to be either asserted or presupposed *together* (with the exception of polarity focus, cp. 6.5.1.), or by specific rhetorical effects. The fact that lexical solidarities are found immediately next to the verb and to the right of it has already been observed in 8.1.1.2.4., with respect to inversion: this is thus not a construction-specific feature of vS, but rather a part of the word order system of the Balkan languages. This systematic nature of the immediately postverbal position is perhaps also a solution to the problem of lexically non-bound elements: with the help of the mechanism of lexical reinterpretation (cp. 8.1.3.), every lexical item placed in the position marked for lexical solidarity gets reinterpreted as forming a quasi-lexical unit with the verb. This explanation would have the advantage of capturing the intuition that *planiraju atentate* in (11-58) looks like a complex predicate to me. If this interpretation is correct, *metavalun ta đeđomena tis ayoras* in (11-57)

would also be a complex predicate: ‘to change the conditions of purchase’ is a kind of action typical for mercantile transaction. This would further explain the air of pragmatic and rhetorical pretense of these sentences, pretense being a general feature of lexical reinterpretations (cp. 4.2.1. for the notion of pretense).¹

Note that the assertive postverbal expressions are formally not distinct from postverbal topic expressions described in 11.2.2.3. The interpretation as assertive or presuppositional is thus a matter of pragmatic construal, on the basis of the clues from the context and the lexical properties of the elements involved.

11.2.3.4. Elements of the focus domain after the subject

In 11.2.1. it was noted that in about 40 clauses in my corpus the subject carrying the nuclear stress is followed by yet another element with an accent which does not seem to be subordinate to that of the subject, i.e. with something like additional nuclear stress. Now, the existence of two nuclear accents is not allowed in the autosegmental model of intonation embraced in this study. I am not sure how to solve this theoretically. Ladd (1996) proposes a metrical model in which the sequence of more equally prominent intonation peaks is treated as a consequence of metrical branching, the latter accent being [+strong] on the higher level than the former. The other possibility would be to postulate the existence of a minimal focus domain, say [verb+subject], [verb+object], etc., which can be expanded to the right by adding further accented elements. This is in line with the findings of Gussenhoven (1983b): The verb-object domain in *John reads a BOOK* may be expanded by a further accented element, as in *John reads a BOOK in the GARDEN*. In the case of verb-subject focus domains, the principle would be absolutely the same, as the following example shows:

(11-59) Alb. (Bishqemi, p. 48)

Në njërin shtrat flinte dikush. E kishte zënë GJUMI me libër HAPUR.

in one bad slept somebody him:CLIT has taken sleep-the with book open

‘Someone was sleeping in one of the beds. He fell asleep with an open book.’

The focus domain [*kishte zënë GJUMI*] (‘has taken sleep-the’) is expanded by a prepositional phrase *me libër HAPUR* (‘with book open’), which carries an additional nuclear stress. Informationally, this could perhaps be interpreted as a gradual addition of elements to the broad assertional scope, with an instruction to the hearer to first assess the propositional

¹ I do not intend to claim that the Balkan languages have a completely productive mechanism of forming complex predicates through loose incorporation on the same level as, e.g., Hungarian (cp. Behrens 1982, É. Kiss 1987, 1994, etc.), but merely that there are clear tendencies, partly integrated into the system, in this direction.

function expressed by the minimal focus domain with respect to the topic referent, and only then the denotatum of the extension. If this hypothesis is correct, sentence (11-59) reads as follows: ‘add the propositional function [was caught by sleep] to your mental representation of the referent [he] first, and then the modification [with an open book] to the ground obtained in this way’.

Like assertional elements between the verb and the subject, the postverbal ones may be divided into two classes. First, the cases in which the subject forms a lexical unity with the verb. In all instances I collected (18 in Alb., 3 in MG, 3 in SC), it is various states of affairs pertaining to body or soul that are thus described, with a body part, a physical state of the body, or emotion as the subject referent, whereas the topic is invariably encoded as a clitic pronoun in the accusative or dative. One example is given in (11-59): *zë* (‘take’) and *gjumë* (‘sleep’) form a quasi-lexical unity meaning ‘fall asleep’ (‘someone is taken by sleep’). Here are some more instances of the same type:

(11-60) Alb. (Bishqemi, p. 4)

Tek eci anës lumit, përtej ... më zënë SYTË një vrimë të FSHEHTË.

while I-went along river-the, on-the-other-side to-me took eyes a hole hidden
‘While I was going along the river, I saw on the other side ... a hidden hole.’

(11-61) MG (Kapandai, p. 29)

Pai ki o navarxos o Kapelo, tu irthe DAMBLAS apo ti SIMFORA.

goes also the captain the C. him:CLIT came heart-failure from the calamity
‘Captain Capello died, too, he got a heart-attack, because of the disaster.’

(11-62) SC (Andrić, p. 112)

Kad to pomislim, udari mi KRV u GLAVU, pa obnevidim

when that I-think, hits me:CLIT blood in head, so I-not-see

‘When I just think of it, I get terribly angry, I can’t see anything (from anger).’

The verb-subject combinations *zënë sytë* (‘take eyes-the’), *erxete o damblas* (‘comes the heart-failure’) and *udari krv* (‘hits blood’) are idiomatic expressions for ‘see’, ‘have a heart-attack’ and ‘get angry’. The fact that Alb. has significantly more such sentences than the other two languages has to do with the frequency with which this type of states of affairs is expressed with this kind of idiomatic expressions in this language (more on this in Section 11.5.2.2.3.).

The other type is the one without a lexical solidarity between the verb and the subject. Alb. has only one such clause, SC five, MG twelve:

(11-63) MG (Matesis, p. 35)

Etsi ke pernusan tris jinekes..., vjenan i ITALI ... sto BALKONI tus

so and passed-by three women went-out the Italians on-the balcony their

‘As soon as three women would pass by..., the Italians would go out on their balcony...’

(11-64) SC (Andrić, p. 18)

Noću se odvode OSUDJENI na izdržavanje kazne ili u PROGONSTVO.

at-night REFL are-taken-away the convicted on serving sentence or in exile

‘At night, the CONVICTS are taken away, to serve their sentence or in exile.’

I cannot see any close semantic tie between the elements of the two above sentences. At least as far as the SC example is concerned, there is, for me, a slight feeling of pragmatic pretense and stylistic oddity (but my judgments may be biased here).

If my interpretation of the clauses with accented postsubjectal elements is correct, then it is no wonder that the close-knit verb-subject combinations so often occur in this construction: it is easier to first add the denotatum of a verb-subject combination which is non-compositional or almost non-compositional, and then add a further modification, than to be compelled to first work out the denotatum of the verb-subject combination. However, as (11-63) and (11-64) show, this is in no way a precondition for the use of extended verb-subject focus domains with postsubjectal accented elements. Let me add at the end that all the clauses with a postsubjectal accented element in my corpus have a broad focus construal.

11.2.3.5. Further elements of the focus domain: A summary

Verb-subject focus domains are, as Table (11-51) shows, as a rule short, consisting only of the minimal elements, the verb and the subject. Only in 4% – 5.5% of all vS clauses is the focus domain extended. Modal adverbs are discourse pragmatically and semantically the least interesting type of extension, and they do not display any kind of construction-specific behavior. The elements between the verb and the subject, which seem to be most frequent in MG¹ may be interpreted as instances of lexical solidarity, sometimes an inherent one, and sometimes brought about through the mechanism of lexical reinterpretation². The elements

¹ The number of instances is so small that the numerical relationships between the languages do not necessarily reflect typological variation, but may be due to pure chance.

² Lexical solidarities seem to influence the possibility of forming extended focus domains even in those languages in which vS construction is much more syntactically restricted than in Alb., MG and SC: Pinto (1997: 22) gives some nice examples for Italian, Culicover and Levine (2001:308-9) for English.

occurring after the verb are more problematic. It is very often the case, especially in Alb., that the verb and the subject form a quasi-lexical unit, usually describing a bodily event or a psychological state, but this need not be the case. My proposal is that we are dealing with a rather minor phenomenon of a gradual increase of the focus domain, but it is only a hypothesis which I am not able to verify at the moment.

11.2.4. Formal properties and information structure: Conclusion

The basic distinctive feature of vS construction are its intonational pattern, with the nuclear accent on the subject, and the ban on inserting the subject between the auxiliary/copula and the participle/predicative or between the matrix verb and the subordinate clause.

vS clauses usually contain, especially in written texts, at least one topic expression, most often in front of the verb, but, especially in MG, also between the verb and the subject. Clitics are also relatively common, in contrast to the topical elements after the subject, which are rare. The basic formal characteristic of the topic expressions in vS is that they are not subjects, but rather objects, free datives, setting adverbials, predicatives and modal adverbs, the last two only in the clauses with a narrow focus construal. It is important to have in mind that there are ways of encoding indirect topics other than as constituents, notably as possessives within the noun phrase.

On the other hand, the possibilities of including other elements into the verb-subject focus domain are rather limited: it is mostly modal adverbs and elements with a close semantic tie to the predicate that are found in this function, though there are enough examples without these two properties.

The sentence template for vS construction may be represented as follows (for the sake of crosslinguistic comparison, some minor issues have been simplified or omitted; the index ‘+/-LS’ marks the presence/absence of a close lexical tie to the verb) :

1. Topical elements:

MG: [NON-RATIFIED TOP.] [verb] [ratified top.] [SUBJECT] [ratified top.]

Alb.: [NON-RATIFIED TOP.] [verb] [ratified top.] [SUBJECT] ([ratified top.])

SC: [NON-RATIFIED TOP./ratified top.] [verb] ([ratified top.]) [SUBJECT]

2. Focal elements:

Alb., MG: [[verb] [modal adverb/X]_{+LS} [SUBJECT] [X]_{+/-LS}]_{Focus Domain}

SC: [[modal adverb] [verb] [X]_{+LS} [SUBJECT] [X]_{+/-LS}]_{Focus Domain}

The main typological difference between the languages is the existence and frequency of postverbal and clause-final non-subject topic expressions, with MG using both positions

productively, Alb. at least the postverbal one, and SC practically neither of them. Consequently, the preverbal topic slot in SC is less pragmatically specified, so that ratified topic expressions may land there too. The positions reserved for the elements of the extended focus domain are basically the same in all three languages, the only relevant difference being the position of modal adverbs.

11.3. Semantic and informational properties of the elements of vS

Much of the discussion on theticity and related issues was based on the semantics of the elements of the construction and on their informational properties (cp. Chapter 2), often with the result that it is the existential or quasi-existential semantics of the verb and the nonagentive character of the subject that is responsible for vS and related constructions, or, alternatively, that it is the newness of the subject and the givenness of the verb, or the newness of both. My findings may at the first sight seem somewhat ambiguous. On one hand, many verbs are indeed existential, etc., many subjects nonagentive, and both are frequently ‘new’. On the other, there are quite a number of instances in which these properties are absent. We are thus dealing with statistical tendencies, not with rules, which are to be explained by another underlying principle (or principles).

My approach to the semantic and informational properties of vS differs from the previous ones in that I shall not confine myself to the verb and the subject. Since direct and indirect topics of vS clauses play an important part in my explanation of vS, various kinds of topic expressions will be taken in account as well.

11.3.1. Semantic and informational properties of topic expressions

In this section, only the semantic and informational properties of *topic expressions* will be scrutinized, not of *topics* (see 4.5.1. for the difference between the two). Recall that I contend that vS clauses often have indirect topics, i.e. those which have to be pragmatically construed by the hearer, on the basis of the lexical material contained in the clause, or merely on the basis of the context. This implies that the topic expression can be semantically (less often informationally) different from the construed topic. For instance, if the topic expression is *then*, as in *Then it started RAINING*, then the construed topic will probably be something like ‘the state of affairs in the current center of attention’. The topic expression is a temporal

notion, the topic itself a state of affairs, two ontologically different things. As I consider it difficult to do a statistical evaluation of pragmatically construed notions, and potentially very arbitrary, I confine myself only to what is visible, i.e. to topic expressions. I have also left out of consideration those indirect topic expressions which are not encoded on the clause, but rather on the phrase level, like possessives (cp. 11.2.2.5).

The number of expressions included in the statistics is considerable: 992 instances in 786 clauses in Alb. (the remaining 236 having no overt topic expressions), 1180 in 820 clauses in MG, and 1204 in 1016 in SC. The first feature to be investigated are the semantic roles of topic expressions. The data are presented in Table (11-65):

(11-65) *Semantic roles of topic expressions in vS*

	Alb.	MG	SC
(1) location/partitive	347 – 34.8%	451 – 38.2%	490 – 40.7%
(2) temporal location	175 – 17.6%	260 – 22.0%	297 – 24.7%
(3) source of perception/emotion	33 – 3.2%	61 – 5.2%	22 – 1.8%
(4) patient/theme	162 – 16.2%	221 – 18.8%	154 – 12.8%
(5) experiencer	92 – 10.0%	107 – 9.1%	73 – 6.0%
(6) interested party	161 – 16.1%	53 – 4.5%	144 – 12.0%
(7) agent/instrumental	5 – 0.5%	4 – 0.3%	9 – 0.7%
(8) nonreferential	17 – 1.7%	23 – 1.9%	15 – 1.3%
Σ	992 – 100%	1180 – 100%	1204 – 100%

Row (1) comprises expressions denoting physical location, like *in the garden*, and different types of ‘symbolic’ location, including partitives, like *among those people*. The expressions like *then, that year, 1982*, etc. are temporal locations (2). Under ‘source of perception or emotion’ (3) I understand expressions like *he impressed me, I know him, I saw him*, etc. The term ‘interested party’ (6) covers such diverse semantic roles as beneficiary, recipient, possessor, and others, often not clearly distinguishable. Nonreferential topics (8) are topicalized predicatives and adverbs (cp. 11.2.2.2.4.). Other terms are self-explanatory.

The first conspicuous feature is the frequency of locationals: taken together, rows (1) and (2), i.e. physical, ‘symbolic’, and temporal location, make up between 50% and 65% of all instances. This is only partly attributable to the fact that this kind of expression most frequently appears in clusters, as witnessed by (11-31); most of the clauses in my corpus have a locational as the sole topic expression. Note that there is some variation among languages: whereas in SC and MG the percentage of locationals ((1)+(2)) is 65.4% and 60.2%, respectively, in Alb. it only amounts to 52.4%.

Rows (3) and (4) contain data on the semantic roles which are usually connected with the lower level of agentivity, humanness, volitionality, etc., in the propositions in which they

occur. MG displays the highest ratio of these expressions ((3)+(4)), 25.0%, tightly followed by Alb. (19.4%), while SC has only 14.6% of the patientlike topic expressions. Note that the term ‘theme’ denotes all kinds of participants whose involuntary movement or position is described.

Experiencers (5) and various semantic roles subsumed under the label ‘interested party’ (6) also show a certain semantic affinity, representing that participant in the state of affairs in whose consciousness the event is reflected. Expressions of this kind are best attested in Alb. (26.1%), which is followed by SC (18.0%) and MG (13.6%). The low ratio of experiencer-like topic participants in MG seems to correlate with the infrequent use of free datives in this language (cp. 11.2.2.1. and 11.2.2.2.).

Agentive and nonreferential topics are extremely weakly represented, the former amounting to approximately 0.5%, the latter to maximally 2% in all three languages.

If finer differences are abstracted away, and the semantic roles defined as locationals, patients and experiencers, the following scales of the frequency of certain topic expression types may be established:

- (11-66) **locationals:** SC > MG > Alb.
patients: MG > Alb. > SC
experiencers: Alb. > SC > MG

In none of the cases is the difference between the three languages large (which is reflected by the absence of the sign ‘>>’), but it is worth noting that it exists: SC uses more topical locationals than the other two, MG more patientlike elements, Alb. more experiencers.

Let us now look at the ontological properties of the topic expressions in vS:

(11-67) *Ontological properties of topic expressions in vS*

	Alb.	MG	SC
(1) thing	87 – 8.7%	95 – 8.1%	80 – 6.7%
(2) thing as location/location	311 – 31.2%	364 – 30.8%	432 – 35.8%
(3) temporal notion	145 – 14.6%	181 – 15.3%	160 – 13.4%
(4) state of affairs/abstraction	156 – 15.7%	216 – 18.3%	298 – 24.6%
(5) emotional/physical state	17 – 1.6%	18 – 1.5%	6 – 0.5%
(6) human [1./2. person]	221 – 22.3% [53 – 5.3%]	209 – 17.7% [39 – 3.3%]	176 – 14.6% [25 – 2.0%]
(7) institution	16 – 1.6%	31 – 2.6%	19 – 1.6%
(8) body part	22 – 2.2%	20 – 1.7%	18 – 1.5%
(9) nonreferential	17 – 1.7%	23 – 1.9%	15 – 1.3%
Σ	992 – 100%	1180 – 100%	1204 – 100%

Like every ontological classification of entities, the one given in (11-67) is in part arbitrary, but it may help the reader get the rough impression on what the assertional bases of vS

clauses are. Only row (2) needs some explanation. The label ‘thing as location’ refers to the fact that most expressions denoting places in the AME languages are ontologically ambiguous between the readings ‘thing’ and ‘location’ (Lyons 1977: 438ff.; Bresnan 1994: 116; more on this later). It is not only the case that *London* is both a ‘thing’ and a ‘location’, but most entities classifiable as things may be interpreted as locations: *table* in ‘I made *a table*’ has a different ontological status than in ‘Put it *on the table*’. Although ‘table’ may retain the same reference in both cases, they seem to represent two different conceptualizations. For this reason, I thought it justified to treat ‘thing’ (1) and ‘thing as location’ (2) separately.

In what follows, only the most conspicuous features will be commented upon. SC seems to use expressions denoting things as topic expressions less often than the other two languages, which is compensated by the more frequent use of expressions denoting states of affairs or abstractions. In Alb., expressions denoting humans are more prominent than in MG and SC. In all three languages, things conceptualized as locations are by far the most prominent topic expression type, followed by expressions denoting humans, temporal notions and states of affairs.

In order to get a more clear picture, one may abstract away the fine conceptual differences depicted above, and lump together the categories (1) and (2) under the label ‘thing’, the categories (3), (4) and (5) under ‘state of affairs’, and (6), (7) and (8) under ‘human’:

(11-68) *Simplified ontology of topic expressions in vS*

	Alb.	MG	SC
thing	398 – 40.1%	459 – 40.8%	512 – 42.5%
state of affairs	318 – 32.1%	415 – 35.2%	464 – 38.5%
human	259 – 26.1%	260 – 22.1%	213 – 17.7%
nonreferential (predic., adv.)	17 – 1.7%	23 – 1.9%	15 – 1.3%
Σ	992 – 100%	1180 – 100%	1204 – 100%

All three languages employ topic expressions denoting things in about 40% of cases; slight differences are observable with respect to states of affairs, with Alb. using them least, SC most, and with expressions denoting humans, where the reverse holds. MG is all three cases somewhere in between. The differences are again not overwhelming, so that sign ‘>>’ is not used.

- (11-69) **thing:** Alb. = MG = SC
state of affairs: SC > MG > Alb.
human: Alb. > MG > SC

The statistical evaluations of the semantic roles and the ontological statuses of the denotata of the topic expressions reveal, apart from minor differences between languages, one important

point. The denotata of the majority of topic expressions in vS are not prototypical topics. In Section 4.3.2. the notion of *statistical topicworthiness* has been introduced, referring to the clear crosslinguistic tendency to use certain types of referents which play certain semantic roles as topics. The basic principle is that the more animate and the more agentive the referent, the greater the probability that it will be chosen as the topic. This tendency has been fully confirmed in this study with respect to VsX construction, where between 65 and 90% of all topic expressions are human or animate (cp. 6.4.1.). Now, the statistical values for the topics of vS construction run contrary to the statistical topicworthiness: agentive referents are almost invisible in the statistics, and human referents, although present, account for approximately one fourth of topics, falling thus far behind such less-topicworthy entities as things and states of affairs.

In Section 4.3.2. yet another notion has been introduced, that of *actual topicworthiness*: it is the actual estimation of the appropriateness of a referent for the topic role relative to the current utterance. Obviously, when using vS clauses, speakers tend to diverge from the statistical topicworthiness, i.e. from the usual choice of the topic referent. This leads me to the first generalization on the semantics of topic expressions in vS: *vS appears to be used in those cases in which there is a mismatch between the statistical and the actual topicworthiness.*

It is extremely difficult to measure actual topicworthiness, since it is dependent on the subjective estimations of the speaker relative to her/his intentions. Some indirect evidence may be gained by exploring informational properties like givenness, the status as the major participant, etc., since given, anaphorically bound, inferable, etc., referents are more likely to be chosen as topics in the actual discourse than those which do not have these properties. It should be kept in mind, however, that this is only an *indirect* evidence: ‘new’, etc., elements may be judged topically relevant as well, if the speaker decides so. Here are the data:

(11-70) *Informational properties of topic expressions in vS*

	Alb.	MG	SC
(1) 1 st and 2 nd person deictic	53 – 5.3%	39 – 3.3%	25 – 2.0%
(2) 3 rd person deictic	45 – 4.4%	21 – 1.8%	24 – 1.9%
(3) deictic locationals	291 – 28.4%	345 – 30.6%	383 – 31.1%
(4) mentioned in the last 1-10 clauses	193 – 18.9%	181 – 16.0%	250 – 20.3%
(5) inferable	211 – 20.6%	321 – 28.5%	313 – 25.4%
(6) major participant	92 – 9.0%	111 – 9.8%	87 – 7.0%

A comparison with the results for topic expressions in VsX (Table (6-34)) reveals the specific features of vS topic expressions. Three features are characteristic: the prominence of deictic elements, the relative rarity of direct previous mentions, and the rarity of major participants.

Taken together, rows (1) – (3) show that deictic elements make up some 35–40% of all topic expressions, which is a significantly higher ratio than in VsX, where the ratio of the deictics amounts to some 10-13%. Previously mentioned topic expressions appear in about 30% of the cases, which is less than in VsX, where they occur in 45-50% of clauses. Finally, major participants of larger stretches of discourse are only exceptionally chosen as topics of vS clauses: only 7-9%, in comparison to 25-35% in VsX. On the basis of these data, the following generalization may be drawn: *The preferred topics of vS clauses are those referents which are of local relevance at one particular point in the discourse, often made relevant by the presence in the physical context.* Referents having a greater textual persistence or inherent relevance are not the typical topics of vS clauses.

In sum: The prototypical topic expressions in vS construction are **untypical topics**. They are more often locationals or patientive than agentive, more often non-human than human, more often locally than generally relevant. The notion of **mismatch between the statistical and the actual topicworthiness** introduced in this section will be one of the keystones of my interpretation of the construction on the whole.

Before taking up the issue of the semantic properties of subjects, a note on the interpretative impact of the topic type on the focus construal. Contrastive topics, and especially nonreferential exclusive contrastive topics (cp. 11.2.2.2.4.), seem to require, or at least greatly facilitate, the narrow focus construal. More on this in Section 11.4.

11.3.2. Semantic and informational properties of the subject

The subjects, or better subject referents, of vS clauses have often been claimed to be obligatorily new in the discourse and preferably nonagentive. Just as with topics, my data only partially confirm these claims.

First the semantic roles. The definitions given for topic expressions in Table (11-65) obtain in (11-71) as well. The only difference between the two tables is that the theme role and the patient role are represented separately, because the former appears to be extremely frequent.

(11-71) *Semantic roles of the subjects in vS*

	Alb.	MG	SC
(1) theme	381 – 37.2%	422 – 37.4%	558 – 45.4%
(2) patient	217 – 21.2%	278 – 24.6%	190 – 15.5%
(3) source of perception/emotion	131 – 12.8%	78 – 6.9%	69 – 5.6%
(4) experiencer	37 – 3.6%	45 – 4.0%	39 – 3.2%
(5) interested party	29 – 2.8%	43 – 3.8%	23 – 1.9%
(6) agent/instrumental	248 – 24.3%	262 – 23.2%	251 – 20.4%
Σ	1023 – 100%	1128 – 100%	1230 – 100%

The roles are ordered (roughly) in the increasing order according to agentivity, volitionality, inherent animateness, etc. The most prominent role seems to be that of theme, i.e. of an entity which is moved to or placed on a certain point with little or no control over the action/state. Theme subjects are especially frequent in SC, where they cover almost a half of all instances. Unlike themes, patient subjects, which amount to 15–25%, are most frequent in MG, tightly followed by Alb., with SC trailing far behind. Both theme and patient subjects are disproportionately well represented in vS, being less than half so frequent in other constructions. The frequency of the subjects denoting the source of perception/emotion, however, especially in Alb., is a true construction-specific feature: in a sample of 1000 clauses from Alb., MG and SC literary texts (Kadare, Kapandai, Kiš), comprising SV, VS, and zeroS, this kind of subject is found only in 0.2% (Kiš) – 0.4% (Kadare) clauses. Taken together, the first three roles in the table (theme, patient, source), occupying the less agentive part of the agentivity scale, make up 71.2% of all vS subjects in Alb., 69.2% in MG, and 66.5% in SC, which is quite an impressive result for nonagentive participants in the subject position.

As with topic expressions, the differences between Alb., MG and SC with respect to the frequency of certain semantic roles are not overwhelming, but they do exist, at least within the group of less agentive participants. The following scheme summarizes the relevant facts:

- (11-72) **theme:** SC > Alb., MG
patient: MG > Alb. > SC
source: Alb. > MG, SC

Note that (11-72) has a close parallel in (11-66), where the ranking of the languages according to the frequency of certain semantic roles with topic expressions is represented. SC has most locational topic expressions and most theme subjects, MG most patient topic expressions and most patient subjects, while in Alb. the topic expressions denoting experiencers and subjects denoting the source of perception/emotion display the highest frequency. The MG situation is not bound to a specific type of situation, but the Alb. and the SC ones are: Obviously, Alb. has most vS clauses denoting perceptive or emotional events with the source as the subject, whereas SC has most descriptive statements with a location as the assertional base.

The remaining three roles, either agentive or inherently animate, represent only about a third of all vS subjects, with similar ratios in all three languages. This is also a construction-specific feature of vS: the prototypical subjects, i.e. the subjects in the majority of constructions, are agentive, or at least exert some control over the situation, or they are

inherently animate (cp. Keenan 1976). However, agentive subjects do occur: in between one fifth and one quarter of instances, the subject referent is an agent or an instrument.

The claim that subjects of vS ('thetic') clauses have to be nonagentive is thus not borne out, at least not entirely: it is not a categorical property of vS, but merely a relatively strong tendency. In view of this fact, a number of linguists (notably Pinto 1997 and Kennedy 1999), trying to salvage the unaccusative account of vS clauses and related constructions, claimed that agentive subjects were compatible only with the narrow focus construal. According to this hypothesis, the broad focus construal, i.e. the true 'thetic' statement, is possible only with nonagentive subjects. I shall try to assess this view with respect to my data in Section 11.4.

The ontology of vS subject referents confirms the impression that we are dealing with non-prototypical subjects here:

(11-73) *Ontology of subject referents in vS*

	Alb.	MG	SC
(1) thing	156 – 15.2%	214 – 19.0%	189 – 15.4%
(2) state of affairs/abstraction	211 – 20.6%	301 – 26.7%	335 – 27.2%
(3) natural force	93 – 9.1%	35 – 3.1%	82 – 6.7%
(4) emotional/physical state	108 – 10.6%	43 – 3.8%	76 – 6.2%
(5) body part	145 – 14.2%	36 – 3.2%	53 – 4.3%
(6) human [1./2. person]	351 – 34.3% [4 – 0.4%]	418 – 37.1% [1 – 0.1%]	403 – 32.8% [2 – 0.2%]
(7) institution	59 – 5.8%	81 – 7.2%	92 – 7.5%
Σ	1023 – 100%	1128 – 100%	1230 – 100%

Human referents are the most prominent single category, but their share in the total number of subjects is two or three times lower than in the canonical clauses with subject topics. For instance, in VsX clauses, human subjects appear in between 66% and 91% of cases (cp. 6.4.1.), whereas in vS their ratio moves from 33% to 37%. It is furthermore conspicuous that the first and second person subjects are almost nonexistent in the construction.

As far as other ontological classes are concerned, two phenomena are worth noting. First, the frequency of the subjects denoting states of affairs, especially in MG and SC. Second, the relatively high ratio of such generally infrequent referents such as natural forces, emotional/physical states and body parts, especially in Alb. In both cases, it is the frequency of certain types of situations that is responsible for the differences between the languages. Recall that Alb. uses most topical expressions denoting humans (11-69). Combined with the subjects denoting emotional states or body parts, these topic expressions are often used in vS clauses denoting emotional or events pertaining to body, which is a situation type already identified as especially frequent in this language (see above). States of affairs, which are often

referred to by the subjects of vS in SC and MG, are usually themes whose existence is asserted with respect to a certain location, a type of situation already identified as typical of these two languages.

In sum: Just as topic expressions in vS are untypical topics with respect to the statistical topicworthiness, the subjects of vS clauses are best defined as **untypical subjects**, both with respect to the frequency of the prototypical semantic roles and to the frequency of the prototypical ontological classes. This, however, does not imply that vS subjects cannot be prototypical subjects: the number of human, agentive, etc. subject referents is the best proof of this.

The other repeatedly adduced feature of vS subjects is their allegedly obligatory ‘newness’. The relevant data are contained in Table (11-74):

(11-74) *Informational properties of the subjects in vS*

	Alb.	MG	SC
(1) 1. and 2. person deictic	4 – 0.4%	1 – 0.1%	2 – 0.2%
(2) 3. person deictic	3 – 0.3%	3 – 0.3%	1 – 0.1%
(3) major participant	41 – 4.0%	84 – 7.4%	52 – 4.2%
(4) mentioned 1-10 claus. ago	88 – 8.6%	177 – 15.7%	104 – 8.5%
(5) inferable	90 – 8.8%	198 – 17.6%	156 – 12.7%
(6) relevant in the follow. text	161 – 15.7%	143 – 12.7%	138 – 11.2%
(7) none of the above	636 – 62.2%	522 – 46.3%	777 – 63.2%

Unlike the topic expressions, the subjects only exceptionally have a deictic reference (rows (1) and (2)). All other types of givenness are weakly represented, too: The subjects of vS seldom have a status of a major participant in the text (3), and are infrequently present in the textual context (4) and (5). Note that in the latter three categories, MG has significantly higher ratios than both Alb. and SC. ‘Old’, contextually given, or textually relevant subjects are thus not typical for vS, amounting to 20-35% (the sum of (3)-(5)), but they do occur, not only with the narrow focus construal, as Pinto (1997) claims, but, as I shall try to show in 11.4, also with broad focus. Be it how it may, vS subjects, as row (7) shows, are dominantly ‘new’, this being a rather strong tendency in Alb. and SC, somewhat weaker in MG, which, as indicated, more easily tolerated given subjects in vS.¹

Row (6) reveals a fact which has not been addressed in the relevant literature: Not only that vS subjects are rarely ‘given’, they are also rarely cataphorically relevant, which is in full

¹ The general tendency to use less contextually given subjects is reflected in their form. Only between 0.5% and 2% of the subjects are pronominal (23 in Alb., 45 in MG, 28 in SC). On the other hand, ‘Heavy’ subjects, i.e. those containing more descriptive material than a noun and, optionally, a one-word modifier, are disproportionally well represented: 51.3% (525 instances) in Alb., 41.1% (464) in MG, 45.1% (555) in SC.

compliance with their not being major participants. In fact, the impression is that the typical vS subjects are only locally relevant, not like vS topic expressions, over a small stretch of discourse, but much more narrowly, only within the clause in which they appear. Note that this is at variance with the ubiquitous claim that the primary function of vS (or ‘thetic’) clauses is to introduce new discourse referents (cp. especially Hetzron’s presentative theory, Section 2.3.): the typical referent of a vS subject is not a discourse referent.

Let me now summarize. The prototypical subject of a vS clause is an **untypical subject** according to the criteria of subjecthood established by Keenan (1976): it is nonagentive and nonhuman, and it is not a discourse referent. This is, however, not to say that it may not have any of these features, or all of them. There is some crosslinguistic variation as to the subject properties investigated, although the differences are much weaker than in the case of VsX and inversion. The variation is partly triggered by the divergent ways of expressing certain types of events, partly by a difference in the degree to which contextually bound subjects are tolerated in vS.

11.3.3. Semantic and informational properties of the verb

The classification of verbs according to semantic and informational classes is a matter even more difficult and potentially more arbitrary than various ontological divisions of referring expressions, not least because verbal expressions are subject to much stronger variation of sense according to the context. It is for this reason that the present section is divided into two parts, the first dealing with meanings and pragmatic features *in abstracto*, the latter with their concrete realizations in the context.

11.3.3.1. Basic features

First, the less problematic issue, that of the informational properties of the verbs in vS. Many have claimed that the verb in a vS clause has to be ‘old’, ‘predictable’, crossreferenced to the preceding text, or at least less new than the subject (see Section 2.2. for details). Apart from the fact that it is unclear how to compare the newness of verbs and subjects, this kind of characterization of vS has only one serious flaw: it does not correspond to facts. Old, etc. verbs do exist in vS, they are even more frequent in this construction than in many others, but they still represent a minority. Consider the following table:

(11-75) *Informational properties of the verbs in vS*

	previously mentioned	inferable	inherently activated	rest
Alb.	38 – 3.7%	66 – 6.5%	31 – 3.0%	888 – 86.8%
MG	14 – 1.2%	112 – 11.8%	53 – 4.7%	949 – 84.1%
SC	35 – 2.8%	151 – 12.3%	78 – 6.3%	1021 – 78.5%

Previously mentioned verbs, i.e. those which literally repeat the verb from one of the previous clauses, are, not surprisingly, infrequent. Under the label ‘inferable verbs’ I understand those predicate expressions whose denotata are linked to the previous discourse by an indirect anaphor. For instance, if a text is about genetic engineering, the verb ‘to clone’ is considered inferable (cp. examples in Section 11.4.). Although there is no acid test for inferability, so that my results are certainly somewhat impressionistic, inferable verbs make up about one tenth of all verbs in vS, which is a lot when compared to the mean value of such predicates in other constructions. Inherently activated, or inherently derivable verbs are the predicate expressions whose denotata represent concepts which, due to their general nature, may be considered continuously present in the interlocutors’ minds on a par with the referents of ‘I’ and ‘sun’, like ‘be caused’, ‘be the first’, ‘suffice’, etc. (more on this in 11.4.). This type is also relatively frequent in vS clauses.

Taken together, the three classes of ‘old’ verbs make up between 15% and 20% of all predicates, which is indeed more than in many other constructions¹, but does not rectify the claim that the verbs in vS are of necessity given.

The semantic properties are more interesting. The following table contains the relevant data (recall that this section the verbs are investigated *in abstracto*, i.e. only their basic meanings are taken into account; the contextually triggered sense changes will be dealt with in 11.3.3.2.):

¹ Accidentally, all three verb-subject constructions have a proportion of contextually bound verbs which is somewhat higher than the mean value, although for different reasons (cp. 6.4.2., 9.1.6. and 9.2.6.).

(11-76) *The semantic classes of the verbs in vS clauses*

	Alb.	MG	SC
(1) existence	167 – 16.3%	159 – 14.1%	237 – 19.2%
(2) lack/nonexistence/need	14 – 1.4%	37 – 3.3%	34 – 2.7%
(3) appearance/disappearance	46 – 4.5%	21 – 1.9%	47 – 3.7%
(4) beginning/end	11 – 1.1%	31 – 2.7%	15 – 1.1%
(5) sequence	13 – 1.3%	17 – 1.5%	28 – 2.2%
(6) position	86 – 8.4%	95 – 8.4%	110 – 8.9%
(7) movement	138 – 13.5%	141 – 12.5%	126 – 10.1%
(8) creation/performance	124 – 12.1%	121 – 10.7%	107 – 8.6%
(9) cognition	7 – 0.7%	18 – 1.6%	17 – 1.3%
(10) perception	61 – 5.9%	41 – 3.6%	53 – 4.2%
(11) emotion	42 – 4.1%	15 – 1.3%	26 – 2.0%
(12) saying	21 – 2.1%	39 – 3.5%	55 – 4.4%
(13) transfer/possession	81 – 7.9%	101 – 8.9%	55 – 4.4%
(14) order/cause	20 – 2.0%	16 – 1.4%	48 – 3.8%
(15) characteristics	33 – 3.2%	41 – 3.6%	37 – 3.0%
(16) none of the above	159 – 15.6%	233 – 21.4%	235 – 19.0%
Σ	1023 – 100%	1128 – 100%	1230 – 100%

Categories (1) and (2) denote existence/nonexistence in the narrower sense of the word: (1) comprises the verbs meaning ‘be’, ‘exist’, ‘remain’, and similar, (2), the verbs meaning ‘lack’, ‘not be’, ‘be needed’, etc. The verbs belonging to categories (3) and (4) denote incipient existence/nonexistence, and correspond to the English verbs ‘appear’, ‘disappear’, ‘begin’ and ‘end’. The verbs marked for sequentiality, like ‘follow’, ‘be joined by’, and similar, are included into category (5), those denoting positional relationships, like ‘be found’, ‘be located’, etc., into category (6). These categories may be subsumed under the label *existentials*, now in the broader sense of the word ‘existence’. The striking feature in the class of existentials is that there are no significant differences among the languages, perhaps with the exception of the fact that SC has somewhat more verbs denoting existence *stricto sensu* than the other two languages. As will become clearer in Table (11-77), the class of existentials is the strongest single category of verbs in all three languages.

Verbs of movement (7) make up about ten percent in all three languages. The same holds for verbs denoting creation of a thing or performance of an action (‘make-verbs’ and ‘do-verbs’), contained in category (8).

The following three classes, (9) – (11), pertain to experiences, cognitive, perceptual, and emotional. The distribution across languages is relatively stable, with Alb. having a slightly higher ratio of verbs of perception and emotion than MG and SC.

More variation is found in the distribution of the verbs belonging to categories (12) and (13), the former comprising verbs of saying, which are most frequent in SC, the latter verbs of transfer and possession ('give', 'bring', 'have', etc.), with a somewhat stronger presence in MG and Alb.

Predicates denoting ordering ('be first', 'rule', etc.) and cause ('be guilty', 'be caused') amount to 2–3% in all languages, those denoting characteristics ('be big', 'be red', 'be friendly', etc.) to 3–4%.

Finally, between 15% and 20% of vS clauses contain verbs which cannot be included in any of the above mentioned categories.

Before evaluating these data, a simplified version of Table (11-76) is given, in which similar categories are grouped together:

(11-77) *Semantic classes of the verbs in vS: a simplified version*

	Alb.	MG	SC
(1)- (6) existentials	337 – 33.0%	362 – 32.1%	471 – 38.3%
(7) movement	138 – 13.5%	141 – 12.5%	126 – 10.1%
(8) creation/performance	124 – 12.1%	121 – 10.7%	107 – 8.6%
(9)-(11) experience	110 – 10.7%	74 – 6.6%	96 – 7.8%
(12) saying	21 – 2.1%	39 – 3.5%	55 – 4.4%
(13) transfer/possession	81 – 7.9%	101 – 8.9%	55 – 4.4%
(14) order/cause	20 – 2.0%	16 – 1.4%	48 – 3.8%
(15) characteristics	33 – 3.3%	41 – 3.6%	37 – 3.0%
(16) none of the above	159 – 15.6%	233 – 21.4%	235 – 19.0%

The first thing to be noted is that the distribution of the predicate classes across languages is surprisingly uniform, with only minor variation. As expected, existentials are extremely well represented, making up about one third or more of all instances. Verbs of motion, often used to mark appearance on the scene, as well as verbs of experience, have also been frequently mentioned as typical for vS in the literature. The same holds for the predicates denoting order or cause.

Less expected is the relatively high frequency of categories (8), (12), (13) and (15). Verbs which denote creation or performance, like 'make' and 'do' (Alb. *bën, krijon*, MG *kano, ftjaxno*, SC *raditi, praviti*) are equally frequent as verbs of motion. Verbs of saying, especially often found in SC, are not less usual than, say, verbs of appearance/disappearance (category (2)), while transfer and possession verbs are even two to three times more frequent than the latter, which are considered to be one of the prototypical lexical fillings of vS crosslinguistically.

The predicates subsumed under the label ‘characteristics’ (category (15)), like Alb. *është i madhë*, MG *ime meyalos*, SC *biti velik* (‘be big’), are not supposed to occur in vS clauses (i.e., in ‘thetic’ statements) at all, being typical individual level predicates (cp. Section 2.6.3.). However, they do, and they are not less well represented than the ‘prototypical’ verbs of appearance and disappearance.

Finally, the class labeled ‘none of the above’ (category (16)) comprises heterogeneous verbs which I am not able to classify. The ratio of these verbs is highest in MG, lowest in Alb.; I am not sure whether this represents a relevant typological fact.

Is there a general conclusion to be drawn from these facts? Similar to the case of the vS subjects, the verbs used in vS clauses only partially fulfill the expectations stemming from the theoretical predictions. In conformance to what is usually claimed, they often have some kind of existential semantics. The semantic fields of movement and experience, also often claimed to be typical of this construction, are well represented, too. But this is only a tendency: Other predicate classes are also allowed, even rather prominent. Some of these ‘unexpected’ verb even explicitly contradict the predictions made by many semantically based explanations of theticity. Whether this picture changes when the contextual data are taken into account will be discussed in 11.3.3.2.

One more thing before turning to the semantics of the vS predicates in context. The verbs of vS clauses have been often defined as ‘weak’ (cp. Section 2, *passim*).¹ Now, the problem is that the notion of weak verbs, however intuitively appealing, becomes extremely difficult when one tries to operationalize it. Though I am not able to offer a receipt on how to discern a weak from a strong verb, I have tried to narrow down the notion by treating as weak all the predicate expressions with a high degree of generality, i.e. those which may be applied to describe a great number of different situations, and which consequently display little or no selectional restrictions. Thus ‘go’ is weak, but not ‘walk’, ‘run’, ‘trot’, etc. I am fully aware that a measurement performed along these lines is still far from objective, but I consider it better to offer a partly impressionistic account than none.

In all three languages, weak verbs cover about a half of all instances, 49.2% (506) in Alb., 51.9% (572) in MG, 56.7% in SC (698). For instance, among verbs of motion, the verbs meaning (roughly) ‘go’ (Alb. *shkon*, MG *pijeno*, SC *íci*), ‘come’ (*vjen*, *erxome*, *doći*), and,

¹ In spite of this, there are not many proposals on how to define ‘weakness’, most authors being satisfied with a vague description referring to little or no semantic content of the verb. The only exception I am aware of is Lambrecht, who, in an early paper (1987), treats as weak all the verbs denoting basic level category concepts within the notional system of Mervis and Rosch 1981.

somewhat surprisingly, ‘pass’/‘pass by’ (*kalon, pernao, proći*), taken together, appear in some 65% of all instances, the remaining 35% being reserved for more specific predicates like ‘run’, ‘float’, ‘turn around’, etc.

Now, it may be the case that this prevalence of ‘weak’ verbs is not a construction-specific matter at all, but merely a function of the general frequency relationships, holding in the whole system, with the more general predicates dominating numerically in all clause types (precisely *because* of their semantic generality). To check this possibility, I counted ‘weak’ and ‘strong’ verbs in continuous passages from *Koha Ditore*, *Elefterotipia* and *Vesti*, 1000 clauses for each language, comprising SV, VS, and zeroS. The results are not equivocal, as I had hoped: Weak verbs are indeed more frequent in vS than the average, but not significantly, the mean value ranging between 42.3% in Alb. and 45.1% in SC, which is only about 10% less than in vS.

Perhaps a larger control corpus would yield more telling results, but, because of the time-consuming nature of the counting procedure, I leave this task to the future research. The only conclusion to be drawn is that ‘weak’, i.e. semantically highly general, verbal concepts are rather frequent in vS clauses, but it is not clear whether this is a specific feature of the construction.

11.3.3.2. Meanings in context

In this section, two questions will be addressed. First, how the context-neutral meanings of the verbs, as enumerated in the preceding section, may change in vS clauses (hereafter **lexical reinterpretation**), and second, what types of lexical relationships between the verbs and the topic expressions or subjects are typically found in vS construction (hereafter **lexical solidarity**).

11.3.3.2.1. Lexical reinterpretation

It has often been claimed that non-existential verbs receive an existential interpretation when used in vS clauses (cp. Section 2.1.), or, differently, that even when an activity verb is used in a vS clause, it is assimilated to the core cases so that ‘the activity serves simply to locate the subject on the scene’ (Bresnan 1994: 85).

This is clearly a case of what I call lexical reinterpretation, i.e. of the adjustment of the meaning of a non-prototypical lexical filling to the meaning of the construction in which it is used, through the mechanism of lexical subordination (Levin and Rapoport 1988; cp. 8.1.3.

for more detail).¹ The resulting meaning is a combination of the prototypical semantic filling of the lexical slot opened by the construction modified by a now subordinate basic meaning of the verb. In the case of existential reinterpretation, a verb like *rise* would be interpreted as ‘be (there) rising’, as in *In front of me, a church rose to the sky*, where the interpretation could be paraphrased as ‘in front of me, there was a church which rose ...’ or ‘in front of me stood a church, rising...’.

Now, in order to see how lexical reinterpretation works, we should know what the prototypical meaning of the construction is. The statistical data presented in 11.3.3.1. are not of much help here: existentials are indeed frequent, but the number of other predicate classes is too large, and their frequency too high, for them to be counted as metaphorical extensions in all instances. I should therefore like to suggest that it is the type of topic expression which determines what the prototypical lexical filling of the predicate slot is. My material allows for two generalizations: (1) Positional/existential predicates are the prototypical verbs in the clauses with locational topics. (2) In clauses with free dative or possessive topic expressions, or without overt topic expressions, the prototypical lexical filling are the predicates denoting inherently dramatic events. In the case of other topic types, I was not able to establish similar preferences.

The case of positionals, i.e. of the verbs of existence used in locational contexts and the verbs of posture proper (‘stand’, ‘sit’, etc.), is the better known one. Basically, my data confirm Bresnan’s definition quoted above: non-positional verbs used in vS with locational topics are reinterpreted as positionals modified by the original meaning of the verb itself.

I have two pieces of evidence for this. The less decisive one pertains to the fact that non-positional verbs used with locational topic expressions can always be paraphrased with a combination of a ‘weak’ positional and the gerund form of the expressed verb, or some other semantically subordinate form, as in the case of the verb *rise* described above. The assumed semantic structure may thus appear on the surface. The following example shows that the paraphrase may also run the other way around: The English original text has a combination of a positional and the gerund form of a non-positional, the SC translation a coordinate structure

¹ Note that the notion ‘existential reinterpretation’ in many cases (the most prominent being the approach advanced by Allerton and Cruttenden (1979)) means that the verb and the subject stand in a close semantic relationship, so that the verb somehow denotes the typical mode of existence of the subject referent. This belongs to the field of lexical solidarity, a phenomenon which I should like to keep apart from lexical reinterpretation, even though no sharp line can be drawn between the two.

functioning in the similar way, but the Alb. and the MG translations use only the finite form of the non-positional verb:

(11-78) Eng.: *And there, out in the snow of their back track, **was** the she-wolf **waiting** for him.* (London, p. 20)

SC: *A tamo dalje, na snijegu predjene prtine, **stajala je** VUČICA i **čekala***
and there further on snow of-passed path stood AUX she-wolf and waited
na njega. (p. 24)
on him

Alb.: *Por atje, në udhën që sapo kish bërë, po e **priste** UJKONJA.* (S.20)
but there in way which just has done PTCL him waited she-wolf

MG: *Ke eki, sti xionismeni ektasi piso tus, ton **perimene** i LIKENA.* (S.26)
and there in snow-covered plain behind them him waited the she-wolf

The other piece of evidence is somewhat more conclusive. A number of dynamic motion verbs receive a stative, positional reading in vS clauses with locational topic expressions. If used in a SV clause, the dynamic reading is the only one possible. Consider the following examples:

(11-79) MG (Eco, p. 47)

*Sto vaθος, aristera tis eklisias, **ipsonotan** to IKOĐOMIMA, pu xorizotan*
in deep left of-the church rose the Building which was-separated
apo tin eklisia me mia ektasi jemati tafus.

from the church with a plain full graves

‘Far behind, to the left of the church, rose the Building, which was separated from the church by a ground full of graves.’

(11-80) SC (Andrić, p. 74)

*Kraj zatvora **raslo je** neko DRVEĆE.*

next-to jail grew AUX some trees

‘Next to the jail grew some trees.’

In these two sentences, the verbs *ipsonome*, ‘rise’, and *rasti* ‘grow’, have a stative meaning, denoting the position of the referents ‘building’ and ‘trees’. If, however, a SV clause were used, only the dynamic motion verb meaning would be possible: *to Ikođomima IPSONOTAN* would then mean that the Building is moving upwards, and *drveće je RASLO* could only be understood as referring to the change of the dimensions of the trees, from small to large plants, not as a description of the position of the trees. What does this have to do with lexical reinterpretation? Motion verbs like *ipsonome* and *rasti*, when used with locational topics in a

vS clause, get reinterpreted as ‘be there rising’ and ‘be there growing’, through lexical subordination. The positional interpretation neutralizes the dynamic component of the verb meaning. Since nothing of the kind takes place in SV clauses, only the basic, dynamic meaning of the verbs is possible there.

In sum: Many predicates used with locational predicates are subject to a lexical reinterpretation process whereby non-positional verbs receive a positional reading, with the original meaning semantically demoted to the status of a modifier.

The second class of topic-triggered prototypical verbs, that of inherently dramatic predicates occurring in the clauses with free dative/possessive topics, or without overt clausal topic expressions, is less clearly delimited semantically. In fact, it is not a semantic class at all, at least not in the sense I usually use the verb ‘semantics’, but rather a preferred interpretation given to the event. In other words, the class ‘inherently dramatic events’ is defined on the basis of the connotations¹ certain events usually have, not on the basis of their common denotational properties.² The notion of lexical reinterpretation is thus slightly extended in this case, referring not only to the truth-conditionally defined meaning, but also to the way certain meanings are interpreted pragmatically.

Let me illustrate what I mean by ‘inherently dramatic’ on the example of a verb often quoted in the context of vS, ‘die’, which has the advantage of being abundantly attested in my corpus.³ Why the event denoted by ‘die’ is inherently dramatic is self-explanatory. The discourse-pragmatic consequences of this connotation are that we (apart from history textbooks and similar contexts) never speak of someone dying/having died as a part of a narrative structure. Thus, a text like the following would be pragmatically deviant: ‘We had breakfast together; then he died; then I went to the greengrocer’s to buy some broccoli’. In the terms of the functions of different focus structures, this means that verbs like ‘die’ are rarely,

¹ Note that I use the term ‘connotation’ in its everyday meaning, as the affective component added to a certain cognitive content, not as a technical term designed to capture the intensional features of referring expressions (cp. Lyons 1977: 175-6).

² The famous dictum of Allerton and Cruttenden (1979) that accented subjects in English (a construction related to vS, see Chapter 2) are, *inter alia*, conditioned by the verbs denoting misfortune, is thus reducible to the pragmatic feature of connotation, i.e. it is not semantic in nature. Furthermore, the term ‘inherently dramatic’ includes also the verbs denoting fortunate events (provided they are unexpected), which, as rightfully pointed out by Sasse (1987), have the same right to be considered ‘prototypical’ vS predicates as the verbs of misfortune.

³ The verbs in question are *vdes* (Alb.), *peθeno* (MG), *umreti* (SC); in SC, there is also a verb denoting forceful death, *poginuti* (‘die in a battle, die in a car accident’).

if ever, used in utterances with a subject topic and a broad focus on the verb (see 4.6. for the functions of broad focus).

Actually, I can think of only three contexts in which ‘die’ is regularly used (two are illustrated by Schmerling’s famous examples *Truman DIED* and *JOHNSON died*, quoted as (4-16) in 4.3.2.). First, less relevant in the present context, dying may be relationally presupposed for some reason (a car accident or a war have been just mentioned, for instance), and the one who died has to be identified (*John and Peter had a car crash. PETER died.*). The information structure used is narrow focus on the subject. Second, someone’s death may be expected (s/he has been ill for a long time, or similar), so that its taking place is confirmed by a clause with ‘die’ (*Truman DIED*). The information structure used is polarity focus, as defined in 6.5.1., the message reading as ‘the event of dying DID take place, as we expected’. Here is a MG example with VsX construction:

(11-81) MG (Kapandai, p. 138)

... pai pja o Filipos, ðen ine pja aftos eðo, ke mono ... permun ... jatri ke

goes more the Ph. not is more he here and only pass-by doctors and

nosokomes... ðen ine pja o babakis tis eki, mono ta matja tu...

nurses not is more the father her there only the eyes his

Ki PEΘANE o babas.

and died the father.

‘Philippos is leaving, he is not here any more, only ... doctors and nurses pass by ... her daddy is not there anymore, only his eyes... And her father DIED.’

The third context in which ‘die’ naturally occurs is that of an unexpected information. This is where the inherently dramatic nature of the event is best seen. Here are two examples:

(11-82) Alb. (Kosovarja, p. 21)

Vdiqën QINDRA shqiptarë.

died hundreds Albanians

‘[A title of a journal article] HUNDREDS of Albanians died.’

(11-83) SC (Andrić, p. 55)

Umrta joj je naprečac ĆERKA.

died to-her AUX suddenly daughter

‘[The woman was traveling from Athens to Smyrna. During the voyage something terrible happened.] Her DAUGHTER suddenly died.’

Both clauses are instance of vS, i.e. of a focus domain consisting of a verb and a subject. In the Alb. example, no overt topic expression is present, while in SC the indirect topic

expression is a free dative. Why is the configuration ‘zero/dative+vS’ so well-suited to express dramatic events? Because both topic types imply the lack of control on the part of the topic referent. When there is no overt topic expression, the best the hearer can do is to construe a very general topic, something like ‘the situation defined by the previous discourse, by the physical context, or by the parameters of »now« and »here«’. The assertion conveyed by the verb-subject combination is then assessed with respect to the topic construed in this way. Since the topic referent is, roughly, a state of affairs, within which the event encoded by vS takes place, it does not have any kind of control over the event. The case of free datives/possessives is simpler. The semantic role of the topic expressions thus encoded (‘interested party’) is explicitly marked for the lack of control: the event happens to a person, s/he does not perform it.

Now, verbs like ‘die’ occur in the two constructions described above especially frequently, by the virtue of their semantics: an inherently dramatic event is by default expressed by an inherently dramatic construction.¹ What is of interest in this section (which, to remind the forgetful reader, is devoted to lexical reinterpretation), is whether other predicates, which are not inherently dramatic, receive the connotation of suddenness when used with a free dative/possessive or without an overt topic expression. The answer is, as could be expected, affirmative. Consider the following sentence, reporting a (presumably) fortunate event:

(11-84) SC (Vesti, p. 21)

Vraća mi se snaga MUŠKARCA.

returns to-me strength of-man

‘[Advertisement for a medicine against impotency: An alleged patient reports what he first noticed after taking the medicine:] My MALE strength is back.’

The verb ‘return’ is not inherently dramatic. In a clause with the topical free dative, however, it is ascribed this connotation, due to the lexical preference such clauses display. Example illustrates the same effect with a topicless clause:

(11-85) MG (Kapandai, p. 32)

Eftasen i ORA.

arrived the hour

¹ Sasse (1996) notices that some predicates, or some verb-subject combinations, occur either in vS clauses or in polarity focus constructions. The notion of inherently dramatic predicate is a part of the answer to the question why this is so; the other part will be described in 11.3.3.2.2.

‘[The Venetians have sent some military help to the Greeks in order for them to fight against the Turks. They are certainly going to free their country now.] The HOUR has come.’

Arriving is not a dramatic action; without an overt topic expression and in a vS clause, it is easily reinterpreted as such, since the topics construed by default tend to take dramatic events as their assertional complements. In the course of Section 11.5., where discourse functions of vS are discussed, the topic of the suddenness effect will be addressed again, from a different point of view.

To conclude: Lexical reinterpretation is indeed present in vS. It is triggered by the topic type chosen. Locational topics facilitate the reinterpretation of non-positional verbs as positionals, while zero topic expressions and free datives/possessives tend to add a dramatic note to the event described by vS. I am not certain that the reinterpretation is obligatory in every single case. It is, however, quite common.

11.3.3.2.2. Lexical solidarities

A very frequent phenomenon in vS clauses, and a very frequently mentioned one, is the fact that the verb and the subject are often somehow semantically close to each other. The terms used for this kind of relationship are multifarious (cp. Sections 2.1. and 2.4.). I shall stick to the term coined by Coseriu (1967) which has already been used in the present study, *lexical solidarity*. My approach to this phenomenon differs from the usual way it is dealt with in three points: I shall deal not only with the relationship between the subject and the verb, but also with the semantic ties between the verb and the topic expression; I shall propose a classification of different types of lexical solidarity; I shall try to explain the phenomenon, as far as it goes.

Let us start with the second issue, that of the classification of the types of lexical solidarity. The semantic closeness between a predicate and an argument may vary in strength. The list that follows presents the types of lexical solidarity in decreasing order, from the strongest to the weakest.

(1) **Idioms, semi-idioms, support verb constructions.** The strongest type of lexical solidarity are the cases in which a verb-noun combination is on the verge of being lexicalized. The semantic compositionality may be completely or partially lost (idioms and semi-idioms), or still transparent (support verb constructions). In the latter case, the semi-lexicalized status of the complex is reflected in the lack of interchangeability of the verb and the noun with near-synonymous terms (you can say *to exert an influence*, but not *?to practice an influence*).

(2) **Rich existential presuppositions.** The existential presupposition of certain entities contains information which is less general than the mere knowledge of existence (cp. 4.2.4.). Thus, in order to existentially presuppose the denotatum of ‘river’, you have to presuppose ‘flowing’, ‘book’ presupposes ‘writing’, etc. The linguistic reflex of rich existential presuppositions are close lexical ties between the words *river* and *flow*, between *book* and *write*, etc.

(3) **Cognitive frames.** Many entities prototypically occur in a rather limited set of situations. For instance, apart from writing, books typically occur in the situational frames of reading and publishing. Consequently, the word *book* is not only closely connected to *write*, but also to *read* and *publish*. In the same way, *door* is close to *open* and *close*, *noise* to *hear* and *disturb*, etc. Note that in the categories (2) and (3) mostly things, or, more broadly, nonanimates have been mentioned. This is not by chance: the less cognitively prominent the entity, the greater the chance that it will be included in one of the stereotypic situational and existential frames (cp. 4.2.4.).

(4) **Ad hoc solidarities.** If a certain type of situation is frequent, or at least noncontroversial, in a certain context, I shall speak of *ad hoc* lexical solidarity (cp. 2.4. for a comparable notion proposed by Fuchs 1980). This is the weakest and the most volatile type of the close semantic relationship between a verb and a noun. For instance, if I am talking to you about a dinner in a restaurant, ‘mutton-chop’ and ‘bring’ may be treated as *ad hoc* lexically tied, as in *Then this French waiter brought me my mutton-chop with an expression of utter disgust*. The action of bringing a mutton chop is a frequent, noncontroversial action in a restaurant, but a rather weird one, say, in a sauna, or in a library. It is in this sense that the words *mutton-chop* and *bring* may be said to be in a relationship of lexical solidarity *ad hoc*, on the occasion of talking about a restaurant, but not systematically, on a par with idioms or other types mentioned above. As indicated, it is not always easy to determine the presence or the absence of a lexical solidarity of this kind, but its existence is, to my feeling at least, beyond doubt.

These four kinds of lexical solidarity are better thought of as points on a scale than as discrete categories. They are rather frequent in my corpus: Taken together, multi-word lexemes consisting of a verb and a subject or of a verb and the topic expression occur in some 50% of vS clauses in all three languages, 52.1% (532) in Alb., 53.0% (598) in MG, and 48.7% (599) in SC.¹ Note that, in order to count as lexically close to a nominal expression, a

¹ As indicated, the fourth class, *ad hoc* lexical solidarities, is not always easy to pin down. In order to avoid an exceedingly generous interpretation of the data, I counted only those cases which seemed unequivocal to me. The number of instances of this phenomenon may thus be somewhat higher than the above figures show.

verb must not be too general, i.e. it may not be ‘weak’ in the sense defined in 11.3.3.1. The only exception to this are support verb constructions, which often contain verbs with highly generalized meanings, as in MG *kano djakopes* (‘I-make holidays’) ‘I am on holiday’.

I do not believe that there is a unified explanation of the phenomenon of lexical solidarity. In some cases, lexical solidarities are triggered by the tendency to conceptualize certain events as unitary notions. In others, the relevant reason is the tendency to think in simplified cognitive schemata, abstracting away from the particularities. Yet in others, lexical solidarities are based on the fact that certain situations often provoke certain events, etc. In order to simplify my way of expression, I shall refer to the verb-noun complexes of the kind presented above as *multi-word lexemes*, even though I am aware that this is, at least in some instances, a rather strong terminological distortion.

As indicated above, both the topic expressions and the subjects of vS clauses may form a lexical solidarity with the verb. Let us look at this in some more detail.

Topic expressions form multi-word lexemes with verbs in some 15% to 22% of vS clauses, most often in MG (21.8%), less in Alb. and SC (17.3% and 16.5%, respectively). All four types are represented in the corpus. Here are some examples:

(11-86) Alb. (Bishqemi, p. 58)

Në orientimin e lakuriqit rol lozin organet e zërit dhe të dëgjimit.

in orientation of-the bat role play organs of-the voice and of-the hearing

‘In the orientation of bats both the voice and the sense of hearing play a role.’

‘To play a role’ is a semi-idiom in Alb. just as it is in English. The object *rol* (‘role’) is topicalized, with the verb *loz* (‘play’) as the other part of a multi-word lexeme.

(11-87) MG (Ciao, p. 11)

To sineðrio ðioryanose to Iðrima ja to peði ke tin IKOJENJA.

the congress organized the institute for the child and the family

‘The congress was organized by the Institute for children and family.’

This is a nice instance of a rich existential presupposition: For a congress to exist, it has to be organized. *To sineðrio* (‘the congress’) is the topic expression, with the verb *ðiorganono* (‘organize’) lexically bound with it.

(11-88) SC (Kiš, p. 173)

Slušalicu je podigla Darmolatovljeva bunovna ŽENA.

receiver AUX raised Darmolatov’s sleepy wife

‘Darmolatov’s sleepy WIFE picked up the receiver.’

Receivers are a sort of thing which occurs in only two or three situations in a human life, the most prominent one being that of picking up the receiver when the telephone rings. This is thus an example of the third type of lexical solidarity, the one based on cognitive frames. Note that the expression *slušalica* ('receiver') is topicalized.

(11-89) SC (Vjesnik, p. 2)

Predsjednikov lijes iznijelo je osam generala HV-a.

presidents coffin took-out AUX eight generals of-HV

'President's coffin was taken out by eight generals of the Croatian Army.'

Finally, the last example illustrates *ad hoc* lexical solidarities. In a description of a funeral, it is a noncontroversial action that the coffin is at a certain moment taken out to be buried. Since *lijes* ('coffin') is the topic expression, this is also an instance of a lexical solidarity between a topic expression and a verb.

The same types of the verb-noun relationship may occur between the verb and the subject. Lexical solidarities of this kind are somewhat more frequent than the ones holding between the topic expression and the verb, amounting to some 30% in all three languages: 34.8% in Alb., 31.1% in MG, and 32.2% in SC. The examples that follow illustrate the four types of lexical solidarity:

(11-90) SC (Andrić, p. 110)

Čovek je odjednom začutao. Izdao ga GLAS.

man AUX suddenly grew-silent betrayed him voice

'The man suddenly fell silent. He had lost his voice.'

This example is a nice instance of an idiom: The verb-subject complex 'the voice betrays' simply means 'to lose one's voice'.

(11-91) Alb. (Bishqemi, p. 624)

Për ne ishte organizuar një pritje e VEÇANTE.

for us was organized a reception the special

'A special reception had been arranged in our honor.'

This sentence is in a way a mirror-image of (11-87), where a topicalized expression meaning 'congress' forms a multi-word lexeme with the verb 'organize'. Here, the focal subject meaning 'reception' stands in the relationship of lexical solidarity with the same verb. The basis of this relationship is the rich existential presupposition of the entity 'reception'.

(11-92) MG (Kapandai, p. 132)

Noris, ke elambe o ILIOS.

early and shone the Sun

‘It was early, the SUN was shining.’

Practically the only situation in which the Sun regularly occurs is that of shining. This example is therefore an instance of a lexical solidarity based on a cognitive frame.

(11-93) MG (Kapandai, p. 21)

Epano sto trapezi ixan ... aploθi ta FAJITA.

above in-the table AUX been-spread the meals

‘The table was set (= the meals were spread on the table)’

There is no context-independent close lexical relationship between *aplonο* (‘spread’) and *fajito* (‘meal’). However, in the context of a description of the preparations for dinner, the complex ‘the meals are spread’ looks like an *ad hoc* lexical solidarity.

Finally, it should be noted that other elements apart from the topic expression and the subject may form a multi-word lexeme with the verb, whereby they are as a rule comprised by the focus domain, as has been elaborated in some detail in Section 11.2.3.3.

11.3.3.3. Semantic and informational properties of the verb: A summary

Is there any conclusion about the nature of the verb in vS that can be drawn from the facts presented in this section? Let me enumerate the highlights:

- (1) The verbs in vS are perhaps more often ‘old’ or ‘given’ than in other constructions, but this is not a dominant feature.
- (2) Many verbs occurring in vS are highly general in meaning, i.e. ‘weak’.
- (3) The existential/positional verbs are the most frequent single class. Furthermore, many other verbs may be assimilated to this core class. However, I have tried to show that existentials/positionals are typical only for clauses with locational topics, and that the lexical reinterpretation is triggered by the topic type rather than by the core status of existentials in the construction as a whole.
- (4) In some other cases, notably in the clauses with no overt topics or with free datives, it is not the semantic properties of the verb that are decisive, but rather its connotational value. If the event the verb denotes can be interpreted as dramatic, it is used in this clause type irrespective of the semantic features it carries.

(5) A very high percentage of verbs form multi-word lexemes with one of the nominal elements.

What emerges as characteristic of vS verbs are two things. First, the dependence of the predicate class on the topic type. Second, a tendency (not a rule!) to diminish the informational loading of the verb, which is ‘old’, or ‘weak’, or loses its independence by participating in a multi-word lexeme. In the following section we shall see how the interaction of weakened or non-weakened verbs with the rest of the clause influences focus construal.

11.4. Semantic/informational properties and focus construal

As has been repeatedly emphasized (4.5., 11.1.), the difference between the narrow focus on the subject referent and the broad verb-subject focus is not a matter of grammar in the Balkan languages, since both are encoded identically, by a focus domain marked by the verb-subject order with the nuclear accent on the subject. The two focus construals are thus a matter of pragmatic inference. The basic trigger for the one or the other construal are the lexical and the informational properties of the elements of the construction, or, more precisely, these properties very often serve as a signal to the hearer which one of the two possible focus construals the speaker has in mind at the moment of the utterance (cp. 4.2.4. on sources of presuppositions). In order to understand how vS clauses function in discourse, we first have to understand how these lexical and informational clues work.

11.4.1. Prototypical verbs and false generalizations

Before turning to the task of clarifying the issue of focus construal, I should like to try to refute a widespread belief according to which only certain, semantically and syntactically restricted, classes of verbs allow for the broad focus construal, whereas the narrow focus construal on the subject does not display any restrictions (cp., e.g., Pinto 1997, Kennedy 1999, and the bulk of literature on stage level predicates mentioned in 2.1. and 2.6.3.). The basic idea is that the verb has to denote temporary properties and to have a nonagentive subject in order to allow for a construal of a broad focus; those verbs that do not have these properties are said to automatically trigger a narrow focus construal. For instance, the Italian equivalent of ‘JOHN blushed’ is claimed by Pinto (1997: 28) to be susceptible only to the narrow focus interpretation on the subject (‘It is John that blushed, not Mary’) because the

verb is not unaccusative and because it has not a covert spatio-temporal argument (see 2.6.3. for the notion).

Now, I do not intend to deny that certain predicates almost invariably allow only for the narrow focus construal, but my contention is that (a) it is a purely pragmatic, not syntactic phenomenon, and (b) the number of such verbs is much smaller than usually assumed, at least in the three languages which are the object of the present study.

Let me to try to demonstrate what I mean by analyzing an example. Verbs of saying and writing, rather frequent in my corpus of vS clauses (11.3.3.1.), should, according to the theory depicted above, allow only for the narrow focus construal, since they have agentive subjects and are not necessarily spatio-temporally bound. This is indeed very often the case, as the following examples show:

(11-94) MG (Elefterotipia, p. 1)

Metaðidi o J. KSENAKIS.

reports the J. Ksenakis

[An announcement of an article on the first page of a daily: ‘A report from the football match between *Olymbiakos* and *Valencia*: A bitter beginning of the season for *Olymbiakos*: page 22.] A report by G. Ksenakis.’

(11-95) SC (Vesti, p. 5)

“Vežbe NATO nisu pretnja ni za koga.”

maneuvers of-NATO are-not threat not for any

To tvrdi neimenovani izvor iz sedišta NATO u BRISELU.

that claims unnamed source from headquarters of-NATO in Brussels

‘»The NATO maneuvers are not intended to threaten anyone.« This is claimed by an anonymous source from the NATO headquarters in Brussels.’

Both examples carry a narrow focus construal, because the fact that a speech has been performed is presupposed, so that the only assertive part of the clause is the accented subject. The structures are: [X reports about the football match]_{Presupposed Propositional Function} [X = J. Ksenakis]_{Focus} and [X claims that ...]_{PPF} [X = an anonymous source...]_{Focus}. The presupposed status of the denotatum of the verb is in both cases triggered by the rich existential presupposition with which every speech act content is connected: If there is a speech act, then someone must have spoken. If in the preceding text it is claimed that a journal article exists, then it lies at hand to presuppose that someone has written it; If a direct speech is quoted, then the fact that someone has spoken is of necessity presupposed in the following clause.

Since verbs of saying are most often used in vS construction when the content of a speech, or the existence of a speech act is presupposed, the most frequent focus construal is that with the narrow focus on the subject. This, however, does not imply that there are no other presuppositional constellations for vS clauses with verbs of saying, in which the broad focus construal is natural:

(11-96) Alb. (Bishqemi, p. 40)

Edhe udhëtimi ynë ka karakter eksperimentues. Së shpejti kjo shpikje do të bëhet e njohur. Do të shkruajnë gazetat e REVISTAT, do të flasë radioja e TELEVISIONI.
known will to write journals and magazines will to speak radio and television
'Our journey has an experimental character. Later, this invention will become public. Journals and MAGAZINES will write (about it), radio and TELEVISION will talk (about it).'

No speech act is presupposed here; the focus construal is broad, comprising the denotata of the verbs 'write' and 'speak' and the referents of the subjects. The indirect topic, construed on the basis of the previous text, is 'our journey' or 'this invention'. The assertional structure is: [our journey]_{Topic} [journals will write (about it), etc.]_{Focus}. The subject referents are agentive, the verbs do not belong to the stage-level class, etc. In spite of this, the focus scope is broad.

The moral of the story is twofold. First, most verbs may appear in vS clauses which are intended to convey broad focus. It is true that nonagentive, stage-level, etc. verbs are more often used in this function, but this is a statistical tendency accountable for in terms of the topic type vS clauses normally have (see 11.6.), not a syntactic rule.

The second point is methodological in nature. As Behrens (2000: 69ff.) has convincingly shown on the example of generic 'beaver' statements, when sentences are construed and judged in isolation, without a context, the danger of positing nonexistent rules is immense. Without additional clues from the context, speakers tend to generalize the most frequent reading as the only one allowed. This explains why so many linguists insist on nonagentivity, spatio-temporal boundness, etc. as the criterial properties of vS: When confronted with clauses like (11-96) taken out of context, most people will favor the interpretation in terms of the prototypical information structure of the vS clauses with the verbs like 'write' and 'speak', viz. the narrow focus construal. As a result, this more common reading is proclaimed the only reading, and a theory is born.

This is not to say that there are no lexical and informational preferences of certain types of predicates, subject, etc. for the one or the other focus construal. As I have shown in 11.3.3.2.,

locational topics are prototypically combined with positional verbs, zero indirect topics and those denoting interested party with the verbs denoting inherently dramatic events. In both cases, the resulting structure is almost regularly the broad focus construal. The number of lexical and discourse contexts which facilitate the narrow focus construal, however, is more numerous. In what follows, I shall first enumerate various triggers for the narrow focus on the subject. The list is followed by a short description of the role lexical solidarities play in focus construal.

11.4.2. Lexical and informational features triggering narrow focus construal

In Section 4.2.4. indications from the context have been presented which facilitate presupposing. Clues from the physical and textual context, as well as those derived from different kinds of discourse frames and lexical information, have been singled out. These clues are of obvious relevance for the topic of this section, in which I shall try to determine what lexical and informational means speakers use to signal to the hearer that the intended focus construal of the underspecified focus structure of vS is the narrow one. The question is which clues help the hearer decode that the denotatum of the verb is to be treated as presupposed.

(1) **Explicit marking**

(1.1.) **Focus particles.** The basic type of explicit marking of the narrow focus construal are focus particles (in the sense of König 1991) like *also, even, only, just*, etc., placed directly in front of the subject in order to mark that only the denotatum of the subject falls under their scope. In the languages of the Balkans, focalized subjects modified with the particles meaning ‘only’ (*vetëm, mono, samo*) are usually fronted (cp. Section 8.3.2.1.), although vS construction is also possible. The most frequent type of focus particle associated with the subjects of vS clauses is that which roughly corresponds to the English particles *also* and *even* (*edhe* in Alb., *ke/ki* in MG, *i* in SC):

(11-97) Eng: *When they halted, it halted ...* (London, p. 16)

Alb: *Kur u ndalën njerëzit, u ndal **edhe** AJO ...* (p. 17)

when PASS stopped people PASS stopped also she

MG: *Otan i andres stamatisan, stamatisë **ki** AFTI ...* (p. 22)

when the men stopped stopped also she

SC: *Kada oni stadoše, zastade **i** ONA ...* (p. 20)

when they stopped stopped also she

Since it is simply a part of the meaning of focus particles to mark certain types of narrow focus unequivocally, this disambiguating mechanism needs no further explanation.

(1.2.) **Setting adverbials with ‘apart from’.** Even though they do not represent a lexicalized/grammaticalized narrow focus marker on a par with focus particles, setting adverbials meaning ‘apart from’ (*përveç, ektos apo, osim/uz*) almost invariably trigger the same kind of interpretation as ‘also’ particles, with which they are often combined. It is for this reason that they are included into the class ‘explicit marking’. Here is an example:

(11-98) SC (Vesti, p. 18)

Uz Vesnu Rivas, primećeni su Era OJDANIĆ, Sanja STEFANI, i DRUGI.

apart-from V. R. noticed are E. O., S. S., and others

‘Apart from Vesna Rivas, Era Ojdanić, Sanja Stefani, and others were seen.’

The use of an apart-from-phrase implies that the situation denoted by the predicate is familiar to the hearer, and that at least one more entity apart from the one contained in the apart-from-phrase is involved in it. The focused subject in a vS clause serves to identify this entity. Note that the meaning conveyed is identical to that of the focus particle ‘also’.

(2) **Informational status**

(2.1.) **Given verbs.** Not surprisingly, if the verb is ‘given’, i.e. either explicitly mentioned in the immediately preceding text, or at least inferable from it, the narrow focus construal on the subject is practically the only interpretation possible. Consider the following example:

(11-99) ... *Shaban Manaj, që na e vranë ... Kur thashti vritet Ismet Rraci,*

Sh. M. whom to-us him they-killed when now is-killed I. Rr.

vritet tërë KOSOVA.

is-killed whole K.

‘[Many were killed.] ... Shaban Manaj, who was murdered ... Now that Ismet Rrac is murdered, it is the whole KOSOVO that is murdered.’

The presupposed status of the denotatum of the predicate in such cases is a direct consequence of the fact that the elements that are somehow ‘given’ are not only easy to presuppose, but also noncontroversial and activated in the mind of the hearer, i.e. potentially topically relevant for the further communication. Note that given verbs are relatively frequent in the Alb., MG, and SC vS clauses (Table (11-75)), a fact which is connected with the elaborative discourse function of narrow focus vS clauses (cp. 11.5.1)

(2.1.) **Given subjects.** In contrast to the verbs, which are relatively often ‘given’, the subjects of vS clauses quite rarely possess a previous relevance in the text (cp. Table (11-74)), probably because ‘old’ narrow foci tend to be fronted in the languages of the Balkans (cp.

8.3.2.1.). However, when they are used in vS, ‘old’ subjects tend to trigger a narrow focus construal. This tendency is, however, less regular than with verbs. Pronominal subjects in vS are the only instance where the broad focus reading is excluded:

(11-100) Alb. (Hoxha, p. 380)

Ne ju kemi konsideruar si miq. Fajin e kam UNË, që ju kam besuar shumë.
 we you have considered as friends guilt it have I as you I-have trusted much
 ‘We treated you as friends. The blame is all MINE, since I trusted you too much.’

(11-101) MG (Matesis, p. 37)

I mitera mu ðen stekotan pote stin ura..., pijena EYO.
 the mother my not stood ever in-the queue went I
 ‘My mother never queued ... I used to go there (= it was me who went)’

In the Alb. example, the fact that someone is to blame is presupposed, the identification of this person is done by the focal pronominal subject. In MG, ‘someone went there (to queue)’ is inferable and therefore presupposed, with the subject *eyo* identifying the person. I was not able to find a single example with a pronominal subject where the broad focus interpretation would be plausible; even worse, I was not even able to invent one. Contextually bound referents, especially those without a descriptive content, seem to be unable to form an integrated reading with the predicate (see 11.6. for the notion of integrated reading).

As far as ‘given’ lexical subjects are concerned, the broad focus construal is not excluded, but it is extremely rare and carries specific cognitive effects. First, the prototypical case, where a textually given subject signals the narrow focus interpretation:

(11-102) SC (Vesti, p. 29)

Krunoslav i Damir A. uhapšeni su pod sumnjom da su ... omogućili šverc kokaina.
 K. and D. A. arrested are under suspicion that AUX enabled smuggle of-cocaine
Čoveka, koga je Damir ... propustio na aerodromu sa 200 grama kokaina ...
 man whom AUX D. let-go on airport with 200 grams cocaine
upoznali su u diskoteci. Većinu droge preprodali su Krunoslav i DAMIR.
 met AUX in discotheque majority of-drug resold AUX K. and D.
 ‘Krunoslav and Damir A. are arrested under the suspicion of taking part in drug dealing. They met the man whom Damir let go (sc. through the control, since he is a policeman) at the airport with 200 grams cocaine in a discotheque. Most of the drug was resold by Krunoslav and Damir.’

The explanation for the preferred narrow focus reading is the same as for pronominal subjects. However, since lexical subjects, even when ‘given’, have a descriptive content, they can

(11-105) Alb. (Kosovarja, p. 4)

Rezonimi ynë shoqëror ka ngelur thellë në baltë. Dhe, natyrisht,
reasoning our social has stuck deep in mud-the and naturally
më së vështiri e kanë FEMRAT.

most of-the difficult-the it have women

‘Our social consciousness is stuck deep in the mud of indifference. Of course, it is the women who suffer most.’

The adverbial expression *më së vështiri* ‘most difficult’ triggers the narrow focus construal on the subject, or, the other way around, it instructs the hearer to treat the predicate as presupposed, by turning it semantically into a set of grades of suffering (‘having it hard’). No hearer is likely to contest that there is little, more, and most suffering, which means that the presupposition of the denotatum of the verb is greatly facilitated.

(3.3.) **Quantified subjects.** Quantified subjects are very often found in the clauses with a narrow focus construal on the subject. Consider the following example, in which the fact that the money was found is clearly presupposed, whereas it is the sum of the money found that represents the focal information:

(11-106) SC (Vesti, p. 28)

Nadjeno je 120.000 lažnih maraka.

found is 120000 false marks

‘[The police have arrested eight persons in possession of counterfeit bank notes.]
120.000 forged deutsch marks have been found.’

The reason for this preference is purely pragmatic, or better, discourse-pragmatic: Quantified subjects are often used in one of the contexts typical for narrow focus utterances, which is labeled *elaborative* in the present study. For more details, see Section 11.5.1.2. That this is not an absolute restriction, but rather a tendency, is witnessed by the following example:

(11-107) SC (Andrić, p. 99)

Vrata su se otvorila ... U sobu su bez šuma ušla dva TAMNA čoveka.

door AUX REFL opened in room AUX without noise entered two dark men

‘The door opened ... Two dark men silently entered the room.’

The assertional structure in (11-107) seems to be something like [room]_{Topic} [entered two dark men]_{Focus}, i.e. the focus domain *ušla dva tamna čoveka* seems to be construed as broad focus. The fact that the subject is quantified thus does not hinder a broad focus construal, it simply makes it less probable.

(4) Discourse-pragmatic factors

(4.1.) **Contrastive topics.** This is actually the only discourse-pragmatic factor facilitating the narrow focus construal for which I have some evidence, though this does not necessarily imply that it is the only one conceivable. The factor I have in mind is the contrastivity of topics: If the topic in a vS clause is contrastive, the preferred, in some cases the only possible, focus construal is the narrow one.

(11-108) MG (Elefterotipia, p. 2)

200.000 drx. kerðizun i ARIΘMI: 1460, etc.; 100.000 drx. kerðizun i ARIΘMI: 2872,
200000 drachms get the numbers ... 100000 drachms get the numbers ...
etc.; 50.000 drx. kerðizun i ARIΘMI: 1332, etc.

50000 drachms get the numbers ...

‘[A report of the lottery:] 200000 drachms go to the numbers: 1460, etc. 100000 drachms go to the numbers: 2872, etc. 50000 drachms goes to the numbers: ...’

(11-109) Alb. (Koha Ditore, p.kult. 6)

E njohur për lexuesin e Beckettit, e më pak për lexuesin shqiptar,

the known for reader-the of-the B. and more little for reader-the Albanian

është PROZA beketiane.

is prose-the of-B.

‘[Beckett’s plays are rightfully famous and known to everyone.] Well known to Beckett’s readership, but less so to the Albanian readership, is Beckett’s PROSE.’

The MG example contains instances of ‘normal’ contrastive topics, i.e. those which receive the contrastive interpretation because they are members of contextually evoked sets. The reason this topic type so often triggers narrow focus construal is that contrastive sets are usually defined through the property which is encoded by the predicate. In (11-108), the topic expressions ‘200.000 drachms’, ‘100.000 drachms’ are contrastive on the basis of the fact that they belong to the set of prizes one can win in a lottery. This is precisely what the predicate *kerðizo*, ‘win, get’, means. Thus, since the numerical expressions are topicalized on the basis of belonging to the set defined by the verb, the verb itself receives a presuppositional interpretation. The result is a narrow focus on the subject.

Sentence (11-109) is different, its topic expression being a predicative of the type I defined above (11.2.2.2.4.) as exclusive contrastive topic. I have argued elsewhere (Matić, 2003) that this topic type invariably requires narrow focus, since it is an operator-like entity which has a presupposed propositional function as its scope. The example adduced thus presupposes that

there exists a thing which is known to Beckett's readers, etc., and identifies it with Beckett's prose.

The list of contextual, lexical, and other factors which facilitate, or even require, narrow focus construal in vS given in this section (certainly not exhaustive) is intended to point to the fact that broad vs. narrow focus construal is not a matter of the semantic structure of predicates as understood in the numerous works on theticity operating with notions like unaccusativity, stage vs. individual level, etc., but rather a matter of the presence of presupposition-facilitating context clues.

11.4.3. Lexical solidarities and focus construal

Although lexical solidarity might have been dealt with in the previous chapter as well, the special status it has in determining the focus construal of utterances makes it a good basis for a special section. The idea I should like to put forward is quite simple: If the verb and the topic are lexically close, the normal interpretation will be a narrow focus on the subject; If it is the verb and the subject that are lexically tied, the preferential focus construal will be the broad one. Schematically (the index LEX.SOL. reads as 'lexical solidarity'):

(11-110) **[[topic + verb]_{LEX.SOL.}]**Presupposed **[subject]_{Focus}**

[topic]_{Presupposed} [[verb + subject]_{LEX.SOL.}]Focus

Let me illustrate what I mean by two sentences which almost form a minimal pair:

(11-111) SC (Kiš, p. 36)

...*knjigu pod naslovom Ireland to Spain...izdala je Dablinska federacija VETERANA*
 book under title Ireland to Spain published AUX Dublin federation of-veterans
 '... the book entitled *Ireland to Spain* ... was published by the Dublin veteran
 federation.'

(11-112) SC (Vesti, p. 14)

... *tim povodom izdata je luksuzna knjiga u tvrdim KORICAMA.*
 with-that reason published is luxurious book in hard cover
 '[A football club is celebrating its 50th anniversary] ... on that occasion, a
 luxurious book in hard cover has been published.'

It is a part of the rich existential presupposition of books that they are published: In order for a book to exist, it has to be published. 'Book' and 'publish' are thus two words standing in the relationship of lexical solidarity. Now, in (11-111), the nominal member of the pair, *knjiga* ('book'), playing the role of the object, is topicalized. The reading resulting from this is

unequivocally a narrow focus on the subject, *Dablinska federacija veterana*. In (11-112), *knjiga* is the subject, the carrier of the nuclear stress within the focus domain. Consequently, the focus construal is broad, comprising the denotata of the verb and the subject. Schematically, this looks as follows:

(11-113) **[[book]_{Topic} + publish]_{LEX.SOL.}]**Presupposed **[the Dublin federation]_{Focus}**
[that occasion]_{Topic, Presupposed} [[publish + book]_{LEX.SOL.}]Focus

In other words, in (11-111), the author asserts about the publication of the book that it was performed by the Dublin veteran federation, in (11-112), about the 50th anniversary of a football club that it was an occasion to publish a luxurious book.

Why is this so? Recall that lexical solidarities *de facto* represent unitary concepts accidentally expressed by more than one word, a kind of multi-word lexeme (cp. 1.3.3.2.2.). Now, if a part of a multi-word lexeme is topicalized, i.e. relationally presupposed, it almost automatically ‘draws’ the remaining part of the lexeme into the presuppositional status. Thus, if I choose the referent ‘book’ for the topic of my utterance, I signal to the hearer that the situation about which I am going to assert something is most probably that of reading, writing, or publishing. This facilitates her/his presuppositional capacities, so that, when I utter that the book has been published, s/he may easily treat the whole complex as presupposed. What remains is to assess the value of the subject referent to the presuppositional basis thus established. Therefore, [book was published] is preferentially interpreted as a relationally presupposed propositional function, about which it is asserted that the one who performed the action is the Dublin veteran association.

If, on the other hand, I decide to assert the content of a part of a multi-word lexeme with respect to some topic, then the remaining part is ‘drawn’ into the assertional status. If my topic in an utterance is ‘the 50th anniversary of a football club’, and I wish to assert about it that a book exists which is relevant to it, the assertion of ‘book’ with respect to ‘anniversary’ I will take as a neutral linking member the appropriate verb forming a multi-word lexeme with ‘book’, in this particular case ‘publish’. The assertion of ‘book’ with respect to a topic will thus almost automatically lead to the inclusion of ‘publish’ into the scope of assertion.

In how far are these processes really automatic? Once again, I should like to argue that they are, being pragmatic in nature, cancelable, if the conditions are met. For instance, if (11-112) were uttered in a corrective context, it would most likely have a narrow focus construal, as in *tim povodom je izdata KNJIGA, ne BROŠURA* (‘on that occasion is published book, not pamphlet’). The fact remains, however, that in most cases the scheme proposed in (11-110) functions in the way I described in this section.

What have we learnt from this lengthy investigation of the mechanisms of focus construal? Basically, the speakers of the languages with (partly) underspecified focus structures, as Alb., MG and SC, have to heavily rely on the semantic and other contextual clues in order to signal to their interlocutors what type of presuppositional/assertional structure they want to convey. More specifically, I have tried to show that there are no semantic laws (often reified as syntactic structures in the literature) which could be made responsible for the one or the other focus construal, but that the whole issue belongs to the field of pragmatics.

11.5. Discourse functions of vS clauses

11.5.1. vS clauses with narrow focus construal

11.5.1.1. Fronted focused subjects and vS

In Section 8.3.3. it has been shown that narrowly construed focus domains may carry the same informational loading as fronted narrow focus expressions, when appropriate conditions are met. Therefore, vS clauses with narrow focus construal may play the same roles in discourse as the clauses with fronted focused subjects: They may occur in corrective contexts, in comments, questions and answers, conclusions and openings, etc. The factors triggering the use of narrowly construed focus domains instead of fronted foci have been identified as the heaviness or cataphoric relevance of the focus expression, or the non-ratified status of the topic. Since all these functions have been described in some detail in 8.3.2., I do not consider it necessary to repeat the whole analysis here. An illustration of one of the typical contexts for fronted focus expressions, the corrective one, will suffice:

(11-114) MG (Ciao, p. 30)

ðen fteo EγO! AFTOS ftei!
 not am-guilty I he is-guilty
 ‘It is not MY fault! HE is to blame!’

(11-115) Alb. (Bishqemi, p. 12)

Shokut tim nuk i pëlqente NËNTOKA, por MBITOKA.
 to-friend my not him:CLIT liked underground but overground
 ‘My friend was not interested in the UNDEGROUND world, but in the world upon the EARTH.’

The MG sentences furnish a beautiful example of how the principle of cataphoric relevance works: In the first sentence, the speaker corrects the assumed belief of the hearer that he is guilty for something. Since the correct identification is to follow immediately, he uses a narrowly construed vS focus domain: [*ðen fteo* [*EγO*]_{Focus}]_{Focus Domain}. The second sentence – *AFTOS ftei* – shows that, when there is no cataphoric relevance, fronted foci are used for corrections. The same holds, *mutatis mutandis*, for the Alb. example.

Similar examples of vS taking over the role of fronted foci can be adduced for all other discourse functions enumerated above. The genuine topic of interest in this section are, however, those contexts which are typical of narrowly focused vS construction only.

11.5.1.2. The narrower domain of narrow foci in vS

Three discourse functions seem to be performed only by narrowly construed vS clauses.

(1) Ordering. If a sequential relationship is to be established, vS clauses containing order/cause predicates, superlatives, and similar expressions, are used. Sentences (11-103), (11-104) and (11-105) illustrate this function of vS in MG and Alb. Here is a SC example:

(11-116)SC (Andrić, p. 47)

*O sebi niko nije rekao ništa ... **Naročito je uzdržljiv bio mladi TURČIN.***

about himself nobody said nothing especially AUX reserved was young Turk

‘Nobody said anything about himself ... Especially reserved was the young TURK.’

(2) Elaboration. vS clauses used as elaborations of the previously introduced state of affairs are a very frequent discourse device, especially in the journalistic discourse, in all three languages. The basic method looks as follows: An underspecified situation is introduced; in the following sentence, the same situation, now presupposed, is encoded again, with the same or different wording, and the information missing from the first mention is introduced via narrow focus. Three basic types of elaboration are identifiable: addition, identification/precision, and exemplification. Let us first consider the most simple type, addition:

(11-117) Alb. (Bishqemi, p. 105)

...gjatë gjithë kohës komandanti rrinte i mbyllur e punonte në kabinetin e tij.

during whole time commander sat the closed and worked in office the his

Me humor të rënë ishte edhe VETIOLA.

with mood the fallen was also V.

‘... the commander was sitting and working in his office all the time. VETIOLË was also in a bad mood.’

The usual way of expressing elaboration through addition is with the focus particle meaning ‘also’, as can be seen from (11-117) and from (11-97). The following example illustrates elaboration through identification/precision:

(11-118) SC (Vesti, p. 28)

*Policija je ... sprečila nelegalnu trgovinu robe ... Roba je trebalo da se nadje na
police AUX stopped illegal trade of-goods goods AUX should that REFL find on
levoj obali Drine. Švercovana je živa stoka, alkohol i benzin.*

left bank of-D. smuggled AUX live cattle alcohol and petrol

‘The police thwarted an illegal trade... The goods were supposed to be transported onto the left bank of the river Drina. CATTLE, ALCOHOL, and PETROL were smuggled.’

First the fact that there was an illegal trade with goods is established, then the situation is elaborated upon by identifying the rather undetermined referent ‘goods’ with cattle, alcohol, etc. A very frequent subtype of elaboration through identification/precision are clauses with quantified subjects, in which the indefinite number of participants in the preceding clause is stated more precisely. Thus, in (11-106), the sentence in which it said that the police arrested some people in possession of counterfeit bank notes is followed by an elaborating vS clause which states that the quantity of the forged money was 120000 deutsch marks.

To elaborate through exemplification means to first introduce a general situation, and then the particular instances of it:

(11-119) MG (Ciao, p. 8)

Poli đjasimi ... exun ... meyales pisines sta spitja tus. Xuliyundjanon

many celebrities have large swimming-pools in houses their of-gigantic

đjastaseon ine i pisina tu epixirimatia Spiru METAKSA ... Poli meyalon

dimensions is the swimming-pool of-the businessman S.M. very of-big

đjastaseon ine ke i pisina tu Vangeli MITILINEU ... Analoyon đjastaseon

dimensions is also the swimming-pool of-the V.M. of-analogous dimensions

ine i pisina tis Anas VISI ...

is the swimming-pool of-the A.V.

‘Many celebrities have ... large swimming pools in their houses. The businessman Spyros Metaxas has a gigantic swimming pool ... Vangelis Mityleneos has a huge swimming pool ... Anna Vissi has a similar swimming pool ...

The situation of many celebrities having large swimming pools is introduced as a general fact and thereafter elaborated upon by exemplifying this general fact.

All types of elaborative contexts seem to stem from the considerations of the economy of presenting information: The speakers who use this construction break the information they want to convey in a number of independent clauses, giving the hearer what Givón (1979) calls one chunk of information per clause.

(3) **Ritual.** Probably the most interesting use of narrowly focused vS clauses is that of presenting a situation as a ritualized sequence of events. If a situation constitutes a cognitive frame (cp. 4.2.4.), its parts may be easily presupposed and treated as presupposed open propositions, with the focus on the most agentive participant. Thus, if there is a frame ‘tea party’ (cp. Van Oosten 1986), then the situations ‘coming’, ‘going’, ‘drinking tea’, ‘receiving guests’, etc. may be easily presupposed in a communicative act dealing with a particular tea party. What the speaker then has to do is to identify the participants of these actions, as in [Mary and John]_{Focus} [came]_{Presupposed Open Proposition}, [Betty]_{Focus} [received the guests]_{Presupposed}, etc. The fact that it is the agentive participants that are most often not presupposed has to do with their being least predictable and most individuated participants, so that their presence in frames is usually marked with empty slots only. The preferred situation type presented in this way are highly ritualized states of affairs, like weddings, funerals, and similar joys of everyday life. Consider the following sequence:

(11-120) SC (Vesti, p. 14)

Crkva u Alteni proslavila je juče svoju slavu Svetog Simeona.

Liturgiju je služio episkop Konstantin sa SVEŠTENSTVOM. Sa vladikom su se

mess AUX served bishop K. with clergy with archpriest AUX REFL

molili jereji Stevan Kolarević i Ljubiša TORNA. ... Uloge kuma prihvatio se

prayed priests S. K. and Lj. T. role of godfather took REFL

Uroš NENADIĆ ... Liturgiju su svojim pevanjem ulepšali članovi hora

U. N. mess AUX with-their singing embellished members of chorus

“Sveti SAVA”...

saint S.

‘The church in Alten celebrated yesterday its Slava [a feast for a patron saint], St. Simeon. The mess was held by bishop Konstantin and his suite. Together with the bishop, young priests S. Kolarević and Lj. Torna held prayers. ... The role of the godfather [a firmly established role in the ritual] was played by U. Nenadić... The mess was adorned with the songs sung by the members of the chorus ‘Saint Sabbas’.

Once that the frame ‘Slava’ (a Serbian religious feast) is introduced, most SC readers will be able to infer the propositional functions ‘the mess was held’, ‘the prayers were held’, ‘the role of godfather was played by someone’, etc. What is left for the speaker is to identify the participants via focusing. Now, rituals are not the only kind of situation frequently described by narrowly construed vS clauses. All kinds of regularly appearing situations are liable to being presented in this fashion:

(11-121) Alb. (Bishqemi, p. 68)

Në pultin e komandimit radhën e shërbimit e kishte VETIOLA.

in panel of-the command order of-the service it:CLIT had V.

‘At the command desk, it was Vetiolë’s turn to do her shift.’

The phrase *ka radhën e shërbimit* (‘to have the order of service’ = ‘to do one’s shift’) conveys a situation which occurs repeatedly, since someone always has to be at the command desk in a space ship. This makes it an ideal candidate to be encoded as a ritualized action, with a narrow subject focus.

The actions presented in this way are, as indicated, repetitive, and as such subject to schematized mental representations. The funny thing about it is that even those actions which are not necessarily ritualized may be presented as such, if narrowly focused vS clauses are used, through a mechanism which resembles lexical reinterpretation, especially its connotative subtype (cp. 11.3.3.2.1): Those situations which are linguistically encoded through a construction specialized for ritual events become ritualized by virtue of being encoded that way. The following example depicts the expectations of a young merchant before a visit to a friend:

(11-122) MG (Kapandai, p. 127)

θa tus perimene o XATZIJANIS, opos panda kapnizondas to narjile tu ...

FUT them awaited the X. as always smoking the nargileh his

θa tus xeretize i TRAKENA, xamojelasti...

FUT them greet the T. smiling

‘They would be received by XATZIGIANNIS, who would be smoking his nargileh, as always, and then ... they would be greeted by TRAKAINA, smiling ...’

None of the functions of the narrowly focused subjects in vS clauses is restricted to subjects only: ordering predications, elaborations, and ritualized actions may, with a higher or lower frequency, be expressed by other constituent types within narrowly construed focus domains as well. There seems to be no significant difference among the Balkan languages with respect to these basic functions, neither in lexical filling nor in frequency. Fronted narrow foci

presumably do not occur in the contexts enumerated above because these contexts do not imply contrast and do not have a contrary-to-expectations interpretative potential.

11.5.2. vS clauses with broad focus construal

In describing discourse functions of vS clauses with broad focus construal, i.e. with the ‘thetic’ interpretation proper, I shall stick to the classification proposed by Sasse (1996), with some minor divergences. The functions Sasse identifies are the **annuntiative, introductive, interruptive, descriptive, and explanative functions**. As will become patent in the course of this and the next sections, my contention is that there are two basic topic types which are responsible for the existence of broad focus vS clauses, although this is not to say that other topic types are excluded from the construction. I am referring to what I shall call **locational** and **situational topics**. According to the presence of the one or the other type in the prototypical vS contexts, the five functions will be assigned to two groups: the situational topic class, with the annuntiative, interruptive, and explanative functions on one hand, and the locational class, with the descriptive and introductive functions on the other. The line between these two classes is not clear-cut, since situations regularly imply locations. Despite that, the division is useful and, I think, justified, in that it helps account for certain clear differences and similarities between the clauses used in the five functions enumerated above.

11.5.2.1. vS clauses with locational topics

Locational topics are generally indirect, i.e. encoded as non-objects and non-subjects. The most frequent topic expressions are prepositional phrases (*in the house, on the hill*), locational adverbs (*here, there*), or there is no overt topic expression at all. In all three cases, the hearer has to construe the topic. If there is a topic expression, only its relevance relationship to the assertion has to be established. If there is none, both the entity to which the assertion relates and the relevance relationship are the work of the hearer (cp. 4.3.3. and 4.5.2.). There are, however, locationals which are encoded as direct topics, notably as objects. This requires a specific verb valence, like in *surround*. The phenomenon of direct locational topics is rather infrequent in Alb., MG, and SC, but it opens the question of definition: What is a location, and how is it discerned from an entity? The question will be answered, at least tentatively, in Section 11.6. For now, I shall confine myself to giving a working definition of locations. Locational expressions are those expressions which denote the place at which an event is

located; Locations are the mental representations of these places. If there is no locational expression, the hearer has to construe a location.

Locational topics may be the only topics in a proposition, but they may also cooccur with other topics, entities or situations. In a vS clause, these multiple topics are usually encoded by means of a PP and a non-subject NP. The relational presupposition then has the form of an incomplete proposition: ‘There is a topically relevant relationship between a location X and an entity Y’, with the assertion identifying this relationship (cp. 4.3.2. on multiple topics).

Let us now take a closer look at the functions of vS clauses with locational topics in discourse.

11.5.2.1.1. Descriptive function

vS clauses with locational topics are probably universally most often used in order to give descriptions of scenes. Here are some more or less beautiful examples:

(11-123) Alb. (Bishqemi, p. 6)

Nga lugina frynte një erë e LEHTË. Në qiell filluan të ndizen YJET.

from valley blew a wind the mild in sky began to shine stars

‘A mild wind was blowing from the valley. On the sky, the stars began to shine.’

(11-124) MG (Fakinou, p. 153)

Sto patoma tis kuzinas tis vriskondan i sakules tis LAIKIS.

in-the floor of-the kitchen her were-found the bags of-the market

‘On her kitchen floor were the bags she had brought from the market.’

Both clauses assert something about a physical entity which is semantically construed as a location: In (11-123), the speaker describes his way home early in the evening by making assertions about the valley and the sky; In (11-124), the speaker describes the scene she has found in her flat by using the kitchen floor as the topic. The location need not be a prototypical ‘place’, like ‘valley’, ‘floor’, etc.; Entities which are primarily interpreted as ‘thing’ freely occur as locations as well:

(11-125) Alb. (Bishqemi, p. 54)

Tejetej nëndeses, për së gjati ... është vendosur një tub i madh

throughout submarine for of-the length is placed a pipe the big

‘Along the whole length of the submarine runs a large tube...’

The topic used here is ‘submarine’, something we usually categorize as ‘thing’, not as ‘place’. However, it is a location (see 11.6.1. for the relationship between locations and things).

It is no great wonder that locational topics are used to describe scenes. It is also no great wonder that the verbs used in this context often have existential/positional semantics (as in (11-124)), or form a multi-word lexeme with the subject (as in (11-123)). This, however, does not have a status of a rule:

(11-126) Eng.: *Down the frozen waterway toiled a string of wolfish dogs.* (London, p. 3)

Alb.: *Nëpër shtratin e ngrirë të lumit po kalonte një KARROÇË*

over bed the frozen of-the river PTCL passed a carriage

që tërhiqej prej qensh. (p. 3)

which was-drawn by dogs

MG: *Tin oxθi tu potamu katiforize mja sira LIKOSKILA, zemena se elkiθro.* (p. 7)

the bank of-the river went-down a row of-wolf-dogs bound in sled

SC: *Niz zaledjeno korito naporno je išao niz pasa VUČJAKA.* (p. 5)

down frozen bed with-toil AUX went row of-dogs German-shepherds

A motion verb like *kalon* ('pass'), *katiforizo* ('go-down'), *içi* ('go') may serve to describe a scene as well as any other.

Descriptive vS clauses often appear in chains, as in (11-123), since a full description of a scene may demand more topics (in (11-123), the valley and the sky), and consequently more vS clauses. In all three languages, this is a very productive pattern. The prototypical case is, as indicated, a description of physical scenes, with subjects denoting physical entities which are located on the scene, performing an action or simply being there. This prototypical case is frequently metaphorically extended to other types of descriptions. First, a subject may be ambiguous between an entity and a state-of-affairs interpretations. This is very often the case with nouns denoting meteorological phenomena:

(11-127) SC (Kiš, p. 170)

Napolju je besnela BURA ...

outside AUX raged storm

'A STORM was raging outside...'

Even more frequent, at least in my corpus, is the extension of the descriptive mechanism to the notions of time. A scene, now temporal, is described as the passing of time:

(11-128) [Two men are waiting.]

Eng: *An hour went by, and a second hour.* (p.4)

Alb: *Kaloi NJË orë, pastaj një TJETËR.* (S.5)

passed one hour then one other

MG: *Perase MIA ora ke meta đEFTERI.* (S.9)

passed one hour and then second

SC: *Prodje JEDAN sat, pa onda i DRUGI.* (S.7)

passed one hour and then also second

One particularly productive extension are the descriptions of complex situations (i.e. not of physical scenes) with a series of vS clauses. Consider the following passage, describing a preparation for a battle:

(11-129) MG (Kapandai, p. 46)

Ap' ta strata mavrize o topos OLOS, Turki ki Arapiđes ..., ki etrizan i ARAMBAđES
of the armies was-black the place whole, Turks and Arabs and ground the carts
i varifortomeni me ta kanonia, ki etrexan zerva-đeksia SPAXIđES me jimna spaθja
the heavily-loaded with the guns and ran left-right spahis with bare swords
sta xerja ke JENITSARI me ta topuzja tus ...

in-the hands and janizeries with clubs their

‘The whole place was black with the armies, Turks and Arabs ..., and the carts, loaded with guns, were grinding everything before them, and the spahis with bare swords in their hands were running around, as well as janizaries with their clubs...’

Note that this is not a narrative passage, but a description, with all the events encoded taking place simultaneously. If such a sequence were construed with VsX clauses, the interpretation would be sequential, i.e. narrative (cp. 6.5.): first the armies made the place black, then the carts, etc. A series of vS clauses is invariably interpreted as a description. The topic is, as indicated, a complex state of affairs taking place on a scene. The probable definition of the topic in this particular case would be ‘preparations for the battle’. A very frequent type of sequenced vS clauses describing a complex situation are those which render habitual events: one thing happens, and another, always in a certain order, or simultaneously:

(11-130) MG (Ciao, p. 42)

Xtipa to TILEFONO ke mu jinete mja epangelmatiki PROTASI.

rings the telephone and to-me happens a professional proposal

‘[–“What do you speak about with the ghost of your late husband?” – “I tell him that I have a lot of problems. At that moment, something always happens.”] The TELEPHON rings, and I am offered a job.’

A specific subtype of the descriptive type extended to describing habitual complex situations are procedural texts describing the way something is done. Again, vS clauses almost invariably appear in a series. Here is a description of a game:

(11-131) SC (Kiš, p. 83)

Postavi se kocka ŠEĆERA pred svakog igrača i čeka se ...

is-put REFL cube of-sugar in-front every player and is-waited REFL

na jednu od kocki sleti MUVA i na taj način odredi dobitnika...

on one of cubes flies-down fly and on that manner determines winner

‘[Following game is played in the South:] A cube of sugar is placed in front of every player and one has to wait ... a fly lands on one of the cubes and in that way determines who the winner is...’

Needless to say, such procedural sequences of vS clauses are the regular way of giving recipes in Alb., MG and SC cookbooks.

Descriptions of complex situations by means of a sequence of vS clauses have an objective, impartial tone. This is a direct function of the fact that there are no entity topics involved, which could represent the source of perspective. The assertional basis is a scene, a locational notion, the one on which the events presented are set.

Now, this objective, non-controllable slant of vS descriptions may be exploited for the purposes other than to describe scenes or complex situations. A very characteristic way of exploiting descriptive vS clauses is illustrated by the following examples:

(11-132) Alb. (Kosovarja, p. 6)

Familjarët vijnë grupe-grupe dhe mbliidhen në vendin e caktuar.

family-members come group-group and gather in place-the the determined

Pritet ora 17. *Të gjithë e kanë të drejtuar shikimin nga lindja*

is-awaited hour 17 the all it have the directed view from east-the

në pritje të mbërritjes së autobusëve.

in expectation of-the arrival of-the buses

‘[Albanian prisoners are released from Serbian prisons. Everybody is expecting them.] Their families are coming in groups, gathering at the place fixed in advance. They are waiting for the five o’clock bell (‘FIVE o’clock is awaited’). Everybody has directed their views towards the East, expecting the buses to arrive.’

(11-133) MG (Kapandai, p. 141)

San itan mikra, tus eleje istories, tis Kirjakes to mesimeri, jiro ap' to trapezi ...

when was small them told stories the Sundays the afternoon around of the table

Kapja kuvenda, kati jeyonos stekotan aformi, kati tu θimizan

some talk some event stood instigation something him reminded

ke arxize i ðijisi, ..., ap' ekso tis ixane maθi ta peðja tis istories.

and started the narration from out them have learned the children the stories

‘[The memories of the main character on her deceased father.] When she was a

child, he would tell them stories, on Sunday afternoons, when they were sitting at

the table... A word or an event would give him a pretext, they would remind him

of something, and there was the STORY, the children had learnt these stories by

heart.’

Descriptive vS clauses are embedded in a narrative context in which one or more human participants perform temporally sequenced actions. More precisely, a sequence of clauses in which the topic is the subject, a human being, is interrupted by a vS clause which seems to convey an assertion about the complex situation, in the same way as (11-129) and (11-131) do. What effect does this have? Descriptive vS clauses do not contain an agent phrase; by virtue of making assertions about scenes, they have an objective note, as if someone from outside were observing the whole situation. Consequently, the effect is that of an action performed by an outer force, although we know that in (11-129) it is the people gathered who are waiting, and that it is the father who began the story in (11-131). The fact that people are waiting is presented as something they cannot influence. The fact that the author's father begins to tell a story from the war is presented as an action nobody can influence, like a rain or a storm: The poor children are exposed to this torture of boredom without the slightest hope to influence it. This effect will hereafter be called **natural force effect**.

Descriptive vS clauses are probably the most frequent and most productive type of vS. They also represent the type which is least subject to the stylistic choice of the speaker. This means that, if a location or a complex situation conceptualized as a scene is to be described, the use of vS is almost automatic in all three languages. The parallel translations I analyzed confirm this in an indirect way. Out of 428 clauses which are vS in at least one language, there are only 63 in which *all* three languages agree in having vS. Out of this number, 54, i.e. 85.7%, are different kinds of descriptive clauses. A conclusion I should like to draw from this is that when a speaker wants to make an assertion about a locational notion, be it a physical, temporal, or situation-like location, s/he uses a vS clause by default, with little or no

crosslinguistic variation. This, of course, does not mean that there are no alternative ways of making assertions about such topics, but they are clearly stylistically marked, often with a flavor of pragmatic pretense¹.

11.5.2.1.2. Introductory function

The first function of vS clauses recognized by the linguists, and the one most frequently mentioned in the literature, is the introductory function: vS clauses serve to introduce new discourse referents. The mechanism of introducing new referents with vS clauses is prototypically illustrated by the beginnings of fairy tales, parallel to the English ‘once upon a time’-clauses, almost invariably construed with an existential verb:

(11-134) MG (Elinika paramiθja, p. 48)

Mja fora ke enan kero itan mia JINEKA, pu ixē sto paraθiro tis mja γλαstra
once and one time was a woman who had in-the window her a flowerpot
me vasiliko ...

with basil

‘Once upon a time, there was a woman, who had a flowerpot with a basil on her window...’

(11-135) SC (Antologija, p. 106)

Bio jedan STARAC i imao puno DJECE.

was one old-man and had many children

‘There was once an old man who had a lot of children.’

This way of introducing referents into the universe of discourse is nowadays obsolete in all three languages, both in written and spoken language. This is not to say that vS clauses are not used as an introductory device any more, but that the structure and the discourse environment of such clauses is different.

The opening clauses of fairy tails have three conspicuous features: (a) They assert the existence of a *major* participant in the discourse, with respect to a very general *temporally* conceptualized scene; (b) They contain an existential verb; (c) They open the whole text. In modern discourse, introductory vS clauses lack at least one of these properties.

¹ Sasse (1996: 38) notes that descriptive statements may be expressed with canonical SV clauses with topical subjects, like in *The sun SET*. This is true, but the utterances of this type do not take a location as their topic, conceptualizing, as will be shown in 11.6.1., potentially locational entities as things. Thus, *The sun SET* is not a descriptive statement about a location, i.e. descriptive in the narrower sense of the word, but rather an entity-related assertion about a thing, ‘the sun’, which is *a posteriori* interpreted as a description of a scene.

The formulaic opening clauses of orally transmitted jokes are closest to traditional fairy tales: they open the text, and introduce the major participant with respect to a temporal scene:

(11-136) Alb. (elicited example, E. T.)

Vjen një BUJK në qytet ...

comes a peasant in town

‘A peasant comes to a town...’

The typical predicates, however, are not those of existence, but rather the verb denoting motion, especially those with *hic*-deixis, like ‘come’, ‘arrive’, etc. The major participant is thus introduced onto a more concrete scene (in the example given above, it is ‘a town’).¹

In modern literary texts, major participants are, as Sasse (1996) rightfully observes, practically never introduced with a vS clause, but rather as objects of transitive verbs, or, even more frequently, via the mechanism of pragmatic pretense, as topical subjects in SV clauses. However, both in spoken and in written language, introductory vS clauses do exist, but they are confined to introducing minor, locally relevant discourse referents, a phenomenon which is in accordance with the prototypical discourse properties of vS subjects in my corpus (cp. 11.3.2.). The context of (11-137), which illustrates the introduction of a minor participant with a vS clause, is as follows: The story deals with a love between a girl and a knight; He is hurt in a battle; She finds him in a monastery, dying; He gives her a ring, so that she does not forget him, and tells her a story of how he got it by saving an old woman accused of being a witch, who gave it to him a sign of gratitude. The old woman, who is a minor character serving only to explain the existence of the ring, is introduced with a vS clause.

(11-137) MG (Kapandai, p. 27)

Mes stin kaliva zuse mja γRIA, jajula EKATOXRONITISA ...

inside in-the hut lived an old-woman granny hundred-years-old

‘In the hut lived an old woman, a hundred years old little grandma...’

Another common type of introductory clauses are those which introduce discourse referents on a particular scene, as opposed to the general scene of the discourse universe, which has been illustrated in the above examples. Consider the following example:

¹ Orally transmitted jokes seem to be the genre which most consistently preserves the introductory function of vS clauses and related constructions crosslinguistically – cp. Önnorfors 1997: 100ff. for their functions in German.

(11-138) SC (Andrić, p. 98)

Jednog jutra, dok je...Kirkor sedeo u dvorištu..., odjednom se pojavio UPRAVNIK...

one morning while AUX K. sat in yard suddenly REFL appeared governor

‘One morning, while ... Kirkor was sitting in the yard ..., the prison governor suddenly appeared ...’

The whole novel from which this example is taken deals with a prison in Istanbul; the governor is one of the major characters, which need not be introduced into the discourse universe. But, in this particular scene, placed on the location definable as ‘the prison yard early in the morning, etc.’, he is introduced, because his presence in this particular part of the discourse universe has not been established yet.

The restrictedness of vS clauses to introducing only minor participants (or major participants in new scenes) is reflected also in the extensions of the prototypical use of this construction, which is, to remind, to introduce participants on the scene. The most important extension is the introduction of states of affairs through an existential vS clause with an abstract noun or a pronoun modified with a relative clause in the subject position:

(11-139) SC (Vesti, p. 7)

Postoji spremnost međunarodne ZAJEDNICE da se dogovore modaliteti

exists readiness of-international community that REFL agree modalities

pružanja pomoći Crnoj Gori.

of-giving help to-Montenegro

‘There is a readiness on the part of the international community to reach an agreement on the modalities of supporting Montenegro.’

The state of affairs ‘the international community is ready to ...’ is first introduced in the text by this existential clause. Its referent represents a minor participant in the journal article in which it appears (the main topic being the relationship between Montenegro and Serbia). The device of presenting states of affairs as (minor) participants in the discourse seems to be most frequent in the journalistic style, especially in SC, where I have some 103 examples of this kind. This clause type is practically restricted to existential verbs meaning ‘be’ (Alb. *është*, MG *ime*, SC *biti*), ‘exist’ (*ekziston*, *iparxo*, *postojati*), ‘happen’ (*ndodhet*, *jinome*, *desiti se*), etc. Its extreme frequency in SC journals and magazines is for a greater part responsible for the slightly higher number of vS in this language. Alb., on the other hand, employs this device somewhat less often than the other two languages (only 32 examples), which is presumably a consequence of the general tendency in this language to avoid the use of abstract nouns (cp. 5.1.).

Introductory vS clauses represent the only vS subtype where the subject does have to be new in a relevant way, either completely unknown to the hearer or not present in the given stretch of discourse. This restriction is triggered by the nature of the speech act of introduction itself: one does not assert the existence of the elements on a particular location if their presence there is already given contextually. However, there is a possibility for (rhetorically skilful) speakers to obviate this restriction in order to achieve a specific effect. Consider the following example:

(11-140) Alb. (Kadare, *Ëndrra mashtruese*, p. 146)

Ne zgjuam herët. ... Jashtë kishitë rënë dëborë. Në jetën tonë po hynte DËBORA.

we woke-up early outside had fallen snow in life-the our PTCL entered snow-the

Ne u veshem dhe dolëm përjashta...

we PASS dressed and went outside

‘We woke up early ... It had snowed outside. The SNOW was entering our lives.

We got dressed and went outside...’

Now, the snow is unequivocally ‘given’ in this passage, but it is nevertheless introduced by an introductory vS clause. The effect the writer achieves is that of direct perception: the reader has the impression that s/he observes the snow entering the lives of the two characters with their eyes, not with her/his own. Let us call this **perceptive effect**. How does this effect come into being? I should like to suggest, as many times in this study, a solution in terms of pragmatic enrichment (Sperber and Wilson 1986). When the narrator introduces an entity which is already there, the hearer assumes that this break of the pragmatic principles of saying no more and no less than necessary (4.1.) implies that s/he has to enrich the message so as to make it informative. In this particular case, two things are done. First, the unitary concept of the first person narrator is cleft into the ‘narrating narrator’ and ‘experiencing narrator’ (cp. Rubovitz 1999 for the notion of narrator deconstruction). The statement in which ‘snow’ first appears is ascribed to the former, the second statement, in which the already given snow is introduced again, to the latter. The effect of immediate perception is thus achieved by forcing the reader to change the perspective, and create the new scene and the new narrator notion. Second, and less important in this context, ‘snow’ receives additional connotations (coldness, desperation, etc.), so that its reintroduction brings about a slightly changed sense.

Introducing discourse referents presupposes the existence of a scene on which they appear (explicitly given or construed by the hearer), which is the reason why I grouped introductory vS clauses together with the descriptive ones, into the class of utterances whose primary topics are locations. The restriction to minor discourse referents, or locally relevant major

discourse referents, reflects the deeper principles of the use of vS clauses. The subject must not be topicworthy, at least not in the sense of actual topicworthiness (cp. 4.2.3.). Since major participants are per definition topicworthy, they are preferably introduced in other ways (see 11.6.2. for the lack of actual topicworthiness of vS subjects).

11.5.2.2. vS clauses with situational topics

In Section 4.4.2., a type of focus labeled **identificational focus** has been introduced. In the assertions conveyed by an identificational focus, the alternative possibilities excluded by the focus are limited as to the ontological class they belong to. The principles according to which this focus type works have been demonstrated on the following example ((4-32), repeated as (11-141)):

(11-141) Peter wanted to go home. But he couldn't get in: *he had lost his KEYS*, and now he was standing helplessly in front of the door.

The presuppositional structure of the italicized sentence is analyzed as follows:

(11-141') [[he]_{Topic} takes part in X]_{Topic} [X = had lost his keys]_{Focus, asserted}

The speaker presupposes that the topic referent takes part in a certain state of affairs (X) and makes her/his assertion by identifying this state of affairs with the propositional function 'lost his keys'. The number of alternatives excluded by the assertion is thus limited as to the type (not the token) to which the focus denotatum belongs: it can be only a state of affairs, not an entity, or a property.

The undetermined element of the relationally presupposed part of the proposition in the scheme given in (11-141'), marked with X, is what I mean when I speak of situational topics. The speaker presupposes that there is a certain situation, a state of affairs; the hearer is instructed to relate the asserted part of the proposition to the mental representation of this situation by identifying the former with the latter. Formally, situational topics usually do not surface by means of a topic expression, although this possibility is not excluded. I shall have more to say about situational topics in 11.6.1.

Utterances with situational topics very often have additional topics, in the way the utterance conveyed by (11-141) has: *he had lost his keys* is, in the given context, meant to assert something about a presupposed situation in which Peter takes part. The relational presupposition in such a constellation is the same as with all multiple topic clauses (cp. 4.3.2.): 'There is a relevant relationship between Peter and a situation'.

Discourse functions performed by the clauses with situational topics are called annuntiative, interruptive, and explanatory in this study. Following Sasse (1996), I shall assume that the difference between the three functions is derivable from the strength of the presupposition of the situational topic, being the weakest in annuntiative contexts and the strongest in the explanatory ones.

11.5.2.2.1. Annuntiative function

It is a commonplace in the linguistic literature on information structure that ‘thetic’ utterances (or however one chooses to call them) are of necessity used in those contexts in which, to put it bluntly, there is no context. The terms used for the assumed lack of a common ground between the speaker and the hearer are numerous, ‘all-new context’ and ‘out-of-the-blue context’ being the most frequent ones. A vS clause uttered out of the blue is felt to announce the existence of a state of affairs. The prototypical instances of annuntiative vS clauses are newspaper article titles, ads, and exclamations (see Sasse 1996). Here are some examples:

(11-142) MG (Kathimerini, p. 7)

Anazopironete o polemos ja ta METALAγMENA

starts-to-burn the war about the changed

‘The war about genetically engineered cotton is breaking out.’

(11-143) Alb. (Koha Ditore, p. 4)

Përsoset logjistika e policisë SHQIPTARE

is-completed logistics the of-police-the Albanian

‘The logistics of the Albanian police on the way of completion’

(11-144) SC (Vesti, p. 18)

Napadnuti RUSI

attacked Russians

‘RUSSIANS attacked’

All three examples are newspaper article titles, according to the common analysis, the prototypical utterances without any context. My analysis is different: When writing a title, journalists rely on the (often erroneous) expectation of the readers that the information furnished by the newspaper is relevant for the current state of affairs in the town, in the country, or in the world. Therefore, every article title is based on the presupposition ‘there is a relevant situation X’. The denotatum of the vS clause is then used to make an assertion about this relevant situation. For instance, (11-142) may be informally paraphrased as ‘as far as the

current situation in Greece is concerned, a war about genetically engineered cotton is starting’.

A similar analysis may be applied to ads, rules, and different official statements:

(11-145) SC (Vesti, p. 21)

Potreban KUVAR

needed cook

‘COOK needed.’

(11-146) MG (Elefterotipia, p. 5)

Apayorevete i singendrosi tu elenxu perisoteron meson enimerosis.

is-forbidden the concentration of-the control of-more means of-information

‘The concentration of more means of public information (in the hands of one person) is forbidden.’

A person reading a help-wanted page in a newspaper ((11-145)), or reading an note attached on a wall, is assumed to have an expectation ‘there is a state of affairs X pertaining to finding a job, an apartment, etc.’; the vS clause used identifies the element ‘X’ in the mental representation of the reader. *Mutatis mutandis*, the same obtains for government decrees, like (11-146), and for other typical instances of annuntiative vS clauses not mentioned here.

My contention is, to make it explicit, that annuntiatively used vS clauses are not without a context and do not represent topicless statements: They are rather a result of a complex negotiating process between the speaker and the hearer, with the speaker adjusting the message to the assumed expectations of the hearer.

A potential problem for this analysis are exclamations and other annuntiative statements uttered outside of the expectation frames established by a newspaper, a law text, etc., since their function is to express a surprising, unexpected event, so that it is difficult to see how they could represent assertions about a presupposed indefinite situation. Consider first some examples:

(11-147) Alb. (Koha Ditore, p. 11)

Kujdes, erdhi RRYMA

attention came electricity-the

‘Watch out, the ELECTRICITY came.’

(11-148) MG (Matesis, p. 37)

Erxete i dosiloji!

comes the collaborator:FEM

‘The COLLABORATOR whose is coming!’

The Alb. example is supposed to be a jocular announcement of the fact that the people of Kosovo, who had not had electricity for two years, were to get it in a month. The MG example is an exclamation by the women standing in a queue for food in a Greek town during the Second World War, uttered when they saw a lover of an Italian officer coming.

Let us first analyze the Alb. sentence. It contains, as many exclamations do, an explicit sign that there is a state of affairs which the speaker is to use as an assertional base to process the statement: *kujdes*, ‘attention’, is an indirect instruction to the hearer to create an expectation of the existence of a relevant state of affairs, and to identify the following proposition with the empty slot in this expectation.

The MG example is not so straightforward. I should like to argue that the women queuing must have signaled in some way that a situation which should be identified was at hand, by looking in the direction of the hapless woman, or by any other means of the body language.

These two cases cover, so I think, the exclamatory and similar subtypes of annuntiative vS clauses: the existence of a situation is evoked in the mind of the interlocutor either by linguistic means (‘attention!’) or extralinguistically, through gestures or otherwise.

The predicates used in annuntiative vS clauses are, as illustrated by the examples, much more diversified than the ones employed when a vS clause has a locational topic, where the existentials/positionals are the dominant type. Most of the predicates used receive a dramatic connotation, due to the type of context in which annuntiatives occur. The subjects may be ‘old’ or ‘new’. The use of the type is productive and frequent in all three languages, or at least I was not able to establish a difference between them.

A short polemical note before turning to the remaining two functions. In the greater part of the literature I sifted through, especially that of generative provenience, out-of-the-blue contexts are treated as the acid test for theticity (cp., e.g., Alexiadou 1996, 2000, Pinto 1997, etc.): If a sentence occurs in such a context (or, better, without a context), it must be ‘thetic’, or ‘sentence focus’, or similar. Now, the relationship between context and presupposition is not one-to-one, as I have repeatedly indicated (cp. 4.2.4., 4.2.5. and 4.6.). Speakers may presuppose more or less than the context allows them to. Consequently, a zero context does not imply zero presuppositions. In fact, as has been shown in Section 8.3.2.2., newspaper article titles, especially in MG, often have very extensive presuppositions, with only one argument in focus. The assumption that every clause in a minimal context has a broad focus construal of the verb-subject complex, to translate it in my terms, is simply wrong.

Broadly construed vS focus domains are only one of the possibilities to encode a proposition in the contexts with a minimal common ground. The decision depends on the

effect one wants to achieve: If the topical situation is to be presented as dramatic, unexpected, etc., a vS clause will be used, for the same reasons as those given in connection with descriptive vS clauses. If the speaker intends to directly relate the assertion to a previous knowledge of the auditor which is more specific than the simple expectation of a generally construed situation, some other construction is chosen.

11.5.2.2.2. Interruptive function

When an event appears on the scene suddenly, interrupting in this way an otherwise undisturbed flow of events, then there is a great deal of probability that a vS clause will be used. The typical context for interruptive vS clauses are larger narrative portions of discourse, with one or more topical discourse referents partaking either in a static situation or in a sequentially ordered states of affairs. The prototypical situations in the modern discourse are a door that opens, a telephone that rings, a person suddenly coming in, appearing, etc.

(11-149) SC (Kiš, p. 33)

Sada su ćutali ... i osluškivali ... talase ... bat koraka ... i dugo klizanje
now AUX were-silent and listened waves ... echo of-steps and long sliding
teških lanaca. Otključaše se vrata. Tri čoveka napustiše svoje boravište ...
of-heavy chains unlocked REFL door three men left their cabin
'Now they were silent ... they listened to the waves ... the echo of the steps ... and
the long gliding of the chains. The door unlocked. Three men left their cabin...'

This is a perfect example of an interruptive vS clause: The situation with three men captured in their cabin listening to the sounds from the dock is interrupted by the action of the door being unlocked. My interpretation of this discourse function of vS is probably already predictable: the preceding state of affairs, the one which is interrupted, creates an expectation on the part of the hearer that there exists a situation following it. Upon hearing (or reading) the sentence 'Now they were silent...', an expectation is created, which can be represented as 'the situation of being silent and listening is followed by X'. The assertion is, as in all contexts with a situational topic, of the identificational type: 'X = the unlocking of the door'.

Now, not all descriptive or narrative passages necessarily open this kind of expectation. A very typical type of those which do is the one represented in (11-149): Static events, which do not seem to bring the narration further, regularly point out to a sequence, by virtue of being an anomaly in a narrative text, which consists of sequentially ordered dynamic events. However, the speaker may also use more explicit markings of the cataphoric value of an event. Frequently, it is a temporal clause, as in the following two examples:

(11-150) Alb. (Kosovarja, p. 44/2)

Ne me shoqen time morëm nga një pije dhe derisa po pinim duhan me të,

we with friend our took from a drink and while PTCL drank tobacco with her

më ra në sy një DJALË, i cili m'u duk si fytyrë e njohur.

to-me fell in eyes a boy who me PASS seemed as face the known

Pas disa minutash ai la shokun e tij vetëm dhe erdhi ...

after some minutes he left friend the his alone and came

‘My friend and I took one drink each and while I was smoking with her, I noticed a BOY, who looked somehow familiar to me. Couple of minutes later, he left his friend and came over...’

(11-151) MG (Matesis, p. 66)

To Fanuli kitakse ti mana mas ... ke proxorise kata tis patates. Ke otan tis pire ke

the F. looked the mother our and went-on for the potatoes and when them took and

tis tris sti xufta tu, saltari kato o xamojelastos JERMANOS ... ke tu varai ... to xeri.

the three in palm his jumps down the smiling German and to-him strikes the hand

‘[A German soldier standing on the truck, smiling, told us not to try to take any potatoes. We stood there motionless while the soldiers were collecting the potatoes. Suddenly, my dear little Phanos broke free from my arms and ran towards the potatoes. We were all paralyzed.] Phanos looked at our mother ... and went on for the potatoes. And when he took three of them in his hand, the smiling German jumps down (from the truck) ... and hits him on the hand.’

Marking a context as potentially cataphoric, i.e. with an expectation of a state of affairs ‘X’ following, is in both these cases made explicit by a temporal clause: ‘while we were smoking’ and ‘when he took three of them..’ are syntactically so geared that they make the hearer expect something, encoded as the main clause, to follow.

One more syntactic context is typical for interruptive vS clauses, *cum-inversum* sentences, a construction defined as assertive (as opposed to presuppositional) in Section 9.2.2. Here is an example from the parallel translations of Jack London’s novel:

(11-152) Eng: *Henry was bending over and adding ice to the bubbling pot of beans when he was startled by the sound of a blow, an exclamation from Bill, and a sharp snarling cry of pain from among the dogs.* (London, p. 10)

Alb: ... *kur befas prapa shpinës së tij u dëgjua një goditje, thirrja e Billit*
when suddenly in-front of-back the his PASS heard a blow cry-the the of-B.
dhe një piskamë e fortë qensh. (p. 11)

and a shrieking the strong of-dogs

MG: ...*otan ton ksafniasan enas γdupos, ena epifonima tu Bil ke mia makrosirti*
when him surprised a blow a shouting of-the B. and a long-drawn
ponemeni kravyi apo tin ayeli ton skilion. (p. 16)

painful cry from the pack of-the dogs

SC: ... *kad ga trže pljesak nekog žestokog udarca i Bilov uzvik, praćen*

when him roused splash of-some strong blow and of-B. cry followed

oštrim urlikom i bolnim režanjem iz sredine pasa. (p. 14)

by-sharp howling and painful snarling from middle of-dogs

Back to the original examples. Sentence (11-150) illustrates one frequent feature of interruptive vS clauses: Apart from a situational topic, these clauses often contain another, explicit topic expression, being thus based on a more complex presupposition than the simple ‘there is a relevant X’. In (11-150), the additional explicit topic expression is the indirect topic *më*, expressed with a clitic free dative, so that the relational presupposition of the clause reads approximately as ‘there is a relevant state of affairs X, which is related to ‘I’’. The assertion is, as always, the identification of this relationship between X and the referent ‘I’.

Yet another important issue is revealed by (11-151): Subjects of interruptive vS clauses, unlike, for instance, those in the descriptive or introductive ones, may be ‘given’: the subject referent, a German soldier who smiles all the time, has been mentioned some five sentences before the vS clause in which it is the focal subject. Major discourse referents also occur: ‘the door’ in (11-149) is a prominent, ‘fourth’ member in the group of three prisoners who are the main characters of the story, being repeatedly locked and unlocked by the guards. This is, of course, not to say that the subjects *must* be given, major participants, etc., but merely that this possibility is applicable without giving rise to special effects like the perceptive effect described in 11.5.2.1.2. This feature, possible also in annuntiatively used vS clauses (though, understandably, less prominent), has to do with the topic type used: Whereas it does not make sense to introduce discourse referents which are already present on the location about which an assertion is made, the identification of a situation expected on the basis of the context clues may be done with both ‘given’ and ‘new’ discourse referents involved in the action. Thus, interruptive and, especially, explanatory vS clauses described below are responsible for the

majority of ‘old’ or ‘major participant’ subject referents found in my corpus (cp. 11.3.2. for the statistics).

As the examples adduced by now show, there are little or no restrictions (apart from those stemming from focus construal) as to the semantic class of the predicate. The only condition a verb has to fulfill to be used interruptively is that it has to be liable to dramatic interpretation, as described in 11.3.3.2.1.

The affective character carried by the vS clauses used interruptively may be exploited for special effects, i.e. for creating an impression of an interruption of a situation which has not been described at all. For instance, the use of a vS clause out of the blue (i.e. without a previous description of a situation which is interrupted) creates the impression that there had existed a nice daily routine before the interruptive event happened. Consider the following example:

(11-153) SC (Kiš, p. 173)

Krajem decembra, dva dana posle hapšenja Novskog, u kući Darmolatova
on-end of-December two days after arrestment of-N. in house of-D.

zazvonio je TELEFON.

rang AUX telephone

‘[The Soviet poet Darmolatov was under informal protection of the hero of the October Revolution, B. Novsky, a fact which can be traced back to their days together in Berlin.] Towards the end of December, two days after Novsky had been arrested, the TELEPHONE rang in the house of the Darmolatov family. [Darmolatov’s wife ... picked up the receiver.]’

There is nothing in this passage that could be disrupted by the ringing of the telephone. The text preceding the vS clause deals with the description of an informal institution of ‘relationship’ between poets and politicians in the early years of the Soviet Union, exemplifying its functioning with the relationship between Novsky and Darmolatov. The use of a vS clause in this context, with a clearly dramatic connotation carried by the verb, implies that there had been a peaceful life going on in the house of Darmolatovs before the telephone rang. The principle according to which this reading comes into being is the same as in all cases of reinterpretation: Through pragmatic enrichment, the reader creates a context in which the interruptive clause has its proper place.

11.5.2.2.3. Explanatory function

Much of the literature dealing with verb-subject order and related constructions mentions the explanatory function as one of the most prominent discourse functions of the construction (McCanna 1973, Wehr 1984, Ulrich 1985, Sasse 1987, 1996). Explanatory vS clauses work in the way typical for identificational foci: A presupposed situation X is identified through the assertion of the information contained in the verb-subject focus domain.

The difference between the explanatory and the previous two types lies in the explicitness of the presupposition. Whereas annuntiative and interruptive vS clauses appear in the contexts where the clues for the relational relevance of the situation X are weak, or even established only through pragmatic enrichment, explanatory contexts give explicit indications to this effect. A useful classification of explanatory contexts is given by Sasse (1996: 36): (1) pragmatically incomplete states of affairs begging for a continuation; (2) questions of the type ‘What happened?’ or ‘Why do you do X?’; (3) extralinguistic means; (4) a combination of the three. The examples that follow are intended to illustrate these presupposition-building devices:

(11-154) SC (Vesti, p. 31)

*Zakukao je posle meča direktor ‘Budućnosti’ Danilo Mitrović, iako je njegova
wailed AUX after match director of-B. D. M. although AUX his
ekipa pobedila. “Zeznuo me je naš trener NIKOLIĆ.”
team won made-a-fool me AUX our trainer N.*

‘The sports director of the basketball team ‘Budućnost’ almost started to whine after the match, even though his team had won. »Our trainer, Nikolić, made a fool of me.«’

(11-155)MG (Ciao, p. 54)

*V.: – “Jati se afise, vre peði mu!?” E.: “Fanike mja ALI.”
why you he-left, PTCL child my appeared one other*

Valeria: “Why did he leave you, you poor thing?” Emmanuela: “Someone else appeared in his life.’

(11-156) Alb. (Bishqemi, p. 8)

*Ndoshta Partinin e pickoi ndonjë PLESHT nga ata të dhelprës sime...?’
maybe Partin-the him stung some flea from those of-the fox my*

[A boy came to school and found out that he had fleas, presumably from a fox he had found the day before. The boy sitting next to him, Partin, starts to behave strangely.] ‘Maybe Partin was stung by one of the fleas from my fox.’

The SC example illustrates the way pragmatically incomplete states of affairs evoke presuppositions: one usually does not begin to wail when one's team has just won a match. Therefore, the situation itself implies that there is a relevant state of affairs X which explains this behavior. The MG sentence represents the second presupposition-building device: a question is asked about a state of affairs. The extralinguistic means of implying the existence of a relevant situation to be identified are illustrated by the Alb. example: when a person acts strangely, there is probably a reason to it.

The second point of difference between the explanatory and the other vS clauses with a situational topic is the type of the presumed relationship between the presupposition-evoking and the presupposition-identifying state of affairs. In interruptively used clauses, for instance, the implication is that of temporal sequence, or of partial temporal overlapping: 'there is an X relevant to the present situation, and this X is in a temporal relationship to it'. In the explanatory clauses, the implied relationship is a causal one: 'there is an X relevant to the present situation, and this X causes it, or at least explains it in causal terms'.

Even more often than the interruptive vS clauses, the explanatory ones contain additional topics, usually human referents denoting an experiencer, a possessor, or even a patient within the state of affairs conveyed. For instance, (11-154) is based on the complex relational presupposition 'there is a causally relevant state of affairs X in which the trainer D. M. is involved', with the referent 'the trainer D. M.' encoded with a clitic pronoun *me* ('me'). A situation type very frequently encoded by the explanatory vS clauses are different psychological, emotional and the events related to the body, like 'I was overcome with tears', 'my foot hurts', etc. The reason for the general frequency of additional human topic, and, in particular, of the 'internal' events, is purely pragmatic in nature: the states of affairs which ask for future clarification are, quite in accordance with the 'egocentric principle' (cp. 4.3.2.), mostly concerned with us, or with the things most similar to us. Here is a nice example of an 'internal' event used in an explanatory context:

(11-157) Alb. (Bishqemi, p. 93)

U këput e ra symbyllur. Dëgjuam gërhimën e tij.

PASS got-tired and fell close-eyed I-heard snorting the his

E kishte kaplluar gjumi.

him had overcome sleep

'Exhausted, he lay down, his eyes shut. I heard him snorting: he had fallen asleep.'

This example illustrates an interesting typological feature of Alb., which sets it apart from the other two languages. A great number of states of affairs denoting some kind of ‘internal’ event are expressed with the topical experiencer/possessor in the object slot. The subject role may be assigned to the half-personified state of affairs itself, which then ‘captures’, ‘overcomes’, etc., the experiencing person, as in (11-157). Thus, the only possible way of saying that a person fell asleep is ‘me took the sleep’ (*më zë gjumi, më kaplloi gjumi*), and the most frequent one for saying that a person woke up is ‘me went-out the sleep’ (*më doli gjumi*). Or the subject role is assigned to a body part, which is presented as performing the action. For instance, it is often not the person that sees something, but rather the eye (‘my eyes saw/caught...’, *më panë/zunë sytë*, cp. (11-60)); the lips laugh, the ears hear, etc. This tendency to express ‘internal’ states of affairs with body parts and event nominalizations as the subjects of vS clauses is reflected in the statistics (cp. 11.3.2.) as a significantly higher number of the subjects of these ontological classes in Alb. than in the other two languages. This is, of course, not to say that MG and SC do not employ this way of encoding ‘internal’ events, but it is restricted to a couple of situations, most notably those of pain (*ponai to KEFALI mu, boli me GLAVA*), whereas in Alb. it seems to be a very productive pattern.

Let us now return to the discourse properties of explanatory contexts. One more point is worth noting. As probably every explanatory device in natural languages, vS clauses may be used not only to explain a state of affairs by identifying its cause, but also in order to explain the reason of the speech act performed (‘I said it because...’), or to make the epistemic source of the speech act explicit (‘I know this because...’). The following example illustrates the latter type:

(11-158) MG (Matesis, p. 65)

O kirjos Ljakopulos molis to idē afto, kati epaθe, ton piyan mesa

the Mr. L. as-soon-as it saw that something suffered him lead inside

sikoton i KORES tu ...

raised the daughters his

‘Mr. Liakopoulos got sick when he saw that, he went inside leaning on his daughters...’

The message here reads approximately as follows: ‘He felt ill’ > ‘You may wonder on the basis of what fact (X) I know it’ (presupposed) > ‘X = his daughters took him in’ (asserted).

Explanatory contexts are a major environment for vS clauses in all three languages. My impression is that Alb. uses vS in this function somewhat more often than MG and SC,

probably due to the above described way of expressing internal events (more on this in 11.6.2.).

At the end of the section devoted to the discourse functions of vS clauses, I should like to note two points. First, the notion of situational topic may seem problematic to some. It is, I hope, indisputable that there is something similar to what I call the presupposition of a relevant situation, X. What is more contestable is my use of the term ‘topic’ to denote this rather abstract entity. I shall try to justify this usage in Section 11.6.1. in more detail. Suffice it for now to say that, within the system outlined in Chapter 4., everything that delimits the set of possible worlds to be eliminated by the assertion counts as topic, which unequivocally qualifies the pragmatically construed element X for a sentence topic.

The second point is more trivial. The postulation of discrete discourse function is to be understood as a heuristic strategy rather than as a description of a true ‘entity’, a ‘thing’. One and the same clause may be used with a number of functions (cp. Sasse 1996: 39 for some nice examples), and it is not always possible to determine what kind of function a clause performs in a particular context. The following example illustrates the latter point:

(11-159) SC (Kiš, p. 31)

Trećeg dana, ... budi se iz košmarnog sna: na uskoj klupi nasuprot njegovom

on-third day wakes REFL from nightmare on narrow bench opposite his

ležaju sede dva ČOVEKA i ćutke ga posmatraju.

bed sit two men and silently him watch

‘On the third day, ... he awakes from a nightmare: two men are sitting/sit on a narrow bench opposite to his bed and watching/watch him silently.’

The uncertainty as to the translation of this sentence (progressive vs. simple present) reflects my uncertainty of how it is to be interpreted, as an explanatory clause (he woke up because they were sitting there) or as a simple descriptive clause (describing the location he sees when he wakes up). Although cases of this kind are not frequent, they confirm that discourse functions are a matter of pragmatics, i.e. of interpreting utterances with the help from the context clues.

11.6. Three essential properties of vS clauses: An attempt of interpretation

In the preceding five sections, the syntactic, semantic, and discourse-pragmatic features of vS clauses have been investigated in some detail. In this section, I shall try to account for them by identifying and describing the properties which seem to be essential for the construction. Since I shall argue that these properties – untypical topics and untypical subjects – are the sufficient, although not necessary, conditions for the appearance of a vS clause in the discourse, this section could also be named **the triggers of vS**. The discussion is confined only to the cases with broad focus construal, both because the narrow focus clauses have been sufficiently dealt with in Section 11.4., and because it is only the instances of vS with broad focus that have given rise to the idea that there is such a thing as theticity.

11.6.1. Untypical topics

In the sections devoted to the formal, semantic, and informational properties of vS topics (11.2.2., 11.3.1.), it has been noted that their basic property is the lack of statistical topicworthiness and the presence of the actual topicworthiness. More simply, the typical situation for a vS clause is that, in which the speaker decides to choose as the topic that kind of discourse referent which s/he normally does not employ in that function. This impression is only confirmed by the statistical tendency of vS topics to be only locally relevant in the text, i.e. not to be major participants. Further, in Section 11.5., it has been postulated that there are two basic topic types appearing in vS clauses, locationals and situations, and that they can be combined with further topics, more often the latter than the former. In what follows, I shall substantiate this claim by going into more detail and demonstrate how locations, situations & Co. determine the choice of vS. A careful reader will recognize the influence of Joan Bresnan's work on locative inversion (Bresnan 1991, 1994), especially in the part devoted to locational topics.

11.6.1.1. Locational topics

The first thing to do in the discussion of the notion of locational topic is to define location as opposed to thing, or entity. Is the difference ontological, or lexical, or is it rather anchored in the propositional semantics? Let me try to answer this question by introducing two important theoretical contributions to the debate.

Lyons (1977) points out that nouns denoting places have two readings, an entity-like one and a location-like one. Thus, the denotation of *London* is ambiguous between ‘thing’ (a rather large set of buildings, streets, squares, etc.) and ‘place’ (a location at which buildings, squares, etc. are situated). Now, it is true that the prototypical cases of this ambiguity are the nouns denoting place names (*London*), landmarks (*hill, street*), and similar, but the fact is that it is found with practically all entity-denoting expressions: *table* may denote a thing which has four legs and serves to be covered with useless piles of paper, or it may denote a place where, say, my red ball-point pen is hidden under the above mentioned piles of paper. This implies that it is impossible to define a location on a purely ontological or lexical basis (place names and landmarks, etc.), since virtually everything can be semantically construed both as an entity and as a location.

The solution I should like to embrace was first proposed by Jackendoff (1987) and further elaborated by Bresnan (1991, 1994). According to Jackendoff and Bresnan, a location is composed “of an object as reference point and a place or path function which takes the reference point as an argument ... [A] place NP designates only the reference point for a semantic place or path function. These functions are either designated by prepositions or incorporated into verb meanings, as with certain intrinsically locative verbs (*enter, climb, pass*)” (Bresnan 1994: 116). Location is, to summarize, a phenomenon defined on the level of propositional semantics: A composite notion in which a reference point, a thing, is related to an abstract place/path function. This straightforwardly accounts for the potential ambiguity of most entity-denoting expressions in the languages of the AME type: Every entity may serve as a reference point bound by a place or path function.

What is even more important, Jackendoff’s and Bresnan’s semantic theory of locatives, combined with Lyons’ lexical ambiguity hypothesis, may help account for a number of issues associated with vS clauses.

(1) **Why indirect topics?** When the speaker decides to make an assertion about a locational topic, s/he cannot encode it as the prototypical direct topic expression, subject – not in the languages under discussion in this study, that is. Alb., MG and SC, like all other familiar European languages, encode locations per default as propositional phrases or, if pronominal, as adverbs.¹ This is a direct consequence of the fact that the semantics of locations is complex, which surfaces in the fact that the reference point is expressed

¹ In the earlier stages of these languages, yet another possibility existed, namely the use of non-core case markings, e.g. locatives. Since this is not a live option in modern Alb., MG and SC, I shall ignore it in what follows.

nominally, whereas the preposition encodes the place/path function; if the reference point has no descriptive content, both components are merged into adverbs like *there, here, thither*, etc. In any case, due to the grammatical interpretation of the semantic structure of locations in this language type, it can practically never be encoded as direct topic, because the place/path function would remain unexpressed.

There are, as the quotation from Bresnan (1994) above shows, some exceptions to this, with the verbs which incorporate the place/path function in their meaning, like *enter, climb*, etc. However, at least as far as Alb., MG and SC are concerned, this is a marginal phenomenon, for two reasons. First, these three languages predominantly belong to the group exemplified by Germanic in Talmy (1985), i.e. to those languages in which the verb does not incorporate the place/path function in its semantic structure. The number of instances is even smaller than in English.¹ Second, even in these few cases, it is invariably the object role that the locations are assigned. Since passivization is excluded, locations are virtually never encoded as subjects, but at best as objects. The objects of verbs like ‘pass by’ or ‘surround’ are thus the only locational topics that may appear as direct topics. The phenomenon is marginal to the extent that it may be neglected *salva veritate*.

(2) **Are there ‘place’ expressions which can be encoded as subjects?** The question refers to the cases like the one quoted in 11.5.2.2.1. (*The sun SET*), and even more to sentences like *London is big* or *My house has small windows*. It is obvious that the semantic structure proposed by Jackendoff (reference point + place/path function) is not given, but the subject referent sounds somehow ‘placy’. The lexical ambiguity hypothesis proposed by Lyons explains these facts straightforwardly (cp. also LaPolla 1995): ‘the sun’, ‘London’ and ‘my house’ are conceptualized as things, not as locations, in the same way ‘the table’ regularly is. The impression that we are dealing with locations here stems from the fact that the referents of ‘London’ and ‘house’ are much more often incorporated into locational structures and thus conceptualized as locations than they are construed as things. The prototypical use is then, via mechanism described in 11.4.1., interpreted as the only one possible. In sum: place names, expressions denoting landmarks, and similar, do not represent locations when encoded as subjects.

(3) **How are locations construed as topics?** The semantic structure assumed for locations explains why they have to be construed: ‘house’ in the sentence *There is a mouse in my house* indeed refers to an entity, but the hearer has to ‘extract’ it out of the locational

¹ I have found only a handful, and even these only occasionally take direct location arguments, e.g. ‘pass by’ (MG *pernao*, SC *proći*), ‘surround’ (Alb. *rrethon*, MG *perivalo*, *peristixizo*), etc.

configuration, in which it is bound by the place function, in order to be able to assess the assertion ‘exists a mouse’ with respect to it.

(4) **What about topical locations without overt topic expressions?** The situation is even simpler here: both the location itself and the entity serving as a reference point (the actual topic) have to be construed by the hearer. If an introductory sentence like *A TALL man came in* is uttered, the hearer first construes a location (for instance ‘in the room’) and then ‘extracts’ the entity out of the locational structure (‘the room’) and uses it as the topic with respect to which the assertion ‘a tall man came in’ is assessed.

(5) **Why are locational topics typologically relevant in assessing the issue of theticity?** As will have probably become clear by now, I wish to argue that one of the reasons why the subjects of vS clauses are included into the focus domain and not chosen as the topic expressions lies in the fact that what is chosen as the topic, a location, cannot be encoded as subject. It is important to understand that this fact is, however self-evident it may seem to a speaker of a European language, a language-specific matter. As Bresnan (1991, 1994) has convincingly shown, Chichewa and other Bantu languages satisfy the semantic structure of locatives in quite a different way, by assigning the nouns which denote the reference point to the locative gender, or class, so that they are still perfectly capable of being encoded as subjects, as in the following example (from Bresnan 1994: 93):

(11-160) <u>M</u> <i>nkhalango</i>	<u>mwa</u> - <i>khal-á</i>	<i>mí-kângo</i> .
18 9.forest	18.SUBJ-PRF-remain-FV	4-lion
‘In the forest remained lions.’		

The underlined morphemes represent the classifier (*m-*) and the verb agreement morpheme (*mw-*), which show that ‘in-forest’ is the canonical subject triggering agreement, even though it is semantically a full locational. Other procedures are conceivable as well. A language may incorporate the place/path function into the semantics of the verbs on a regular basis. One procedure to do this is to ‘raise’ locational expressions to the subject position through creation of complex verbs via incorporation, so that the complex verb now obviously also incorporates the place/path function. This seems to be a rather regular pattern in Chukchee, if I interpret the data from Polinsky (1990, 1993) right.

This admittedly modest evidence points out to a possible solution of the theticity problem, or at least a part of it: Instead of postulating new cognitive categories, it is perhaps more reasonable to look for the explanation of the ‘marked’ constructions like vS in the language-specific solutions to specific problems. One of the problems vS solves in the languages of the

Balkans is how to make assertions with respect to locations; other languages may have other solutions, as the evidence from Africa and Siberia seems to show.

11.6.1.2. Situational topics

The notion of situational topic may, as noticed in the concluding note to Section 11.5., seem a strange and inappropriate extension of the already overloaded term ‘topic’. In this section I shall argue that it is in full compliance with the notion of topic embraced in this study, and that its semantic properties, or rather its lack of semantic properties, makes it a good candidate for a trigger of vS.

Recall that topics are those elements of the common ground which delimit the portion of the common ground which is to be changed by the assertion, and which are assumed by the speaker to be the optimal instruction to the hearer what the actual purpose of the utterance is. Recall also that topics need not be definite or generic: it suffices for the information processing that the hearer is ready and able to accept them for the purposes of the current information exchange (cp. Section 4.3.). The element X, a situation whose existence and relevance are presupposed by the speaker in the way which has been described in detail in 11.5.2.2., has all these properties: It gives an instruction to the hearer to what part of her/his background knowledge the assertion is to be related, and it reveals to her/him what intention the speaker is lead by in uttering a sentence. The speaker sends a signal which reads approximately as follows: ‘There is a situation which is relevant with respect to a point of current interest’. The context clues, as described in 11.5.2.2., give more precise instructions on the way the situation presupposed is relevant to the point of interest, i.e. whether it explains it, or temporally follows it, or interrupts it, or some other type of relationship is involved. As indicated above, the fact that the presupposed situation has only a vague denotation, the only explicit semantic feature being its status as a state of affairs (as opposed to entity) does not diminish its capability of being topical.

As a matter of fact, speakers tend to combine situational topics with at least one more topic, usually denoting a human, or an animate referent, so that a multiple topic construction is created. In a sequence like *What’s the matter? – Më dhimbet KOKA* (Alb. ‘me hurts head-the’), *Ponai to KEFALI mu* (MG ‘hurts the head my’), *Boli me GLAVA* (SC ‘hurts me head’), the underscored experiencer/possessor expressions are indirect topic expressions referring to additional topics. The presuppositional structure is ‘There is a situation X which is relevant to the referent “I”’. By using the device of the multiple topic construal, the speakers achieve two things. First, they narrow down the reference of the situational topic, since X is not any old

situation, but the one in which ‘I’ is involved. Second, they make the instruction of the way X is relevant more explicit: the situation is already on the presuppositional level of processing related to an entity which is of current interest.

Let us now turn to the consequences the existence of this kind of topic has for vS clauses. Much of it has already been said with respect to locational topics, so that only what is specific for situational topics will be elaborated upon.

(1) **Why indirect topics?** The answer to this question is even more obvious than in the case of locational topics. Most situational topics do not belong to the proposition whose assertional basis they represent, i.e. most of them are not terms bound by the predicate at all. They are rather extrapropositional entities which are construed as mental representations on the basis of the context clues. Consequently, in most cases, they may not be encoded linguistically at all – not as subjects, not as objects, not even as indirect topic expressions.

However, there are instances in one of the three languages this study deals with, SC, in which a syntactically free element occurs which seems to represent a linguistic encoding of situational topics. The term ‘syntactically free’ refers to the fact that these elements cannot be ascribed any syntactic function (*pace* Progovac 1998), which is presumably a reflection of their complete lack of a semantic role. The element in question has the form of a demonstrative pronoun in the neuter: *to*. Consider the following examples:

(11-161) *Ko to kuca?*

who that knocks

[On hearing a knocking on a door] ‘Who’s knocking?’ (= ‘Who’s there?’)

(11-162) “*Šta je ovo?*” – “*To mi se prosulo MLEKO.*”

what is this that to-me REFL spilt milk

[Seeing a stain on the carpet] “What’s this?” ‘Oh, I spilt some milk.’ (‘Milk spilt to me’)

(11-163) *Čuo sam nešto: to je padala KIŠA.*

I-heard AUX something that AUX fell rain

‘I heard something: it was raining.’

The element *to* in these examples seems to exactly correspond to what I call situational topic (note that it occurs not only with vS clauses, like 0 and (11-163), but also in questions, like

(11-161)). Being void of semantic and syntactic functions, it is, of course, an instance of an indirect topic expression.¹ Nothing comparable to *to* is found in Alb. and MG.

In sum: situational topics are practically never direct topics, i.e. they are not encoded as subjects or objects, because they generally do not represent a term of the proposition.

(2) **Are there situational topics which can be encoded with direct topic expressions?** In one specific use of non-syntactic pronouns (cp. 6.4.1.), i.e. of those pronouns which do not take their morphological properties from an antecedent, since there is no such a thing, situational topics seem to surface in all three languages as topical subjects. After a series of states of affairs have been presented, often in the form of a sequence of descriptive vS clauses, a presupposition of the following form may come into being: ‘The situations A, B, C, etc., represent a unified situation X’. The assertion that follows identifies the element X, invariably with a stative, usually nominal predicate. The subject of this identifying clause, referring directly to X, is encoded as a non-syntactic pronoun (obligatory in SC, less so in Alb. and MG), or as zero (often in MG, less so in Alb., never in SC). Here is an example:

(11-164) Eng.: *Dark spruce forest frowned on either side the frozen waterway. The trees ... seemed to lean toward each other, black and ominous ... A vast silence reigned over the land. ... There was a hint in it of laughter, but of a laughter more terrible than any sadness ... **It was the Wild, the savage, frozen-hearted Northland Wild.*** (London, p. 3)

Alb: ***Kjo** ishte një vend i humbur, Shkretëtira e Veriut, e ngrirë gjer në palcë* (p. 3)
that was a place the lost, desert the of-north-the the frozen till in core

MG: *Itan i Ayria Fisi, i protoyoni ke payomeni mexri ta mixja tis Ayria Fisi*
was the wild nature the primitive and frozen till the inmost her wild nature
tu Vora (p. 7)
of-the north

SC: *Bila je **to** Divljina, pusta Divljina Severa, ledena srca.* (p. 5)
was AUX it wild waste wild of-north of-frozen of-heart

In Alb. and SC, the situational topic is encoded as a non-syntactic pronoun (*kjo*, *to*) and as a zero in MG. In all three languages it is the subject. This is probably the only context in which situational notions of the kind described above are assigned a grammatical role: the assertion pertaining to them, being inherently identificational (copula + predicative), is such that it does

¹ The conditions under which *to* appears and those under which it does not appear are not entirely clear. It certainly has a ‘perceptive’ flavor, as Progovac (1998) observes, since it is most frequently found in the context of situations which are directly observable. I am at present not able to say whether this is the whole story.

not require its arguments to have a semantic role. In all other cases, to repeat the rule, situational topics are either not encoded at all or only as indirect topics without a syntactic role.¹

This section has been devoted to presenting evidence for the existence of an entity I call *situational topic*, and to its ontological justification. The relevant conclusion, i.e. relevant for the issue of vS clauses, is that such topics are indeed operative in the language, and that they are, with the exception of identificational predicates, never encoded as direct topics, indeed only exceptionally encoded at all.

Excursus: Am I re-inventing the spatio-temporal argument?

In Section 2.6.3. a number of approaches to the vS phenomenon and related constructions has been mentioned which operate with the notion of spatio-temporal boundness. Some of them transpose the whole issue to syntax, positing a covert argument for certain predicates which denotes time and place. My locational and situational topics, and the indirect topics in general, may in certain respects resemble these notions, but they essentially represent a completely different kind of animal.

Ontologically, locational and situational topics are mental representations which come into being on the level of interpretation of utterances through the mechanism of pragmatic construal, with the hearer using context clues in order to construe an assertional basis to which s/he is to relate the asserted proposition. Locational and situational topics are thus **extralinguistic entities**, which *may*, but *need not* have a concrete linguistic realization: A sentence may contain an indirect topic expression giving a clue to the hearer what kind of topic s/he should construe, but in many cases it does not. In contrast to this, the spatio-temporal argument is a **syntactic entity**. The difference could not be greater.

The pragmatic interpretation of utterances with locational and situational topics is only partly similar to the one which would necessarily follow from the existence of a spatio-temporal argument. Namely, as indicated in 2.6.3., the ‘thetic’ clauses with a spatio-temporal argument are, not surprisingly, said to be spatio-temporally bound, which then implies that only those predicates which denote temporary properties may be used. Now, this is indeed the

¹ A little speculation: Bolinger (1977) has argued that the so-called dummy or expletive *it* found with meteorological and similar predicates in English is in fact a meaningful element, *referring* to the situation, or ‘ambient’ (the term used by Bolinger is *ambiental ‘it’*). I wonder if many ‘impersonal’ expressions in the Balkan languages, like *vrexī* (‘rains’) and *kani krio* (‘does cold’) in MG, could be interpreted as having a zero subject referring to ‘situation’, i.e. the element X in my notation.

preferential lexical filling of vS clauses with broad focus construal, but I should like to contend that this is merely due to the type of situations which usually occur when an assertion is made with respect to a location or to a situation. There is no *a priori* reason to exclude predicates denoting more time-stable eventualities. The only condition a predicate has to fulfill in order to be used in a vS clause with broad focus construal is that it has to be interpretable as relevant for the given topic. Thus, to repeat the example given in 2.6.3., the question ‘What’s wrong with my looks?’ may be answered with *T’ është e madhe KOKA* (Alb. ‘to-you is the big head-the’), *Ine meçalo to KEFALI su* (MG ‘is big the head your’), *Velika ti je GLAVA*’ (SC ‘big to-you is head’). The permanent property of having a big head is used to identify the presupposed situational topic X (‘there is an X such that it is relevant to me and it does not make me exactly enchanting’). Thus, the predictions following from the existence of locational/situational topics are different from those which would have to be generated by the spatio-temporal argument. My predictions are, to put it mildly, more realistic.

Thus, the answer to the (obviously rhetoric) question in the title of this section, whether I am re-inventing the spatio-temporal argument, is unequivocally no.

11.6.1.3. Other untypical topics

The investigation of the ontological properties of topic expressions in vS (Section 11.3.1.) has shown that, apart from locational, state-of-affairs, and similar topics, vS clauses contain a relatively high percentage of topic expressions referring to things and humans (Table (11-67)). The semantic roles the referents of these expressions play in the proposition are most commonly those of a patient/theme, experiencer, or an interested party of any kind (Table (11-65)). Syntactically, they surface as free datives, possessives, or objects (Section 11.2.2.). Only in the last case, they may be considered direct topics, since selected by the verb and thus explicitly marked for the relevance relationship to it. The first two types are the prototypical instances of indirect topics, whose relationship to the asserted proposition has to be construed by the hearer.

The way these additional topics (i.e., additional with respect to situational topics, with which they cooccur) work within discourse has been demonstrated in some detail in the sections devoted to interruptive and explanatory vS clauses: the existence of a relevant state of affairs is presupposed, with the additional topic narrowing down its potential field of reference and giving a further instruction to the hearer how to process the utterance. In what follows, I shall therefore concentrate only on one feature of theirs, that of their linguistic encoding.

Like locationals and situations, the additional topics have to fulfill the condition of not being assigned the subject role. Statistically, this is obviously the function of their semantic roles: patients, experiencers and other roles denoting an interested party are in the languages under consideration typically not encoded as subjects.

There is much literature on the mapping of semantic roles onto grammatical relations (e.g. Dowty 1991, papers in Van Valin 1993, Bresnan 1994, Levin and Rappaport Hovav 1995, to mention just a few), which I cannot summarize here. The basic tenet of all the work done in this field is that in a situation, it is as a rule the more agentive element of the two (if there are two) that is promoted to the subject position, at least in the language type to which the Balkan languages belong. Consequently, the topics of vS clauses are not subjects, since they are either the less agentive participants (patients, experiencers), or not involved into the competition for the subject position at all ('interested party'). This is an important point, since, if the additional topic qualifies for the subject role, by virtue of its agentivity or for some other reason, vS construction is not used. Let me illustrate this with an example:

(11-165) MG (elicited)

[What is the matter with you?]

(a) *Mu xiθike to γALA*

to-me was-spilt the milk:NOM

(b) *ke vromisa to XALI.*

and I-made-dirty the carpet:ACC

'I spilt some milk ('to me was spilt the milk') and made a mess out of the carpet.'

In both sentences, there is a presupposed situation X ('there is a relevant situation in which I am involved and which makes me unhappy'), but it is only in (a) that a vS clause is used. The reason is the semantic role of the additional topic. In (a), it is an 'interested party', encoded as a free dative (actually a genitive), so that it is not a subject, which is therefore left free for the formation of the focus domain. In (b), it is a more agentive participant of the two involved ('I' is more agentive than 'the carpet'), so that it is assigned the subject role (marked on the verb morphology). Instead of a vS clause, a canonical clause with a subject topic appears.

The moral of the story is that the locational or situational nature of the primary topic is not a sufficient condition for the use of vS. If there is an additional topic, it has to be a non-subject, by virtue of its semantic role or due to the valence of the predicate used. Obviously, this is potentially another field of crosslinguistic variation. The idiosyncratic features of the valence of predicates may in one language assign the subject role to a participant, and demote it to an object, or even to some minor grammatical relation, or possessor, in another. Very

often, it is different psych-verbs that display this kind of variation. An example was given in 11.5.2.2.2., where it was mentioned that in Alb. the ‘internal’ events are more often encoded with the experiencer-topic in the object position or as a free dative than in the other two languages. Another nice example is the verb meaning ‘remember’, which is in Alb. construed with the experiencer in the dative, and with the usual AME pattern (experiencer = subject) in MG and SC:

(11-166) Eng: *In the meantime, Bill had bethought himself of the rifle.* (London, p. 21)

MG: *Sto metaksi o Bil sillojotat to tufeki tu.* (p. 27)

in meantime the Bill remembered the rifle his

SC: *Za to vrijeme Bil se bješe sjetio puške.* (p. 25)

for that time Bill REFL AUX remembered rifle

Alb: *Ndërkaq Billit iu kujtua pushka.* (p. 20)

in-meantime to-Bill to-him-PASS remembered rifle-the

The consequence of this grammatical difference is the crosslinguistic variation in the use of vS. This again points out to the language-specific nature of vS, in that their occurrence vs. non-occurrence is a matter of the overall grammatical structure of the language. This structure-dependency is not something one is entitled to expect if the reason for the existence of vS would be an underlying universal category, be it ‘theticity’ or something else.

In sum: The most important condition for the occurrence of vS is that **the topic is not encoded as a subject in the given language**. In Alb., MG and SC, this is regularly the case with locational topics. When a situational topic is used, it either has to be the sole topic, or it has to be combined with an additional topic which, for the reasons of its semantic role and/or language-specific valence restrictions, does not qualify for the subject role.

11.6.2. Untypical subjects

The second feature which appears to be relevant in connection with vS is much more simple and much less potentially contestable than the first one: It is the lack of topicworthiness of the subjects used in this construction. They tend either not to have the status of a discourse referent at all, or at least to be ‘new’, and display a very low degree of textual presence in general. They are relatively rarely agentive, the predominant roles being more on the patientive side of the scale. And they are significantly more often non-human than is the case in all other constructions.

When all these facts are taken together, the picture that emerges is that vS is used every time the element of the proposition which has to be assigned the subject role, for the reasons which have to do with the prototypical valence of the predicates in the languages of the AME type (see below), does not qualify for the topic role. To take one example, from a conversation which I witnessed on the isle of Angistri (07.06.2001):

(11-167) SC

Vidi, zbrčkali su mi se PRSTI!

look wrinkled AUX to-me REFL fingers

‘Look, my FINGERS have wrinkled!’

This utterance was made by a friend of mine after he had spent some two hours swimming in the sea. Now, in my analysis, this is an annuntiatively used vS clause, with an explicit announcement of the existence of a state of affairs to be identified (*vidi* ‘look’) and an additional topic in the form of a free dative (*mi* ‘to-me’). The presuppositional structure is thus [there is X, relevant to ‘I’]_{Presupposed, Topic} [X = fingers are wrinkled]_{Asserted, Focus}. So far, so good. But why did the speaker decide not to choose ‘fingers’ as the topic, and thus form a canonical SV clause with a subject topic?

The answer lies in the function of topics in discourse: They are there in order to delimit the sequence of possible worlds against which the assertion is to be assessed, in order to make the intentions of the speaker as transparent as possible. The intention of this utterance is obviously to give some information about the consequences of long bathing on one’s own body, and, *in ultima analysi*, to emphasize the heroic nature of the person who can endure so much. The best way to convey this message is to give an instruction to the hearer to concentrate on those possible worlds in which ‘I’ am involved in a certain situation, which, as the context clues indicate, has to do with swimming, which is precisely what the speaker had chosen to do.

If ‘fingers’ had been chosen as the topic, with the resulting sentence *Prsti su mi se ZBRČKALI* (‘Fingers AUX to-me REFL wrinkled’), the message would not correspond to the speaker’s intention. Topics carry a relational presupposition, which indicates to the hearer that what is chosen as the topic is in some non-trivial way relevant for the discourse. If the referent ‘fingers’ were topical, the message would not be related to ‘I in a situation X which has to do with swimming’, but rather to ‘the fingers in a situation X...’. The hearer would have to conclude that it is ‘the fingers’ that should reveal to her/him the intentions of the speaker, and would probably construe a context in which the fingers are a relevant entity, for instance ‘the

speaker wants to give his contribution to the debate on the behavior of fingers in the water', or some other similar absurdity.

Now we have the whole picture. Most subject used in vS clauses are like the fingers in the example above: they are simply no good in revealing the speaker's intentions, either because they are inherently irrelevant and not an appropriate subject for a conversation, like fingers¹, or because they in the given moment do not correspond to that what the speaker wants to say.

The problem with such entities is that the languages of the AME type, including here Alb., MG and SC, tend to have an obligatory sentence slot for subjects. In other words, the overwhelming majority of predicates have a subject function in their valence, so that one of the terms of the proposition has to be assigned this function. As we have seen in the previous section, the typical topics in vS – locations, situations, and the additional topics with some kind of patient- or experiencer-like role – are not normally eligible for subjects, each of them for a special reason. Thus only that element which is not topical is left, and receives the subject role despite its nontopical status.

The second factor which triggers vS in the Balkan languages is thus **the non-topicality of the subject.**²

11.7. Theticity and vS

Let us first summarize what has been concluded about the reasons for the existence of vS:

(1) **Structural reasons:** In the languages of the Balkans, the grammatical (i.e. conventional) means of marking the scope of assertion is the focus domain construction, which is a syntactic expression of an unequivocal focus structure with a broad and a narrow focus construals. The Balkan focus domain is marked in the sentence template by means of

¹ This is, of course, not to say that even such apparently untopicalizable entities like 'fingers' cannot, under appropriate conditions, receive their share of discourse relevance and be used as topics. The idea is simply that such conditions are extremely seldom met in the normal human discourse.

² In Sections 2.5.3. and 2.6.1. a question has been raised on how our knowledge of a non-topical referent is increased by an utterance, given that it is the topic the assertion is ascribed to as a property. For instance, how come it that, on hearing a sentence like (11-167), the hearer not only learns about the actual topic, 'I in a situation X', that his fingers are wrinkled, but also about the non-topical subject, 'the fingers', that they are wrinkled. My answer to this is given in Section 4.3.1., where a distinction is made between assertion and predication, the former being applicable only to topics, the latter to all the terms of the proposition. For more details, see the above mentioned sections.

word order and intonation: the potentially assertive material is placed after the verb, with the last accented non-verbal element representing the right border of the domain, the verb itself its left border. Thus, when it is the verb-subject complex that is to be interpreted as a focus domain, the natural consequence of the grammatical structure of these languages is the placement of the subject after the verb and the assignment of the nuclear stress to the subject. The resulting construction, vS, is structurally identical to other focus domain types, like v[erb]O[bject], v[erb]ADV[erb], etc. There is no need to postulate a specific, sentence focus structure for vS, which would then be different from the predicate focus structure (vO, vADV, etc.), because there is no need to put the equality sign between focus structure and constituency (cp. 2.6.2.).

(2) **The nature of the topic.** The topics of vS clauses in the Balkan languages cannot be encoded as subjects, for the reasons which have to do with the lexical structure of these languages, with the semantic structure of the topic itself, and with the rules which map semantic roles to grammatical roles.

(3) **The nature of subjects.** The subjects of vS clauses cannot be interpreted as topical, because this would lead to aberrant interpretations. They are nevertheless encoded as subjects because of the overall grammatical structure of the languages in question.

All other features which have been claimed to be the *raison d' être* of vS and related constructions are derivable from these three properties (cp. Section 3.4. for a summary of these claims).

First, the ones pertaining to the predicates. The verbs are so often existential/positional, or interpreted as such, since the assertions about locations must have a positional interpretation because of the meaning of the topic phrase, just like assertions about books generally have to pertain to reading, writing and publishing (and in some extreme cases, to burning). The dramatic, 'misfortunate', connotation of many predicates used in vS is triggered by the nature of discourse environments in which situational topics are of relevance, as shown in Sections 11.4.2. and 11.5.2.2. The frequency with which the subject and the verb stand in some kind of lexical solidarity is explainable in terms of the focus domain formation, with semantically close-knit verb-noun combinations being easier to construe as a unified piece of assertional material, a broad focus, than the ones in which two denotata have to be processed separately. The same holds, *mutatis mutandis*, also for the fact that the verbs are so often 'weak', i.e. general in meaning (cp. 2.4. for the analysis along these lines proposed by Fuchs). A feature related to the latter two, the tendency to use only monoargumental predicates in the construction, is also a function of the focus domain formation: as indicated in

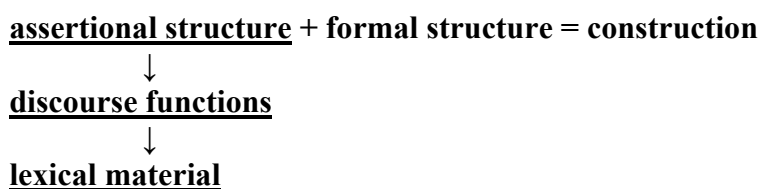
Section 11.2.3., the more elements have to be included into the focus domain, the more difficult the construction of a unified assertional reading, i.e. of broad focus.

The second group of features often mentioned in connection with vS are related to the subjects. They are very often nonagentive, which has lead many to assume that vS has to do with unaccusativity. A more elegant solution is at hand: the subjects in vS are so often nonagentive because, when they are agentive, they are much more often pragmatically construed as topics. The result is then a canonical SV sentence (or VsX, for that matter). The other very frequent claim is that it is the ‘newness’ of the subjects that triggers vS: as I have repeatedly pointed out throughout this study, ‘newness’ is only one of the symptoms of non-topicality, and a very weak one: there are many topics which are ‘new’, and there are many non-topics which are ‘old’, as my statistics has shown (cp. especially 11.3.2.).

All the specific interpretational properties of vS – the perceptive effect, the natural force effect, the objective effect, even the much quoted spatio-temporal boundness of vS clauses – are easily derivable from the nature of topics in vS clauses, as has been shown in this chapter.

The hierarchical scheme which I have proposed for the representation of discourse-pragmatically marked constructions in Section 6.7 (scheme (6-84)) with respect to VsX thus seems to poignantly describe the situation in vS as well. I repeat it here as(11-168):

(11-168)*Hierarchical structure of constructions*



On the highest level, a certain assertional structure is mapped onto a certain formal structure, which is the basis of every construction. The assertional structure determines which discourse functions the construction can perform; the discourse functions, in their turn, determine the preferred lexical and informational filling.

As has been noted in 6.7., this scheme represents the ideal case, i.e. the case in which the construction is fully productive in a language, so that no lexicalizations and discourse conventionalizations disrupt the primacy of expressing the assertional structure. In Alb., MG and SC, vS construction is certainly closest to this ideal picture of all three VS constructions, being fully productive in all three languages. However, as has been indicated sporadically throughout this chapter, even in this construction there are a lot of idiomatic and semi-idiomatic verb-subject combinations which are per default expressed with vS, and there are

discourse functions which are almost automatically performed by this focus domain type. On the whole, in spite of these minor irregularities triggered by various conventionalizations, vS is a highly productive pattern, far from the moribund character of VsX in Alb. and SC, and far from the uncertain fate of inversion.

Now, the final question. Do the properties of vS in Alb., MG and SC presented in this study justify the idea that there is a primitive, nonanalyzable category ofthetic statements, placed on the level of cognition, or on the level of discourse pragmatics? The answer is unequivocally no. vS clauses are a language-specific solution for a mismatch between the grammatical structure and the needs of the discourse in certain contexts. The semantic, pragmatic, informational, etc., features which have been assumed to be the *differentia specifica* ofthetic statements turn out, on the analysis proposed in this study, to be merely a consequence of the assertional structure of this particular clause type. As the scheme given in (6-84) shows, the real home of ‘theticity’, i.e. of the meaning conveyed by vS, is the interface of information structure, discourse, and lexicon.

Sources

- Andrić, Ivo. 1981. *Prokleta avlija*. Sarajevo: Svjetlost.
- Antologija narodnih pripovedaka*, ed. Vojislav Djurić. 1960. Novi Sad: Matica Srpska.
- Avji*. Athens. 21.11.2000.
- Bishqemi, Astrit. 1987. *Udhëtim i jashtëzakonshëm*. Tirana: Naim Frashëri.
- Camaj, Martin. 1978. *Rrathë*. München: im Selbstverlag.
- Ciao*. Athens. 05.06.2001.
- Eco, Umberto. *To onoma tu roðu* (= U.E., *Il nome della rosa*), transl. E. Kaliphatidi. Athens: Ellinika Grammata.
- Elefterotipia*. Athens. 13.09.2000.
- Elinika laika paramiþja. Griechische Volksmärchen*. 1992. München: DTV.
- Fakinou, Eugenia. 2001. *I meçali prasini*. Athens: Kastanioti.
- Hoxha, Enver. 1975. *Vepra*, vol. 19. Tirana: Mihal Duri.
- Kadare, Ismael. 1980. *Kronikë në gur*. Prishtina: Rilindja.
- Kadare, Ismael. 1991. *Ëndrra mashtruese*. Tirana: Naim Frashëri.
- Kaþimerini*. Athens. 30.05.2001.
- Kapandaj, Ismini. 1989. *Epta fores to ðaxtiliði*. Athens: Estia.
- Kiš, Danilo. 1986. *Grobnica za Borisa Davidoviça*. Belgrade: BIGZ.
- Koha Ditore*. Prishtina. 21.10.2000.
- Kosovarja*. Prishtina. Nr. 10, 2001.
- London, Jack. 1994. *White Fang. The call of the wild*. London: Penguin Classics
- London, Dzak. 1997. *O Asproðondis*, transl. P. Mosxopoulou. Athens: Anastasiadi.
- London, Džek. 1966. *Bijeli Očnjak*, transl. M. Djordjević. Sarajevo: Svjetlost.
- London, Xhek. 1999. *Dhëmbi i bardhë*, transl. P. Xhixha. Tirana: Mësonjëtorja e pare.
- Matesis, Paulos. 1990. *I mitera tu skilu*. Athens: Kastanioti.
- Mandeljštam, Nadežda. 1984. *Uspomene i razmišljanja* (= Nadežda Mandel'stam, *Vtoraja kniga*), transl. A. Demetrović-Matijašević. Belgrade: Prosveta.
- Mbledhës të hershëm të folklorit shqiptar (1635-1912)*. Vol. III: *Materiale folklorike të botuara në vëll. Valët e detit të Spiro Dines, Sofie 1908*, ed. Z. Sako et al. 1962. Tirana: Institut i folklorit.
- NIN*. Belgrade. 13.06.1997.

Pavić, Milorad. 1995. *Predeo slikan čajem*. Belgrade: Prosveta.

To fos ton spor. Athens. 21.11.2000.

Vjesnik. Zagreb. 14.12.1999.

Vesti. Frankfurt a.M. 04.03.2000.

References

- Abbott, Barbara. 1994. "Referentiability, specificity, strength, and individual concepts." *WCCFL* 12: 473-483.
- Abbott, Barbara. 2000. "Presuppositions as nonassertions." *Journal of Pragmatics* 32: 1419-1437.
- Adamec, Přemysl. 1966. *Porjadok slov v sovremennom ruskom jazyke*. Praha: Academia.
- Agouraki, Georgia. 1990. "On the projection of maximal categories: The case of CP and FP in Modern Greek." *UCL Working Papers in Linguistics* 2: 183-200.
- Agouraki, Georgia. 1991. "A Modern Greek complementizer and its significance for universal grammar." *UCL Working Papers in Linguistics* 3: 1-24.
- Aissen, Judith and Jorge Hankamer. 1972. "Shifty subjects: A conspiracy in syntax?." *Linguistic Inquiry* 3: 501-504.
- Alexiadou, Artemis. 1996. "Aspectual restrictions on word order." *Folia Linguistica* 30: 35-46.
- Alexiadou, Artemis. 2000. "Some remarks on word order and information structure in Romance and Greek." *ZAS Papers in Linguistics* 20: 119-136.
- Alexiadou, Artemis and Elena Anagnostopoulou. 2000. "Greek syntax: A principles and parameters perspective." *Journal of Greek Linguistics* 1: 171-222.
- Allerton, David J. and Alan Cruttenden. 1979. "Three reasons for accenting a definite subject." *Journal of Linguistics* 15: 49-53.
- Alisova, Tatiana. 1972. *Strutture semantiche e sintattiche della proposizione semplice in italiano*. Firenze: Sansoni.
- Ambar, Manuela. 1992. *Para uma sintaxe da inversão sujeito-verbo em Português*. Lisboa: Colibri.
- Ambar, Manuela. 1998. "Word order effects of a split CP hypothesis." Paper presented at the Workshop on Romance Inversion, University of Amsterdam.
- Arnold, Jennifer E.; Thomas Wasow; Anthony Losongco; and Ryan Ginstrom. 2000. "Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering." *Language* 76: 28-55.
- Atlas, Jay D. and Stephen C. Levinson. 1981. "It-clefts, informativeness, and logical form: Radical pragmatics (revised standard version)." In: Cole (ed.), 1-61.

- de Bakker, Cecile. 1997. *Germanic and Romance inversion in French: A diachronic study*. Den Haag: Holland Academic Graphics.
- Barbosa, Pilar. 1998. "On inversion in wh-questions in Romance." Paper presented at the Workshop on Romance Inversion, University of Amsterdam.
- Basilico, David. 1999. "Pronoun positioning." *Lingua* 109:155-181.
- Behagel, Otto. 1932. *Deutsche Syntax. IV: Wortstellung, Periodenbau*. Heidelberg: Winter.
- Behrens, Leila. 1982. *Zur funktionalen Motivation der Wortstellung: Untersuchungen anhand des Ungarischen*. München: Finnisch-ugrisches Seminar.
- Behrens, Leila. 2000. *Typological parameters of genericity*. Institut für Sprachwissenschaft, Universität zu Köln [*Arbeitspapier* 37 (NF)].
- Behrens, Leila and Hans-Jürgen Sasse. 2003. *The microstructure of lexicon-grammar interaction: A study of 'gold' in English and Arabic*. München: Lincom.
- Belletti, Adriana. 1988. "Unaccusatives as case assigners." *Linguistic Inquiry* 19: 1-34.
- Bellver, Phyliss and Laura A. Michaelis. 1999. "What is the information structure-syntax interface in Basque?." *BLS* 25: 26-37.
- Bennett, David. 1987. "Word order change in progress: The case of Slovene and Serbo-Croat and its relevance for Germanic". *Journal of Linguistics* 23: 269-287.
- Bentivoglio, Paola. 1983. "Continuity and discontinuity in discourse: A study on Latin-American spoken Spanish." In: Givón (ed.) 1983a: 255-311.
- Berneker, Erich. 1900. *Die Wortfolge in den slavischen Sprachen*. Berlin.
- Birner, Betty J. 1994. "Information status and word order: An analysis of English inversion." *Language* 70: 233-259.
- Birner, Betty J. 1996. *The discourse function of inversion in English*. New York: Garland.
- Birner, Betty J. and Shahrzad Mahootian. 1996. "Functional constraints on inversion in English and Farsi." *Language Sciences* 18: 127-138.
- Birner, Betty J. and Gregory Ward. 1992. "On the interpretation of VP inversion in American English." *Journal of Linguistics* 28: 1-12.
- Birner, Betty J. and Gregory Ward. 1993. "There-sentences and inversion as distinct constructions: A functional account." *BLS* 19: 27-39.
- Birner, Betty J. and Gregory Ward. 1996. "A crosslinguistic study of postposing in discourse." *Language and Speech* 39: 111-140.
- Birner, Betty J. and Gregory Ward. 1998. *Information status and noncanonical word order in English*. Amsterdam: Benjamins.
- Blake, Barry J. 2001. "Global trends in language." *Linguistics* 39: 1009-1028.

- Blinkenberg, Andreas. 1928. *L'ordre des mots en français moderne*. København: Høst.
- Bolinger, Dwight L. 1954a. "English prosodic stress and Spanish sentence order." *Hispania* 37: 152-156.
- Bolinger, Dwight L. 1954b. "Meaningful word order in Spanish." *Boletín de Filología* 8: 45-56.
- Bolinger, Dwight L. 1971. "A further note on the nominal in the progressive." *Linguistic Inquiry* 2: 584-586.
- Bolinger, Dwight L. 1977. *Meaning and form*. London: Longman.
- Bonnot, Ch. and I. Fougeron. 1982. "L'accent de phrase initial en russe est-il toujours un signe d'expressivité ou de familiarité?." *Bulletin de la Société de linguistique de Paris* 77: 309-330.
- Bonnot, Ch. and I. Fougeron. 1983. "Accent de phrase non final et relations interénonciatives en russe moderne." *Revue des études slaves* 55: 611-626.
- Bosch, Peter and Rob van der Sandt (eds.). 1999. *Focus: Linguistic, cognitive, and computational perspectives*. Cambridge: CUP.
- Bošković, Željko. 2002. "On multiple *wh*-fronting." *Linguistic Inquiry* 33: 351-383.
- Bresnan, Joan. 1991. "Locative case vs. locative gender." *BLS* 17: 53-68.
- Bresnan, Joan. 1994. "Locative inversion and the architecture of Universal Grammar." *Language* 70: 72-131.
- Braun, M. 1962. "Zur Frage der Wortfolge im Russischen." *Slavica Pragensia* 4: 289-296.
- Brody, Michael. 1990. "Some remarks on the focus field in Hungarian." *UCL Working Papers in Linguistics* 2: 201-225.
- Brown, Gillian and George Yule. 1983. *Discourse Analysis*. Cambridge: CUP.
- Brustad, Kristen E. 2000. *The syntax of spoken Arabic: A comparative study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects*. Washington, DC: Georgetown UP.
- Buchholz, Oda and Wilfried Fiedler. 1987. *Albanische Grammatik*. Leipzig: VEB Verlag Enzyklopädie.
- Büring, Daniel. 1999. "Topic." In: Bosch and van der Sandt (eds.), 142-165.
- Büring, Daniel. 2000. "Kontrastive Topiks in einem formalen Diskursmodell." Paper presented at 11. Wuppertaler Linguistisches Kolloquium: Topik-Kommentar-Struktur, 17-18.11.2000.
- Burton-Roberts, Noel. 1989. *The limits to debate*. Cambridge: CUP.
- Burton-Roberts, Noel. 1999. "Presupposition-cancellation and metalinguistic negation: A reply to Carston." *Journal of Linguistics* 35: 347-364.

- Burzio, Luigi. 1981. *Intransitive verbs and Italian auxiliaries*. Ph.D. diss., MIT, Cambridge, Mass.
- Burzio, Luigi. 1986. *Italian syntax*. Dordrecht: Reidel.
- Byrne, William. 1997. "The Thetic/Categorical Distinction and Bare Nominals in Spanish." In: *BLS* 23: 26-37.
- Carlson, Gregory N. 1977. *Reference to kinds in English*. Ph.D. diss., University of Massachusetts, Amherst.
- Carston, Robyn. 1998. "Negation, 'presupposition' and the semantics/pragmatics distinction." *Journal of Linguistics* 34: 309-350.
- Carston, Robyn. 1999. "Negation, 'presupposition' and metarepresentation: a response to Noel Burton-Roberts." *Journal of Linguistics* 35: 365-389.
- Caton, Charles E. 1981. "Stalnaker on pragmatic presupposition." In: Cole (ed.): 83-100.
- Catsimali, Georgia. 1991. "On the placement of Modern Greek on the +/- configurational spectrum." *Glossologia* 7-8 (1988-1989): 73-86.
- Chafe, Wallace L. 1974. "Language and consciousness." *Language* 50: 111-133.
- Chafe, Wallace L. 1976. "Givenness, contrastiveness, definiteness, subjects, topics and point of view." In: Li (ed.): 25-56.
- Chafe, Wallace L. 1980. "The deployment of consciousness in the production of a narrative." In: W. Chafe (ed.). *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*: 9-50. Norwood, NJ: Ablex.
- Chafe, Wallace L. 1987. "Cognitive constraints on information flow." In: R. Tomlin (ed.). *Coherence and Grounding in Discourse*: 21-51. Amsterdam: Benjamins.
- Chafe, Wallace L. 1992. "Information flow in speaking and writing." In: P. Downing; S. D. Lima; and M. Noonan (eds.). *The Linguistics of Literacy*: 17-29. Amsterdam: Benjamins.
- Chafe, Wallace L. 1994. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.
- Chierchia, Gennaro. 1992. "Anaphora and dynamic binding." *Linguistics and Philosophy* 15: 111-184.
- Chomsky, Noam. 1972. *Studies on semantics in the generative grammar*. Den Haag: Mouton.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Cinque, Guglielmo. 1993. "A null theory of phrase and compound stress." *Linguistic Inquiry* 24: 239-297.

- Collins, Chris and Phil Branigan. 1997. "Quotative inversion." *Natural Language and Linguistic Theory* 15: 1-41.
- Čistjakova, A.L. 1954. "Vzaimoporjadok podležaštegogo i skazuemogo v povestvovatel'nom predloženi." *Russkij jazyk v školah* 6: 5-10.
- Clancy, Patricia M. and Pamela Downing. 1987. "The use of *wa* as a cohesion marker in Japanese oral narratives." In: Hinds, Maynard and Iwasaki (eds.): 3-56.
- Clark, Herbert H. 1977. "Bridging." In: P. Johnson-Laird and P. C. Wason (eds.). *Thinking: Readings in cognitive science*: 411-420. Cambridge: CUP.
- Cole, Peter (ed.). 1978. *Pragmatics*. New York: Academic Press. [*Syntax and Semantics* 9].
- Cole, Peter (ed.). 1981. *Radical pragmatics*. New York: Academic Press.
- Combettes, B. 1992. *L'organisation du texte*. Metz: Centre d'analyse syntaxique de l'Université de Metz.
- Comrie, Bernard. 1988. "Topics, grammaticalized topics, and subjects." *BLS* 14: 265-279.
- Cooreman, Ann. 1983. "Topic continuity and voicing system of an ergative language: Chamorro." In: Givón (ed.) 1983a: 425-489.
- Cooreman, Ann. 1987. *Transitivity and discourse continuity in Chamorro narratives*. Berlin: Mouton de Gruyter.
- Cooreman, Ann. 1992. "Word order variation in Chamorro." In: Payne (ed.): 263-243.
- Coseriu, Eugenio. 1967. "Lexikalische Solidaritäten". *Poetica* 1: 293-303.
- Costa, João. 1996. "Positions for subjects in European Portuguese." *WCCFL* 15: 49-63.
- Costa, João. 1998. "Marked vs. unmarked inversion and Optimality Theory." Paper presented at the Workshop on Romance Inversion, University of Amsterdam.
- Culicover, Peter W. and Robert D. Levine. 2001. "Stylistic inversion in English: A reconsideration." *Natural Language and Linguistic Theory* 19: 283-310.
- Culicover, Peter W. and Louise McNally (eds.). 1998. *The Limits of Syntax*. New York: Academic Press. [*Syntax and Semantics* 29].
- Culicover, Peter W. and Michael Rochemont. 1983. "Stress and focus in English." *Language* 59: 123-165.
- Dahlgren, Sven-Olof. 1998. *Word order in Arabic*. Göteborg: Acta Universitatis Gothoburgensis. [*Orientalia Gothoburgensia* 12].
- Daneš, František. 1957. *Intonace a věta ve spisovné češtině*. Praha: Academia.
- Darnell, Michael. 1995. "Preverbal nominals in Colville-Okanagan." In: Downing and Noonan (eds.): 85-104.

- Davison, Alice. 1984. "Syntactic markedness and the definition of sentence topic." *Language* 60: 797-846.
- Delin, Judy and Jon Oberlander. 1995. "Syntactic constraints on discourse structure: The case of *it*-clefts." *Linguistics* 33: 465-500.
- Dellbrück, Berthold. 1900. *Vergleichende Syntax der indogermanischen Sprachen. 3. Theil.* Straßburg: Trübner.
- Devine, A. M. and Laurence D. Stephens. 2000. *Discontinuous syntax. Hyperbaton in Greek.* Oxford: OUP.
- Diesing, Molly. 1992. *Indefinites.* Cambridge, Mass.: MIT Press.
- Dik, Simon, et al. 1981. "On the typology of focus phenomena." In: T. Hoekstra, H. van der Hulst, M. Moortgat (eds.), *Perspectives on FG*, 41-74. Dordrecht: Foris.
- Dittmar, Norbert (ed.). 1992. *Topic. From grammar to discourse.* New York: Mouton de Gruyter. [= *Linguistics* 30.1]
- Dobrovie-Sorin, Carmen. 2001. "Head-to-Head Merge in Balkan Subjunctives and Locality." In: Rivero and Ralli (eds.), 44-71.
- Donnellan, Keith S. 1981. "Intuitions and presuppositions." In: Cole (ed.): 129-142.
- Dover, Kenneth J. 1960. *Greek word order.* Cambridge: CUP.
- Downing, Pamela and Michael Noonan (eds.). 1995. *Word order in discourse.* Amsterdam: Benjamins.
- Dowty, David. 1991. "Thematic proto-roles and argument selection". *Language* 67: 547-619.
- Drubig, H. Bernhard. 1992. "Zur Frage der grammatischen Repräsentation thetischer und kategorischer Sätze." In: Jacobs (ed.): 142-195.
- Dryer, Matthew S. 1996. "Focus, pragmatic presupposition, and activated propositions." *Journal of Pragmatics* 26: 475-523.
- Eguzkitza, Andolin and Georg A. Kaiser. 1999. "Postverbal subjects in Romance and German: Some notes on the Unaccusative Hypothesis." *Lingua* 109: 195-219.
- Emonds, Joseph. 1970. *Root and structure preserving transformations.* Bloomington, Indiana: IULC.
- Epstein, Richard. 1999. "Roles, Frames and Definiteness." In: K. van Hoek, A. Kibrik and L. Nordman (eds.). *Discourse studies in cognitive linguistics.* Amsterdam: Benjamins.
- Erteschik-Shir, Nomi. 1997. *The dynamics of focus structure.* Cambridge: CUP.
- Etchegoyhen, Thierry and George Tsoulas. 1998. "Thetic and categorical, attributive and referential. Towards an explanation of definiteness effects." In: A. Schwegler; B.

- Tranel; and M. Uribe-Etxebarria (eds.). *Romance linguistics: Theoretical perspectives*: 81-95, Amsterdam: Benjamins.
- Fillmore, Charles J. 1982. "Frame semantics". In: Linguistic Society of Korea (ed.), *Linguistics in the morning calm*, 111-137. Seoul: Hanshin Publishing Co.
- Fillmore, Charles J. 1999. "Inversion and constructional inheritance." In: G. Webelhuth; J.-P. Koenig; and A. Kathol (eds.), *Lexical and constructional aspects of linguistic explanation*, 113-128. Stanford: CSLI.
- Filppula, Markku. 1999. *The grammar of Irish English*. London: Routledge.
- von Fintel, Kai. 1994. *Restrictions on quantifier domains*. Ph.D. diss., University of Massachusetts at Amherst.
- Firbas, Jan. 1964. "On defining the theme in functional sentence analysis." *Travaux linguistiques de Prague* 1: 267-280.
- Firbas, Jan. 1966. "Non-thematic subjects in contemporary English." *Travaux linguistiques de Prague* 2: 239-256.
- Firbas, Jan. 1992. *Functional sentence perspective in written and spoken communication*. Cambridge: CUP.
- Floqi, S. 1969. "Mbi disa çështje të rendit të fjalëve në shqipen e sotme." *Studime filologjike* 1969-2: 125-143.
- Floqi, S. 1976. "Rreth modeleve të fjalisë në shqipën e sotme letrare." *Studime filologjike* 1976-4: 25-51.
- Fortescue, Michael. 1996. "Grammaticalized focus in Yukagir." In: E. Engberg-Pedersen; M. Fortescue; P. Harder; L. Heltoft; and L. Falster Jakobsen (eds.), *Content, expression and structure. Studies in Danish Functional Grammar*, 17-38. Amsterdam: Benjamins.
- Fournier, Nathalie. 1997. "La place du sujet nominal dans les phrases à complément prépositionnel initial." In: Fuchs (ed.): 97-132.
- Fox, Barbara. 1985. "Word order inversion and discourse continuity in Tagalog." In: T. Givón (ed.). *Quantitative studies in discourse*: 138-159. [= *Text* 5.1].
- Frey, Werner. 2000. "Über die syntaktische Position der Satztopiks im Deutschen." *ZAS Papers in Linguistics* 20: 137-172.
- Frisk, Hjalmar. 1932. *Studien zur griechischen Wortstellung*. Göteborg: Wettergren & Kerbers. [*Göteborgs högskolas årsskrift* 39].
- Fuchs, Anna. 1976. "'Normaler' und 'kontrastiver' Akzent." *Lingua* 38: 293-312.

- Fuchs, Anna. 1980. "Accented subjects in 'all-new' utterances." In: G. Brettschneider and C. Lehmann (eds.), *Wege zur Universalienforschung. Festschrift für H.-J. Seiler*, 449-461. Tübingen: Narr.
- Fuchs, Anna. 1987. "Pragmatics as against grammatical factors in the determination of accent placement." In: J. Verschueren; and M. Bertucelli-Papi (eds.), *The pragmatic perspective*, 421-434. Amsterdam: Benjamins.
- Fuchs, Catherine (ed.). 1997. *La place du sujet en français contemporain*. Louvain-la-Neuve: Duculot.
- Garcia, Erica C. 1979. "Discourse without syntax." In: Givón (ed.): 23-49.
- Gasde, Horst-Dieter. 1999. "Are there 'topic-prominence' and 'subject-prominence' along the lines of Li & Thomson (1976)." Paper presented at the 21. Jahrestagung der DGS, Februar 1999, Konstanz.
- Gazdar, Gerald. 1979. "A solution to the projection problem." In: Oh and Dinneen (eds.), 57-89.
- Gensler, Orin. 1977. "Non-syntactic antecedents and frame semantics." *BLS* 3: 321-334.
- Georgieva, Elena. 1974. *Slovoled na prostoto izrečenie v bălgarskija knižoven ezik*. Sofia: Bălgarska Akademija na Naukite.
- Giannakidou, Anastasia. 1997. *The landscape of polarity items*. Ph.D. Diss., University of Groningen.
- Giannakidou, Anastasia. 2000. "Crosslinguistic semantics and the study of Greek." *Journal of Greek Linguistics* 1: 223-261.
- Givón, Talmy. 1977. "The drift from VSO to SVO in Biblical Hebrew: The pragmatics of tense-aspect." In: Ch. N. Li (ed.). *Mechanisms of syntactic change*: 181-254. Austin: University of Texas Press.
- Givón, Talmy. 1979a. *On understanding grammar*. New York: Academic Press.
- Givón, Talmy (ed). 1979b. *Discourse and syntax*. New York: Academic Press. [*Syntax and Semantics* 12].
- Givón, Talmy. 1979c. "From discourse to syntax: Grammar as a processing strategy." In: Givón (ed.): 81-112.
- Givón, Talmy. (ed.) 1983a. *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam: Benjamins.
- Givón, Talmy. 1983b. "Topic continuity and word order pragmatics in Ute." In: Givón (ed.): 141-214.

- Goldberg, Adela E. and Farrel Ackerman. 2001. "The pragmatics of obligatory adjuncts." *Language* 77: 798-814.
- Grabski, Michael. 2000. "Satztopik und Diskurstopik in Elaboration-Kontexten." *ZAS Papers in Linguistics* 20: 173-207.
- Green, Georgia. 1980. "Some wherefores of English inversions." *Language* 56: 582-601.
- Grenoble, Lenore A. 1998. *Deixis and information packaging in Russian discourse*. Amsterdam: Benjamins.
- Grice, Paul H. 1975. "Logic and conversation." In: P. Cole and J. L. Morgan (eds.). *Speech acts*: 41-58. New York: Academic Press. [*Syntax and Semantics* 3].
- Grice, Paul H. 1978. "Further notes on logic and conversation." In: Cole (ed.): 113-128.
- Grice, Paul H. 1981. "Presupposition and conversational implicature." In: Cole (ed.): 183-198.
- Grupo di Padova. 1974. "L'ordine degli sintagmi nella frase." In: Medici and Sangregorio (eds.) 1:147-161.
- Gundel, Jeanette K. 1985. "'Shared knowledge' and topicality." *Journal of Pragmatics* 9: 83-107.
- Gundel, Jeanette K. 1988. "Universals of topic-comment structure". In: M. Hammond, E. Moravcsik, J. Wirth (eds.), *Studies in syntactic typology*, 209-239. Amsterdam: Benjamins.
- Gussenhoven, Carlos. 1983a. "Focus, mode, and nucleus." *Journal of Linguistics* 19: 377-417.
- Gussenhoven, Carlos. 1983b. "Testing the reality of focus domains." *Language and Speech* 26: 61-80.
- Haberland, Hartmut. 1994. "Thetic/Categorical distinction." In: R. E. Asher and J. M. Y. Simpson (eds.), *The Encyclopedia of language and linguistics* 9: 4605-4606. Oxford-New York-Tokyo: Pergamon Press.
- Haberland, Hartmut and Johan van der Auwera. 1990. "Topics and clitics in Greek relatives." *Acta linguistica Hafniensia* 22: 127-157.
- Haebler, Claus. 1957. "Eine albanische Wortstellungsregel." *Münchener Studien zur Sprachwissenschaft* 11: 49-61.
- Haiman, John. 2002. "Systematization and the origin or rules: The case of subject-verb inversion in questions". *Studies in Language* 26: 573-593.
- Hajičová, Eva and Petr Sgall. 1982. "Functional sentence perspective in the Slavonic languages and in English." *Južnoslovenski filolog* 38: 19-34.

- Halliday, Michael A.K. 1967. "Notes on transitivity and theme in English, Part 2." *Journal of Linguistics* 3: 199-244.
- Halliday, Michael A.K. 1970. "Language structure and language function." In: J. Lyons (ed.), *New Horizons in Linguistics*, 140-165. Harmondsworth: Penguin.
- Hartvigson, Hans and Leif Jakobsen. 1974. *Inversion in the present-day English*. Odense: Odense University Press.
- Hatcher, Anna Granville. 1956. *Theme and underlying question. Two studies in the Spanish word order*. New York: The Linguistic Circle of New York. [Supplement to *Word*, Monograph 3].
- Hawkins, John A. 1995. *A performance theory of order and constituency*. Cambridge: CUP.
- Heim, Irene. 1988 [=1982]. *The Semantics of Definite and Indefinite Noun Phrases*. New York: Garland. [= Ph.D. Diss., University of Massachusetts, Amherst].
- Heim, Irene. 1983. "On the projection problem for presuppositions." *WCCFL* 2: 114-126.
- Hendriks, Herman. 2002. "Links without locations. Information packaging: from cards to boxes." In: F. Hamm and T. E. Zimmermann (eds.). *Semantics*: 173-197. Hamburg: Buske. [*Linguistische Berichte*, Sonderheft 10].
- Herring, Susan C. and John C. Paolillo. 1995. "Focus position in SOV languages." In: Downing and Noonan (eds.): 163-198.
- Hesseling, Dirk Christiaan. 1927. "Het nieuwgriekse *poû*." *Neophilologus* 12: 218-223.
- Hetzron, Robert. 1971. "Presentative function and presentative movement." *Studies in African Linguistics*, Supplement 2: 79-105.
- Hetzron, Robert. 1975. "The presentative movement, or why the ideal word order is V.S.O.P.." In: Li (ed.): 346-388.
- Hinds, John; Senko Maynard; and Shoichi Iwasaki. (eds.). 1987. *Perspectives on topicalization. The Case of Japanese 'wa'*. Amsterdam: Benjamins.
- Hockett, Charles F. 1958. *A course in modern linguistics*. New York: Macmillan.
- Höhle, Tilman L. 1992. "Über Verum-Fokus im Deutschen." In: J. Jacobs (ed.), 112-141.
- Holton, David; Peter Mackridge; and Irene Philippaki-Warbuton. 1997. *Greek. A comprehensive grammar of the modern language*. London: Routledge.
- Hopper, Paul J. 1975. *The syntax of the simple sentence in Proto-Germanic*. The Hague: Mouton.
- Hopper, Paul J. 1979. "Aspect and foregrounding in discourse." In: Givón (ed.):213-241.

- Hopper, Paul J. 1986. "Discourse function and word order shift: A typological study of the VS / SV alternation." In: W. P. Lehmann (ed.). *Language typology 1985*: 123-140. Amsterdam: Benjamins.
- Hopper, Paul J. 1987. "Stability and change in VN / NV alternating languages. A study in pragmatics and linguistic typology." In: J. Verschueren and M. Bertucelli-Papi (eds.). *The pragmatic perspective*: 455-476. Amsterdam: Benjamins.
- Hopper, Paul J. and Sandra A. Thompson. 1980. "Transitivity in grammar and discourse." *Language* 56: 251-299.
- Horn, Laurence R. 1981. "Exhaustiveness and the semantics of clefts." *NELS* 11: 125-142.
- Horn, Laurence R. 1984. "Toward a new taxonomy for pragmatic inference: Q-based and R-based implicatures." In: *Georgetown University Round Table on Languages and Linguistics*, 11-42. Washington, DC: Georgetown UP.
- Horn, Laurence R. 1985. "Metalinguistic negation and pragmatic ambiguity." *Language* 61: 121-174.
- Horn, Laurence R. 1986. "Presupposition, theme and variations." *CLS* 22,2: *Papers from the parasession on pragmatics and grammatical theory*: 168-192.
- Horn, Laurence R. 1989. *A natural history of negation*. Chicago: University of Chicago Press.
- Horn, Laurence R. 1996. "Presupposition and implicature." In: Lappin (ed.), 299-319.
- Horrocks, Geoffrey. 1983. "The order of constituents in Modern Greek.." In: G. Gazdar, E. Klein, and G. Pullum (eds.), *Order, concord and constituency*, 95-111. Dordrecht: Foris.
- Horrocks, Geoffrey. 1990. "Clitics in Greek: A diachronic review." In: M. Roussou and S. Pandeli (eds.), *Greek outside Greece II*, 35-52. Athens.
- Horrocks, Geoffrey. 1994. "Subjects and configurationality: Modern Greek clause structure." *Journal of Linguistics* 30: 81-109.
- Horrocks, Geoffrey. 1997. *Greek. A history of the language and its speakers*. London: Longman.
- Horvath, Julia. 1985. *Focus in the theory of grammar and the syntax of Hungarian*. Dordrecht: Foris.
- Hubbard, Philip L. 1985. *The syntax of the Albanian verb complex*. New York: Garland.
- Huckin, Terry N. 1977. *An integrated theory of English intonation*. University Microfilms International: Ann Arbor.
- Itagaki, Nobuya and Gary D. Prideaux. 1985. "Nominal properties as determinants of subject selection." *Lingua* 66, 135-149.

- Jacennik, Barbara and Matthew Dryer. 1992. "Verb-subject order in Polish." In: Payne (ed.): 209-241.
- Jackendoff, Ray S. 1987. "The status of thematic relations in linguistic theory." *Linguistic Inquiry* 18: 369-411.
- Jackendoff, Ray S. 1990. *Semantic structures*. Cambridge, Mass.: MIT Press.
- Jacobs, Joachim. 1991. "Focus ambiguities." *Journal of Semantics* 8: 1-36.
- Jacobs, Joachim. (ed.). 1992. *Informationsstruktur und Grammatik*. Opladen: Westdeutscher Verlag [*Linguistische Berichte*. Sonderheft 4].
- Jacobs, Joachim. 1993. "Integration." In: Reis (ed.), 63-116.
- Jacobs, Joachim. 1999. "Informational autonomy." In: Bosch and van der Sandt (eds.), 56-81.
- Jacobs, Joachim. 2001. "The dimensions of topic-comment." *Linguistics* 39: 641-681.
- Jäger, Gerhard. 1997. "The stage/individual contrast revisited." *WCCFL* 15: 225-239.
- Jäger, Gerhard. 2001. "Topic-comment structure and the contrast between stage level and individual level predicates." *Journal of Semantics* 18: 83-126.
- Joseph, Brian D. 1983. *The synchrony and diachrony of the Balkan infinitive. A study in areal, general, and historical linguistics*. Cambridge: CUP.
- Joseph, Brian D.; Geoffrey Horrocks; and Irene Philippaki-Warbuton (eds.). 1998. *Themes in Greek linguistics II*. Amsterdam: Benjamins.
- Justus, Carol F. and Edgar C. Polomé (eds.). 1999. *Language change and typological variation. In honor of W. P. Lehmann*. Washington D.C.: Institute for the Study of Man [*JIES Monographs* 31].
- Kallulli, Dalina. 2001. "Direct object clitic doubling in Albanian and Greek". In: Rivero and Ralli (eds.), 127-160.
- Karttunen, Lauri. 1973. "Presuppositions of compound sentences." *Linguistic Inquiry* 4: 169-193.
- Karttunen, Lauri. 1976. "Discourse referents." In: J. McCawley (ed.). *Notes from the linguistic underground*, 363-385. New York: Academic Press. [*Syntax and Semantics* 7].
- Kayne, Richard S. and Jean-Yves Pollock. 1978. "Stylistic inversion, successive cyclicity and move NP in French." *Linguistic Inquiry* 9: 595-621.
- Kayne, Richard S. and Jean-Yves Pollock. 1998. "New thoughts on stylistic inversion." Paper presented on the Workshop on Romance Inversion, University of Amsterdam.
- Keenan, Edward L. 1976. "Towards a universal definition of subject". In: Li (ed.), 303-333.

- Keijsper, Cornelia E. 1985. *Information structure (with examples from Russian, English and Dutch)*. Amsterdam: Rodopi.
- Kendall, Martha B. 1980. "The unethical dative". In: K. Klar et al. (eds.), *American Indian and Indoeuropean Studies. Papers in Honor of M. S. Beeler*, 383-394. The Hague: Mouton.
- Kenesei, István; Róbert Vago; and Anna Fenyvesi. 1998. *Hungarian*. London: Routledge.
- Kennedy, Becky. 1999. "Focus constituency." *Journal of Pragmatics* 31: 1203-1230.
- Kieckers, Ernst. 1911. *Die Stellung des Verbs im Griechischen und in den verwandten Sprachen*. Strassburg: Trübner.
- Kiefer, Ferenc. 1994. "Aspect and Syntactic Structure." In: F. Kiefer and K. É. Kiss (eds.) 415-464.
- Kiefer, Ferenc and Katalin É. Kiss (eds.). 1994. *The sentence structure of Hungarian*. New York: Academic Press. [*Syntax and Semantics* 27]
- Kim, Alan Hyun-Oak. 1988. "Preverbal focusing and type XXIII languages." In: Michael Hammond; Edith Moravcsik; and Jessica Wirth (eds.), *Studies in Syntactic Typology*, 147-169. Amsterdam: Benjamins.
- Kiss, Katalin É. 1987. *Configurationality in Hungarian*. Dordrecht: Reidel.
- Kiss, Katalin É. 1994. "Sentence structure and word order." In: F. Kiefer and K. É. Kiss (eds.), 1-90.
- Kiss, Katalin É. (ed.). 1995. *Discourse configurational languages*. Oxford: OUP.
- Kiss, Katalin É. 1998a. "Discourse-configurationality in the languages of Europe." In: Siewierska (ed.): 681-727.
- Kiss, Katalin É. 1998b. "Identificational focus versus information focus." *Language* 74: 245-273.
- Kompeer, Kees. 1992. "A note on word order and its meaning: *spit bábuška* versus *bábuška spit*." In: A. A. Barentsen, B. M. Groen, R. Sprenger (eds.). *Studies in Russian linguistics*: 215-225. Amsterdam: Rodopi.
- König, Ekkehart 1991. *The meaning of focus particles: A comparative perspective*. London: Routledge.
- Koopman, Willem and Wim Van Der Wurff. 2000. "Two word order patterns in the history of English: Stability, variation, and change." In: R. Sornicola, E. Poppe and A. Shisha-Halevy (eds.), *Stability, variation and change of word-order patterns over time*, 259-283. Amsterdam: Benjamins.

- Kovtunova, Irina I. 1965. "O ponjatii inversii." In: *Problemy sovremennoi filologii*: 167-171. Moskva.
- Kovtunova, Irina I. 1976. *Sovremennyj russkij jazyk. Porjadok slov i aktual'noe členenie predloženiya*. Moskva: Prosveščeniye.
- Krapova, Iliyana. 2001. "Subjunctives in Bulgarian and Modern Greek." In: Rivero and Ralli (eds.), 105-126.
- Kratzer, Angelika. 1995. "Stage-level and individual-level predicates." In: G. N. Carlson and F. J. Pelletier (eds), *The generic book*, 125-175. Chicago: University of Chicago Press.
- Krifka, Manfred. 1991. "A compositional semantics of multiple focus constructions." *SALT* 1: 127-158.
- Krylova, Olga A. and S. A. Khavronina. 1986. *Porjadok slov v russkom jazyke*. Moskva: Ruskij Jazyk.
- Kuno, Susumu. 1972. "Functional sentence perspective: A case study from Japanese and English." *Linguistic Inquiry* 3: 269-320.
- van Kuppevelt, Jan. 1995. "Discourse structure, topicality and questioning." *Journal of Linguistics* 31: 109-147.
- van Kuppevelt, Jan. 1996. "Inferring from topics: Scalar implicature as topic-dependent inferences." *Linguistics and Philosophy* 19: 393-443.
- Kuroda, Sige-Yuki. 1972. "The categorical and the thematic judgment." *Foundations of Language* 9: 153-185.
- Kuroda, Sige-Yuki. 1976. "Subject." In: M. Shibatani (ed.). *Japanese generative grammar*: 1-16. New York: Academic Press. [*Syntax and Semantics* 5].
- Kuroda, Sige-Yuki. 1987. "A study of the so-called topic *wa* in passages from Tolstoj, Lawrence and Faulkner (of course, in Japanese translation)." In: Hinds, Maynard and Iwasaki (eds.): 143-161.
- Kuroda, Sige-Yuki. 1990. "The categorical and the thematic judgment reconsidered." In: K. Mulligan (ed.). *Mind, meaning and metaphysics*: 77-88. Dordrecht: Kluwer.
- Ladd, D. Robert. 1996. *Intonational phonology*. Cambridge: CUP.
- Ladusaw, William A. 2000 [=1994]. "Thematic and categorical, stage and individual, weak and strong." In: L. R. Horn and Y. Kato (eds.). *Negation and polarity. Syntactic and semantic perspectives*: 232-242. Oxford: OUP. [= *SALT* 4, New York: Cornell University].
- Lambrecht, Knud. 1986. *Topic, focus, and the grammar of spoken French*. Ph.D. diss., University of California: Berkeley.

- Lambrecht, Knud. 1987. "Sentence focus, information structure, and thethetic-categorical distinction." *BLS* 13: 366-381.
- Lambrecht, Knud. 1988. "There was a farmer had a dog: Syntactic amalgams revisited." *BLS* 14: 319-339.
- Lambrecht, Knud. 1994. *Information structure and sentence form*. Cambridge: CUP.
- Lambrecht, Knud. 1995. "The pragmatics of case: On the relationship between semantic, grammatical, and pragmatic roles in English and French." In: M. Shibatani and S. Thompson (eds.), *Essays in Semantics and Pragmatics. In Honor of Charles J. Fillmore*, 145-190. Amsterdam: Benjamins.
- Lambrecht, Knud. 2000. "When subjects behave like objects: An analysis of the merging of S and O in sentence-focus constructions across languages." *Studies in Language* 24: 611-682.
- Lambrecht, Knud. 2001. "Varieties of cleft constructions and why they exist", paper presented at University of Düsseldorf 08.06.2001.
- Lambrecht, Knud and Maria Polinsky. 1997. "Typological variation in sentence-focus constructions." *CLS* 33, II: *Panels*: 189-206.
- Lambrecht, Knud and Laura A. Michaelis. 1998. "Sentence accent in information questions: Default and projection." *Linguistics and Philosophy* 21: 477-544.
- Lappin, Shalom. (ed.). 1996. *The handbook of contemporary semantics*. Oxford: Blackwell.
- LaPolla, Randy J. 1995. "Pragmatic relations and word order in Chinese." In: Downing and Noonan (eds.): 297-329.
- Lascaratou, Chryssoula. 1989. *A functional approach to constituent order with particular reference to Modern Greek*. Athens: Parousia Journal Monograph Series.
- Lascaratou, Chryssoula. 1998. "Basic characteristics of Modern Greek word order." In: Siewierska (ed.), 151-171.
- Le Bidois, Robert. 1952. *L'inversion du sujet dans la prose contemporaine (1900-1950)*. Paris: d'Artey.
- Leech, Geoffrey. 1983. *Principles of pragmatics*. London: Longman.
- Levin, Beth and Tova Rapoport. 1988. "Lexical subordination." In: *CLS* 24: 275-289.
- Levin, Beth and Milka Rappaport Hovav. 1995. *Unaccusativity: At the syntax-lexical semantics interface*. Cambridge, Mass.: MIT Press.
- Levinson, Stephen C. 1983. *Pragmatics*. Cambridge: CUP.
- Lewis, David. 1979. „Scorekeeping in a language game“. *Journal of Philosophical Logic* 8: 339-359.

- Li, Charles N. (ed.). 1975. *Word order and word order change*. Austin: University of Texas Press.
- Li, Charles N. (ed.) 1976. *Subject and topic*. New York: Academic Press.
- Li, Charles N. and Sandra A. Thomsson. 1976. "Subject and topic. A new typology of languages." In: Li (ed.): 457-490.
- Löbner, S. 1990. *Wahr neben Falsch. Duale Operatoren als die Quantoren natürlicher Sprache*. Tübingen: Niemeyer.
- Loepfe, Alfred. 1940. *Die Wortstellung im griechischen Sprechsatz, erklärt an Stücken aus Platon und Menander*. Freiburg i.d. Schweiz: Paulusdruckerei.
- Longacre, Robert E. 1995. "Left shifts in strongly VSO languages." In: Downing and Noonan (eds.), 331-354.
- Longacre, Robert E. 1999. "A footnote to Lehmann's OV/VO typology." In: Justus and Polomé (eds.), 374-381.
- Longobardi, Giuseppe. 2000. "«Postverbal» subjects and the mapping hypothesis." *Linguistic Inquiry* 31:691-701.
- Lonzi, Lidia. 1974. "L'articolazione presupposizione–asserzione e l'ordine V-S in italiano." In: Medici and Sangregorio (eds.) 1: 197-215.
- Lötscher, Andreas. 1992. "The pragmatics of nonreferential topics in German (and other languages)." In: Dittmar (ed.), 123-145.
- Luraghi, Silvia. 1995. "The pragmatics of verb initial sentences in some ancient Indo-European languages." In: Downing and Noonan (eds.): 355-386.
- MacCanna, P. 1973. "On Celtic word order and the Welsh 'abnormal' sentence." *Ériu* 24: 90-120.
- Mackridge, Peter. 1985. *The Modern Greek language*. Oxford: OUP.
- Mackridge, Peter. 1993. "An editorial problem in medieval Greek texts: The position of the object clitic pronoun in the Escorial Digenes Akrites." In: N. M. Panayotakis (ed.), *Arxetis neolinikis loyotexnias / Origini della letteratura neogreca*, 325-342. Venezia: Istituto Ellenico di Studi Bizantini e Postbizantini di Venezia.
- McCawley, James D. 1993. *Everything that linguists have always wanted to know about logic, but were ashamed to ask*. Second edition. Chicago: Chicago University Press.
- McCawley, Noriko A. 1973. "Boy, is syntax easy!." *CLS* 9: 369-377.
- McNally, Louise. 1994. "Adjunct predicates and the individual/stage distinction." *WCCFL* 12: 561-576.

- McNally, Louise. 1997. "On recent formal analyses of topic." In: J. Groenendijk, D. de Jongh and M. Stokhof (eds.), *The Tbilisi Symposium on language, logic, and computation: Selected papers*, 149-162. Stanford-Cambridge: CSLI and CUP.
- McNally, Louise. 1998. "Stativity and theticity." In: S. Rothstein (ed.), *Events and grammar*, 293-307. Dordrecht: Kluwer.
- Marty, Anton. 1918. *Gesammelte Schriften II.1: Schriften zur deskriptiven Psychologie und Sprachphilosophie*. Halle: Niemeyer.
- Maslova, Elena. 1995. "VS/SV opposition in Russian and the thetic/categorical distinction." In: Matras and Sasse (eds.), 106-124.
- Maslova, Elena. 1997. "Yukagir focus in a typological perspective." *Journal of Pragmatics* 27: 457-475.
- Maslova, Elena. 2003. *A grammar of Kolyma Yukaghir*. New York: Mouton de Gruyter.
- Mathesius, Vilém. 1929 = 1983. "Functional linguistics". In: J. Vachek (ed.), *Praguiana*, 121-142. Amsterdam: Benjamins.
- Mathesius, Vilém. 1939. "O tak zvaném aktuálním členění větném." *Slovo a slovesnost* 5: 171-184. [English translation: "On information-bearing structure of the sentence." In: S. Kuno (ed.): *Harvard Studies in Syntax and Semantics* 1: 467-480]
- Mathesius, Vilém. 1975. *A Functional analysis of present-day English on a general linguistic basis*. transl. L. Dušková, ed. J. Vachek. Prague: Academia.
- Matić, Dejan. 2002. "Topik, Fokus und Parameter der Wortstellungsvariation." MS, University of Cologne.
- Matić, Dejan. 2003. "Topic, focus, and discourse structure: Ancient Greek word order." *Studies in Language* 27: 573-633.
- Matras, Yaron and Hans-Jürgen Sasse (eds.) 1995. *Verb-subject order and theticity in European languages*. Berlin: Akademie-Verlag. [*Sprachtypologie und Universalienforschung* 48:1/2].
- Matras, Yaron. 1995. "Connective (VS) order in Romani." In: Matras and Sasse (eds.), 189-203.
- Medici, Mario and Antonella Sangregorio (eds.). 1974. *Fenomeni morfologici e sintattici nell'italiano contemporaneo. Atti del sesto congresso internazionale di studi*. Roma: SLI/Bulzoni.
- Mervis, C.B. and Eleanor Rosch. 1981. "Categorization of natural objects." *Annual Review of Psychology* 32: 89-115.
- Meyer-Lübke, Wilhelm. 1899. *Romanische Syntax*. Leipzig: Reisland.

- Michaelis, Laura A. and Knud Lambrecht. 1996. "The exclamative sentence type in English." In: Adele Goldberg (ed.), *Conceptual structure, discourse and language*. Stanford: CSLI.
- Milsark, Gary 1974. *Existential sentences in English*. Ph.D. Diss. MIT, Cambridge, Mass.
- Milsark, Gary 1977. "Toward an explanation of certain peculiarities of the existential construction in English." *Linguistic Analysis* 3: 1-29.
- Mistrík, Jan. 1966. *Slovosled a vetosled v slovenčine*. Bratislava.
- Mohanan, K. P. and Tara Mohanan. 1999. "Representing presuppositions." In: T. Mohanan and L. Wee (eds.). *Grammatical semantics*: 217-259. Stanford and Singapore: CSLI and National University of Singapore.
- Molnár, Valéria. 1993. "Zur Pragmatik und Grammatik des TOPIK-Begriffes." In: Reis (ed.), 155-201.
- Molnár, Valéria. 1998. "Topic in focus." *Acta Linguistica Hungarica* 45: 89-166.
- Morocho Gayo, Gaspar. 1985. "Prosa griega y orden de palabras: una aproximacion." In: G. Morocho Gayo (ed.). *Estudios de prosa griega*: 141-177. Universidad de León.
- Motapanyane, Virginia 1997. "Preverbal focus in Bulgarian." *Journal of Slavic Linguistics* 5: 265-299.
- Müller-Hauser, M.L. 1943. *La mise en relief d'une idée en français moderne*. Genève: Droz.
- Mulligan, K. (ed.). *Mind, meaning and metaphysics*. Dordrecht: Kluwer.
- Myhill, John. 1984. *A study of aspect, word order, and voice*. Ph.D. Diss., University of Pennsylvania.
- Myhill, John. 1985. "Pragmatic and categorial correlates of VS word order." *Lingua* 66: 177-200.
- Myhill, John. 1986. "The two VS constructions in Rumanian." *Linguistics* 24: 331-350.
- Myhill, John. 1992a. "Word order and temporal sequencing." In: Payne (ed.): 265-278.
- Myhill, John. 1992b. *Typological discourse analysis*. Oxford: Blackwell.
- Nikolaeva, Irina. 2001. "Secondary topic as a relation in information structure." *Linguistics* 39: 1-49.
- Ocampo, Francisco. 1990. "The pragmatics of word order in constructions with a verb and a subject." *Hispanic Linguistics* 4: 87-127.
- Ocampo, Francisco. 1995. "The word order of two-constituent constructions in spoken Spanish." In: Downing and Noonan (eds.): 425-447.
- Ogawa, Yoshiki. 2001. "The stage/individual distinction and (in)alienable possession." *Language* 77: 1-25.

- Oh, C.-K. and D. Dinneen (eds.). 1979. *Presupposition*. New York: Academic Press. [*Syntax and Semantics* 11].
- Önnerfors, Olaf. 1997. *Verb-erst-Deklarativsätze. Grammatik und Pragmatik*. Stockholm: Almqvist & Wiksell.
- van Oosten, Jeanne. 1986. *The nature of subjects, topics and agents: A cognitive explanation*. Bloomington, Indiana: IULC.
- Ortiz de Urbina, Jon. 1989. *Parameters in the grammar of Basque*. Dordrecht: Foris.
- Papafragou, Anna. 2000. "On speech-act modality." *Journal of Pragmatics* 32: 519-538.
- Partee, Barbara. 1975. "Montague grammar and transformational grammar." *Linguistic Inquiry* 6: 203-300.
- Pasch, Renate. 1981. "Die Konzeption der "kommunikativen Dynamik" (CD) von Sätzen und die Grammatiktheorie." In: F. Daneš and D. Viehweger (eds.). *Satzsemantische Komponenten und Relationen im Text*: 64-87. Praha: Academia.
- Payne, Doris L. (ed.). 1992. *Pragmatics of word order flexibility*. Amsterdam: Benjamins.
- Penhallurick, John. 1984. "Full-verb inversion in English." *Australian Journal of Linguistics* 4: 33-56.
- Penčev, Jordan. 1980. *Osnovni intonacionni konturi v bälgarskoto izrečenie*. Sofia: Bälgarska Akademija na Naukite.
- Perlmutter, David. 1978. "Impersonal passives and the Unaccusative Hypothesis." *BLS* 4: 159-189.
- Philippaki-Warburton, Irini. 1985. "Word order in Modern Greek." *Transactions of the Philological Society* 1985: 113-145.
- Philippaki-Warburton, Irini. 1987. "The theory of empty categories and the pro-drop parameter in Modern Greek." *Journal of Linguistics* 23: 289-318.
- Philippaki-Warburton, Irini. 1992. "I sindaktiki taftotita tu na." *Meletes ja tin Eliniki ylosa / Studies in Greek Linguistics* 13: 255-274.
- Pierrehumbert, Janet and Julia Hirschberg. 1990. "The meaning of intonational contours in the interpretation of discourse." In: P. R. Cohen; J. Morgan; and M. E. Pollack (eds.). *Intentions in communication*, 271-323. Cambridge, Mass.: MIT Press.
- Pinto, Manuela. 1997. *Licencing and interpretation of inverted subjects in Italian*. Utrecht: Utrecht Institute of Linguistics.
- Polinsky, Maria. 1990. "Subject incorporation: Evidence from Chukchee." In: K. Dziwirek et al. (eds.), *Grammatical relations*, 349-364. Stanford: CSLI.

- Polinsky, Maria. 1993. "Subject inversion and intransitive subject incorporation." *CLS* 29: 342-362.
- Popović, Ljubomir. 1997. *Red reči u rečenici*. Beograd: DSJKS.
- Portner, Paul and Katsuhiko Yabushita. 1998. "The semantics and pragmatics of topic phrases." *Linguistics and Philosophy* 21: 117-157.
- Prince, Ellen F. 1978a. "A comparison of *wh*-clefts and *it*-clefts in discourse." *Language* 54: 883-906.
- Prince, Ellen F. 1978b. "On the function of existential presupposition in discourse." *CLS* 14: 362-376.
- Prince, Ellen F. 1981. "Toward a taxonomy of given / new information." In: Cole (ed.): 223-254.
- Prince, Ellen F. 1985. "Fancy syntax and shared knowledge." *Journal of Pragmatics* 9: 65-81.
- Prince, Ellen F. 1986. "On the syntactic marking of presupposed open propositions." *CLS* 22, 2: *Papers from the Parasession on Pragmatics and Grammatical Theory*, 208-222.
- Prince, Ellen F. 1992. "The ZPG letter: Subjects, definiteness, and information status." In: S. Thompson and W. Mann (eds.). *Discourse description: Diverse analyses of a fund-raising text*: 295-325. Amsterdam: Benjamins.
- Prince, Ellen F. 1998. "On the limits of syntax, with reference to left-dislocation and topicalization." In: Culicover and McNally (eds.), 281-302.
- Progovac, Ljiljana. 1998. "Event pronominal *to*". *Journal of Slavic Linguistics* 6: 3-39.
- Quine, Willard Van Orman. 1960. *Word and object*. Cambridge, MA.: MIT Press.
- Raymond, W.R and K. Homer. 1996. "The interaction of pragmatic roles and thematic structure in the selection of question form." *BLS* 22.
- Raynaud, Savina. 1992. "Anton Marty (1847-1914)." In: M. Dascal; D. Gerhardus; K. Lorenz; and G. Meggle (eds.). *Handbücher zur Sprach- u. Kommunikationswissenschaft, 7.1: Sprachphilosophie*: 445-467. Berlin-New York: de Gruyter.
- Reinhard, Tanya. 1982. *Pragmatics and linguistics: An analysis of sentence topics*. Bloomington, Indiana: IULC.
- Reis, Marga (ed.). 1993. *Wortstellung und Informationsstruktur*. Tübingen: Niemeyer.
- Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.
- Rivero, María-Luisa. 1993. "Bulgarian and Serbo-Croatian yes-no questions: V^0 -raising to *-li* versus *-li* hopping." *Linguistic Inquiry* 24: 567-575.
- Rivero, María Luisa and Angela Ralli (eds.). 2001. *Comparative syntax of Balkan languages*. Oxford: OUP.

- Robblee, Karen E. 1994. "Russian word order and the lexicon." *Journal of Slavic Linguistics* 2: 238-267.
- Roberts, Craige. 1996. "Information structure in discourse: Towards an integrated formal theory of pragmatics." *Ohio State University WP in Linguistics* 49: 91-136.
- Roberts, Craige. 1998. "Focus, the flow of information, and Universal Grammar." In: Culicover and McNally (eds.), 109-160.
- Rochemont, Michael and Peter Culicover. 1990. *English focus constructions and the theory of grammar*. Cambridge: CUP.
- Rooth, Mats. 1992. "A theory of focus interpretation." *Natural Language Semantics* 1: 75-116.
- Rooth, Mats. 1996. "Focus." In: Lappin (ed.), 271-297.
- Rooth, Mats. 1999. "Association with focus or association with presupposition?." In: Bosch and van der Sand (eds.), 232-244.
- Rosengren, Inger. 1997. "Thethetic/categorical distinction revisited once more." *Linguistics* 35: 439-479.
- Rothenberg, M. 1971. "Les propositions relatives à antécédent explicite introduites par des présentatifs." *Études de linguistique appliquée* 2: 102-117.
- Roussou, Anna. 2000. "On the left periphery. Modal particles and complementisers." *Journal of Greek Linguistics* 1: 65-94.
- Roussou, Anna. 2001. "Control and raising in and out of subjunctive complements." In: Rivero and Ralli (eds.), 74-104.
- Rubovitz, Tali. 1999. "Evidential-existentials: The interaction between discourse and sentence structure." *Journal of Pragmatics* 31: 1025-1040.
- Rude, Noel. 1992. "Word order and topicality in Nez Perce." In: Payne (ed.): 193-208.
- Rushi, Th. 1983. "Rreth rendit të gjymtyrëve ne fjalën dëftore." *Studime filologjike* 1983-1: 71-85.
- Rushi, Th. 1984. "Funksioni gramatikor i rendit të fjalëve në gjuhën shqipe", *Studime filologjike* 1984-4: 47-56.
- Rushi, Th. 1985. "Çështje të normës letrare në rendin e fjalëve në shqipen e sotme", *Studime filologjike* 1985-1: 117-121.
- Rushi, Th. 1988. "Rendi i gjymtyrëve të shkallës së dytë", *Studime filologjike* 42: 105-112.
- Saeed, John I. 1984. *The syntax of topic and focus in Somali*. Hamburg: Buske.
- Safir, Kenneth. 1985. *Syntactic chains*. Cambridge: CUP.
- Saltarelli, Mario. 1981. "Post-verbal subjects in Italian." *CLS* 17: 361-367

- Samek-Lodovici, Vieri. 1998. "When preverbal focus is clause-final: A unified analysis of contrastive focus in Italian." Paper presented at the Workshop on Romance Inversion, University of Amsterdam.
- van der Sandt, Rob. 1988. *Context and presupposition*. London: Routledge.
- van der Sandt, Rob. 1992. "Presupposition projection as anaphora resolution." *Journal of Semantics* 9: 333-373.
- Sasse, Hans-Jürgen. 1982. "Subjektprominenz." In: H. Sieglinde and U. Wandruszka (eds.), *Festschrift für H. Stimm zum 65. Geburtstag*, 267-285. Tübingen: Narr.
- Sasse, Hans-Jürgen. 1984. "The pragmatics of noun incorporation in Eastern Cushitic languages." In: F. Plank (ed.), *Objects*, 243-268. London: Academic Press.
- Sasse, Hans-Jürgen. 1987. "Thethetic/categorical distinction revisited." *Linguistics* 25: 511-580.
- Sasse, Hans-Jürgen. 1995a. "A contrastive study of VS clauses in Modern Greek and Hungarian." In: Matras and Sasse (eds.), 142-188.
- Sasse, Hans-Jürgen. 1995b. "'Theticity' and VS order: a case study." In: Matras and Sasse (eds.), 3-31.
- Sasse, Hans-Jürgen. 1995c. "Prominence typology." In: J. Jacobs, A. von Stechow, W. Sternefeld and T. Vennemann (eds.), *Handbücher zur Sprach- u. Kommunikationswissenschaft, Bd. 9.2: Syntax*, 1065-1075. Berlin: de Gruyter.
- Sasse, Hans-Jürgen. 1996. *Theticity*. Institut für Sprachwissenschaft, Universität zu Köln [Arbeitspapier 27 (NF)].
- Schachter, Paul. 1977. "Reference-related and role-related properties of subjects." In: P. Cole and J. M. Sadock (eds.), *Grammatical relations*, 279-305. New York: Academic Press. [Syntax and Semantics 8].
- Schlobinski, Peter and Stephen Schütze-Cobrun. 1992. "On the topic of topic and topic continuity." In: Dittmar (ed.), 89-121.
- Schmerling, Susan F. 1976. *Aspects of English sentence stress*. Austin: University of Texas Press.
- Schroeder, Christoph. 1995. "Post-predicate arguments in Turkish." In: Matras and Sasse (eds.), 204-218.
- Seifert, Soula. 1984. *Om ordföljen i nygrekiska bisatser*. Stockholm: Institutionen for klassiska språk, Stokholms universitet.
- Selkirk, Elisabeth. 1984. *Phonology and syntax*. Cambridge, Mass.: MIT Press.
- Seuren, Pieter A.M. 1998. *Western linguistics. An historical introduction*. Oxford: Blackwell.

- Seuren, Pieter A.M. 1999. "Topic and comment." In: Justus and Polomé (eds.), 348-373.
- Seuren, Pieter A.M. 2000. "Presupposition, negation and trivalence." *Journal of Linguistics* 36: 261-297.
- Shank, Roger C. and Robert P. Abelson. 1977. *Scripts, plans, goals, and understanding*. Hillsdale, NJ: Lawrence Erlbaum.
- Shibatani, Masayoshi. 1991. "Grammaticalization of topic into subject." In: E. C. Traugott and B. Heine (eds.). *Approaches to grammaticalization*, 2: 93-133. Amsterdam: Benjamins.
- Siewierska, Anna. 1987. "Postverbal subject pronouns in Polish in the light of topic continuity and the topic/focus distinction." In: J. Nuyts and G. de Schutters (eds.), *Getting one's word in line: On word order and Functional Grammar*, 147-161. Dordrecht: Foris.
- Siewierska, Anna. 1990. "The source of the dative perspective in Polish pseudo-reflexives". In: M. Hannay and E. Vester (eds.), *Working with FG: Descriptive and computational applications*, 1-15. Dordrecht, Foris.
- Siewierska, Anna (ed.). 1998. *Constituent order in the languages of Europe*. Berlin: de Gruyter.
- Simspon, Jane. 1999. "Comments on K.P. and T. Mohanan." In: T. Mohanan and L. Wee (eds.). *Grammatical semantics*: 261-275. Stanford and Singapore: CSLI and National University of Singapore.
- Sirotnina, Olga B. 1974. *Sovremennaja razgovornaja reč' i ee osobennosti*. Moskva: Prosveščeniye.
- Skënderi, Arion. 1997. *Functional sentence perspective in Albanian texts*. Wien: WUV.
- Soames, Scott. 1982. "How presuppositions are inherited: A solution to the projection problem." *Linguistic Inquiry* 13: 483-545.
- Sperber, Dan and Deirdre Wilson. 1981. "Irony and the use-mention distinction." In: Cole (ed.), 295-318.
- Sperber, Dan and Deirdre Wilson. 1986. *Relevance: Communication and cognition*. Oxford: Blackwell; Cambridge, Mass.: Harvard University Press.
- Stalnaker, Robert C. 1974. "Pragmatic presuppositions." In: M. K. Munitz and P. K. Unger (eds.). *Semantics and philosophy*. New York: New York University Press.
- Stalnaker, Robert C. 1978. "Assertion." In: Cole (ed.): 315-332.
- Stavrou, Melita. 1984. "I klitiki andinomia stis perioristikes anaforikes protasis me eksartisi amesu andikimenu pu isayonde me to pu." *Meletes ja tin Eliniki ylosa / Studies in Greek Linguistics* 5: 121-136.

- Strawson, Peter F. 1950. "On referring." *Mind* 59: 320-344.
- Strawson, Peter F. 1964. "Identifying reference and truth values." *Theoria* 30: 96-118.
- von Stechow, Arnim. 1991. "Current issues in the theory of focus." In: A. von Stechow and D. Wunderlich (eds.), *Handbücher zur Sprach- und Kommunikationswissenschaft 6: Semantics*, 804-825. Berlin: de Gruyter.
- Szabolcsi, Anna. 1980. "Az aktuális mondattagolás szemantikájához." *Nyelvtudományi Közlemények* 82: 59-83.
- Szabolcsi, Anna. 1981. "The semantics of topic-focus articulation." In: J. Groenendijk; T. Janssen; and M. Stokhof (eds.), *Formal methods in the study of language 2*: 513-540. Amsterdam: Mathematisch Centrum.
- Szabolcsi, Anna. 1983. "Focusing properties, or the trap of first order." *Theoretical Linguistics* 10: 125-145.
- Szabolcsi, Anna. 1994. "All quantifiers are not equal: The case of focus." *Acta linguistica Hungarica* 42: 171-187.
- Szober, Stanisław. 1933. "Edin osnoven tip nareždane dumite v bǎlgarskoto izrečenie." Reprinted in: K. Popov. 1979. *Pomagalo na bǎlgarski sintaksis*: 275-285. Sofia: Nauka i izkustvo.
- Svalberg, Agneta. 1992. "Functions and encodings of complements in Modern Greek from a crosslinguistic perspective." *Meletes ja tin Eliniki γλoσa / Studies in Greek Linguistics* 13: 295-309.
- Thumb, Albert. 1910. *Handbuch der neugriechischen Volkssprache*. Straßburg: Trübner.
- Tomlin, Russell S. 1995. "Focal attention, voice, and word order: An experimental, cross-linguistic study." In: Downing and Noonan (eds.): 517-554.
- Tortora, Christina M. 1999. "The post-verbal subject position of Italian unaccusative verbs of inherently directed motion." In: E. Treviño and J. Lema (eds.), *Semantic Issues in Romance Syntax*, 283-298. Amsterdam: Benjamins.
- Tosco, Mauro. 2002. "Incorporation in Somali (and beyond)." Paper presented at University of Cologne 14.05.2002.
- Trávníček, František. 1939. "Slovosled při důrazu." *Slovo a slovesnost* 5: 131-144.
- Trávníček, František. 1951. *Mluvnice spisovné češtiny. Část 2: Skladba*. Praha.
- Trávníček, František. 1961. "O tak zvaném aktuálním členění větěném." *Slovo a slovesnost* 22: 163-171.
- Trypanis, C.A. 1960. "Early medieval Greek *hina*." *Glotta* 38: 312-313.

- Tsimpli, Ianthi Maria. 1990. "The clause structure and word order of Modern Greek." *UCL Working Papers in Linguistics* 2: 226-255.
- Tsimpli, Ianthi Maria. 1995. "Focusing in Modern Greek." In: É. Kiss (ed.): 176-206.
- Tsimpli, Ianthi Maria. 1998. "Individual and functional readings for focus, wh- and negative operators: Evidence from Greek." In: Joseph, Horrocks, and Philippaki-Warbuton (eds.): 197-227.
- Turano, Giuseppina. 1993. "Subjunctive constructions in Arbëresh and standard Albanian." *Rivista di Grammatica Generativa* 18: 101-133.
- Turano, Giuseppina. 1994. "Elementi *wh* e indefiniti in albanese." In: G. Borgato (ed.), *Teoria del linguaggio e analisi linguistica. XX Incontro die Grammatica Generativa*, 409-424. Padova: Unipress.
- Turano, Giuseppina. 1995. *Dipendenze sintattiche in albanese*. Padova: Unipress.
- Tzanidaki, Dimitra Irini. 1996. *The syntax and pragmatics of subject and object position in Modern Greek*. Ph.D. Dissertation, London University.
- Tzanidaki, Dimitra Irini. 1998. "Clause structure and word order in Modern Greek." In: Joseph, Horrocks and Philippaki-Warbuton (eds.), 229-254.
- Tzartanos, Achileus. 1963². *Neoliniki sindaksi*. Vol. 2. Athens: OEDV.
- Ueno, Noriko Fujii. 1987. "Functions of the theme marker *wa* from synchronic and diachronic perspectives." in: Hinds, Maynard and Iwasaki (eds.): 221-263.
- Ulrich, Miorita. 1985. *Thetisch und Kategorisch*. Tübingen: Narr.
- Vachek, Josef (ed.). 1964. *A Prague School Reader in Linguistics*. Bloomington, Indiana: Indiana University Press.
- Valioui, Maria. 1994. "Postverbal subjects, reduplicated objects or retopicalization." In: I. Philippaki-Warbuton, K. Nikolaidis and M. Sifianou (eds.), *Themes in Greek Linguistics I*, 85-92. Amsterdam: Benjamins.
- Vallduví, Enric. 1992. *The informational component*. New York: Garland.
- Vallduví, Enric. 1994. "The dynamics of information packaging." In: E. Engdahl (ed.). *Integrating information structure into constraint-based and categorial approaches*: 1-27, Amsterdam: University of Amsterdam.
- Vallduví, Enric and Elisabeth Engdahl. 1996. "The linguistic realization of information packaging". *Linguistics* 34: 459-519.
- Vallduví, Enric and Maria Vilkuña. 1998. "On rheme and kontrast (sic!)." In: Culicover and McNally (eds.), 79-108.

- Van Valin, Robert. (ed.) 1993. *Advances in Role and Reference Grammar*. Amsterdam: Benjamins.
- Wackernagel, Jakob. 1892. "Über ein Gesetz der indogermanischen Wortstellung." *Indogermanische Forschungen* 1: 333-436.
- Wandruszka, Ulrich. 1982. *Studien zur italienischen Wortstellung. Wortstellung – Semantik – Informationsstruktur*. Tübingen: Narr.
- Ward, Gregory. 1990. "The discourse functions of VP preposing." *Language* 66: 742-763.
- Ward, Gregory. 1998. "A comparison of postposed subjects in English and Italian." In: A. Kamio, K.-I. Takami (eds.). *Function and structure. In honor of Susumu Kuno*: 3-21. Amsterdam: Benjamins.
- Ward, Gregory and Betty J. Birner. 1995. "Definiteness and the English existential." *Language* 71: 722-742.
- Ward, Gregory and Ellen F. Prince. 1991. "On the topicalization of indefinite NPs." *Journal of Pragmatics* 16: 167-177.
- Watkins, Calvert. 1963. "Preliminaries to a historical and comparative analysis of the syntax of the Old Irish verb." *Celtica* 6: 1-49.
- Wehr, Barbara. 1984. *Diskursstrategien im Romanischen*. Tübingen: Narr.
- Wehr, Barbara. 2000. "Zur Beschreibung der Syntax des *français parlé* (mit einem Exkurs zu 'thetisch' und 'kategorisch')." In: B. Wehr and H. Thomaßen (eds.). *Diskursanalyse: Untersuchungen zum gesprochenen Französisch*: 239-288. Frankfurt a. M.: Lang.
- Weil, Henry. 1844. *De l'ordre des mots dans les langues anciennes comparées aux langues modernes*. Paris: Joubert.
- Wilson, Deirdre and Dan Sperber. 1979. "Ordered entailments: An alternative to presuppositional theories." In: Oh and Dinneen (eds.), 299-323.
- Wu, Guo. 1992. "Thethetic-categorical distinction as expressed by subject-predicate sentences in Chinese." *Australian Journal of Linguistics* 12: 271-296.
- Xu, Liejiong. 2000. "The topic-prominence parameter." *ZAS Papers in Linguistics* 20: 21-42.
- Zubizarreta, Maria Luisa. 1998. *Prosody, focus, and word order*. Cambridge, Mass.: MIT Press.