

Szűcs Péter

Left peripheral constructions in English and Hungarian

*A perspective from
Lexical-Functional Grammar*



Szűcs Péter

**LEFT PERIPHERAL CONSTRUCTIONS
IN ENGLISH AND HUNGARIAN**

A perspective from Lexical-Functional Grammar

DUPress

DUPress

Szűcs Péter

**LEFT PERIPHERAL CONSTRUCTIONS
IN ENGLISH AND HUNGARIAN**

A perspective from Lexical-Functional Grammar

DUPRESS



Debreceni Egyetemi Kiadó
Debrecen University Press
2020

Lektorálta:

ALBERTI GÁBOR
Pécsi Tudományegyetem

SURÁNYI BALÁZS
Pázmány Péter Katolikus Egyetem

© Debreceni Egyetemi Kiadó Debrecen University Press,
beleértve az egyetemi hálózaton belüli elektronikus terjesztés jogát is

ISBN 978-963-318-885-9

Kiadta: a Debreceni Egyetemi Kiadó Debrecen University Press
Felelős kiadó: Karácsony Gyöngyi
Nyomdai munkálatokat
a Debreceni Egyetem sokszorosítóüzeme végezte 2020-ban
dupress.unideb.hu

TABLE OF CONTENTS

ABSTRACT	7
ACKNOWLEDGEMENTS	9
LIST OF ABBREVIATIONS	11
CHAPTER 1 INTRODUCTION	12
1.1 Aims of the research	12
1.2 Lexical-Functional Grammar	15
1.3 Information structure	21
1.3.1 Features of information structure.....	22
1.3.2 Categories of information structure	29
1.4 Conclusion to chapter 1	37
CHAPTER 2 LFG PERSPECTIVES ON HUNGARIAN SYNTAX	38
2.1 The simple clause in Hungarian	38
2.2 Subordinate clauses in Hungarian	44
2.2.1 General properties of Hungarian subordinate clauses	46
2.2.2 The expletive-demonstrative debate	52
2.3 Conclusion to chapter 2	67
CHAPTER 3 CONSTRUCTIONS ON THE ENGLISH LEFT PERIPHERY	72
3.1 Topicalization in English.....	72
3.1.1 Properties of topicalization	72
3.1.2 Analysis of topicalization	77
3.2 Clause-initial adjuncts in English.....	79
3.2.1 Properties of clause-initial adjuncts.....	80
3.2.2 Analysis of clause-initial adjuncts	82
3.3 Left dislocation in English.....	85
3.3.1 Properties of left dislocation in English.....	85
3.3.2 Analysis of left dislocation in English.....	89
3.4 Conclusion to chapter 3	91
CHAPTER 4 CONSTRUCTIONS ON THE HUNGARIAN LEFT PERIPHERY	92
4.1 Left dislocation in Hungarian	92
4.1.1 Properties of topic left dislocation	93
4.1.2 Analysis of topic left dislocation	100
4.1.3 Properties of free left dislocation.....	107
4.1.4 Analysis of free left dislocation	108
4.2 Operator fronting in Hungarian	109
4.2.1 Properties of operator fronting.....	110

4.2.2 Analysis of operator fronting in Hungarian	122
4.2.3 On split operator fronting	139
4.2.4 On long topicalization	141
4.3 Conclusion to chapter 4.....	142
CHAPTER 5 CROSSLINGUISTIC AND THEORETICAL PERSPECTIVES	143
5.1 Perspectives on fronting.....	144
5.2 Perspectives on left dislocation.....	148
5.3 Perspectives on prolepsis	153
5.3.1 Proleptic structures cross-linguistically	153
5.3.2 Theoretical perspectives on prolepsis	164
5.4 Control in the finite domain	175
5.4.1 Finite raising.....	177
5.4.2 Finite equi.....	183
5.4.3 Summary of finite control	191
5.5 Conclusion to chapter 5.....	192
CHAPTER 6 CONCLUSION.....	195
REFERENCES.....	198

ABSTRACT

This monograph investigates discourse-related clause-initial, left peripheral constructions within the framework of Lexical-Functional Grammar (LFG). The following structures are to be scrutinized: English topicalization (TOP), clause-initial adjuncts (CIADJ) and left dislocation (LD-Eng); Hungarian left dislocation (LD-Hun) and operator fronting (OF). The overall goal is to provide an account for these structures, especially with regard to their syntactic and information structural properties and to put them into cross-linguistic and theoretical perspective.

The proposals build on several earlier approaches, expands and improves them in various ways. As for information structure, it is argued that no satisfactory account for the understanding and representation has been proposed and a new framework is therefore put forward. The proposal is an amalgamation and enhancement of earlier feature-based approaches. I argue that the features *NEW*, *DISCOURSE-STRUCTURING* and *CONTRASTIVE* are well-definable and able to account for the existing information structural categories. In syntax, I rely on Laczkó's works (2014a, 2014b, 2015) on Hungarian and supplement them in two ways. First, I incorporate my proposals about information structure into them and second, I expand them to cover subordinate clauses as well. For this latter part, I investigate the theoretical and empirical landscape and argue extensively that contrary to the standard view (Kenesei 1992/1994) the pronouns associated with Hungarian subordinate clauses *az* 'that' are semantically contentful and not expletives.

Once the theoretical background is laid, I discuss the target-constructions in detail. After examining their syntactic and information structural properties, I argue that these structures necessitate analyses where the left-peripheral constituents are integrated into the sentence-structure in various ways and to different degrees. I support my analyses with a range of evidence and provide a LFG-theoretic formal account for each of the constructions.

In English, TOP-Eng is argued to be a proper "fronting" construction, where the topicalized element is functionally identified with a clause-internal function. At information structure, TOP is always contrastive. Some CIADJs also follow this pattern, but there also exist CIADJs that are "base-generated" neutral topics, without a link to the inner part of the sentence. Left-dislocated entities in English are even less integrated: they are argued to be "syntactic orphans" which are not related to the host-sentence at the level of syntax.

In Hungarian, two types of left dislocation are distinguished: topic left dislocation (TLD) and focus left dislocation (FLD). While the first one is a integrated structure (involving an anaphoric dependency between the discourse-prominent element and the associate pronoun), FLD is not and is argued to involve a sentence-external left-peripheral entity, akin to LD-Eng. Operator fronting is also divided into two types. The first type is a proper fronting construction, sharing many properties with TOP-Eng, while the second type is a representative of "prolepsis", where the fronted element becomes a thematic argument of the main verb and is anaphorically linked to a clause-internal function.

The last part of the monograph is concerned with the theoretical and cross-linguistic ramifications of the findings and proposals put forward in the monograph. I show that "fronting", "left dislocation" and "proleptic" structures exist in various languages and substantial parallelisms may be observed with regards their properties. I compare and contrast these properties and outline a taxonomy. As prolepsis is a relatively little-studied phenomenon,

I put special emphasis on its place in syntactic typology. I argue that it may be regarded as a type of finite control and substantiate this claim with providing an overview of such constructions. I also show the place of prolepsis in the typology of control.

Overall, the monograph achieves two goals: it offers comprehensive analyses for various clause-initial discourse-related linguistic structures and it also provides a wider perspective by investigating the theoretical and cross-linguistic places for the ideas put forward.

DUPress

ACKNOWLEDGEMENTS

This monograph is an improved, updated and streamlined version of my PhD-dissertation “On clause-initial discourse-related constructions in English and Hungarian”, which I defended on 23 May 2018 at the University of Debrecen.

During the past few years, there were times when the completion of this monograph seemed so distant that I doubted that I would ever achieve it. That these doubts have remained unfounded is to a great extent a result of me being supported by a number of wonderful people, to whom I cannot be grateful enough.

First and foremost, I would like to express my gratitude towards my supervisor, Tibor Laczkó. He introduced me to the world of linguistics and LFG in particular, so but for him, the idea of me entering PhD-studies would have never occurred to me. During my research, he was always there for me to discuss my half-baked ideas, gently push me in the right directions, away from the wrong ones and he also helped me out in various formal aspects of the monograph. Without his support and guidance, the previously mentioned doubts would surely have been proven.

Rákosi György is another vital member of my close linguist community. His sharp observations and insightful suggestions have formed and improved my ideas tremendously. Also, he has been an outstanding motivator for me to keep on going with my projects at times when my internal drives were running low.

I could not wish for a better person to share an office with than Kardos Éva, whose academic attitude and scientific determination and supply of snacks for the department staff have been a major source of inspirational energy for me.

I must also individually mention Réka Jurth, Anna Ware and John Ware, whose help in the linguistic and typographic revision of the text was invaluable. Without *them*, the my monograph may look like this “sentence’ looks..

I would also like to thank the entire Institute of English and American Studies at the University of Debrecen for providing a superb background for me. This community of people, instructors, students and administrative personnel included, has been truly stimulating both professionally and personally.

I also owe a lot to the reviewers of my dissertation, on which this monograph is based on. The comments, criticisms and corrections from Gábor Alberti and Balázs Surányi have vastly improved the content.

The annual LingDok conferences in Szeged were the highlights of the academic year for me. The outstanding community of linguists I met there has really been formative for me. I hope that their knowledge, skills and attitudes have made a permanent mark on my scientific progress. I am forever indebted to the organizers for putting these conferences together and to all the participants for their insightful remarks and also for the sense of linguist-community that I have experienced there. In particular, I thank Zsuzsanna Gécseg for very helpful comments about the material that I presented at the LingDok conferences.

A similar thanks should go to the organizers and participants of all the other conferences and other linguistic events that I attended throughout the years: the LFG conferences, CECIL’S 2013, ConSOLE XXIII, DELITE PhD Symposium, Budapest Linguistics Conference,

SinFonIJA 6, EGG 2011 and 2017. I am thankful to life for making it possible for me to participate in these. These have greatly contributed to my professional and personal growth.

I am especially honored to be part of the international LFG community, whose insights have helped me in more than one way. I must mention Louisa Sadler, Louise Mycock, Miriam Butt, Rachel Nordlinger, Ron Kaplan and Tracy Holloway King specifically, who gave invaluable pieces of advice about various aspects of my research.

This section would not be complete without mentioning István Kenesei, whose constructive criticism has made me think much harder about numerous aspects of my proposals, greatly improving the end result.

Last but by no means least, I am immensely grateful to my family who supported me with a solid background throughout the years. A very special thanks goes to my father, Sándor Szűcs, whose scientific mind, academic rigor and superb work-ethic has always inspired me to bring the best out of myself.

Needless to say, all remaining inadequacies are my responsibility.

Debrecen, October 22, 2020

The research has been supported by Project no. 111918 (New approaches in the description of the grammar of Hungarian pronominals), which has been implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the K funding scheme.

LIST OF ABBREVIATIONS

Throughout the example sentences in the monograph, if not indicated otherwise, nominative case, present tense, singular number, indicative mood and neutral intonation are assumed.

ABL	ablative
ACC	accusative
AUX	auxiliary
BI	background information
CF	contrastive focus
CI	completive information
CL	clitic
COND	conditional
CT	contrastive topic
DAT	dative
DEF	definite
DEL	delative
DIST	distal
FOC	focus
FUT	future
IF	information focus
INDEF	indefinite
INF	infinitive
MASC	masculine
NOM	nominative
PASS	passive
PL	plural
POSS	possessive
PROX	proximal
PRT	particle
PV	preverb
Q-PRT	question particle
SBJV	subjunctive
SG	singular
SM	subject marker
that(c)	complementizer <i>that</i>

CHAPTER 1

INTRODUCTION

1.1 Aims of the research

The left periphery of the clause is commonly viewed as one of the most prominent areas for the interaction of syntax and information structure. This is true not only for the so-called discourse-configurational languages (where the primary structuring force behind the order of constituents is their discourse/information structural status, see e.g. É. Kiss 1995), but also in conventionally configurational ones like English, where the phrase structure rules impose a more rigid syntactic hierarchy and order of constituents. The aim of this monograph is to investigate constructions, which deviate from what is considered as “normal” or “basic” sentence structure. More precisely, in the structures to be scrutinized, either some elements occur in a left-peripheral position, not in the one that they thematically belong to (“fronting-”-type constructions) or there is some extra constituent in the preverbal field, associated with a left-peripheral entity (“left-dislocations”). The non-canonical structure in each case has information/discourse-structural effects. That is, the prominent constituents receive a specific discourse-function, which would not necessarily be the case if they were not involved in the configurations to be discussed.

The monograph investigates such discourse-related left-peripheral constructions in two typologically distant languages, English and Hungarian. The structures under investigation are as follows. (1)-(3) show the English constructions: (1) illustrates topicalization (TOP), (2) shows a clause-initial adjunct (CIADJ) and (3) demonstrates left dislocation in English (LD-Eng). In examples (4)-(5) we see the Hungarian repertoire: (4) displays versions of Hungarian left dislocation (LD-Hun) and (5) showcases what I label Hungarian operator fronting (OF). In every case, the names of the constructions should simply be viewed as convenient labels, not as theoretical/analytical suggestions:

- (1) *John, I like.*
- (2) *In New York, there is always something to do.*
- (3) *John, I like him.*
- (4) a. *Jánost, azt Kati szereti.*
John.ACC that.ACC Kate likes
'As for John, Kate likes him.'
- b. *Jánost, őt Kati szereti.*
John.ACC him Kate likes
'As for John, Kate likes him.'
- (5) *János(t) mondtam, hogy jön a partira.*
John(ACC) said.1SG that(c) comes the party.to
'(Of) John I said that he will come to the party.'

Throughout the monograph, expressions like “fronting”, “dislocation” and “extraction” will be extensively used. As with the names of the constructions, such usage should be viewed as

descriptive labelling for the configurations/processes, not as theoretical/analytical commitments on my part. These terms are deeply embedded in the linguistic canon and my concern here is not terminological innovation so I stick with tradition in this respect.

The over-riding question of my inquiry may be put simply:

What is the relationship between the left-peripheral expression and the rest of the sentence?

Elaborating on this, the answers that I seek are related to two levels of linguistic analysis: sentence structure and information structure.

The first level is concerned with the thematic/argument structural, formal and functional links between the left-peripheral expression and the host-sentence. As we will see, there are different ways to incorporate the fronted constituent into the sentence and these different mechanisms are detectable in how the fronted element interacts with the composition of the sentence. Part of this are “connectivity-effects”: formally observable markings that appear as if the fronted element were plugged into its canonical position. These connectivity-effects vary in the constructions shown in (1). Such variability will be shown to be demonstrable in various syntactic phenomena including formal features (agreement), binding patterns and interaction with syntactic islands. We will see that the nature of the link between the fronted elements and their intra-sentential correlates may range from a strict syntactic dependency through a looser semantic/anaphoric one to a mere pragmatic inference.

As for information structure, I will seek to establish what kind of communicative purpose is served by fronting the element from its “normal” position. I aim to do so in a systematic and principled framework. The construction of such a framework is essential as despite a lot of work in the area, no consensus has been reached about the proper view of information structure. While it is clear that some sort of a feature-based proposal is favorable, there are substantially different understandings of what the relevant features are, what they really mean, how they interact and what the categories are that these features should ultimately build. Once an adequate picture of information structure is constructed, it will become available for us to understand how these constructions interact with the discourse.

It is clear that the proper analysis of these constructions involves various levels of linguistic theory. As such, the framework of Lexical-Functional Grammar (LFG), with its parallel levels of representation is a suitable choice for the investigation. A related goal of the monograph is to contribute to the LFG-theoretic analysis of Hungarian, as not much work has been done in these respects on the area of long-distance dependencies and on subordinate clauses.

The structure of the monograph is as follows. The remainder of **chapter 1** introduces the relevant aspects of LFG, the formal framework. The various levels of representation will be showcased and of these levels, information structure will get the greatest emphasis. After providing a brief overview of the field I will put forward a new information structural taxonomy. The proposed system is an amalgamation of previous approaches from several theoretical backgrounds. It follows the research trend of analyzing information structural categories in terms of features and it will be shown that the features +/-NEW, +/-D-STRUCTURING and +/- CONTRASTIVE are necessary and sufficient for the proper classification of the information structural categories.

Chapter 2 gives an overview of Hungarian sentence structure. This is necessary for the subsequent analyses of the structures in (4)-(5). I follow the LFG-theoretic approach to

Hungarian of Laczkó (2014a, 2014b, 2015), which in turn is influenced by É. Kiss (1992). As operator fronting involves subordinate-clauses, this chapter also provides a summary of the main issues concerning them. I will confront the influential idea presented in Kenesei (1992/1994) that the pronoun associates of Hungarian subordinate clauses are expletives and I will subscribe to the alternative view in which they are contentful pronouns as argued by Tóth (2000), Rákosi & Laczkó (2005). I will support my position with theoretical, empirical and typological data.

Chapter 3 and 4 present the analyses of the constructions. **Chapter 3** deals with the English constructions. For each of them, I will first survey their empirical properties in a theory-neutral manner and then I will propose an LFG-theoretic analysis. TOP-Eng will be shown to be a prototypical fronting construction, with strong connectivity effects and, as such, it will be analyzed as a strict functional dependency between the fronted element and its canonical position. Some CIADJs in English also follow this pattern but there are CIADJs that lack such clause-internal ties and are best viewed as “base-generated” at the left-periphery. Finally, LD-Eng will be argued to be so disconnected from the host-sentence that they call for an extra-syntactic, “orphan” analysis. Information structurally, TOP will be shown to be a contrast-marker while the other two constructions are related to the category of neutral Topics.

Chapter 4 is about the Hungarian constructions: left dislocation in Hungarian and operator fronting. Following the pattern of the previous chapter, first a survey of empirical properties will be provided and then the details of my analyses will be presented. Two subtypes of LD-Hun will be distinguished: “topic left dislocation” and “free left dislocation”. It will be argued that free left dislocation should be regarded as a phenomenon lying outside syntax proper and in this respect it should be treated on par with left dislocation in English. The other left dislocation in Hungarian is part of core syntax, with a sentence-internal anaphoric link. Operator fronting will also be shown to bifurcate into two sub-categories. The fronted element may be part of a traditional “fronting”/ long-distance dependency construction, in which case the clause-initial element maintains strong ties with its canonical position (so connectivity effects are present). The other possibility is what will be labelled as “prolepsis”: the fronted element becomes an object or an oblique argument of the main verb via an argument-structural process and is only semantically (anaphorically) linked with an embedded grammatical function. A formal analysis for both the English and the Hungarian constructions will be provided with reference to various levels of linguistic representation.

Chapter 5 puts the constructions discussed throughout the monograph into cross-linguistic and theoretical perspectives. It will be demonstrated that many languages possess similar constructions. We will see how the various fronting configurations relate to each other and what the main similarities and differences are. Notably, “topic left dislocation” will be shown to be a mixture of the properties of Germanic left dislocation and clitic left dislocation configurations. Furthermore, particular attention will be paid to the exploration of prolepsis and how it can be put into the theoretical space provided by LFG. In particular, I will argue that there is a natural link between prolepsis and “control”-constructions.

Chapter 6 is a conclusion to the monograph. I briefly restate the questions and topics raised by my research and summarize the answers and claims that I have provided for them. Some avenues for future research will also be put forward.

Overall, the monograph provides an investigation of a number of clause-initial, discourse-related constructions in a considerable depth and while doing so it fully makes use of and in some respects expands the theoretical space provided by Lexical-Functional Grammar.

1.2 Lexical-Functional Grammar

In this section I give a brief overview of the main theoretical framework of this monograph, Lexical-Functional Grammar (LFG), a model of the generative grammar tradition.¹ The first letter in the acronym “LFG” stands for “lexical”. In essence, this means that LFG holds that instead of transformations, it is the (mental) lexicon that is the engine behind most grammatical phenomena. Thus LFG is a non-transformational/non-derivational generative framework.

LFG subscribes to the Lexical Integrity Hypothesis: “morphologically complete words are leaves of the constituent-structure tree and each leaf corresponds to one and only one constituent-structure node” (Falk: 2001:26).

As a result of the Lexical Integrity Principle, “empty” or “zero” categories (traces, *pros*, PROs, silent copies) are generally avoided in constituent-structures in LFG (although some LFG researchers have suggested that under some very restricted circumstances, they may be resorted to, see Falk 2007).

The second letter in the acronym stands for “functional”. This means that instead of assuming that traditional grammatical functions like “subject” or “object” are only configurationally derived concepts that are defined in terms of particular tree-structural positions (e.g. an element is the “subject” if it occupies the Spec/IP position), LFG assumes that these categories are syntactic primitives.

In LFG linguistic expressions are analyzed via several interconnected and parallel structures. Every representational level has its own “vocabulary” and rules and various correspondences mediate the relations between them. A sentence is grammatical if it conforms to the rules of every level of representation. As such, LFG is a member of the class of constraint-based, representational theories, where grammaticality is dependent on conforming to constraints on representations and not on properties of derivations. The main levels of representations are the following:²

- **Constituent-structure:** simple and flexible tree-structures to represent constituency and the linear order of words.
- **Functional-structure:** attribute-value matrices which represent the grammatical and functional relations within sentences.
- **Argument-structure:** the module where various diatheses and lexical semantic information are represented.
- **Information structure:** the component where information-packaging properties of sentences are represented.
- **Prosodic-structure:** the layer where phonological information (stress, intonation) is represented.
- **Semantic-structure:** the structure containing the logical-semantic analysis of sentences. Most often these utilize linear logic and Glue Semantics.

¹ For more comprehensive introductions the reader is referred to Falk (2001), Bresnan et al. (2016), Dalrymple, Lowe & Mycock (2019) and Börjars, Norldinger, & Sadler (2019).

² Besides these, there have been proposals in the literature about a separate structure for representing (some) morpho-syntactic information (inflection, case, agreement), see Butt et al. (1999) and Falk (2006a).

Out of these the first four are the most relevant for our purposes, so a more detailed description of them is provided below. The interested reader is referred to Dalrymple, Lowe & Mycock (2019, Chapter 8 and 9) for further information about semantic and prosodic-structure.

Constituent-, or c-structures are syntactic trees. Their role is to represent constituency and linear word order. The guiding principle for c-structures in LFG is the principle of Economy of Expression (see Bresnan 2001:91, Dalrymple 2019:85, Falk 2001:33). Here, I quote Falk's formulation:

- (1) **Economy of expression:** all syntactic phrase structure nodes are optional and are not used unless required to create a well-formed functional-structure or to add semantic content.

LFG's approach to phrase structure has its roots in the X'-theoretic approach of Chomskyan models (Government-Binding Theory (GB, Chomsky 1981), the Minimalist Program (MP, Chomsky 1995, etc.), with some important differences, which make c-structures in LFG more compact in general than the syntactic trees of Chomskyan approaches. Firstly, the Lexical Integrity Hypothesis and Economy of Expression greatly reduces the number of syntactic nodes: abstract, purely functional, empty nodes are generally avoided. Secondly, there is no strict binary-branching requirement. Lastly, as all nodes are optional. Therefore, if there is no overt element in a position, the corresponding node is simply not projected. This may result in "headless" constructions.

Functional structures (f-structures) are the other side of syntactic descriptions in LFG. They are attribute-value matrices which represent information about grammatical and functional relations within sentences.

LFG standardly recognizes the following grammatical functions: subject (SUBJ), object (OBJ), oblique (OBL), secondary object/thematically restricted object (OBJ_θ), closed complement (COMP), open complement (XCOMP), possessor (POSS), closed adjunct (ADJ), open adjunct (XADJ).

Out of the list above, subjects, objects and possessors are standard and well-known grammatical functions. The OBL function is assigned to elements that are associated with a unique semantic role by their predicates and usually bear idiosyncratic case. These are marked by an adposition (e.g. *John appeals to Mary*) or case affixes. Because of their association with specific semantic roles, obliques are commonly referred to as OBL_θ. Secondary objects are objects that are specified for some semantic role, like dative shifted theme objects in English (*John gave Mary an apple*) or some applicative objects (e.g. in Bantu languages, see Kibort 2008).

COMP and XCOMP are the functions of clausal arguments. Clauses that host their own subject are standardly assumed to have the closed complement function COMP (*I hope that John passes the exam*), while clauses which lack an independent subject are analyzed as having the open complement function XCOMP. For instance, these include infinitival clauses of "raising"-sentences (*John seems to be happy*) and other predicative complements (*Mary didn't sound ashamed of herself*). The open adjunct function XADJ is parallel to this on the realm of adjuncts (e.g. *Mary arrived drunk*).

Some researchers have expressed doubts about the necessity of the COMP and the XCOMP function and it has been proposed that they should be abandoned. From this

perspective, clauses that had been analyzed as bearing these functions are actually OBJs, OBJ_{θs} or OBL_{θs}. However, arguments for the original approach have also been evoked.³ Here I am going to follow the conservative approach, keeping the (X)COMP functions. Nothing crucial hinges on this for the analyses presented in the subsequent chapters.

Originally LFG also posited that “topic” and “focus” are grammaticalized discourse functions and they are represented at f-structure. However, later research has gone in a direction where their representation is entirely relocated to the separate information structure. Some researchers (Alsina 2008, Asudeh 2011) have suggested that dislocated/extracted elements associated with discourse functions should receive some general information structural label at f-structure. In what follows I will adopt Asudeh’s (2011) Unbounded Dependency Function (UDF) for this purpose.

UDF, ADJ and (X)ADJ are somewhat different from the other functions mentioned, as they are not uniquely instantiated. Since a clause may well contain multiple topics, foci and adjuncts, these grammatical functions are usually represented as sets.

All well-formed f-structures conform to the three well-formedness constraints of LFG (there are several formulations of the constraints; here I quote that of Dalrymple 2001):

- **Completeness:** an f-structure is *locally complete* if and only if it contains all the governable grammatical functions that its predicate governs. An f-structure is *complete* if and only if it and all its subsidiary f-structures are locally complete.
- **Coherence:** an f-structure is *locally coherent* if and only if all the governable grammatical functions that it contains are governed by a local predicate. An f-structure is *coherent* if and only if it and all of its subsidiary f-structures are locally coherent.
- **Consistency:** in a given f-structure a particular attribute may have at most one value.

An issue that must receive a mention regarding functional structure is that LFG posits mechanisms that make it possible to state identity relations between distinct f-structural entities. There are two such mechanisms: functional and anaphoric identification.⁴ Functional identification is a strict, formal identity relation which allows one element to satisfy two grammatical functions at the same time. This happens for example in the case of wh-questions. In (2), for example, according to the LFG-analysis, the question word simultaneously serves as the OBJ of the sentence and it is also an Unbounded Dependency Function (UDF). This is graphically represented as a solid line between the two functions in the f-structure.

(2) *Whom did he assassinate?*

³ See Dalrymple & Lødrup (2000), Alsina, Mohanan & Mohanan (2005) and Szűcs (2018) for various perspectives.

⁴ The terms “functional control” and “anaphoric control” are also used. I prefer “identification” so as to avoid any collision with other uses of the term “control” (see section 5.3.2).

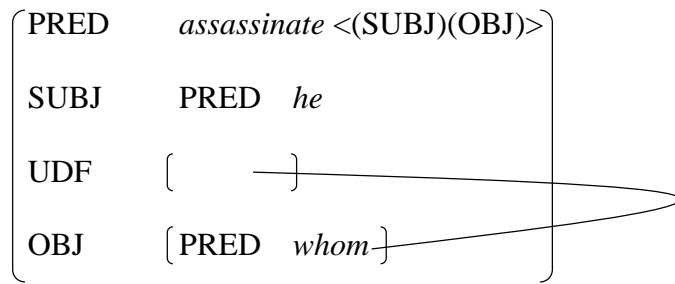


Figure 1.
F-structure for (2)

Functional identification is also utilized in “raising” sentences. In (3), *Gavrilo Princip* is the syntactic subject of the main (*seem*) and the embedded predicate (*assassinate*) at the same time, while the culprit is thematically related only to the embedded predicate.

(3) *Gavrilo Princip seems to have assassinated the emperor.*

Anaphoric identification is about the referential identity of two f-structural elements, without assuming a syntactic identity. It is primarily assumed in the analysis of “equi”-type control sentences, like (4). (A detailed characterization of “raising” and “equi” structures will be provided in section 5.3.2)

(4) *Gavrilo Princip agreed to assassinate the emperor.*

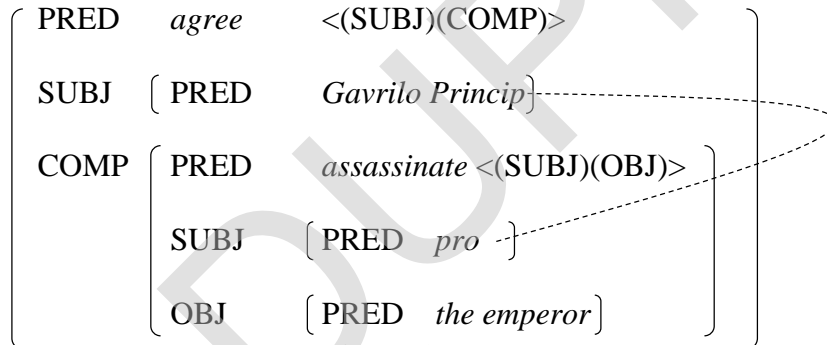


Figure 2.
F-structure for (4)

Unlike (3), here *Gavrilo Princip* is thematically related to both predicates. The unexpressed embedded subject is represented in f-structure as a “*pro*”,⁵ and the main clause subject is constrained to be referentially identical with it. This is informally represented here with the dotted line. Note that the dotted line is only used here for expository purposes; it is not a standard LFG representation. (The standard LFG representation displays INDEX attributes with identical values for the identified elements.)

⁵ Note that unlike the Chomskyan treatment this “*pro*” is only an f-structural entity. In accordance with Economy of Expression, no zero c-structural node is assumed.

Argument structure (a-structure) is the locus of morphosyntactic and morphosemantic operations on lexical entries, so it may be regarded as a core mechanism in the semantics-syntax interface in LFG. Although it was not part of the original theory of LFG as it was put forward in Bresnan ed. (1982), it is now an indispensable component of the architecture. The approach to characterizing the properties and mechanisms of argument-structure is called Lexical Mapping Theory (LMT).

Bresnan and Kanerva (1989) proposed that grammatical functions can be distinguished via the features $+/-o$ (object-like) and $+/-r$ (semantically restricted). The feature-system builds on the intuition that SUBJ and OBJ may be associated with any semantic role (or even with semantically empty idiom chunks as in *The cat seems to be out of the bag* and *I believe the cat to be out of the bag.*), while OBL_{θ} and OBJ_{θ} are always associated with a specific semantic role. On the other hand, OBJ and OBJ_{θ} are both object-functions while SUBJ and OBL_{θ} are not. The following taxonomy is the result:

	$-o$	$+o$
$-r$	SUBJ	OBJ
$+r$	OBL_{θ}	OBJ_{θ}

Table 1.

LMT's feature decomposition of basic grammatical functions⁶

The arguments of the predicates are associated with these features and then mapping principles determine the grammatical functions of these elements. In Kibort's (2007) approach to LMT, which I will adopt, the arguments of a predicate are taken to have a set of entailments, as the participants of the event denoted by the predicate. These are associated with a fixed and universal valency-template, shown in (5).

$$(5) \quad \begin{array}{cccccc} -o/-r & -r & +o & -o & \dots & -o \\ \text{arg1} & \text{arg2} & \text{arg3} & \text{arg4} & & \text{arg}_n \end{array}$$

Once the semantic role entailments of the predicate are associated with the argument positions offered by the template, grammatical function assignment takes place according to the following principle (Kibort 2007):

- (6) a. **Mapping Principle:** The ordered arguments are mapped onto the highest (i.e. least marked) compatible function on the markedness hierarchy.
Markedness: having positive feature-specification.
- b. **Markedness hierarchy:** SUBJ > OBJ, OBL_{θ} > OBJ_{θ}

There are two kinds of operations that may affect the argument-structure of a predicate. When a morphosyntactic operation like passivization takes place, it is modelled as increasing the

⁶ The status of COMP and XCOMP in LMT depends on one's theoretical stance about them, see footnote 3.

markedness (adding positive specification) of the given slot of the valency-template. Let us use (7) as an illustration for which (8) contains the argument-structural representation.

- (7) a. *Jack kissed Monica.*
 b. *Monica was kissed by Jack.*
- (8) a. *kiss* agent patient
 b. -o -r ← universal valency frame
 arg1 arg2
 c. SUBJ OBJ ← active mapping
 d. -o/+r -r ← increasing markedness (adding a +r specification)
 e. OBL_θ SUBJ ← passive mapping

Apart from morphosyntactic argument-structure changes, which only affect the mapping of grammatical functions, but not the semantics of predicates, there also exist morphosemantic operations, which do have effects on the meaning of the predicates. This involves the alternative alignment of the semantic participants on the valency template. One example is the English dative shift. Although the prepositional dative and the double object dative are obviously related lexical entries, there is a meaning difference between them. For instance, it is well-known that the double object dative is incompatible (at least in standard dialects) with an abstract beneficiary, as per (9a) vs. (9b).

- (9) a. *John sent a letter to England.* b. **John sent England a letter.*

According to Kibort (2007) such meaning differences arise because the valency slots are associated with different meaning entailments, e.g. the arg1 slot is for proto-agents, arg2 is for proto-patients and arg3 is for proto-themes. These morphosemantic variations are modelled in this version of LMT as different alignments of the arguments of the predicate on the valency frame. So in the prepositional version the beneficiary is mapped to the 4th argument slot, while double object dative maps it to the 2nd position.⁷

- (10) a. *send* agent patient beneficiary
 b. -o -r -o ← universal valency frame
 arg1 arg2 arg4
 c. SUBJ OBJ OBL_θ ← prepositional dative (9a)
 d. *send* agent beneficiary patient
 e. arg1 arg2 arg3 ← universal valency frame
 -o -r +o
 f. SUBJ OBJ OBJ_θ ← double object dative (9b)

⁷ The 3rd position for the beneficiary is not available in standard English, aside from some marginal examples like *??You can give it me back*. It is productive in other languages, e.g. Polish (Kibort 2008).

Also note that in the double object dative, the patient receives a +o specification, which makes it impossible for it to map to SUBJ. Hence, only the beneficiary can be a passive SUBJ (**A letter was sent Mary* vs. *Mary was sent a letter*).

1.3 Information structure

Information structure (i-structure, IS) constitutes an important aspect of our subsequent investigations, so I will explore this level of representation in more detail.

I-structure is conceptualized as an autonomous level of linguistic representation, which, following Krifka and Musan (2012: 1), may be defined as those aspects of linguistic representation which concern not the information-content of a sentence itself, but rather the way it is presented, “packaged”, and also “helps speakers to take into consideration the addressee’s current information state, and hence to facilitate the flow of communication”. The main purpose of IS is the structuring of the discourse: to mediate and assist the integration of the sentence into the conversation. Thus it will be argued that the nature of the relationship between the elements of an utterance and the wider discourse is a vital aspect of i-structure.

Information structure originally was not part of the machinery of LFG. As noted earlier, information structural notions like TOPIC and FOCUS were used as f-structure components. The addition of i-structure as an independent label was motivated by two factors. On the one hand, it seemed that a number of phenomena pertaining word-order variations may be traced back to different information structural status of the elements involved. This was the primary drive behind Choi’s (1996) analysis of German and Korean scrambling and behind Butt & King’s (1996) treatment of word-order variations in Urdu and Turkish. On the other hand, King (1997) showed that there may be mismatches between f-structure and i-structure. This, given that in LFG f-structures feed the semantic module, could lead to interpretational anomalies. An illustration from King (1997) is the case of *it*-clefts.

- (11) A: *Was it the ex-convict with the red SHIRT that he was warned to look out for?*
 B: *No, it was the ex-convict with the red TIE that he was warned to look out for.*

Since syntactically, the entire phrase *the ex-convict with the red SHIRT* is highlighted (as an extraposed clause), the constituent would have to bear the FOCUS grammaticalized discourse-function at f-structure. But this is cannot be right, as only the piece of clothing is in focus. Without i-structure, the mismatch cannot be resolved. As a solution, a new level of representation was posited, where bare PRED values could be specified as having discourse functions.

Since then, several frameworks have been put forward about the representation of IS in LFG. A common thread running through them is the decomposition of information structural categories with features (see Choi 1996, Butt & King 1996, Cook & Payne 2006, Gazdik 2011). The advantage of such approaches is that it is possible to capture the commonalities among these notions. However, three issues obscure the view. First, it is not clear how many IS-categories should be distinguished. Second, it is not clear what features should be used. Finally, it is often inadequately defined what particular features really mean. In what follows I present my views on these topics and outline my proposal about the organization of information structure.

In this section I aim to construct an IS-taxonomy that builds on previous approaches and also improves upon them. The improvement is manifested in two aspects: first, my proposal will include all and only the well-established IS-categories, and second, I will attempt to provide a clearer definition of the features than the existing frameworks do. The advocated approach is shown in Table 2.

		+NEW	–NEW
+D-STRUCTURING	+CONTRASTIVE	contrastive focus	contrastive topic
	(–CONTRASTIVE)	information focus	topic
–D-STRUCTURING		completive information	background information

Table 2.
The proposed IS-taxonomy

In the following sections, I will investigate the content of these features and the properties of the IS-categories.

1.3.1 Features of information structure

In the proposed system, six discourse functions are distinguished with the features NEW, D-STRUCTURING and CONTRASTIVE.

I would like to note that although present in several earlier taxonomies (Choi 1996, Butt & King 1996, Gazdik 2011), “+/- PROMINENT” as a feature is not used here. At the heart of the issue is that defining prominence is very problematic as it is an inherently subjective notion that can only be stated in relation of two entities. What is more, prominence is a gradient property so using it as a +/- attribute seems off the track. Next it should be pointed out that i-structure is a level of linguistic representation and prominence as such is not something that seems like an exclusively linguistic property: sounds, visuals and even ideas may also be prominent in some sense. Even in the realm of linguistics prominence is ambiguous, as it has phonological, syntactic and pragmatic aspects and these aspects do not necessarily behave in a parallel fashion. Ultimately, I argue that the distinctions that the PROM feature makes are either unnecessary or are storable in terms of the features that I use, so PROM itself is superfluous. It is best understood as a derivative property might piggyback on the more fundamental features that I utilize. As we will see, the concepts of newness, discourse-structuring and contrastiveness can be defined more satisfactorily and are able to provide us with the necessary distinctions.

Let us start with the feature +/-NEW. As Gundel (1999) points out much terminological confusion stems from the failure to distinguish between two senses of newness: the referential and the relational sense.⁸ The referential sense is about “a relation between a linguistic expression and a corresponding non-linguistic entity in the speaker/hearer’s mind” (Gundel 1999). The cognitive status of some entity may be characterized in various ways (completely novel, identifiable, familiar, etc.) and these statuses may have consequences for the linguistic encoding of the entities but the cognitive statuses themselves are not inherently linguistic.⁹

⁸ A similar point is made by Lambrecht (1994: 48).

⁹ For approaches about the effects of referential givenness see e.g. Gundel, Hedberg & Zacharski (1993), Prince (1981), Ariel (1988) and Lambrecht (1994: 165-172).

Relational newness on the other hand makes a directly linguistic distinction. It “involves two complementary parts, X and Y, of a linguistic or conceptual representation, where X is given in relation to Y, and Y is new in relation to X” (Gundel 1999). Consider the following example, from Lambrecht (1994: 48).

- (12) A: *When did you move to Switzerland?*
 B: *When I was sixteen.*

As Lambrecht (1994: 48) notes, “what constitutes the information conveyed by this answer is not the fact that at some point in his life the speaker was sixteen (...) but the RELATION [emphasis by Lambrecht] established between an act of moving to Switzerland, the person involved in that act, and the time at which the moving occurred”. In other words, the answer provides a value for X in the proposition evoked by the question: “I was X (years old) when I moved to Switzerland”.

What Gundel (1999) refers to as X and Y is termed by Lambrecht (1994: 52) as the “pragmatic presupposition” (X) and “pragmatic assertion” (Y) of the sentence. The pragmatic presupposition of a sentence is “the set of propositions lexicographically evoked in a sentence which the speaker assumes the hearer already knows or is ready to take for granted at the time the sentence is uttered”. The pragmatic assertion on the other hand is “the proposition expressed by a sentence which the hearer is expected to know or take for granted as a result or hearing the sentence uttered”.

What is part of the presupposition and what is part of the assertion may be told apart by using the “lie-test” (Lambrecht 1994: 52). Consider Lambrecht’s example in (13).

- (13) *I’ve finally met the woman who moved downstairs.*

Imagine someone challenges the claim in (13) by saying “That’s not true!”. This would be understood as rejecting the claim that the meeting took place, not that someone did move downstairs or that the speaker has a neighbor.

Let us now illustrate that the referential and the relational senses are independent with two examples. Consider the following piece of conversation, from Gundel (1999).

- (14) A: *Who called?*
 B: *Pat said SHE called.*

Here *SHE* in the answer has a high level of referential givenness: it refers to a proper name, a specific individual that had just been mentioned and is in the center of attention. However, it still provides a new piece of information, by providing a value for the presupposed open proposition, “X called”.

In turn, some referentially new material may still be part of the pragmatic presupposition. Consider the following exchange.

- (15) A: *What’s that big red spot on your arm?*
 B: *A mosquito bit me.*
 A: *That cannot be true! Mosquito-bites aren’t like that!*

In B's response, the noun phrase *a mosquito* refers to a nonspecific entity, newly introduced to the discourse. However, as A's reply shows, applying the lie-test may leave the existence of the mosquitoes intact, which means that it is part of the pragmatic presupposition.

Out of these two notions of newness it is the second, relational sense that I take to be relevant for information structure. A working-definition is provided in (16).

- (16) +/- NEW: A linguistic entity with a +NEW feature at information structure provides relationally new information by being part of the pragmatic assertion of the sentence. A linguistic entity with a -NEW feature at information structure is relationally given and is part of the pragmatic presupposition of the sentence.

Let us now turn to the feature +/-DISCOURSE-STRUCTURING. Gazdik (2011), building on the theories of Büring (2003) and Asher & Lascarides (2003), distinguishes between those information structural categories which are "D-LINKED"¹⁰ and those that are not: the ones of the first type "link the sentence to the discourse by introducing a subtopic of the discourse topic or reshaping the discourse topic". I find the idea attractive but Gazdik's implementation unsatisfactory so I make a proposal in the same spirit, but with some important modifications.

I relabel the feature as +/-DISCOURSE-STRUCTURING (abbreviated as D-STR) for two reasons. One, I believe that "discourse-structuring" describes what I want to express more precisely than "discourse-linking". Two, I would like to avoid causing terminological confusion often found in the literature, where the same word is used with a number of related but still different definitions (as in the situation of "old information" and "new information", or the term "topic"; actually, Gazdik's (2011) use of the expression is already not exactly what Pesetsky 1987 means by it, see footnote 10).

To understand this feature let us first review the theory of Büring (2003), on which Gazdik (2011) builds. The starting point is that all discourse is structured along a hierarchical web of questions and subquestions (an earlier formulation of the idea is to be found in Roberts 1996). The top of the hierarchy may be conceptualized as the (poetic-sounding) Big Question: "What is the way things are?". Questions that are of more immediate concern are Questions Under Discussion (QUD). The structure of the discourse may be represented with a discourse-tree. Consider the conversation in (17) and its representation in Figure 3 (slightly modified from Büring 2003: 4).

As we will see in connection with the contrastive IS-categories, there are also cases where a subquestion is only implicitly present in the discourse-tree.

- (17) A: *How was the concert? Was the sound good?*
 B: *No, it was awful.*
 A: *How was the audience?*
 B: *They were enthusiastic.*
 A: *How was the band? How was the drummer?*
 B: *Just fantastic.*
 A: *And what about the singer?*

¹⁰ The term originates from Pesetsky (1987), for whom some question words are D-LINKED and others are not. For example, according to Pesetsky (1987), the wh-phrase in *Which book did you read?* is D-LINKED as it may be only felicitously answered if the speaker and the hearer already have a representation of a set of books in their minds.

B: *Better than ever.*
 A: *Did they play old songs?*
 B: *Not a single one.*
 A: *So what did you do after the concert?*

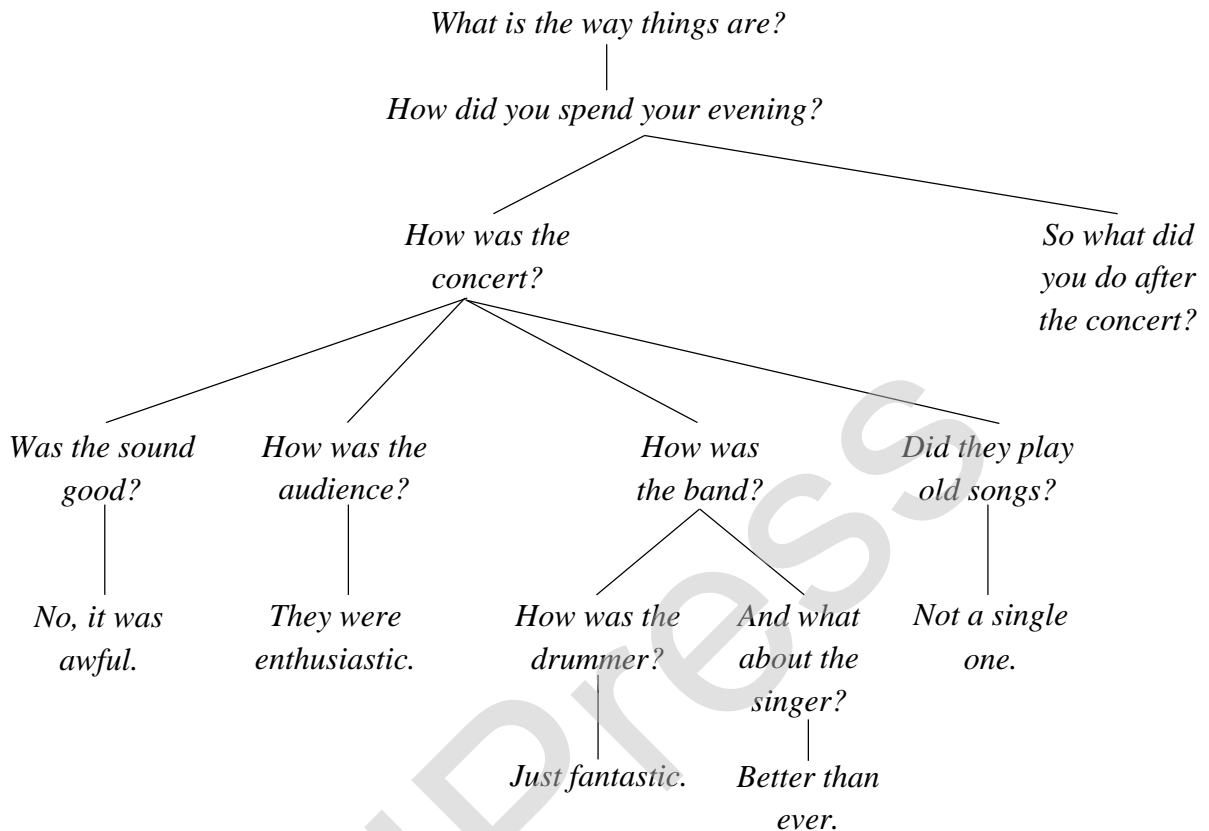


Figure 3.
 Discourse-tree for (17)

Gazdik (2011) also draws from Asher & Lascarides (2003) who analyze the semantics-pragmatics interface via rhetorical/discourse relations: narration, background, result, continuation, parallel, contrast, alternation, dialogue, correction elaboration, topic, explanation, consequence, and various question types. Gazdik (2011) then correlates these relations with the i-structural categories, as in Table 3.

I have already noted that I sympathize with the spirit of this approach, but I find several problematic aspects in the way it is fleshed out. My primary concern is with the featural composition of IS categories. Remember that according to Gazdik (2011: 284), +D-LINKED means that the element “links the sentence to the discourse by introducing a subtopic of the discourse topic or reshaping the discourse topic”. Given this, it is not clear to me why a focus is not D-LINKED. Giving an answer for a question, offering a correction, or providing contrast does seem to me to be like things that reshape the discourse topic. Consider (18).

(18) A: *What concert did you attend yesterday?*
 B: *Pink Floyd.*
 A: *Oh, I like them too. What is your favorite album?*

Information structure role	Information structure features	Discourse relation(s)
Thematic shifter	+PROM, +D-LINKED	narration, topic
Contrastive topic	+PROM, +D-LINKED	implicative answer
Focus	+PROM, –D-LINKED	answers, parallel, contrast, correction
Hocus	+PROM, –D-LINKED	narration
Question phrases	+PROM, +/-D-LINKED	question-answer pairs
Completive information	–PROM, –D-LINKED	any
Background information	–PROM, D-LINKED	any

Table 3.

Gazdik's (2011) view of i-structure roles and discourse relations

Here B's answer is a sentence including only a focus. The rest of the conversation is about the band Pink Floyd, so the answer has introduced a new subtopic to the conversation and has reshaped the discourse topic. In fact, focussed elements prototypically function as such.

On the other hand, background information is claimed to be +D-LINKED, which I find highly counterintuitive. Why would pieces of information, which are already known, repeated and clearly out of the center of attention participate in restructuring the discourse? Their D-LINKEDNESS only related to the fact that they are discourse-anaphoric (that is: referentially old, see the previous section), but this is not what Gazdik (2011) seems to argue D-LINKEDNESS to be. If discourse-anaphoricity were to be equated with DISCOURSE-LINKEDNESS, then a corrective focus like the one in (19/B) would also be D-LINKED, contrary to Table 3.

- (19) A: *I heard that you like [Deep Purple]_{FOCUS} the most.*
 B: *I like Deep Purple, but it is not [Deep Purple]_{FOCUS} that I like the most. It's Pink Floyd.*

However, notice that background information and completive information share a property: they may be present in any discourse relation. This is because they do not participate in the structuring of the conversation in the sense that they are not directly related to Questions Under Discussion of the conversation, they are not involved in resolving these. Topics and foci directly structure the discourse: the former by being the entity that a new piece of information is predicated of, the latter by providing the new pieces of information that are needed for the resolution for a QUD. Both aspects may be stated in terms of the subquestions of the discourse, as we will see in section 1.3.2. This I claim is the core of the concept of being a +D(ISCOURSE)-STRUCTURING IS category.

- (20) +/-DISCOURSE-STRUCTURING: a linguistic entity with a +DISCOURSE-STRUCTURING feature at information structure participates in the structuring of the conversation by being directly involved in the formation or resolution of questions under discussion in the discourse.

A linguistic entity with a –DISCOURSE-STRUCTURING feature at information structure is not directly involved in the formation or resolution of questions under discussion in the discourse.

Finally, let us take a look at +/–CONTRASTIVE. There have been debates about the status of “contrast” in the linguistic literature. The central issue is whether contrast should be regarded as a concept with direct and independent linguistic relevance.

On one end of the scale is Lambrecht (1994: 290) who asserts that contrastiveness is a mere pragmatic effect (a conversational implicature) arising “from particular inferences which we draw on the basis of given conversational contexts” and it is not a category of grammar at all.

Another view (Roth 1985, Büring 1997, Krifka 2008) is that contrast may be regarded as a concept with direct linguistic relevance, but it is not an independent notion: it is a consequence of what the i-structural category Focus does. That is, every focus is contrastive in some sense (more on this below).

The third view, which is the closest to mine, is the view of Molnár (2002, 2006) and Titov (2013). Once the necessary distinctions are made, contrast may be regarded as an independent notion of information structure (and thus, of grammar in a looser sense).¹¹ Evidence for this view is the existence of linguistic structures which are associated with contrast, independent of focus, some of which will be discussed in this monograph. For example, A-scrambling in Dutch is possible with elements that bear contrastive IS-categories, contrastive topics and contrastive foci, see (21), from Neeleman et al. (2009).

- (21) a. *Ik geloof dat [alleen DIT boek]_{CONTRASTIVE FOCUS} Jan Marie gegeven heft.*
 I believe that only this book John Mary given has
 ‘I believe that John has given ONLY THIS BOOK to Mary.’ (and not another book)
- b. *Ik geloof dat [zo’n boek]_{CONTRASTIVE TOPIC} alleen JAN Marie gegeven heeft.*
 I believe that such-a book only John Mary given has
 ‘I believe that such a book, only John has given to Mary.’ (others may have given other books)

So what does it mean to be +CONTRASTIVE? It is clear that contrast is related to a set of alternatives. However, the set of alternatives may be defined on two levels. As Titov (2013) points out, referring to a set of alternatives is taken to form the basis of semantic computation for any kind of focus (as e.g. in “alternative semantics”, descending from Roth 1985). Consider example (22) and a semantic interpretation of it in (23), from Kenesei (2006: 144-145).

- (22) A: *Where did you go in the summer?*
 B: *I was [in Italy]_{FOCUS}.*
- (23) a. For which x, x a country, you went to x in the summer?
 b. $\exists x$, such that $x \in \{\text{countries I went to in the summer}\}$, and x is Italy

¹¹ However, given LFG’s modular architecture, contrastiveness is not a syntactic feature, as it is in some cartographic approaches (e.g. Molnár & Winkler 2010). See Neeleman et al. (2009) for a detailed critique for the cartographic approach.

As is clear from (22)-(23), answering the question does involve the semantic computation with reference to a set (a set of countries in this case). However, this should be distinguished from what Titov (2013) calls the “pragmatic set of alternatives”, and what I also believe to be the relevant concept for *i*-structure. These are entities that are active in the minds of the hearer and the speaker once a contrastive element is uttered. That is, contrastive elements “indicate either through a link to the context or within the utterance itself that the set to which the focused constituent belongs indeed contains alternative members that are relevant for the discourse at hand” (Titov 2013: 422). So *in Italy* in (22) is not contrastive because there it is not indicated that other countries are specifically relevant for the question under discussion. In (24), they are, so *in Italy* in (24) is contrastive.

- (24) A: *Did you go to Spain in the summer?*
 B: *No, I was IN ITALY. (not in Spain)*

It is sometimes claimed (e.g. É. Kiss 1998) that the open or closed nature of the referent set is a crucial factor. However, Repp (2009) calls attention to Krifka’s (2008) observation that the answer in (25a) feels no more contrastive than the one in (25b), despite the fact that in (25a) an explicitly restricted set of alternatives is provided.

- (25) a. A: *What do you want to drink, tea or coffee?*
 B: *Tea.*
 b. A: *What do you want to drink?*
 B: *Tea.*

So when exactly these contrastive interpretations are licensed is highly dependent on the context and the communicative purpose of the interlocutors. That is, it is not enough that potential alternatives are present in the discourse, they have to be salient from the perspective of the interlocutors. Certain contexts may of course facilitate or hinder the obtaining of the contrastive interpretation but it is never going to be obligatory. For researching the conditions, a framework like Asher & Lascarides (2003) could prove to be a fruitful research avenue.

For present purposes, the relevant part of the discussion is that contextually salient alternatives become active in the consciousness of the interlocutors. This may also be represented via Büring (2003)-style discourse-trees. I propose that the discourse-structuring feature is –CONTRASTIVE by default, but it may receive a positive specification from an additional +CONTRASTIVE feature, superimposed on the base concept: it manages the questions under discussion, by activating alternative nodes in the discourse-tree. Note that the +/- CONTRASTIVE is strictly subordinated to the DISCOURSE-STRUCTURING feature so “contrast” is not applicable in the realm of –DISCOURSE-STRUCTURING information structural categories (these are –CONTRASTIVE by definition).¹²

¹² This is comparable to Molnár’s & Winkler’s (2010) C(ohere)nce-feature, which if present (positively specified, in other words) may be further specified as C-continuity and C-Contrast. However, there are two important differences. One, the C-feature is just described as being “responsible for the formally marked type of discourse linking, i.e. for cohesion in discourse” (Molnár & Winkler 2010: 1396), while I strive for a characterization in terms of discourse-trees, which I believe to be a) more precise b) able to capture the distinction between Topics and Background information (I do not see how Molnár & Winkler would distinguish these categories). Two, for Molnár & Winkler such discourse-features are parts of syntax, while my approach is more

- (26) +/-CONTRASTIVE: a linguistic entity with a +CONTRASTIVE feature at information structure participates in discourse-structuring by evoking a contextually salient pragmatic set of alternatives.
 A linguistic entity with a –CONTRASTIVE feature at information structure does not evoke a contextually salient pragmatic set of alternatives.

Now that the features of my proposed information structural taxonomy have been adequately described, I can now proceed to the IS-categories themselves.

1.3.2 Categories of information structure

In the following sections, I survey the i-structural categories that I posit. I repeat Table 7 here for convenience.

		+NEW	–NEW
+D-STRUCTURING	+CONTRASTIVE	contrastive focus	contrastive topic
	–CONTRASTIVE	information focus	topic
–D-STRUCTURING		completive information	background information

Table 2.

The proposed IS-taxonomy

I would like to highlight at this point that the taxonomy is aimed to characterize the information structural properties of these notions. As I will point out, there are some properties of these categories that may only be adequately described with reference to other representational levels (semantics is a prime candidate for this). So, while I will describe several of these aspects in the following sections, I do not aspire to derive each of them from the information structural taxonomy itself. Care will be taken to delimit the scope of the system.

Topic, contrastive topic (CT) and background information (BI) are all –NEW, so they are part of the pragmatic presupposition of the sentence. The first two participate in discourse-structuring, while Background information does not. Let us investigate some properties of these notions.

The common wisdom about the topics is that “the topic of a sentence is the thing the proposition expressed by the sentence is about” (Lambrecht 1994: 118).¹³ However, it is challenging to give this “aboutness”-relationship substantial content, other than by intuition. Consider (27).

in line with Horváth (2010) and Neeleman & van de Koot (2008), who argue that such notions are best captured as interface phenomena.

¹³ See also Reinhart (1981). Note that this definition does not include “discourse topics”, as its scope is limited to the sentence-level.

(27) *Gavrilo Princip assassinated the emperor.*

Under default assumptions, the topic of the sentence is *Gavrilo Princip*. However, one can argue that the sentence is not just about *Gavrilo Princip*, but also about *the emperor*, or about history in general. There are some heuristics that have been proposed as empirical tests for this kind of aboutness-relationship. Prince (1999), citing Gundel (1974) and Reinhart (1981), lists 3 such tests.

- The “as for X”-test: Can the sentence be plausibly paraphrased with an initial “as for X”-phrase, where X is the supposed topic expression?
- The “what about X”-test: Can the sentence plausibly answer a “what about X”-question, where X is the supposed topic expression?
- The “say about X that...”-test: Could the sentence be plausibly reported about by using an initial “Y said about X that...”-phrase, where X is the supposed topic expression?

If we use these tests on (27), we may get some evidence that the sentence is about *Gavrilo Princip*. Nevertheless, this is weak evidence as it would be difficult to argue that (28a’-c’) are absolutely infelicitous paraphrases for (27).

- (28) a. *As for Gavrilo Princip, he assassinated the emperor.*
a’. *#As for the emperor, Gavrilo Princip assassinated him.*
b. A: *What about Gavrilo Princip?*
B: *Gavrilo Princip assassinated the emperor.*
b’. A: *#What about the emperor?*
B: *Gavrilo Princip assassinated the emperor.*
c. *He said about Gavrilo Princip that Gavrilo Princip assassinated the emperor.*
c’. *#He said about the emperor that Gavrilo Princip assassinated the emperor.*

Another reflex of this aboutness-relationship is the often mentioned semantic requirement that topics be referential, as only referential entities can be targets for predication. For instance, non-referential entities, e.g. quantifiers, cannot be in the topic-position in Hungarian (see more about the structure of Hungarian in chapter 2), as in (29) (from É. Kiss 2002: 10).

- (29) **[Kevés várat]_{TOPIC} meg-védtek a zsoldosok a törökök ellen.*
few forts.ACC PV-defended.3PL the mercenaries the Turks against
‘Few forts were defended against the Turks by the mercenaries.’

A complicating factor is that in the case of CTs, the referentiality constraint seems to be relaxed. Certain nonreferential elements may also be contrastive topics, as in (30), where the CT-elements would be ungrammatical as neutral topics (30b-c are from Gécseg 2001).

- (30) a. *[Kevés várat]_{CT} a zsoldosok védtek meg, sokat pedig a hősök.*
 few castles the mercenaries defended.3PL PV many conversely the heroes
 ‘Few forts were defended by the mercenaries; many were defended by the heroes.’
- b. *[Úszni]_{CT} nem tudok.*
 swim not can
 ‘Swim, I cannot.’
- c. *Szépnek_{CT} szép a kutyád, habár nem okos.*
 pretty.DAT pretty the dog.POSS.2SG though not clever
 ‘Pretty, your dog (in fact) is, it is not clever though.’

At any rate, the concept of aboutness clearly sets background information apart from topics, as the former is definitely not what the sentence is about. It only provides the contextual frame of the utterance without being involved in the informational dynamics of the conversation.

Apart from these semantic requirements, some pragmatic constraints have also been proposed for topics. These are related to the referential newness/givenness dichotomy, as described earlier.

It has been suggested that the entity that is denoted by the topic expression should be accessible in the discourse universe. Gundel (1985) calls this the “familiarity condition” on topics. Consider (31), from Lambrecht (1994: 159), which could be the beginning of a telephone conversation, where someone had dialled the wrong number.

- (31) A: *Is Alice there?*
 B1: *#Alice isn't here.*
 B2: *There is no Alice here.*

Even though *Alice* is a referential, definite expression, and is clearly discourse-old by the time of the answer, reply B1 in (31) is undoubtedly strange. The problem in B1 is that *Alice* is not properly established in the universe of the discourse, since the one who replies doesn't know which *Alice* the questioner could refer to. The way to circumvent this problem is to remove *Alice* from the position where she is interpreted as a topic, as in B2.

Another example for such issues is that the referential newness/givenness status of the topic entity may influence the choice of the form of the topic expression. Newly introduced topics (“thematic shifters”, “shifting topics”) are often distinguished from “continuing topics” (Frascarelli 2007, Gazdik 2011). Consider the Hungarian example in (32), which is based on Gazdik (2011: 168)

- (32) *Tamás szeret olvasni. (#Ő) intelligens, szorgalmas és sokra fogja vinni.*
 Thomas likes read.INF he intelligent hard-working and much will.3SG
 reach.INF
 ‘John likes reading. He is intelligent, hard-working and he will achieve a lot.’

In (32) the pronoun is pragmatically anomalous. This shows that pronominal continuing topics are preferably dropped in Hungarian.

The semantic and discourse-pragmatic aspects of topics are certainly important in a full characterization of IS-notions but my primary concern here is information structure, so I will now turn to those properties of Topic, Contrastive topic and Background information that are relevant from the perspective of the features that I have introduced in section 2.6.1.

Topic as an entity of IS has the feature composition –NEW, +D-STRUCTURING. In the case of contrastive topics, the +CONTRASTIVE feature is added. Background information is –D-STRUCTURING (and thus –CONTRASTIVE by definition).

Being –NEW, a topic is part of the pragmatic presupposition. As such, it is outside the scope of sentence negation. Lambrecht (1994: 52) illustrates this with the following example:

- (33) a. *John is my friend.*
b. *My friend is John.*

(33a) is assumed to be about *John*, so *John* is the topic. Imagine someone challenges the claim in (33a) by saying “That’s not true!” This would be understood as claiming “John is NOT your friend,” but the existence of *John* would still be taken for granted. Since it is presupposed, it is outside of the scope of sentential negation. In fact, the denial could be felicitously complemented with the presupposition-cancelling utterance “you don’t have any friends,” which indicates that only the existence of the topic (*John*) is presupposed, while the content of the comment is not. Conversely, uttering “That’s not true!” in response to (33b) where the topic is *my friend*, would still presuppose that I have a friend (just not *John*).¹⁴ This is a shared property of topics and background information.

As for the DISCOURSE-STRUCTURING feature, from my perspective, its conceptualization should boil down to the capacity of topics for structuring of the conversation by being directly involved in the formation and resolution of questions under discussion of the discourse. Consider the conversation in (34).

- (34) A: *What car did Jack buy?*
B: [*Jack*]_{TOPIC} [*bought*]_{BACKGROUND INFORMATION} [*a Volvo*]_{FOCUS}.

I will consider the discourse-structuring of the focus (and other +NEW entities) in the next section. It seems straightforward that background information does not participate in the structuring of the conversation. It is referentially old and is part of the presupposition but it is completely unaffected by the informational dynamics of the conversation: it does not provide new questions or answers for the discourse, nor does it direct the attentions of the participants.¹⁵

It is less straightforward exactly how topics are discourse-structuring. It is intuitively clear and commonly accepted that topics serve as “links” to the previous discourse, i.e. they are related to the discourse topic. In Gazdik’s (2011) framework their role introduce new subtopics.

¹⁴ Note that the test even works if we replace *John* with a definite expression like *the king of France*. Of course one could say something like “That’s not true, because the king of France doesn’t even exist!”, but that is an explicit modification of the presupposition.

¹⁵ According to Mycock (2013:426), it does not “establish new pragmatic relations”.

It is also intuitively clear that they play a central role in the information structure of the sentence by virtue of being the targets of predication in the sentence.

How could this intuition be adapted to a discourse-tree-based conceptualization? One should realize that (34) must be part of a wider discourse. Somehow the interlocutors must have reached this point of the conversation from the Big Question (“What is the way things are?”). For a sentence like (34/B) to be uttered, speaker B has to assume that A wants to know about Jack. That is to say that, at least implicitly, the discourse tree for B must contain a node with an instruction “Tell A about Jack!”. As such, *Jack* in the answer directly corresponds to an instruction in the discourse-tree, as illustrated in Figure 3. If this is on the right track, it can be justified how topics (but not Background information) do participate in the formation and the resolution of the questions under discussion in the conversation.

What is the way things are?

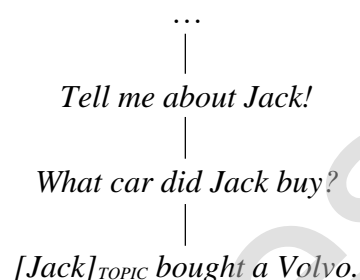


Figure 4.

Topic in a discourse tree

This may well be connected to at least two dimensions of topics, as mentioned earlier. One is the straightforward conceptual connection to the aboutness-property put forward Reinhart (1981) and Lambrecht (1994). The other one is the +PROMINENT feature specifications of topics in some earlier taxonomies of IS. Under the assumption that topics are implicitly but directly introduced by a process illustrated in Figure 4, it is easy to see why they are felt to be prominent.

In the case of contrastive topics, a contextually salient pragmatic set of alternatives is evoked. This could be represented, following Büring (2003), by positing that the original question is split into subquestions. This extends the discourse-tree horizontally and clearly is a discourse-structuring move.¹⁶

¹⁶ Because of this, Titov (2013) argues that CTs are –PRESUPPOSED, as they introduce new material. In her framework they are eventually equated with contrastive foci, their distinction only being the result of the different configurations they occur in. I do not agree with this position, as it may be true that CTs introduce referentially new material, but it is not necessarily the case. Consider (i):

- i) A: *Did Jack buy a Volvo or an Audi, or both?*
 B: *A [Volvo]_{CT} he did buy, I don't know about an Audi.*

Also, even though referentially new material may be introduced via CT-formation, the CT (simply because it is a topic itself) is still part of the presupposition, thus they are –NEW (or +PRESUPPOSED, in reversed terminology).

Furthermore, equating CTs and CFs would make it hard to understand why certain languages have a specific morphological marker for neutral topics and contrastive topics (*wa* in the Japanese example in (ii) below), with the exclusion of CFs (Neeleman et al. 2009).

focus” and “information focus”.¹⁸ An example is (36), note that the translations are from É. Kiss (1998).

- (36) A: *Hol jártál a nyáron?*
 where went.2SG the summer.on
 ‘Where did you go in the summer?’
- B1: [*Olaszországban*]_{IDENTIFICATIONAL FOCUS} *jártam.*
 Italy.in went.1SG
 ‘It was Italy where I went.’
- B2: *Jártam [Olaszországban]*_{INFORMATION FOCUS}
 ‘I went to Italy (among other places).’

According to É. Kiss (1998), while the identificational focus in (36/B1) exhaustively¹⁹ identifies the places where the speaker had gone, the Information focus in (36/B2) merely mentions one such place.

While the recognition of this distinction is very influential, there are certain aspects of É. Kiss’s (1998) approach that have come under criticism. Kenesei (2006) criticizes É. Kiss (1998) on the grounds that both types of foci are “identificational” in the sense that they identify a member from a set of countries that B may have visited. The difference, according to Kenesei, is that while information focus operates on a subset relation, identificational focus involves a *proper* subset relation, that is at least one more member of the set is involved in the semantic computation. Thus Kenesei (2006) suggests that the label “contrastive focus” should replace identificational focus.

Returning to the answers in (36), one must observe that while in (36/B1), *Olaszországban* ‘in Italy’ strictly corresponds to the *wh*-phrase in (36/A) (so it is +D-STRUCTURING), an answer like the one in (36/B2) may also answer a question like “What did you do in the summer?”, where the answer is the entire VP, not the preverbal constituent (Gazdik 2011: 209). *Olaszországban* is still NEW information and part of the pragmatic assertion. However, it is not directly linked to the question under discussion. In such a scenario, it should be regarded as –DISCOURSE-STRUCTURING, thus having the status of completive information. Another example for this IS-notion is (37), from Butt & King (1996). Again, the focus is the entire sentence (which identifies where Nadya is coming from), *toffee* merely provides circumstantial information.

- (37) A: *Where is Nadya coming from?*
 B: *She was just buying [toffee]_{CI} at the market.*

As for the label “information focus”, I would like to reserve it for cases of new information that directly answer questions under discussion, so in the context of (36) both instances of

¹⁸ At this point, I would like to emphasize that the way that I use the term “information focus” does not correspond to É. Kiss’s terminology.

¹⁹ Also, it must be pointed out that the “exhaustive” nature of identificational focus is not an inherent property of the IS-category itself. Rather, it seems to be a property of certain constructions, in which foci is manifested in various languages, the Hungarian preverbal “focus-position” being the prime candidate for it. For É. Kiss (1998) this focus is +EXHAUSTIVE and +CONTRASTIVE, there is an ongoing debate about this issue, see e.g. Onea (2009), Wedgwood (2009), Geröcs, Babarczy & Surányi (2014), Pintér (2016).

Olaszországban count as IF. (That is, “identificational focus” in Kenesei’s (2006) sense.²⁰). Let me repeat example (34) and its discourse-tree from the previous section, in the light of this discussion, but now let us zoom in on the analysis of the focal part of the sentence.

- (38) A: *What car did Jack buy?*
 B: *Jack bought [a Volvo]_{IF}.*

(*What is the way things are?*)

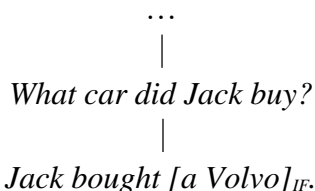


Figure 6.

Information focus in a discourse tree

Additionally, the focus may also be +CONTRASTIVE, in the sense defined in 1.3.1. Similarly to contrastive topics, I assume that in this case new subquestions are added to the discourse-tree. The difference is that in the case of CF, at least one of the new subquestions is answered in an opposite manner compared to the original question²¹ – see (39) and its discourse-tree in Figure 7.

- (39) A: *What car did Jack buy? An Audi?*
 B: *No, [A Volvo]_{CF} [Jack]_{TOPIC} [bought]_{BACKGROUND INFORMATION} (and not an Audi).*

(*What is the way things are?*)

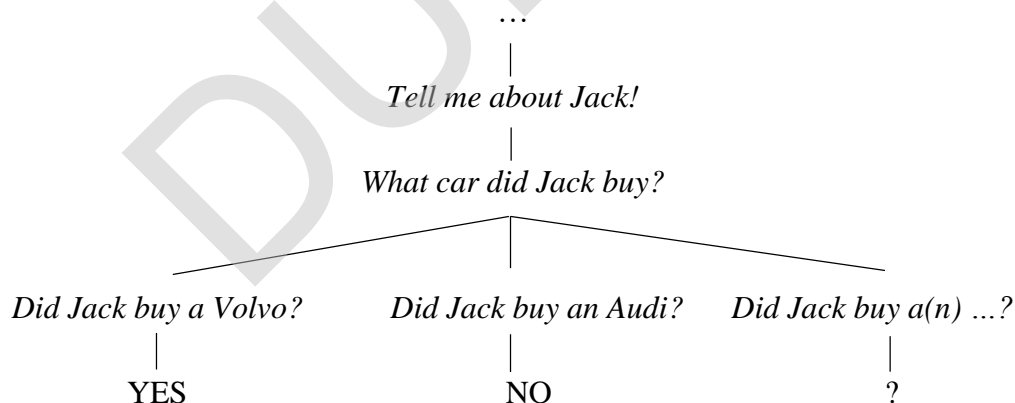


Figure 7.

Contrastive focus in a discourse-tree

²⁰ Titov (2013) uses the term “new information focus” (NIF), with the same justification.

²¹ How this opposing truth value is added to the semantic computation is an important question for further research. In Titov’s (2013) system it is a consequence of an additional contrastive verum focus.

1.4 Conclusion to chapter 1

In this chapter, I first outlined the scope of the research to be presented. We will investigate the syntactic and information structural properties of the following constructions: Topicalization in English (TOP), Clause-initial adjuncts in English (CIADJ), left dislocation in English (LD-Eng), left dislocation in Hungarian (LD-Hun) and Hungarian Operator fronting (OF). It is going to be shown that these are integrated into the sentence structure in various ways and to different degrees. Typological considerations will also be provided.

Next, I gave a general outline of the theoretical framework of the monograph, Lexical-Functional Grammar. LFG is a lexicalist, constraint-based theory with a parallel architecture. An overview of the main levels of representational levels was provided (constituent-, functional-, argument- and information structure). The most attention was given to information structure as the currently available frameworks are argued to be deficient in various respects. I have proposed a new taxonomy for information structure, which builds on the insights of previous approaches but is more advanced in comparison to them because it includes all and only the IS-categories that are well established and it utilizes the i-structural features in a principled and theoretically sound way.

CHAPTER 2

LFG PERSPECTIVES ON HUNGARIAN SYNTAX

In this section, I lay the theoretical and empirical groundwork for the analysis of the Hungarian fronting constructions to be discussed in Chapter 4. First, I will outline the basic structure of Hungarian simple sentences and then I will offer some insights into subordinate clauses as well. Following the general framework of this monograph, I will take the perspective of Lexical-Functional Grammar (LFG).

2.1 The simple clause in Hungarian

Hungarian is standardly assumed to be a discourse-configurational language. Its phrase structure is determined by discourse considerations to a large extent and information structural categories like topic and focus play a vital role in the word order of the language (see e.g. É. Kiss 1995, chapter 1). Furthermore, it is generally recognized that the Hungarian sentence shows a duality in terms of its phrase structure. While the preverbal field is hierarchically structured, the primary organizing forces being scope and discourse-functions, the postverbal area is usually assumed to be flat.²²

There have been a number of proposals that have been put forward about the LFG-theoretic analysis of the Hungarian phrase structure: Mycock (2006), Gazdik (2011) and Laczkó (2014a, 2014b, 2015, 2017). In this monograph I chose Laczkó's works as a baseline, as these constitute the most articulated framework and they have also been designed to be compatible with LFG's computational platform, the Xerox Linguistic Environment (XLE). This approach is inspired by É. Kiss's (1992) GB account. As Laczkó's works mostly concentrate on syntax, I will supplement them with my proposals regarding information structure.

Figure 1 shows the basic structure of Hungarian, as assumed in this monograph. I have simplified the diagram, leaving out those annotations which are not directly relevant for the purposes of the monograph. The interested reader is referred to the cited works of Laczkó. For ease of interpreting the information structural notations, let me also repeat the proposed taxonomy from section 1.3.

		+NEW	-NEW
+D-STRUCTURING	+CONTRASTIVE	contrastive focus	contrastive topic
	(-CONTRASTIVE)	information focus	topic
-D-STRUCTURING		completive information	background information

Table 1.
The proposed IS-taxonomy

²² While this picture might still be called the standard view (see e.g. É. Kiss 2002), there have been proposals that argue for a hierarchical postverbal field, see Surányi (2006b).

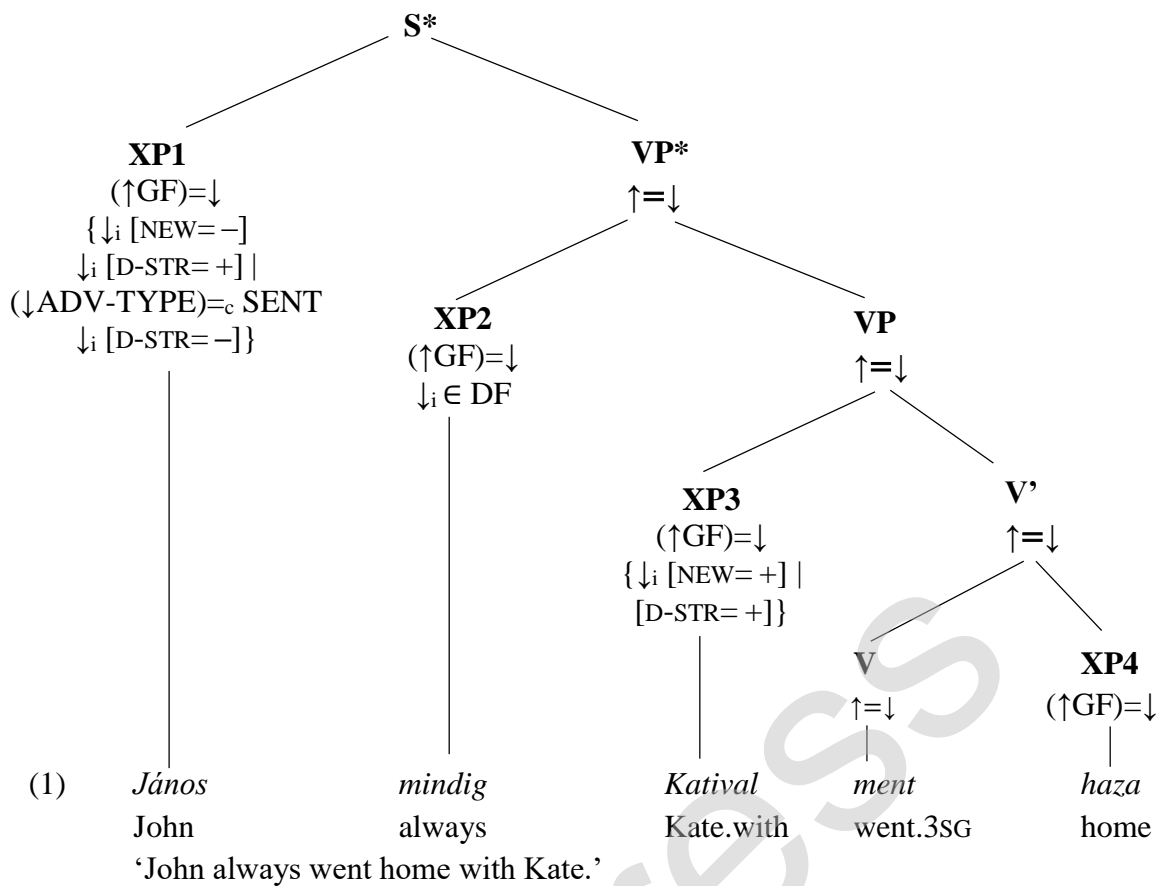


Figure 1.
Basic Hungarian sentence structure.

The sentence is headed by the exocentric S node. (The IP label is also possible if one assumes this category for Hungarian. However, see Laczkó (2014a: 333-336) for an argumentation on the side of S.) This node dominates what is traditionally called the “topic-field” of the Hungarian sentence.²³ This is labelled as XP1 in Figure 1. The arrows in the annotations are “metavariables”. ↑ refers to the mother-node’s functional structure while ↓ refers to the node’s own f-structure. So the (↑GF)=↓ should be read as “my grammatical function is the same as my mother node’s grammatical function” and this notation indicates that any grammatical function (GF) may occupy this position, including adjuncts. Arguments are either topics or contrastive

²³ The semantic properties of the topic field are not in the focus of this monograph, so they are not explicitly present here. It is generally assumed that topical entities are referential, although contrastiveness complicates the issue, as noted in connection with example (30) in Chapter 1. É. Kiss (2002: 10) claims that topics in Hungarian must also be specific, so the indefinite *egy autó* ‘a car’ in (ia) must refer to a member of a previously introduced set. However, Gécseg & Kiefer (2009) point out that in the proper context, subjects similar to the ones in (ia) may introduce brand-new referents, as in (ib). This is possible because we can easily accommodate the referents by contextual inference.

- (i) a. *Egy autó megállt a házunk előtt.*
one car stopped.3SG the house.POSS.1PL in.front.of
‘A car stopped in front of our house.’
- b. *Képzeld, mi történt tegnap.* (Guess what happened yesterday.)
Egy gyerek leesett a villamosról, de szerencsére nem sérült meg.
one child fell.off.3SG the tram.from but luckily not injured.3SG PV
‘A child fell off the tram but luckily, he wasn’t injured.’

topics at information structure (the feature CONTRASTIVE is not specified). Most adjuncts (ones with referential interpretation, e.g. temporal, locative, instrumental) behave in a parallel fashion. Sentence-adverbs (*luckily*, *hopefully*, etc.) however, seem to reject DISCOURSE-STRUCTURING IS-categories. This may be seen from (2a), where it is impossible to interpret *valószínűleg* ‘probably’ as a contrastive topic. Since restrictions are looser on CTs than they are on neutral topics (see example 30, Chapter 1), we may assume that *valószínűleg* is not a neutral topic either (it would also be incompatible with the aboutness-relation associated with topics). *Valószínűleg* may not be a focus (2b) either.

- (2) a. #*János valószínűleg Katival ment haza, valószínűtlenül pedig*
 John probably Kate.with went.3SG home improbably conversely
Annával
 Anne.with
 ‘Probably, John went home with Kate and improbably, he went home with Ann.’
- b. #*János valószínűleg ment haza (és nem biztosan).*
 ‘It was probably that John went home (not surely).’

However, such adjuncts may still provide some new information (about the mental disposition/judgment of the speaker in (3)) or be part of the background, so I assume that the non D-STRUCTURING categories are available for them.

- (3) A: *Hol van (valószínűleg) János?*
 where is probably John
 ‘Where is John (probably)?’
- B: *János [valószínűleg]_{CI/BI} az egyetemen van.*
 John probably the university.on is
 ‘John is probably at the university.’

This possibility is restricted to sentential adverbs with the help of a “constraining equation”. As opposed to regular (defining) equations in LFG annotations, which simply specify a value (making it exist), constraining equations check whether the specifications hold, without making them exist themselves. So (\downarrow ADV-TYPE)_{=c} SENT is satisfied if the adverb occupying that position is specified as sentential, as a lexical property.

The link between canonical topics and sentence adverbs that makes this constellation possible could be their capacity to restrict the domain for which the predicate holds. Both could be conceptualized as what Chafe (1976) (quoted in Maienborn 2001: 229) calls “Chinese-style topics: they limit the applicability of the main predication to a certain restricted domain. [...] Typically, it would seem, the topic sets a spatial, temporal, or individual framework within which the main predication holds.” While standard “aboutness”-topics restrict the application of the comment to a certain entity, a sentential adverb restricts it in a way that is related to the cognition or the assessment of the speaker (e.g. a probability-judgment in 3). In the long run, it may turn out that the disjunction in (1) is unnecessary and the position in question may simply be regarded as a topic-position. To establish this, more work on the semantic aspects of information structure should be done. Until that is carried out, I maintain this conservative

position. Since a sentence can contain several such entities and they freely intermingle,²⁴ this node is iterative.

The next position in the Hungarian sentence is the quantifier-field, labelled as XP2 in Figure 1. This is also iterative and the order of elements affects the meaning of the sentence, linear order being tied to scope.²⁵ This field is unspecified with respect to information structure, as can be seen from the following examples.

- (4) a. A: *Hova mentek a gyerekek nyaralni?*
 where went.3PL the kids be.on.vacation.INF
 ‘Where did the kids go for a vacation?’
 B1: *[Minden gyerek]_{TOPIC} a Balatonhoz ment nyaralni.*
 every kid the Balaton.to went.3SG be.on.vacation.INF
 ‘Every kid went to Balaton for vacation.’
 B2: *[Néhány gyerek]_{CF} a Balatonhoz ment (néhány pedig a Mátrába).*
 some kid the Balaton.to went.3SG some conversely the Mátra.in
 ‘Some kids went to the Balaton (and some went to the Mátra).’
- b. A: *Hova mentek mindig nyaralni a gyerekek?*
 where went.3SG always be.on.vacation.INF the kids
 ‘Where did the kids always go for vacation?’
 B: *A gyerekek [mindig]_{BI} a Balatonhoz mentek nyaralni.*
 ‘The kids always went to the Balaton for vacation.’
- (5) a. A: *János soha nem ment Katal haza?*
 John never not went.3SG Kate.with home?
 ‘Did John never go home with Kate?’
 B: *Nem, János [MINDIG]_{CF} Katal ment haza.*
 no John always Kate.with went.3SG home
 ‘No, John ALWAYS went home with Kate.’
- b. A: *Mikor ment haza János Katal?*
 When went.3SG home John Kate.with
 ‘When did John go home with Kate?’
 B: *János [mindig]_{IF} Katal ment haza.*
 ‘John always Kate.with went home.’
- c. A: *Kivel ment haza János?*
 with.whom went.3SG home John
 ‘Who did John go home with?’
 B: *János [mindig]_{CI} Katal ment haza.*

²⁴ Certain preferences may apply, see Gécseg (2001), É. Kiss (2005).

²⁵ Kálmán (2001) identifies three subfields, the *is* (‘too’)-field, the *minden* (‘every’)-field and the *sok* (‘many’)-field.

The next position is the widely discussed “focus-position” of Hungarian, which is Spec/VP in this framework.²⁶ It can be easy to see if something occupies this position as such elements trigger the inversion of the verb and the preverb (if there is one): preverbs normally occupy this position immediately in front of the verb, see (6)-(7) below. This is the primary place for information foci (answers to wh-questions) and for contrastive foci (e.g. corrections, emphasis, etc.).

- (6) *János meg-vette az autót.*
 John PV-bought.3SG the car.ACC
 ‘John bought the car.’
- (7) a. A: *Mit vett meg János?*²⁷
 what bought PV John?
 ‘What did John buy?’
 B: *János [az autót]_{IF} vette meg.*
 ‘John bought the car.’
- b. A: *János a motort vette meg?*
 John the motorbike.ACC bought PV
 ‘Did John buy the motorbike?’
 B: *Nem, János [AZ AUTÓT]_{CF} vette meg.*
 no John the car.ACC bought.3SG PV
 ‘No, John bought the CAR.’

IF and CF are +NEW, +D-STRUCTURING. However, the preverbal constituent is not necessarily related to a discourse subquestion. For example, (6) could be an answer for a question like “What happened?”, where the entire sentence is the answer. Also, non-focussed bare nouns may also be positioned here, in contexts like (8). These are still part of the pragmatic assertion so they are +NEW. As such, they should be seen as instances of Completive information (+NEW, –D-STRUCTURING).

- (8) A: *Mit csinál János?*
 what.ACC does John
 ‘What is John doing?’
 B: *János [fát]_{CI} vág.*
 John wood.ACC cuts
 ‘John is cutting wood.’ (‘John is wood-cutting.’)

This seems to suggest that the Spec/VP is the locus of +NEW IF-categories. This is not the entire picture though. In certain contexts, even topical entities may be in Spec/VP, as in (9). For more on this, see Gécseg (2013).

²⁶ It had been Spec/VP in the Chomskyan analyses (e.g. É. Kiss 1992) as well before Brody’s (1990) FocusP-analysis gained wide acceptance.

²⁷ Question words may also be analyzed in terms of IS-categories. A wh-item in a standard question is the counterpart of information focus. In the case of multiple-wh questions, e.g. *Who bought what?*, one of the question words acts as a “sorting key”, a link to a known set of entities with which new information is to be associated with, making it a counterpart for topic. See Mycock (2013) for discussion.

- (9) A: *Ki Szentgyörgyi Albert?* ‘Who is Szentgyörgyi Albert?’
 B: [*Szentgyörgyi Albert*]_{TOPIC} *fedezte fel a C-vitamint.*
 Szentgyörgyi Albert discovered.3SG up the vitamin.C
 ‘Szentgyörgyi Albert (was the one who) discovered vitamin C.’

So it seems that the only IF-category excluded from the Spec/VP slot is background information, the i-structural category with full negative specification. As we have seen sentence adverbs cannot be topics, so *valószínűleg* is ungrammatical in (10a/B). It would be marginally acceptable as focus (10b/B).

- (10) a. A: *Ki Szentgyörgyi Albert valószínűleg?*
 who Szentgyörgyi Albert probably
 ‘Who is Szentgyörgyi Albert probably?’
 B: **Szentgyörgyi Albert [valószínűleg]_{BI} fedezte fel a C-vitamint.*
- b. A: *Én biztos vagyok benne, hogy Szentgyörgyi Albert fedezte*
 I certain am that.in that(c) Szentgyörgyi Albert discovered.3SG
fel a C-vitamint.
 up the vitamin.C
 ‘I am certain that Szentgyörgyi Albert discovered vitamin C.’
 B: *??Szerintem csak [valószínűleg]_{CF} fedezte fel a C-vitamint*
 in.my.opinion only probably discovered.3SG up the vitamin.C
 ‘I think that he only PROBABLY discovered vitamin C.’

How such constructions should be analyzed is still under research. It seems from this discussion that the Spec/VP position in Hungarian is not particularly unique in terms of information structure, rather it is quite underspecified. There is no IS-related category that is exclusively associated with this slot. Rather, what seems to be peculiar about the Hungarian “focus position” has to do with its semantics. These additional semantic restrictions are revolving around exhaustivity although the empirical and theoretical landscape is debated (see the references in footnote 19).

While I acknowledge the significance of these semantic restrictions and their necessity for the full characterization of Spec/VP, in my present analysis I do not include them, as information structure is defined here as a level for representing those aspects of linguistic representation which concern not the information-content of a sentence itself, but rather the way it is presented. Exhaustivity is a concept clearly belonging to truth-conditional semantics and is not simply a packaging phenomenon.

Finally, the main verb itself may either be information/contrastive focus (11) or background information (12) in the sentence. That is, it is either the main assertion of the sentence or entirely part of the background.

- (11) a. A: *Mit csinál János?*
 what does John
 ‘What is John doing?’

B: [*Eszik*]_{IF}.
 eats.3SG
 ‘He is eating.’

b. A: *János eszik?*
 John eats
 ‘Is John eating?’

B: *Nem, [ISZIK]*_{CF}.
 no drinks.3SG
 ‘No, he is DRINKING.’

(12) A: *Mit eszik János?*
 what.ACC eats John?
 ‘What is John eating?’

B: *Halat [eszik]*_{BI} *János.*
 fish.ACC eats.3SG John
 ‘John is eating fish.’

As already mentioned, the word-order in the postverbal area is generally free and different word-order permutations do not result in different interpretations. This part of the Hungarian sentence lies outside the scope of our current investigation, so I finish this section off and go on to show the properties of subordinate clauses.

2.2 Subordinate clauses in Hungarian²⁸

According to É. Kiss (2002), the structure of subordinate clauses essentially parallels that of main clauses, so they contain the same structural positions. In her account, there are only two differences: one, subordinate clauses may be introduced by the complementizer *hogy* ‘that(c)’ and two, they are often associated with a pronoun. In this monograph, I subscribe to the generally accepted view of É. Kiss (2002), that is, that the internal structure of subordinate clauses is assumed to be similar to that of main clauses, as described in the previous section. A sketch is presented in Figure 2.

The point where debates are to be found in the literature is about the nature of the relationship between the matrix verb, the associate pronoun and the subordinate clause. The questions are the following: What is the nature of the associate pronoun? How is the pronoun associated with the subordinate clause.

²⁸ The scope of this discussion only extends to finite subordinate clauses that are associated with some argument function. Consequently, cases like (i), where the predicate is fully saturated without a subordinate clause will not be discussed. Relative clauses are also set aside.

(i) *János boldog volt, (hogy nyert).*
 John happy was that(C) won.3SG
 ‘John was happy (that he won).’

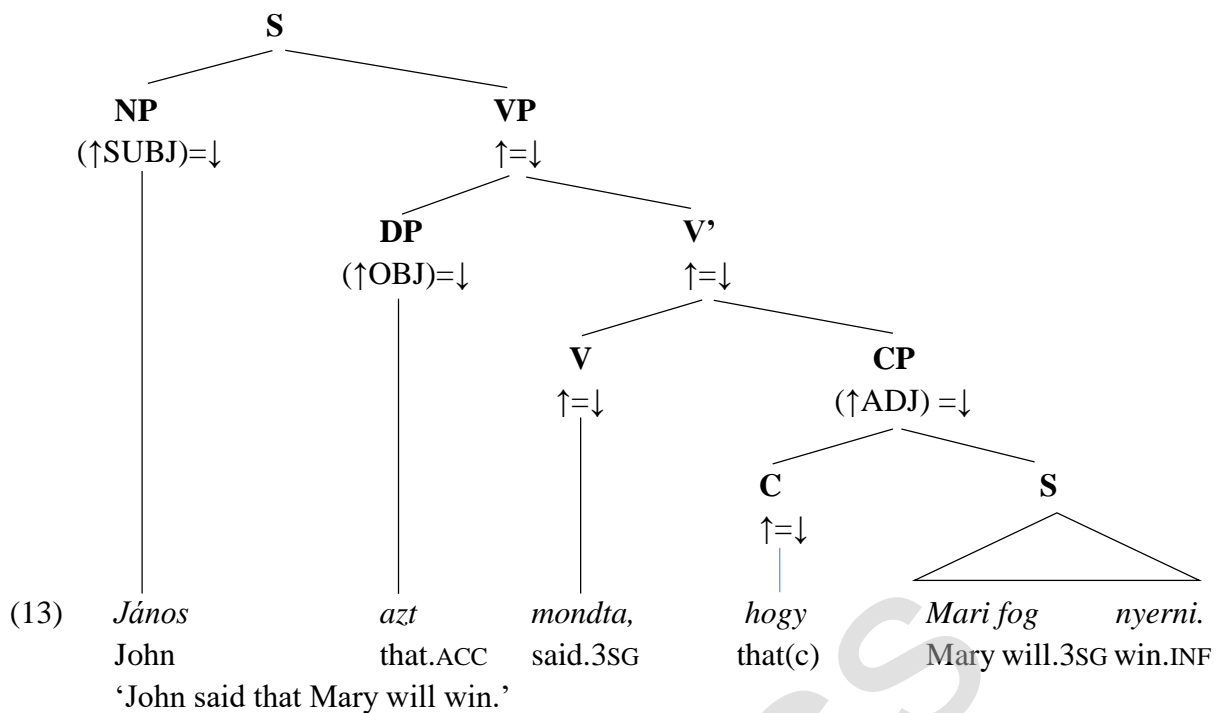


Figure 2.

A Hungarian sentence with a subordinate clause

These questions are crucial for this monograph since the proper analysis of one of the target-constructions, the operator fronting (OF) construction, cannot be carried out without clarifying these issues. One example for OF is shown in (14b): some constituent which is thematically related to an embedded predicate surfaces in a matrix sentence operator position (the detailed discussion of OF will take place in 5.2). What is important here is that the verbs that can participate in OF all take subordinate clauses (15c-d). On the other hand, elements undergoing OF are in complementary distribution with the associate pronoun: if one is present, the other one cannot appear, as in (15).

(14) a. *Az-t* *mondta,* *hogy jön János.*
 that-ACC said.2SG that(c) comes John.
 ‘You said that John comes.’

b. *János(-t)* *mondta,* *hogy jön.*
 John(-ACC) said.2SG that(c) comes
 ‘John you said that he will come.’

c. **Az-t futottad,* *hogy jön János.*
 that-acc ran.2SG that(c) comes John.

d. **János-t futottad,* *hogy jön.*
 John-ACC ran.2SG that(c) come’

(15) a. *Jánost *(azt) mondta *(azt), hogy jön.*

b. *Párizsba *(azt), mondta *(azt), hogy megy.*
 Paris.to that.ACC said.2SG that.ACC that(c) go.2SG

In the following sections, I will go into the details of the aforementioned issues. In brief, I will argue that contra Kenesei's (1992/1994) influential proposal, the associate pronouns are not expletives but contentful demonstrative pronouns. When such pronouns are present, the subordinate clause itself is an adjunct. When the pronoun is not present, the embedded clause itself picks up the grammatical function otherwise designated to the pronoun. As the complementizer is also optional, following LFG's Economy of Expression principle (see section 1.2), a CP projection for the subordinate clause is only assumed when *hogy* ('that(C)') is present. In other cases, the subordinate clause should be analyzed as an S. The constraints on the distribution of the pronoun and the complementizer then naturally follow from the proposed system and from general principles.

2.2.1 General properties of Hungarian subordinate clauses

The structure of Hungarian subordinate clauses is described in Kenesei (1992/1994) in detail so this section draws on Kenesei's work with respect to empirical data, although the subsequent analysis will be considerably different. In the most basic form of the Hungarian subordinate clause construction, there is a main verb, an associate pronoun and a subordinate clause itself. The pronoun is a demonstrative one in form and bears some grammatical function determined by the verb: it can be a subject (16a), an object (16b) or an oblique complement (16c).

- (16) a. *Az valószínű, hogy János fog nyerni.*
 that likely that(c) John will win.INF
 'It is likely that John will win.'
- b. *Az-t mondtam, hogy János fog nyerni.*
 that-ACC said.1SG that(c) John will win.INF
 'I said that John would win.'
- c. *Arra számítok, hogy János fog nyerni.*
 that.onto expect.1SG that(c) John will win.INF
 'I expect that John will win.'

Although such pronouns are commonly present in complex sentences containing subordinate clauses, they are optional in many circumstances. So in sentences like (16a-b), the demonstrative could be absent without a change in meaning. There is variation with regard to the droppability of pronouns bearing oblique case (as in 17c and 17d).

- (17) a. *Valószínű, hogy János fog nyerni.*
 likely that(c) John will win.INF
 'It is likely that John will win.'
- b. *Mondtam, hogy János fog nyerni.*
 said.1SG that(c) John will win.INF
 'I said that John would win.'
- c. **(Arra) számítok, hogy János fog nyerni.*
 that.onto expect.1SG that(c) John will win.INF
 'I expect that John will win.'

- d. (*Arról*) *meggyőződtem, hogy János a győztes.*
 that.from ascertained.1SG that(c) John the winner
 ‘I ascertained that John was the winner.’

In these examples, the pronoun is always preverbal. Moreover, it is associated with some “prominent” discourse-function, IF, CF or CT. Interpreting them as standard topics is not possible. Thus sentence (18) is ungrammatical with a falling intonation, it must receive the fall-rise contour that characterizes contrastive topics. (The focus interpretation is excluded since the Spec/VP position is occupied by the preverb *meg*.)

- (18) *Az-t meg-mondtam, hogy János fog nyerni.*
 that-ACC PV-said.1SG that(c) John will win.INF
 ‘That John will win, I did say.’

This could just be a side-effect of the event-structural properties of the verbs in question. É. Kiss (2005) observes that although in principle any argument of a predicate may be positioned in the topic-field, not all possibilities are equally neutral information-structurally. Take change-of-state verbs as an example. Sentence (19a) with the predicate *nyír* ‘cut’ is entirely natural with a standard falling intonation but in (19b) *a füvet* ‘the grass.ACC’ must be either information/contrastive focus or contrastive topic.

- (19) a. *Mari nyírja a füvet.*
 Mary cuts the grass.ACC
 ‘Mary is cutting the grass.’
 b. *[A füvet]_{#TOP/CT/IF/CF} nyírja Mari.*
 ‘Mary is cutting the GRASS.’/ ‘The grass, Mary is cutting.’

According to É. Kiss (2005) the reason for this is that depending on the event-type they express, predicates lexically select some of their arguments as “main news” and these must be in the “predicate part” of the sentence: the postverbal area or the focus position. For a change-of-state verb like *nyír* ‘cut’, the theme is the “main news”. So in (19a) it is postverbal, which is the default mapping. If it is not postverbal, like in (19b), the informativeness of the predicate part must be strengthened via some other means, like the focussing of the theme (which makes it part of the predicate part). Other methods are also available: focussing some other element (an adjunct in 20a), creating a contrastive topic (20b), or temporally modifying the predicate (20c).

- (20) a. *[A füvet]_{TOPIC} tegnap nyírta le Mari.*
 the grass.ACC yesterday cut.PAST.3SG down Mary
 ‘Mary cut the grass yesterday.’
 b. *[A füvet]_{CT} nyírja Mari, de a bokrot nem.*
 the grass.ACC cuts Mary but the bush.ACC not
 ‘The grass, Mary is cutting, but the bush, she is not.’
 c. *[A füvet]_{TOPIC} már nyírja Mari.*
 the grass.ACC already cuts Mary
 ‘Mary is already cutting the grass.’

The verbs under discussion behave similarly. The main news for them is whatever semantic role is associated with the respective grammatical functions in (21-24) (theme/proposition), so this is postverbal by default. If it is preverbal, their i-structural behavior parallels (20).

- (21) a. *Mari mondott egy történetet.*
 Mary said.3SG a story.ACC
 ‘Mary told a story.’
- b. *[Egy történetet]_{#TOP/CT/IF/CF} mondott Mari.*
- (22) a. *Mari fontolgatja a lemondását.*
 Mary contemplates the resignation.POSS.3SG.ACC
 ‘Mary is contemplating her resignation.’
- b. *[A lemondását]_{#TOP/CT/IF/CF} fontolgatja Mari.*
- (23) a. *Mari sérelmezi a cselekedeteidet.*
 Mary resents the acts.POSS.2SG.ACC
 ‘Mary resents your acts.’
- b. *[A cselekedeteidet]_{#TOP/CT/IF/CF} sérelmezi Mari.*
- (24) a. *Meg-győződtem János győzelméről.*
 PV-ascertained.1SG John victory.POSS.3SG.FROM
 ‘I ascertained that John won.’
- b. *[János győzelméről]_{IF/CF} győződtem meg.*
- c. *[János győzelméről]_{#TOP/CT} meggyőződtem.*

The pronoun behaves in a similar fashion, with the addition that some verbs, e.g. *mond* ‘say’, *gondol* ‘think’ and *állít* ‘claim’ disprefer to be a neutral initial element in the “predicate part”.

- (25) *?Mondtam/gondoltam/ állítottam azt, hogy János fog nyerni.*
 said.1SG thought.1SG claimed.1SG that.ACC that(c) John will win.INF
 Intended: ‘I said/thought/claimed that John would win.’

Thus, these verbs require some additional weight in this configuration, e.g. adding a quantifier (26a), a verum focus (26b), or filling the Spec/VP slot with a particle (26c). This is probably connected to the fact that while lexical noun objects are indeed possible for these verbs (as in 21, to be also discussed in 2.2.2.), they are often restricted.

- (26) a. *Mindig mondtam azt, hogy János nagyon okos.*
 always said.1SG that.ACC that(c) John very smart
 ‘I have always said that John is very smart.’
- b. *El-mondtam azt, hogy szerintem János fog nyerni.*
 away-said.1SG that.ACC that(c) in.my.opinion John will win.INF
 ‘I told (them) that in my opinion John would win.’

- c. #*Mondtam / MONDTAM azt, hogy János fog nyerni.*
 said.1SG that.ACC that(c) John will win.INF
 ‘I DID say that John would win.’

If the verb is not of this kind, the pronoun may naturally be postverbal, like in (27).

- (27) a. *Fontolgom azt, hogy lemondok.*
 contemplate.1SG that.ACC that(c) resign.1SG
 ‘I am contemplating that I resign.’
- b. *Sérelmezem azt, hogy korán hazamentetek.*
 resent.1SG that.ACC that(c) early home.went.2PL
 ‘I resent that you went home early.’

What I conclude from this is that the placement and the discourse function of the pronoun associated with the subordinate clause are dependent on the semantic type of the verbs in question and are not special properties of the construction with the pronoun.

Another optional element in such sentences is the complementizer *hogy*. Following the principle of Economy of Expression, this means that the subordinate clause can be expressed either as a CP (if there is a complementizer) or as an S (when there is not). (28) here shows the complementizer-less versions of (16).

- (28) a. *?(Az) valószínű, János fog nyerni.*
 that likely John will.3SG win.INF
 ‘It is likely that John will win.’
- b. *Azt mondtam, János fog nyerni.*
 that-ACC said.1SG John will.3SG win.INF
 ‘I said that John would win.’
- c. *Arra számítok, János fog nyerni.*
 that.onto expect.1SG John will.3SG win.INF
 ‘I expect that John will win.’

The droppability of the complementizer is subject to a number of lexical and structural constraints, as discussed in Kenesei (1992/1994). One of the most prominent restrictions is that “focusing in the main clause blocks the deletion of the embedded complementizer” (Kenesei 1994: 336), as in (29b-c). However, it must be added that in the case of information focus, only lexical noun foci block the omission of the complementizer. If the IF is a pronoun, it may be dropped (29a/B). The situation is similar with contrastive topics (29d-e), but it is important to recall that CTs are always associated with a focus so the complementizer must be present, independent of the presence of the CT.

- (29) a. A: *Mit mondtál?*
 what.ACC said.2SG
 ‘What did you say?’
- B: *[Azt]_{IF} mondtam, (hogy) János fog nyerni.*
 thatACC said.1SG that(c) John will win.INF

‘I said that John would win.’

- b. A: *Kinek mondtad, hogy János fog nyerni?*
who.DAT said.2SG that(c) John will win.INF
‘To whom did you say that John would win?’

B: *[Marinak]_{IF} mondtam, *(hogy) János fog nyerni.*
Mari.DAT said.1SG that(c) John will win.INF
‘I said to Mary that John would win.’

- c. A: *Annának mondtad, hogy János fog nyerni?*
Anne.DAT said.2SG that(c) John will win.INF
‘Did you say it to Anne that John would win?’

B: *Nem, [MARINAK]_{CF} mondtam, *(hogy) János fog nyerni.*
No Mari.DAT said.1SG that(c) John will win.INF
‘I said to Mary that John would win.’

- d. *[Azt]_{CT} MONDTAM, *(hogy) János fog nyerni, de azt nem*
that.ACC said.1SG that(c) John will win.INF but that.ACC not
hogy ma.
that(C) today
‘I did say that John would win, but I did not say that he will win today.’

- e. *[Marinak]_{CT} MONDTAM, *(hogy) János fog nyerni de Annának*
Mari.DAT said.1SG that(c) John will win.INF but Ann.DAT
nem.
not
‘To Mary, I did say that John would win, but I did not say it to Anne.’

It has been claimed that dropping both the pronoun and the complementizer leads to unacceptability (Jánosi 2014: 104-107). Example (30) and the judgment are from her work.

- (30) *Mondta, *(hogy) új autót vett.*
said.3SG that(C) new car.ACC bought.3SG
‘He said that he had bought a new car.’

However, a closer investigation reveals that the generalization is premature. For example, as noted in Kenesei (1992: 620/1994: 337), matrix verbs that take an interrogative complement readily allow dropping both the pronoun and the complementizer.

- (31) *Kérdezte, mikor vettem új autót.*
asked.3SG when bought.1SG new car.ACC
‘He asked when I had bought a new car.’

I think that the weirdness of the complementizer-less version of (30) stems from the fact that without the pronoun, the matrix predicate *mond* is interpreted as an emphatic element in the sentence. In other words, in isolation, *mond* in (30) is interpreted as the focus of the sentences (‘He DID say that he bought a new car’). If that is the case, the restriction shown earlier is the

culprit: the main clause focus prevents dropping the complementizer. If a proper context is construed, where it is ensured that no emphasis is put on the predicate itself, dropping the pronoun and the complementizer becomes possible. Consider (32), which could be a reply to (31), or (33). In these cases, the information-content of the embedded clause is much more salient, allowing for an unstressed main clause.

- (32) *?Mondtam, nem vettem én újat, csak átfestettem a régit.*
 said.1SG not bought I new.ACC just repainted.1SG the old.ACC
 ‘I said that I hadn’t bought a new one, I just repainted the old one.’
- (33) A to B: *Fordulj balra.* [B turns right.] *Mondom fordulj balra.*
 turn.IMP.2SG left.to say.1SG turn.IMP.2SG left.to
 ‘Turn left. Left I say.’

Other predicates, like *kérdez* ‘ask’, are less prone to this analytical aspect, and the result is acceptable sentences like (31). Other examples for the same phenomenon include *ígér* ‘promise’, *gondol* ‘think’, *remél* ‘hope’ or *gyanít* ‘suspect’, as in (34).

- (34) *Ígérem,/ Gondolom,/ Remélem,/ Gyanítom, János fog nyerni.*
 promise.1SG think.1SG hope.1SG suspect.1SG John will.3SG win.INF
 ‘I promise/think/hope/suspect John will win.’

Nevertheless, there does seem to be a considerable amount of lexical variation in the omission possibilities. Nominal predicates do not allow the complementizer to be dropped (35a). A number of verbal predicates allow only either the pronoun or the complementizer to be dropped (35b), conforming to Jánosi’s (2014) claim. No context could save the ones in (35b).

- (35) a. *Kár/ Szerencse, *(hogy) János nyert.*
 pity luck that(c) John won.3SG
 ‘It’s a pity/luck that John won.’
- b. **(Azt) jószoltam/ vallottam/ képzeltem, János nyert.*
 that.ACC predicted.1SG testified imagined John won.3SG
 ‘I predicted/ testified/ imagined that John won.’

A full account of the omission possibilities and constraints lies beyond the scope of this investigation, but the phenomenon itself will be featured in section 4.2, where we are going to discuss the argument structural aspects of Operator Fronting. My general position is that dropping these elements is possible by default but syntactic (word class)/semantic (verb-type)/information structural (focus or contrast) factors may block the omission.

As a last general point about subordinate clauses, it is important to mention that with some verbs, it is possible to use another demonstrative pronoun, for example *úgy* (‘so.DIST’). Sometimes it is the only option, or it may alternate with *azt* (‘that.ACC’). We will return to these in the next section.

- (36) a. *Az/ *Úgy tűnik, hogy János fog nyerni.*
 that so.DIST seems that(c) John will.3SG win.INF
 ‘It seems that John will win.’
- b. *Azt/ Úgy gondolom, hogy János fog nyerni.*
 that.ACC so.DIST think.1SG that(c) John will.3SG win.INF
 ‘I think that John will win.’

2.2.2 The expletive-demonstrative debate

There are two main approaches to the characterization of the pronoun and its relationship to the embedded clause. According to the first one (put forward by Kenesei 1992/1994, also accepted by Lipták 1998 and Gervain 2002), the pronoun is an expletive²⁹ and the real, semantic argument of the main verb is the propositional complement clause. The opposing view holds that the pronoun is semantically contentful and it is associated with the complement clause via adjunction (Tóth 2000, Rákosi & Laczkó 2005) or complex NP-formation (É. Kiss 2002).

Let us take a closer look at the chronologically first view, which holds that the pronouns in question are expletives. As noted, according to Kenesei (1992/1994), verbs taking a subordinate clause in Hungarian strictly subcategorize for a CP as the realization of their propositional argument. Kenesei supports this view with examples like (37-40), in which replacing the CP with a semantically related DP results in ungrammaticality.

- (37) a. *(Az) szerencse volt, hogy idejekorán vettek fel kölcsönt.*
 that luck was that(c) in.time bought.3PL up loan.ACC
 ‘It was lucky that they had taken out a loan in time.’
- b. **A kölcsön felvétele szerencse volt.*
 the loan taking.out.POSS.3SG luck was
- (38) a. *(Az) jó volt, hogy már tegnap megérkeztetek.*
 that good was that(c) already yesterday arrived.2PL
 ‘It was good that you had already arrived yesterday.’
- b. **A tegnapi megérkezésetek jó volt.*
 the yesterday arrival.POSS.2PL good was
- (39) a. *Mari azt hitte, hogy jókor szólalt meg.*
 Mary that.ACC believed.3SG that(c) at.a.good.time spoke.3SG PV
 ‘Mary believed that she had spoken at the right time.’
- b. **Mari a jókori megszólalást hitte.*
 Mary the at.a.good.time speaking.ACC believed.3SG

²⁹ An expletive is a semantically empty dummy pronoun, whose presence is required by some syntactic principle. Some standard examples are weather *it* (*It is raining*), extraposition *it* (*It seems that John is happy*) or extraposition *there* (*There are three people in the room*). Note that the expletive-status of these pronouns is not uncontested, see e.g. Hedberg (2000), Moro (1997), Tortora (1997).

- (40) a. *Azt mondta, hogy tudja a választ.*
 that.ACC said.3SG that(c) knows the answer.ACC
 ‘S/he said that s/he knew the answer.’
- b. **A válasz tudását mondta.*
 the answer knowing.ACC said.3SG

Furthermore, Kenesei (1992/1994) proposes that subordinate clauses introduced by the complementizer *hogy* ‘that(C)’ cannot tolerate case-marking. So there is a “division of labor” between the subordinate clause and the pronoun: the first receives theta-marking, while the second carries the case-feature. From this it follows that the pronoun, without a thematic role, is an expletive, forming a syntactic chain with the subordinate clause.

From an LFG-viewpoint, this means that these predicates have a lexical entry along the lines of (41), where the pronoun is an athematic argument. This is conventionally represented as placing it outside the angled brackets in (41). (The angled brackets contain the grammatical functions mapped onto semantic arguments of the verbs.) (41a) illustrates a subject-expletive (for a predicate like for *szerencse* ‘luck’ in 37), while (41b) is for an object-expletive (like *mond* in 40).

- (41) a. *predicate* <(COMP)> (SUBJ)
 b. *predicate* <(SUBJ)(COMP)> (OBJ)

The alternative view (Tóth 2000, É. Kiss 2002, Rákosi & Laczkó 2005) holds that the pronoun is not an expletive, but a real, referring element and the clause is associated with it via complex NP-formation (É. Kiss) or adjunction (Tóth, Rákosi & Laczkó). One of the main arguments for this view is that the pronoun does not display the behavior of expletives in general: it may be questioned, explicitly focussed or quantified. Expletives, being semantically vacuous, are expected not be compatible with such operations.

- (42) a. *Mit mondott Mari?*
 what.ACC said.3SG Mary
 ‘What did Mary say?’
- b. *CSAK AZT mondtad, hogy János fog nyerni. (Azt nem, hogy ennyire fölényesen)*
 only that.ACC said.2SG that(c) John will.3SG win.INF
 ‘You only said that John will win. (And not that he will do so by such a large margin.)’
- c. *Azt is mondtad, hogy a verseny után elmegyünk ünnepelni.*
 that.ACC too said.2SG that(c) the competition after away.go.1PL
 celebrate.INF
 ‘You also said that after the competition we’ll go to celebrate.’

- (43) a. **What seems?*
 b. **ONLY IT seems that John is happy.*
 c. **Even it seems that John is happy.*

Furthermore, expletives cross-linguistically occur as subjects,³⁰ while for example in (40a) and (42b), *azt* clearly functions as an object. However, Kenesei (1992/1994) proposes that the reason for this state of affairs is that expletives behave in different ways in typologically different languages. The aforementioned restriction (barring nonsubject functions) holds in configurational languages like English. In such languages expletives really just serve as fillers for specific syntactic slots (e.g. Spec/IP). In discourse-configurational languages, including Hungarian, they have a different role: they represent clauses in configurations which are unavailable for the clauses themselves, namely, when they would bear certain discourse functions.³¹ The positions in (44a-b) are indeed unsuitable for *that*(C)-clauses in Hungarian.

- (44) a. **Csak hogy János fog nyerni mondtad.*
 only that(c) John will.3SG win.INF said.2SG
 b. **Hogy a verseny után elmegyünk ünnepelni is mondtad.*
 that(C) the competition after away.go.1PL celebrate.INF too said.1SG

According to Kenesei (1992/1994), the source of the restriction is phonological: the focus and the verb must form a phonological phrase and such a phrase cannot be larger than a finite clause. If this is correct, then this kind of expletive is different from standard ones in terms of its origin of motivation: syntax vs. phonology. However, it is not clear why a phonological restriction would necessarily result in inserting a semantically empty pronoun. It seems much more plausible that phonology should not care about the semantic nature of the pronoun.

It must be added that a leakage of this generalization is that from its perspective, one would expect that a clause will not appear preverbally in a position that can also be occupied by a pronoun. To put it differently, if the function of the pronoun is to represent a clause in a position that is barred for the clause, then the clause should definitely not be able to surface in the barred position: we would expect the complementary distribution of the pronoun and a

³⁰ It has been suggested (e.g. Postal & Pullum 1988) that the sentences like (i) contain the expletive pronoun *it* as an object

(i) *I regretted it every time that I had dinner with John.*

The proper analysis of such sentences is debated. Rothstein (1995) argues that even though the semantics of such pronouns is bleached, they are in fact not expletives. This can be seen for example from the fact that removing the pronoun results in a slightly altered meaning. Thus, while (i) means that every event of dinner was matched by an event of regret, the pronoun-less version would mean that there was only one regretting event (for example, some incident makes me reinterpret my evaluation of the past dinners with John, which may have seemed happy at those times). Another counter-argument is that Postal's & Pullum's (1998) verbs do occur with uncontroversially semantic objects.

(i) *I regretted my decision.*

³¹ "Rather than being required by the extended projection principle to fill in an empty subject, expletives in Hungarian can occur in quantifier field or topic positions where the clauses are blocked or have decreased acceptability" (Kenesei 1994:324).

clause in a given position. Although this is borne out in the case of foci, as shown in (45), both a pronoun or a clause may serve as a contrastive topic.

- (45) a. *[Hogy nyerhet]_{CT}, NEM gondolta János.*
 that(c) win.COND.3SG not thought.3SG John.
 ‘That he could win, John didn’t think (of).’
- b. *[Azt]_{CT} NEM gondolta János, hogy nyerhet.*
 that.ACC not thought.3SG John that(c) win.COND.3SG
 ‘That he could win, John didn’t think (of).’

What I am suggesting is that while the data in (44) may be explained along the lines of Kenesei’s (1992/1994) typological proposal, the alternative theory is not excluded by it: Hungarian may also utilize another strategy (real pronoun + adjunction/complex-NP formation) to bypass the restriction illustrated in (44), without the postulation of a new kind of expletive. In other words, the data in (44) are not conclusive, as they can be interpreted from several analytical perspectives. Hence it is other empirical/theoretical/typological considerations that should decide the issue.

Another argument for the view where the pronoun is a genuine contentful demonstrative is the fact that the complement clause may be dropped, leaving the pronoun stranded. A sentence like (46) is problematic under the expletive-view since the pronoun does not form a chain with a semantically contentful constituent. In the GB-framework of Kenesei (1992/1994), (46/B) would violate the Chain condition: every nominal must be a part of chain (which may be a one-membered a chain) which contains a case-position and a θ -position. There is no θ -marked position to license the supposedly expletive pronoun in the case position in the sentence (46/B).

- (46) A: *Szerintem János okos.*
 in.my.opinion John smart
 ‘In my opinion, John is smart’
- B: *Én is azt/ ezt mondom.*
 I too that.ACC this.ACC say.1SG
 ‘I also say so.’ (‘that John is smart’)

Furthermore, Kenesei’s (1992/1994) typological generalization seems not to be very well supported by cross-linguistic data. Although there is a considerable body of research about non-configurational languages, to my best knowledge no other language exhibits the pattern proposed by Kenesei (1992/1994). For instance, Finnish, a language related to Hungarian (which is also discourse-configurational), has an expletive which behaves in an entirely orthodox way: according to Nikanne & Holmberg (2002), Holmberg (2005) and Holmberg (2010), the expletive pronoun *sitä* must be inserted to Spec/TP (and to nowhere else) if nothing else occupies it (47c). No object-expletives are reported in Finnish.

- (47) a. *Minulle sattui onnettomuus.*
 to-me happened accident
 ‘An accident happened to me.’

- b. **Sattui minulle onnettomuus.*
happened to-me accident
- c. *Sitā sattui minulle onnettomuus.*
 EXPLETIVE *happened to-me accident*
 ‘There happened an accident to me.’

Other non-configurational languages either have no expletives or they have expletives that behave like English expletives.

Furthermore, Hungarian has another candidate for an element being expletive-like. As already mentioned, some verbs associate the pronoun *úgy* ‘so.DIST’ with their subordinate clause. The Hungarian equivalent of *seem* is one such verb. Here *úgy* behaves just like English expletives: it cannot be questioned (48b), focussed (48c), or quantified (48d) and it also cannot be omitted (48e). As no other subject is allowed in such sentences, it is also most probably the grammatical subject of the sentence (48f). Thus, if anything is an expletive in Hungarian, it should be *úgy* and not *azt*, and the typological division put forward by Kenesei (1992/1994) should be discarded.

- (48) a. *Úgy tűnik, hogy János fog nyerni.*
 so.DIST seems that(c) John will.3SG win.INF
 ‘It seems that John will win.’ Lit.: ‘So seems that John will win.’
- b. **Hogy tűnik János?*
 how seems John
- c. **ÚGY tűnik, hogy János fog nyerni, (és nem úgy, hogy Péter).*
 intended: ‘What seems is that John will win, not that Peter will do so.’
- d. *Úgy tűnik, hogy érdekes lesz a verseny. *Ezenkívül úgy is tűnik, hogy János fog nyerni.*
 intended: ‘It seems that the race is going to be interesting. What also seems is that John will win.’
- e. **Tűnik, hogy János fog nyerni.*
- f. **János úgy tűnik, hogy ő fog nyerni.*
 John so.DIST seems that(c) he will.3SG win.INF

A related argument has to do with the form of the demonstrative pronoun. In most examples given so far the demonstrative pronoun was in its distal form, which is the usual case. However, in certain contexts the proximal version can also appear. Intuitively, this is related to the discourse-status of the information presented in the associated clause, on which *ezt* ‘this.ACC’ imposes a stricter restriction than *azt* ‘that.ACC’. *Is* ‘too’ signals in (49) that the proposition had already been evoked in the discourse and in this case the CP may be dropped. If the CP is there, the absence of *is* ‘too’ would make (49) worse. If the pronoun was semantically empty, we would not expect such discourse considerations to have any effect on its form.

- (49) *Én is ezt mondom, hogy János fog nyerni.*
 I too this.ACC say.1SG that(c) John will.3SG win.INF
 ‘I also say that John will win.’

Importantly, the form of *úgy* ('so') (formally it is also a distal demonstrative) seems to be unaffected by the very same discourse considerations – the proximal counterpart is always quite degraded, even in contexts that otherwise license this form of the pronoun. This suggests that the argument is valid in Hungarian and *úgy* is more plausibly analyzed as an expletive than *azt/ezt*.

- (50) *?#Nekem is úgy tűnik/látszik, hogy János fog nyerni.*
 me.DAT too so.PROX seems that(c) John will.3SG win.INF
 'It seems to me too that John will win.'

Some verbs which occur with the object accusative pronoun *azt* can also alternatively select *úgy*. These verbs show a mixed behavior in these tests. On the one hand, they may be questioned and replaced with a proximal counterpart. On the other hand, focussing and quantifying are still ungrammatical.

- (51) a. *Azt/ úgy gondolom, hogy János fog nyerni.*
 that.ACC so.DIST think.1SG that(C) John will.3SG win.INF
 'I think that John will win.'
- b. *Te hogy gondolod?*
 you how think.2SG
 'What do you think?' Lit.: 'How do you think?'
- c. **CSAK ÚGY gondolom, hogy János fog nyerni, úgy nem, hogy nagy fölényel.*
 only so.DIST think.1SG that(c) John will.3SG win.INF so.DIST not that(C)
 large margin.with
 intended: I think that John will win, but not that he will do so by a large margin.'
- d. **Úgy is gondolom, hogy János fog nyerni.*
 so.DIST too think.1SG that(c) John will.3SG win.INF
 intended: What I also think is that John will win.'
- e. *Én is így gondolom, hogy János fog nyerni.*
 I too so.PROX think.1SG that(c) John will.3SG win.INF
 'I also think that John will win.'

Another problem with the expletive analysis of the pronouns associated with Hungarian subordinate clauses is that they occur not only as grammatical functions associated with structural case (subject, object), but also as complements of predicates assigning inherent case. This is problematic for an expletive-analysis regardless of one's theoretical persuasion, even if one allows for object-expletives. Inherent case is always associated with the idiosyncratic meaning of the predicate.³² From an LFG viewpoint, such elements bear the OBL(ique) grammatical function, which in LMT is specified as "semantically restricted" (see section 1.2). This is obviously incompatible with an analysis positing a lack of semantic content. Moreover the CP-complements of such predicates are always replaceable with case-marked DPs. The

³² Chomsky (1986) formulates this as the Inherent case condition: if A is an inherent case assigner, then A assigns case to an NP if and only if A theta-marks the NP.

apparent lack of such DPs in sentences like (37-40) is one of Kenesei's (1992/1994) main arguments for the expletive-analysis, so their systematic availability is surprising for such an approach. This had led Lipták (1998), who otherwise subscribes to the expletive-analysis in the case of nominative and accusative pronouns, to abandon this aspect of Kenesei's theory and refer to such inherently case-marked pronouns as "argumental referring words".

- (52) a. *János büszke volt arra, hogy nyert.*
 John proud was that.onto that(c) won.3SG
 'John was proud that he had won.'
- b. *János büszke volt a győzelmére.*
 John proud was the victory.POSS.3SG.onto
 'John was proud of his victory.'
- (53) a. *János attól tartott, hogy veszít.*
 John that.from was.afraid that(c) loses
 'John was afraid that he might lose.'
- b. *János a vereségétől tartott.*
 John the defeat.POSS.3SG.from was.afraid
 'John was afraid of his (possible) defeat.'

Based on these considerations, the alternative theories of É. Kiss (2002), Tóth (2000) and Rákosi & Laczkó (2005) seem to be a viable alternative. As has been mentioned, in these approaches the pronouns in question are genuine demonstrative pronouns and the clauses themselves are associated with them via adjunction or complex-NP formation.

According to these proposals the relevant predicates, instead of having lexical entries like (41), have entries like (54).

- (54) a. *predicate <(SUBJ)>*
 b. *predicate <(SUBJ)(OBJ)>*

The first challenge for this approach is the pattern illustrated in (37-40): predicates occurring with a pronoun and a subordinate clause often reject a single DP subject or object. I would like to argue that while there are indeed certain restrictions, the generalization as a whole does not hold water. My first observation is that even in Kenesei's original examples some judgments are debatable. For instance, (37b) seems to be relatively acceptable to me (it is even better if one takes a modified version like *nagy szerencse* 'big luck'). Also, a web-search results in several examples where *szerencse* 'luck' takes a DP subject.

- (55) a. *Merkel: micsoda szerencse volt a békés kelet-európai rendszerváltás.*
 Merkel: what luck was the peaceful eastern.European change.of.system
 'Merkel: what a luck was the peaceful change of the political system in Eastern Europe.'
 from: http://hvg.hu/vilag/20140227_Merkel_micsoda_szerencse_volt_a_bekes_ke

- b. *Óriási szerencse volt számunkra a PC-k gyors elterjedése.*
 huge luck was for.us the PC-PL fast spread.POSS.3PL
 The fast spread of the PCs was a huge luck for us.
 from: Hungarian National Corpus
- c. *Szerinte inkább véletlen szerencse volt a gép lelövése.*
 according.to.him rather random luck was the plane shooting.POSS.3SG
 ‘According to him, the plane’s shooting down was sheer luck.’
 from: Hungarian National Corpus

Secondly, the landscape of pronoun-distribution is more complex than the picture presented in Kenesei (1992/1994). There are predicates which can never occur with a pronoun, only with a clause. Such predicates are plausibly analyzed as subcategorizing for CPs. However, here the expletive-demonstrative question does not arise in the first place. *Muszáj* ‘must’ is a predicate of this kind (Kálmán 2001: 170).³³

- (56) a. (*Az) *muszáj (*az), hogy elgyere.*
 that must that that(c) come.2SG. SBJV
 ‘It is a must that you come.’

With other predicates, the pronoun is optional but its presence or absence clearly modifies the meaning of the sentence. *Van* ‘be’ is one such predicate: without *az*, the sentence has an existential reading. With it, the sentence describes a particular situation.

- (57) a. *Van, hogy János nyer.*
 be.3SG that(c) John wins
 ‘It happens that John wins.’
- b. *Az van, hogy János nyert.*
 that be.3SG that(c) John won.2SG
 ‘The situation is that John won.’

That the meaning difference between (57a) and (57b) is not simply because of the focus discourse function associated with the pronoun is evidenced by the fact that a distinct focussed version of (57a) actually exists, but with a different pronoun, *olyan* ‘such.DIST’.

- (58) *Csak olyan van, hogy János győz.*
 only such.DIST be.3SG that(c) John wins
 ‘Only that John wins happens.’

³³ Jánosi (2013: 61) claims that *szól* ‘tell’ is also a predicate subcategorizing directly for a CP. However, it can alternatively take an inherently case-marked DP, which makes it similar to *tart* ‘be.afraid’ in (53).

- (i) *Szólt (arról), hogy sikerült a vizsga.*
 told.2SG that.from that(c) succeeded.3SG the exam
 ‘S/he told us that s/he had passed the exam.’

If the presence of *az* causes such a difference in meaning, it is unlikely that it should be analyzed as an expletive. If *az* in (57b) is not an expletive, we then expect it to behave like normal pronouns and unlike expletives. This prediction is borne out: a direct question, using the corresponding question word *mit* ('what') can only be formed about (57b) and not (57a). Hence, only B1 is a proper response in (59).

- (59) A: *Mi van?*
 what is
 approx.: 'What's up?'/ 'What is the situation?'
- B1: *Az van, hogy János nyert.*
 that be.3SG that(C) John won.3SG
 '(The situation is that) John won.'
- B2: *#Van, hogy János nyer.*
 be.3SG that(c) John wins
 'It happens that John wins.'

Interestingly, *úgy* 'so.DIST', a pronominal element which I have argued to behave much like ordinary expletives, can also follow the existential sense of *van*: (60) is synonymous with (59a). However, as expected if *úgy* is an expletive, no question can be formulated about it, see (61). *Olyan* cannot be questioned either, which leaves us uncertain about its status.

- (60) *Van úgy, hogy János nyert.*
 be.3SG so.DIST that(c) John wins.
 'It happens that John wins.'
- (61) **Hogy/ Milyen van?*
 how/ like.what be.3SG

What this could mean is that there are at least two separate lexical entries for *van*, one which subcategorizes for a CP (as in Kenesei's 1992/1994 proposal), and a separate one that accepts a DP subject (which could be the pronoun). Therefore, there is no "division of labor" between the clause and the pronoun: *van* in (57a) takes the clause as a subject and no expletive is needed in the structure.³⁴

Another group of predicates may occur with a DP or a CP without a change in meaning. For example *fontos* 'important' may take a subject pronoun and a clause or just a DP. Thus, it directly contrasts with *szerencse* 'luck' in (37a-b).

- (62) a. *(Az) fontos (az), hogy János nyert.*
 that important that that(c) John won.3SG
 'It's important that John won.'

³⁴ Alrenga (2005) uses a similar argumentation to distinguish between two senses of *happen* and *appear*. For example, *appear* in (ia) is synonymous with *seem*, while in (ib) it means approximately "became visible".

i) a *It appears that John will win.* b *The sun appeared on the horizon.*

- b. *Fontos János győzelme.*
 important John victory.POSS.3SG
 ‘John’s victory is important.’

There are also a lot of examples where an object pronoun (*azt* ‘that.ACC’) can be replaced with a single DP. (63)-(69) represents the capacity of subordinating verbs taking CPs and DPs alternatively, a pattern that is claimed to be ungrammatical by Kenesei (1992/1994). These examples stand in direct contrast with (37)-(40).

- (63) a. *A miniszter azt fontolgatja, hogy lemond.*
 the minister that.ACC contemplate.3SG that(C) resign.3SG
 ‘The minister is contemplating that he may resign.’
- b. *A miniszter a lemondást fontolgatja.*
 the minister the resignation.ACC contemplate.3SG
 ‘The minister is contemplating about resignation.’
- (64) a. *Azt jósolom, hogy János fog nyerni.*
 that.ACC predict.1SG that(c) John will.3SG win.INF
 ‘I predict that John will win.’
- b. *János győzelmét jósolom.*
 John victory.POSS.3SG.ACC predict.1SG
 ‘I predict John’s victory.’
- (65) a. *Valótlanul állítottuk azt, hogy jó idő lesz.*
 wrongly claimed.1PL that.ACC that(c) good weather be.FUT
 ‘We wrongly claimed that there will be good weather.’
- b. *Valótlanul állítottuk a jó időt.*
 wrongly claimed.1PL the good weather.ACC
 ‘We wrongly claimed that there will be good weather.’ Lit: ‘We wrongly claimed good weather.’
 from: http://index.hu/belfold/2015/05/27/idojaras_eforejelzes_meteorologia/
- (66) a. *Az orvos azt tanácsolja, hogy sokat mozogjak.*
 the doctor that.ACC recommend.3PL that(c) lot.ACC move.IMP.1SG
 ‘The doctor recommended that I exercise a lot.’
- b. *Az orvos sok mozgást tanácsol.*
 the doctor lot movement.ACC recommend.1SG
 ‘The doctor recommends a lot of exercise.’
- (67) a. *Furcsállom azt, hogy János vesztett.*
 find.strange.1SG that.ACC that(c) John lost.3SG
 ‘I find it strange that John had lost.’
- b. *Furcsállom János vereségét.*
 find.strange.1SG John defeat.POSS.3SG.ACC
 ‘I find John’s defeat strange.’

- (68) a. *Azt firtatta, hogy miért távoztam korán.*
 that.ACC asked.3SG that(c) why left.1SG early
 ‘S/he pumped me for why I had left early.’
- b. *A korai távozásom okát firtatta.*
 the early leave.POSS.1SG reason.POSS.3SG.ACC asked.3SG
 ‘S/he pumped me for the reason for me leaving early.’
- (69) a. *Azt sérelmezem, hogy korán távoztatok.*
 that.ACC resent.1SG that(c) early left.2PL
 ‘I resent that you had left early.’
- b. *A korai távozásotokat sérelmezem.*
 the early leave.POSS.2PL.ACC resent.1SG
 ‘I resent you leaving early.’

It is true that some of the most frequent subordinating verbs like *mond* ‘say’ and *hisz* ‘believe’ are usually not grammatical with an object that is the result of a nominalization of a clause, like in (39) and (40), but in the light of (63)-(69) that seems to be a lexical restriction on them rather than a substantive generalization.

Furthermore, even *mond* is capable of taking DP objects if these are not clause-nominalizations, but simple nouns with a propositional meaning.

- (70) *Mondott nekem egy viccet/ hírt/ történetet/ három dolgot.*
 said.3SG me.DAT one joke.ACC news.ACC story.ACC three thing.ACC
 ‘S/he told me a joke/ a piece of news/ story/ three things.’

One might argue that (70) represents another lexical entry. Although immediate semantic intuition would not support this idea, it would not be very far-fetched to hold such a claim. Actually, this is the line of argumentation that I have pursued in connection with *van* ‘be’ earlier. However, I think that a good case can be made against it in the case for *mond* ‘say’. The objects in (70) are manifestations of the same object function that can also be fulfilled by the pronoun. Evidence for this comes from coordination facts.

To understand the reasoning, let us take the English verb *believe* as an introductory example. It can occur with a thematic object as in (71a) or a nonthematic one as in (71b).³⁵ The relevant LFG-style lexical entries are shown in (71a’) and (71b’).

- (71) a. *I believe the story.* a’. *believe* <(SUBJ)(OBJ)>
 b. *I believe John to be happy.* b’. *believe* <(SUBJ)(XCOMP)>(OBJ)

Although they contain an object, the lexical entries cannot be mixed so the objects cannot be coordinated (see also Rothstein 1995: 505).

- (72) a. **I believe the story and John to be happy.*

³⁵ I do not claim that the lexical entries are entirely unrelated. What I say is that on the syntactic level of f-structure, they clearly have different subcategorizations, while the connection may be represented at the argument-structure.

Mond ‘say’ and *hisz* ‘believe’ also have a usage like *believe* in (71b), where they take a nonthematic object and a non-finite complement.

- (73) *Jánost mindenki okosnak mondja/ hiszi.*
 John.ACC everyone smart.DAT say.3SG believe.3SG
 ‘John is said/believed to be smart by everyone’ Lit.: ‘Everyone says/believes John to be smart.’

What I propose is that while the usage in (73) indeed represents a separate lexical entry (which is like (71b’)), an object like the one in (70) and a pronoun in a subordinating sentence are manifestations of the OBJ function of the same lexical entry.

Thus, I expect that the object of (70) and a pronoun-clause complex can be coordinated, but an object like the one in (73) and a single DP or a pronoun-clause complex cannot. This prediction is borne out, see (74a). Similar sentences may be construed about the other examples in (63)-(69).

- (74) a. *Tomi éppen mondta a viccet és azt, hogy hol hallotta, amikor elment az áram.*
 Tom just said.3SG the joke.ACC and that that(c) where heard.3SG when away.went the electricity
 ‘Tom was telling the joke and telling about where he had heard it, when a power outage occurred.’
- b. *Valótlanul állítottuk a jó időt és azt, hogy lehet majd kirándulni.*
 wrongly claimed.1PL the good weather.ACC and that.ACC that(c) possible then make.a.trip.INF
 ‘We wrongly claimed that there will be a good weather and that one can make a trip.’
- c. *Fontolgom a visszavonulásomat és azt, hogy ezt hamar megteszem.*
 contemplate.1SG the resignation.POSS.1SG.ACC and that.ACC that(c) this.ACC soon do.1SG
 ‘I’m contemplating about my resignation and that I do this soon.’
- d. *Furcsállom János vereségét és azt, hogy ez mintha nem is érdekelné.*
 find.strange.1SG John defeat.POSS.3SG.ACC and that.ACC that(c) this as.if not even interest.COND.3SG
 ‘I find John’s defeat and that it doesn’t seem to bother him strange.’
- e. *Sérelmezem a korai távozásotokat és azt, hogy nem is köszöntetek el.*
 resent.1SG the early leave.POSS.2PL and that.ACC that(c) not even greeted.2PL away
 ‘I resent you leaving early and that you hadn’t even said goodbye.’

On the other hand, the lexical entry illustrated in (73) cannot be coordinated with either a pronoun-clause complex or a simple DP object.

- (75) a. **Jánost okosnak és azt mondtam, hogy nyerni fog.*
 John.ACC smart.DAT and that.ACC said.1SG that(c) win.INF will.3SG
- b. **A viccet/ hírt/ történetet és Jánost okosnak mondtam.*
 the joke.ACC news.ACC story.ACC and John.ACC smart.DAT said.1SG

Hisz ‘believe’ is not really productive in contemporary Hungarian with a simple DP object. However, in an artistic/archaic style, it is capable of taking a DP object (76a-b). Occasionally, one may come across contemporary examples as well, showing that the structure is not entirely obsolete (76c). If one is willing to take these examples into consideration, *hisz* behaves exactly like *mond* in (73) and (74): the single DP object can coordinate with a pronoun-clause complex but the athematic object cannot, see (76) and (77).

- (76) a. ?*Apám hitte a szavak igazát.*
 father.POSS.1SG believed.3SG the words truth.ACC.POSS.3SG
 ‘My father believed (in) the words’ truth.
 from a Hungarian pop song, <http://www.zeneszoveg.hu/dalszoveg/173/zoran/apam-hitte-zeneszoveg.html>
- b. ?*Hiszem a római katolikus anyaszentegyházat.*
 believe.1SG the roman catholic holy.church.ACC
 ‘I believe (in) the Holy Catholic Church’
 from: the Compendium of the Catechism of the Catholic Church
http://www.vatican.va/archive/compendium_ccc/documents/archive_2005_compendium-ccc_hu.html
- c. *Hiszem Isten jóságát.*
 believe God goodness.ACC.POSS.3SG
 ‘I believe (in) God’s goodness.’
 from: interview with Géza Röhrig in *Heti Válasz* (Hungarian political magazine), 2015-June-18.
- (77) a. *Hiszem apám igazát/ Isten jóságát*
 believe.1SG father.POSS.1SG truth.ACC.POSS.3SG God goodness.ACC.POSS.3SG
és azt, hogy ez a helyes cselekedet.
 and that.ACC that(C) this the right deed
 ‘I believe my father’s truth/ God’s goodness and that this is the right deed.
- b. **Jánost okosnak és apám igazát/ Isten jóságát hittem.*

In the light of such counterexamples, I would like to argue that Kenesei’s (1992/1994) proposal, which is based on the apparent scarcity of DP complements of subordinating verbs, cannot be maintained. As the typological picture seems to weigh rather against than for the expletive-analysis of the pronouns in question, I conclude that an alternative theory, where these demonstratives (*az* and its case-marked derivatives) are referring, is to be preferred.

In É. Kiss’s (2002) version of this approach, the pronoun and the clause form a complex noun phrase. Although it is not stated explicitly, most probably É. Kiss assumes that the

pronoun-clause complex is base-generated postverbally and that the pronoun may be moved to preverbal operator positions afterwards. This framework predicts that “movement” out of subordinate clauses is ungrammatical since complex noun phrases are islands (this is Ross’s 1967 Complex Noun Phrase Constraint, CNPC). See (78a) and (79a) below.

As Rákosi (2006b) notes, what is problematic about this account is that it cannot straightforwardly account for cases when “movement” does take place, namely, when the pronoun is not explicitly present (78b and 79b). É. Kiss (2002) herself also acknowledges this.

(78) a. **János mondtad azt, hogy jön.*
John said.2SG that.ACC that(c) comes

b. *János mondtad, hogy jön.*
John said.2SG that(c) comes
'(Of) John you said that he will come.'

(79) a. **Mitől mondtad azt, hogy fél János?*
what.from said.2SG that.ACC that(c) fears John?

b. *Mitől mondtad, hogy fél János?*
what.from said.2SG that(c) fears John?
'What did you say that John fears?'

To overcome this problem, É. Kiss (2002: 253) has to stipulate “that a projection containing no phonologically realized material is transparent for subjacency. Then the noun phrase subsuming the argument clause would activate the CNPC only when its nominal head is spelled out phonologically.”

As already mentioned, this kind of data is also problematic for Kenesei’s (1992/1994) theory, and for a similar reason: both É. Kiss (2002) and Kenesei (1992/1994) would posit zero pronominal heads in (90b) and (91b). In a Chomskyan framework, this would mean that the extraction possibilities are determined only at PF (den Dikken 2010, footnote 6). Without further elaboration, such an approach is unprecedented and thus has reduced plausibility. A further burden for the expletive-approach is that expletive *pros* are theoretically problematic: a fundamental issue from a GB/MP perspective is that such elements would contribute neither semantic (LF) nor phonological (PF) data to the linguistic computation.³⁶

There is no need for such stipulations in the proposal of Tóth (2000), to which Rákosi & Laczkó (2005) also subscribe. In these frameworks, a verb like *mond* ‘say’ takes a simple DP object and the clause itself is an adjunct. (78a) follows, since adjuncts are also islands. This is attested in Hungarian as well.

(80) a. *Jöttem, hogy János lásson engem is.*
came.1SG that(c) John see. SBJV.3SG me too.

'I came so that John can see me too.'

b. **János jöttem, hogy lásson engem is.*
John came.1SG that(c) see. SBJV.3SG me too

³⁶ See Biebraurer (2008) for a detailed argumentation against expletive *pros*.

- c. **Kit jöttél, hogy lásson János?*
 whom came.2SG that(c) see.SBJV.3SG John
 intended: ‘Whom did you come so that John can see?’

The verb *jön* ‘come’ in (80a) does not subcategorize for a proposition at any level so the clause must be an adjunct, expressing purpose. If we try to extract the subject of the clause, the sentence becomes ungrammatical as “movement” out of adjuncts is impossible. The same explanation goes for (78a) and (79a).

As for (78b) and (78b), I assume along with Tóth (2000) and Rákosi & Laczkó (2005) that the clause itself functions as the object of the predicate, and as such, it is an argument, from which extraction can take place.

Let us conclude this section with some remarks about the place of Hungarian subordinating constructions in the cross-linguistic palette. The idea that a propositional theta-role may be assigned to a pronoun is not unprecedented in the literature. In Dutch, Hoekstra (1983) and Bennis (1986) argued that *het*, which had often been taken to be an expletive, is in fact a referring pronoun. *Het* occurs as a subject of weather-verbs (81a) or as subject/object of some verbs taking propositional complements (81b-c). This second use of *het* is directly comparable to the Hungarian situation.

- (81) a. *Het regent.*
 it rains
 ‘It rains.’
- b. *Het wordt betreurd dat Jan ziek is.*
 it is regretted that(c) John ill is
 intended: ‘It is regretted that John is ill.’
- c. *Jan betreurde het dat hij ziek was.*
 John regretted it that(c) he ill was.
 ‘John regretted (it) that he had been sick.’

Just like in the case of Hungarian associate pronouns, *het* in (81b-c) is optional. However, when it is present, extraction is impossible from the subordinate clause, see (82b) and (82d). Hoekstra’s (1983) and Bennis’s (1986) explanation for this is the same as my explanation was for (78)-(79): when the pronoun is present, it is the argument of the main predicate and the clause itself is an adjunct and thus a syntactic island.

- (82) a. *Wat wordt door iedereen betreurd dat Jan gelezen heeft?*
 what is by everyone regretted that(c) John read has
 intended: ‘What is regretted by everyone that John read?’
- b. **Wat wordt het betreurd dat Jan gelezen heeft?*
 what is it regretted that(c) John read has
- c. *Wat betreurde jij dat hij gezegd had?*
 What regretted you that(c) he said had
 intended: ‘What do you regret that he has said?’

- d. **Wat betreurde jij het dat hij gezegd had?*
 what regretted you it that(c) he said had

The situation is similar in German. Berman (2001) analyzes the pronoun in (83a) as being the object argument of *sagen* ('say'). Just like in Hungarian and Dutch, the presence of the pronoun is optional, but when it is present, it blocks extraction out of the embedded clause (83b).

- (83) a. *weil er (es) gesagt hat, dass Hans krank ist*
 because he it said have that(c) Hans ill is
 'because he said that Hans is ill.'
- b. *Was hat er (*es) gesagt, dass er gelesen hat?*
 what has he it said that(c) he read has
 'What did he say that he read?'

Finally, the object pronouns in (96) which were claimed to be expletives by Postal and Pullum (1988) but are argued to be real pronouns by Rothstein (1995) also show this pattern.

- (84) a. *What do you believe (*it) that John will do?*
 b. *A full compensation, I strongly demand (*it) that I get.*

The emerging pattern may also be related to research done in the area of CP-licensing. Since Stowell (1981) it has been noted that there are two basic ways of integrating sentential arguments:³⁷ as direct arguments or as appositions to a nominal head (see also Synder 1992). These patterns seem to be a subtype of this latter relation, although the exact relationship between adjunction and apposition is yet to be investigated. What seems to be unique about the Hungarian pattern is that the apposition/adjunction configuration occurs across the board, with a wide range of verbs of saying and cognition, while in other languages this is much more limited.

2.3 Conclusion to chapter 2

Let us present the overall picture that emerges out of this view of subordinate clauses. Some predicates realize their propositional argument exclusively with a CP, as shown in in (85).

- (85) *(*Az) Muszáj, *(hogy) elgyere.* *muszáj <(SUBJ)>*
 that must that(c) come. SBJV.2SG
 'You must come.'

Other predicates can take either a nominal or a clause as a realization of their propositional argument. Functionally, these may be either subjects, objects or obliques. Categorially, a nominal may be realized as a lexical noun (86) or a demonstrative pronoun (87). By default, a

³⁷ This may be elaborated in terms of CP-licensing or the licensing of propositional arguments. In a Minimalist framework a null C head is hypothesized even in the absence of an overt head. In LFG, in accordance with the Economy of Expression, such cases are treated as bare IPs or exocentric S heads.

clause may be either a CP (if the complementizer is present) or an S (if the complementizer is absent). If the propositional argument is realized as a demonstrative, the clause is an adjunct. If the clause is present but a nominal is not, then the clause is the SUBJ/OBJ/OBL_θ argument of the predicate (88).

- (86) a. *A vereséged nagy kár számunkra. kár <(SUBJ)>*
 the defeat.POSS.2SG great pity for.us
 ‘Your loss is a great pity for us.’
- b. *A lemondásomat fontolgom. fontolgat <(SUBJ)(OBJ)>*
 the resignation.ACC.POSS.1SG contemplate.1SG
 ‘I’m contemplating my resignation.’
- c. *A lemondásomon gondolkodom. gondolkodik<(SUBJ)(OBL_θ)>*
 the resignation.POSS.1SG.on thinking.1SG
 ‘I’m thinking about my resignation.’
- (87) a. *Az kár, hogy vereséget szenvedtél.*
 that pity that(c) defeat.ACC suffered.2SG
 ‘It’s a pity that you have been defeated.’
- b. *Azt fontolgom, (hogy) lemondok.*
 that.ACC contemplate.1SG that(c) resign.1SG
 ‘I’m contemplating (the issue) that I may resign.’
- c. *Azon gondolkodom, (hogy) lemondok.*
 that.on thinking.2SG that(c) resign.1SG
 ‘I’m thinking about (the issue) that I may resign.’
- (88) a. *Kár, hogy vereséget szenvedtél.*
 b. *Fontolgom, hogy lemondok.*
 c. *Gondolkodom, hogy lemondok.*³⁸

To round up this chapter, let us present the simplified c- and f-structural analyses of the (86b), (87b) and (88b), typical representatives of the subject matter of this chapter.

³⁸ It is possible that the CP is an adjunct in this case, as *gondolkodik* (‘think’) may be used intransitively. Similar verbs seem to require the oblique arguments to be nominals (see the examples below). This suggests that the realization of the OBL_θ in these cases is restricted to casemarked NPs or DPs.

- (i) **(Attól) tartok, hogy nem én fogok nyerni.*
 that.from be.afraid.1SG that(C) not I will.3SG win.INF
 ‘I am afraid that I may not win.’
- (ii) **(Azzal) számolok, hogy én fogok nyerni.*
 that.with count.1SG that(C) I will.3SG win.INF
 ‘I expect that I will win.’

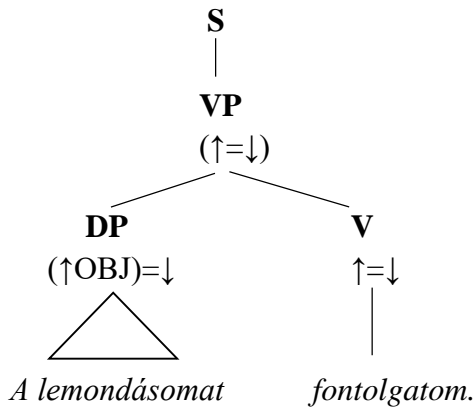


Figure 3a.
C-structure of (86b)

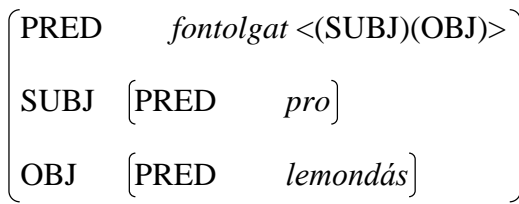


Figure 3b.
F-structure of (86b)

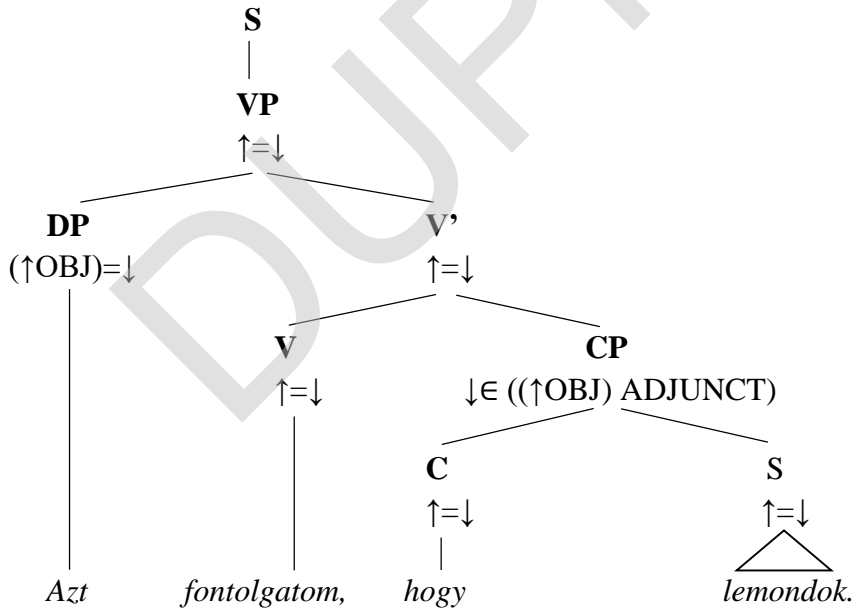


Figure 4a.
C-structure for (87b)

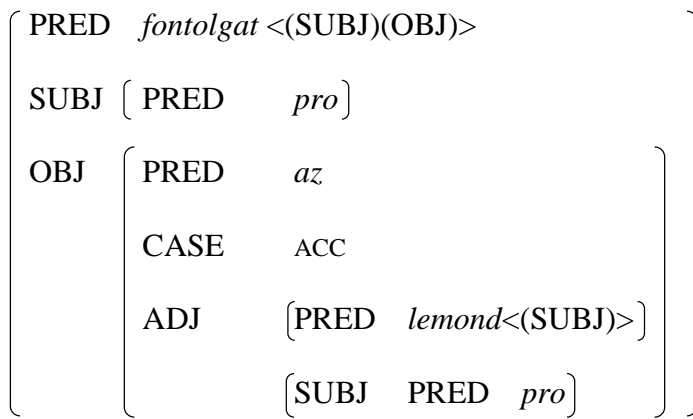


Figure 4b.

F-structure of (87b)

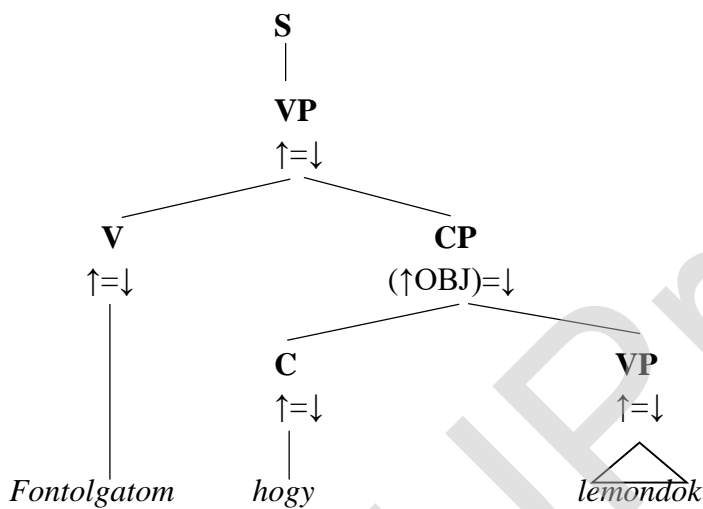


Figure 5a.

C-structure for (88b).

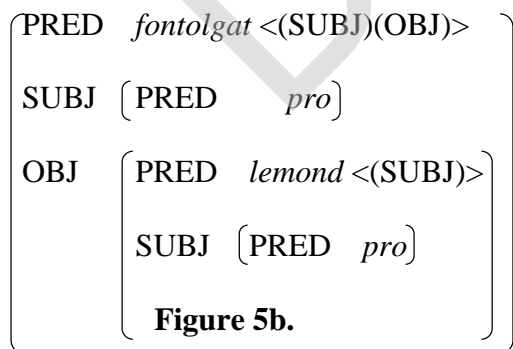


Figure 5b.

F-structure of (88b)

In this chapter I have provided a general overview of Hungarian sentence-structure. I have taken the work of Laczkó (2014a, 2014b, 2015) as a baseline and elaborated it in two areas. First, information structural considerations were added, utilizing the framework elaborated in Chapter 1, section 2.6. Second, I also investigated the structures of subordinate clauses. I have shown

how the arguments of these verbs are mapped to grammatical functions and categories. I have argued that the optional *az* ('that') pronoun and its case-marked and deictic variants in Hungarian sentences with subordinate clauses should receive an analysis where they are contentful demonstrative pronouns, and not expletives. This approach was shown to be preferable on theoretical, empirical and cross-linguistic grounds as well.

DUPress

CHAPTER 3

CONSTRUCTIONS ON THE ENGLISH LEFT PERIPHERY

Although English is considered to be a syntactically configurational language, it does not mean that the left-periphery of clauses is inactive for discourse purposes. The three structures that we are going to investigate are topicalization (TOP), clause-initial adjuncts (CIADJs) and left dislocation (LD). (1) provides an illustration of each type.

- (1) a. *John, I like.* → TOP
b. *Yesterday, John met Mary.* → CIADJ
c. *John_i, I like him_i.* → LD

In each of these cases, some element is placed at the left-periphery of the clause and they bear some sort of discourse prominence. The syntactic status of the clause-initial element is different in each case. In (1a), an argument (an object) is fronted. In (1b), the fronted element is an adjunct. In (1c), the syntactic status of the fronted element itself is not straightforward. It is coreferent with an object pronoun, which is indicated by the indexing.

In this chapter, I will investigate these left-peripheral constructions. In each subchapter first I will show that the constructions each have distinct syntactic and information-structural properties and then I will propose LFG-based analyses for them.

3.1 Topicalization in English

The term “topicalization” originates in Ross (1967). Here it is defined as to those constructions in English where semantically obligatory elements, such as an argument or a predicate is located at a clause-peripheral position on the left edge, like *John* in (1a) above. (Clause-initial adjuncts will be discussed in the next section.)

3.1.1 Properties of topicalization

(1a) has already illustrated topicalization with a direct object NP. However, topicalization is not restricted to this functional-categorial constellation. As (2) shows it can occur with a range of functions and grammatical categories.

- (2) a. *To John, I gave an apple.* (OBL_θ, PP)
b. *That John would win, I would have never said.* (COMP, CP)
c. *To arrive in time, John at least tried.* (XCOMP, IP/CP³⁹)
d. *Surrender, John never will.* (PRED, VP)

³⁹ In the Chomskyan tradition, infinitival *to* is standardly assumed to be heading an IP (or more recently a TP). LFG is less committed on this issue. Falk (2001: 139-140) argues that *to* is a complementizer for non-finite clauses.

- e. *Happy, John will never be.* (PRED, AP)

Topicalization may be rightfully called a long-distance dependency, as the canonical position of the fronted element can be embedded at arbitrary depth. (Stylistic and processing factors may put a limit to this, but not syntax *per se*.)

- (3) a. *John, Mary said that she saw.*
b. *John, Mary said that John believed that she saw.*

Although the depth of the canonical position is unlimited, the path to the embedded position can be constrained by syntactic factors. Dalrymple (2001: 392-394) discusses some of these (the examples in 4-6 are hers). One such constraint is that an intervening non-bridge⁴⁰ predicate can disrupt the dependency.

- (4) **John, Mary whispered that she saw.*

Furthermore, Ross's (1967) island constraints also affect topicalization. (5) illustrates this with a complex noun phrase-island (5a), a subject-island (5b), an adjunct-island (5c) and a wh-island (5d).

- (5) a. **John, I doubt the claim that Mary saw.*
b. **John, that you saw surprised me.*
c. **John, we think that David laughed when we selected.*
d. **John, I can't guess why you like.*

It must be noted that in contrast with (5c) TOP may target an argument inside an untensed adjunct. This is shown in (6), where the fronted phrase is the prepositional object of the locative adjunct.

- (6) *?That room, Chris teaches his classes in.*

Facts like this indicate that the topicalized constituent maintains a strong relationship with its canonical position. The presence of the topicalized constituent can also be detected with its interaction with binding theory:

- (7) a. *Himself_i, John_i likes.*
b. **Him_i, John_i likes.*
c. *His_i mother, John_i likes.*
d. **John_i, he_i likes.*

⁴⁰ Traditionally, a "bridge verb" is one that allows extraction from a complement clause. So for example, while *say* and *think* are bridge verbs, *whisper* and *yell* are not.

- (i) *Who did you say/think/*whisper/*yelled Bill saw?*

(7) is classic binding-theoretic data. While the technical implementation of these generalizations is different in different frameworks (Minimalism relies on tree-structures, while LFG mainly utilizes f-structures), the basic idea is the same: a reflexive must be bound in its binding domain, a pronoun must be free in its binding domain and an R-expression must be free everywhere. The pattern in (7) is clear if it is assumed that the topicalized phrases in some sense are also tied to their canonical positions. (For a detailed discussion on binding theory in LFG see Dalrymple 1993, for an introduction, see Falk 2001, chapter 7).

Deploying multiple TOP-Engs in a sentence causes ungrammaticality.⁴¹ Combining them with questions leads to similar results.

- (8) a. **An apple, John, I gave.*
- b. **An apple, whom did you give?*
- c. **Whom, an apple did you give?*
- d. **Did an apple, you give John?*
- e. **An apple, did you give John?*

Let us now turn to the information structural properties of TOP-Eng. For a start, let us review the IS-taxonomy that I have argued for in 1.3. Given this, a precise formulation of the information structural properties of TOP-Eng will be possible.

		+NEW	-NEW
+D-STRUCTURING	+CONTRASTIVE	contrastive focus	contrastive topic
	(-CONTRASTIVE)	information focus	topic
-D-STRUCTURING		completive Information	background information

Table 1.

My proposed information structural taxonomy

As pointed out by Prince (1999), despite its name, the construction is not about standard, neutral topics. Instead, the topicalized constituent may actually have two distinct functions: it can be interpreted as some kind of focus or a topic of a particular kind. What I would like to argue for is that TOP-Eng is a structure that is tied to the +CONTRASTIVE feature, so topicalized elements are interpreted either as Contrastive focus or Contrastive topic: the set of alternatives becomes active in the discourse at the point the sentence containing the contrastive element is uttered. This position is not unprecedented (see e.g. Chafe 1976 or Molnár & Winkler 2010), but the concise summary from this perspective has not been provided for it in the literature.

Regarding the focus-like reading, Choi (1997), referring to Ward (1988), asserts that the fronted phrase actually refers to two discourse elements: one, a set or a scale, and two, a

⁴¹ There has been some debate in the literature about the grammaticality of multiple topicalizations. For a discussion about such cases, see Breul (2004: 199-205).

specification of a value or an element in that set or scale. In (9a) this would mean that the sentence evokes a set of dog names that one may choose and picks *Fido* from that set. The evoked alternative set is lacking in the non-preposed version (9b).

- (9) a. *Fido I named my dog.*
 b. *I named my dog Fido.*

Zimmerman (2008) notes that for a wh-question, a topicalized answer is only acceptable if it is against the expectations of the hearer, not as a neutral answer (10/B2). Being unexpected involves that there is some other entity to which the expectedness is compared. So being unexpected may be interpreted as being a surprising pick from a set of alternatives.

- (10) A: *What did you eat in Russia?*
 B1: *Caviar.*
 B2: *#Caviar we ate!*

A frequent context for this use of TOP-Eng is corrections, which is obviously contrastive in the sense that the speaker rejects a wrong answer and at the same time provides a correct one. Thus multiple discourse questions are involved.

- (11) A: *What did you eat in Russia? Pelmeni?*
 B: *No, caviar we ate!*

In the discourse-tree representation I adopted in 1.3, (11) may be represented as Figure 1.

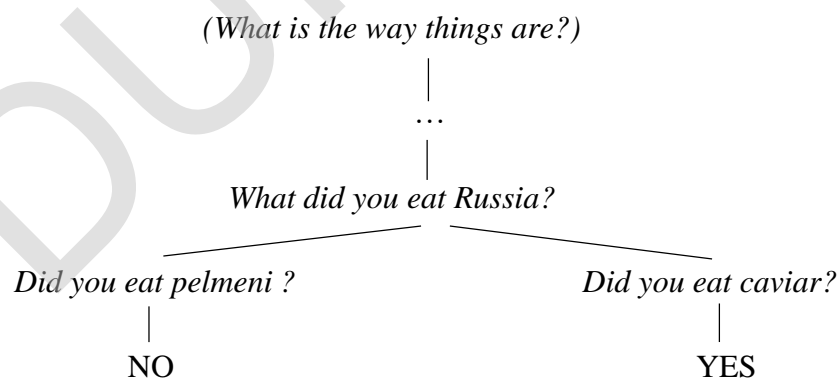


Figure 1.
 Discourse-tree for (11)

Birner & Ward (1998) mention further contexts where a TOP-Eng structure is used. For example, it is used in “proposition assessment”, that is, when the speaker expresses some emphatic stance with regards to the uttered proposition. The assessment may be affirmation (12a), suspension (12b) or denial (12c).

- (12) a. *At the end of the term I took my first schools; it was necessary to pass, if I was to stay at Oxford, and pass I did.*

- b. *Mark submitted his report late, if submit it he did.*
- c. *The NBA's new collective-bargaining agreement sounds as though it was written by the same people who put together the Internal Revenue Service's long form. Simple it is not.*

In each of these cases the proposition is contrasted to some alternative state of affairs: (12a) evokes the scenario of not passing, (12b) the scenario of not submitting the report and (12c) a number of other characteristics that the agreement could be assessed with. These interpretations are actually closer to the CT interpretation as there is no clear indication of the truth-status of the evoked subquestions. For example, (12c) may be represented with the discourse-tree shown in Figure 2.

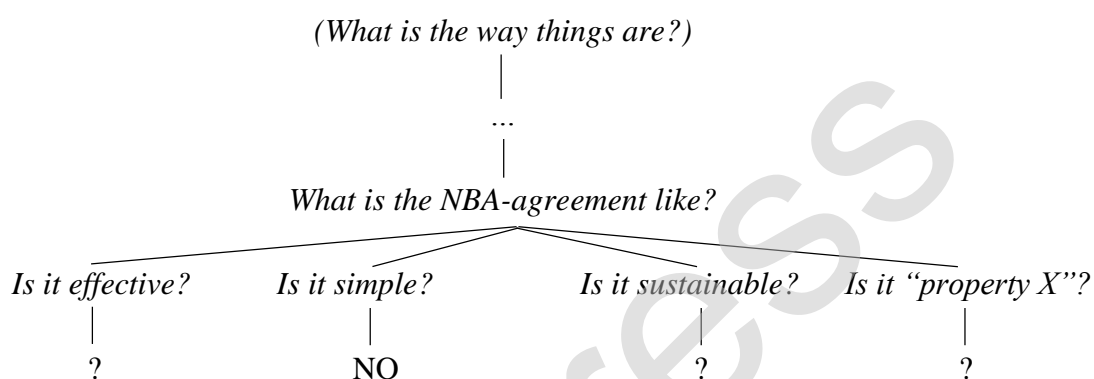


Figure 2.
Discourse-tree for (12c)

It may happen that the evoked alternative is not an opposing situation, but a lesser version of the event denoted by the predicate. Birner & Ward (1998) label this as “scalar affirmation”. This is an indication that the event denoted by the predicate happens to a large degree (it occupies a high position on a hypothetical scale which measures the intensity/frequency of the event). One example that Birner & Ward (1998) mention is (13), which expresses that riding took place in some spectacular manner.

- (13) *Asked what he thought about during today's race on a sultry day, [Tour de France winner Greg LeMond] said: "I didn't think. I just rode." Ride he did.*

I must add that Birner & Ward (1998) actually argue against the contrastive nature of topicalization. However, they point their refutation against Chafe's (1976) conception of contrast, which is “an assertion on the part of a speaker that one of ‘a limited number of candidates’ is ‘correct’”. Birner & Ward (1998) point out that there are many instances of TOP, where an interpretation resting on a “correct selection” is implausible. However, my notion of contrast is less specific than Chafe's, as mine involves the evocation of discourse-salient alternatives (which are modelled as being linked to subquestions). Thus, it is possible for me to subscribe to the contrastive interpretation of TOP, even in the light of Birner & Ward's (1998) criticism towards Chafe (1976).

The claim that topicalization marks contrastive categories also sheds some light on why it can be used with nonreferential expressions, as demonstrated in (2d-e). As discussed in 1.3.2,

the referentiality requirements on Contrastive topics are much looser than the ones on regular topics.

In conclusion, Topicalization behaves as a regular long-distance dependency, marking contrastive information structural categories.

3.1.2 Analysis of topicalization

Let us summarize the properties of topicalization:

- the fronted element may be an argument or a predicate
- the functional and categorial status of the fronted element is flexible
- the fronted element maintains strong ties to its canonical position (seen in island- and binding-data)
- the fronted element is interpreted contrastively (CT or CF)
- at most one instance of TOP is possible

The analysis I advocate is based on Dalrymple (2001, Chapter 14). Since topicalized entities follow complementizers and precede subjects in complex sentences as (14) attests, a natural place for them is an IP-adjoined position.

(14) *John said that Mary, he likes.*

TOP's sensitivity to islands and its "reconstruction" with respect to binding proves that a strong link with the canonical position should be maintained. This strong link has been modelled as movement in Chomskyan frameworks: the topicalized phrase is base-generated clause-internally and it undergoes A-bar movement to some left-peripheral position. LFG's functional identification serves essentially the same purpose. Functional identification immediately accounts for the binding data, as the fronted element is fully present in the embedded f-structure as well.

Based on the data and such considerations, the following ID-rules are needed for TOP:

(15)
$$\begin{array}{l} \text{IP} \rightarrow \quad \text{XP} \quad \text{IP} \\ \downarrow \in (\uparrow \text{UDF}) \\ (\uparrow \text{TOPPATH}) = \downarrow \\ \downarrow_i [\text{D-STR} = +] \\ \downarrow_i [\text{CONTRASTIVE} = +] \end{array}$$

A couple of notes are in order regarding (15). The XP node stands for the variety of grammatical categories that may participate in TOP-Eng (NP, DP, VP, AP, PP, AdvP). TOPPATH is a shorthand for the possible path of identification. Dalrymple (2001:396) defines it as follows:

(16)
$$\begin{array}{l} \text{TOPPATH} \equiv \{ \text{XCOMP} \mid \text{COMP} \mid \text{OBJ} \}^* \quad \{ (\text{ADJ} \in) \quad (\text{GF}) \mid (\text{GF}) \} \\ (\rightarrow \text{LDD} \neq -) \quad (\rightarrow \text{TENSE}) \quad \neg(\rightarrow \text{TENSE}) \end{array}$$

In the words of Dalrymple (2001: 396), “this expression allows the within-clause grammatical function of the UDF to be arbitrarily deeply embedded inside any number of properly constrained XCOMP, COMP, or OBJ functions, and optionally to appear as an untensed member of the ADJ set of such a function, or as an argument of the ADJ.” So this specification covers all cases of argument and predicate fronting, as well as those cases of clause-initial adjuncts where the adjunct is interpreted contrastively (TOP with an adjunct). Note that the first part of the equation ($\{XCOMP \mid COMP \mid OBJ\}^*$) does not include the SUBJ function, as extraction out of subjects is generally ungrammatical.

(17) **Mary, that John saw is true.* (c.f. *That John saw Mary is true.*)

In (16) we find a number of constraints about the possible path to the original position of the topicalized element. This is the TOPPATH in (16). The $(\rightarrow LDD \neq -)$ notation is meant to capture the fact that nonbridge verbs block the association of the topicalized phrase with its canonical position. It should be decoded as: “the path does not contain an element which is negatively specified for the LDD (long-distance dependency) feature” (nonbridge verbs are negatively specified in this respect as a part of their lexical entry). The $(\rightarrow TENSE)$ notation means that the OBJ function that the TOPPATH goes through must have a tense value. The function of this is to ensure that those verbs that realize their CP as an OBJ can also participate in TOP. This is needed because CPs may actually bear either the COMP function or they may be OBJs, as mentioned in section 1.2 (see also Szűcs 2018 and references therein). Such objects have a tense value. They contrast with objects acting as heads of complex noun phrases, which lack such a value. These are islands and cannot participate in TOP. Finally, $\neg(\rightarrow TENSE)$ makes sure that the topicalized entity can appear as a grammatical function only inside an untensed ADJUNCT function (recall the contrast between 5c and 6). These constraints block island-violations. The annotation about information structure ensures that the topicalized phrase is interpreted as one of the contrastive categories, CT or CF.

As noted earlier, the arrows in the annotations are “metavariables”. \uparrow refers to the mother-node’s functional structure, while \downarrow refers to the node’s f-structure. So $\downarrow \in (\uparrow UDF)$ in (15) should be read as “this node is a UDF in the functional structure of the IP”. As an example, Figure 1a and 1b show the constituent- and functional structure of a sentence like (1a) (*John, I like*). For brevity, I only annotate the parts of the c-structure that are directly relevant for the current issues.

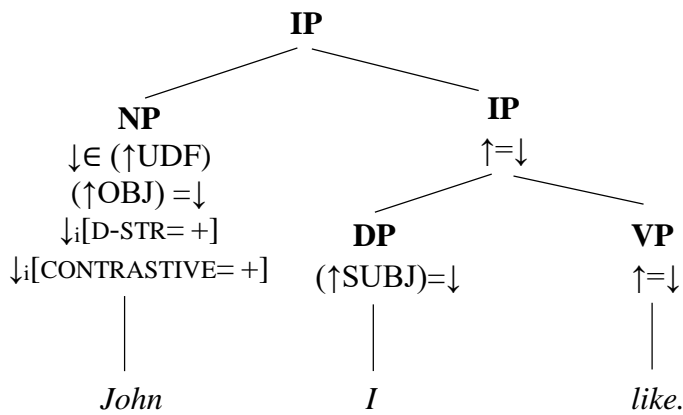


Figure 1a.

Constituent-structure of (1a)

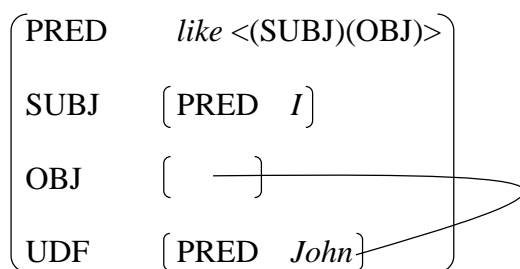


Figure 1b.

Functional structure of (1a)

What is not clear at this point is what causes the uniqueness of this IP-adjoined position. That only one functionally identified element may be IP-adjoined seems to be part of some wider conspiracy with parametric variation in languages, see Engels (2012: 171-176), who gives an Optimality-theoretic⁴² account of the data. Such ambitions are beyond the scope of this monograph so I leave this aspect of the research for future investigations.

3.2 Clause-initial adjuncts in English

While TOP-Eng may involve adjuncts, in this section I will show that clause-initial adjunct of the kind shown in (1b) may participate in a different construction as well. A deeper investigation to be presented here will reveal several crucial differences, as regards both syntax and information structure.

It should be taken into consideration that the class of adjuncts is huge, and they do not behave in a homogenous manner. I will restrict this discussion to those aspects of such adjuncts that are directly relevant for the purposes of this monograph and put many fine details aside. For such details, the interested reader is referred to the references mentioned in this section.

⁴² Prince & Smolensky (1993).

3.2.1 Properties of clause-initial adjuncts

From a syntactic point of view, the most important distinction that we have to make is that adjuncts differ according to what we might call as their canonical position. While certain adjuncts are “base-generated” at the left periphery and have no ties to any clause-internal position, other adjuncts maintain ties to such positions (Frey 2003). This is mirrored in what part of the event expressed by the sentence these adjuncts modify. Following Frey (2003), I will refer to the first type of adjuncts as “frame-” and “event external adjuncts”, while I will call the second type “event-internal adjuncts”. As an illustration, consider (18) and (19).

- (18) a. *In New York, there is always something to do.*
b. *These days it rains a lot.*
- (19) a. *In the box, John found a hammer.*
b. *With his metal detector, John found some interesting items in the garden.*

In (18a), *in New York* and in (18b), *these days* provide a spatial/temporal frame for the event. That is, they do not modify the contents of the event but restrict the domain for which the claim holds. On the other hand, *in the box* in (19a) specifies the place of the hammer, *with a metal detector* in (19b) specifies the tool of the process. These provide more information about some internal part of the event.

(18) and (19) show that the event external-event internal distinction is not based on grammatical category or semantic type (as both *in New York* and *in the box* are locative PPs), but on interpretation. As we will see, this distinction has consequences for the syntax and information structure of these adjuncts. Event-internal adjuncts show syntactic properties that liken them to TOP-Eng: they are sensitive to island-constraints (20a) and show principle C effects (20b).

- (20) a. **In the box, we heard the claim that John had found a hammer.*
b. **In Beni's office, he_i lay on the desk.*

Haegeman (2003) and Frey (2003) also demonstrates that CIADJs are interpreted at a distance (“long adverbial fronting” in Haegeman’s words) pattern with TOP (argument fronting) with respect to a number of syntactic phenomena. For instance, an event-external adjunct may alleviate a “that-trace” effect (21a), but an event-internal (21b) may not (21c is a topicalization example, for comparison).

- (21) a. *Who do you think that yesterday had a great time?*
b. **Who do you think that in the box found a hammer?*
c. **Who_i do you think that, John, ____i likes? (TOP)*

These CIADJs also seem to follow TOP with respect to their distribution: only one may be present and they cannot co-occur with TOP, as in (22), see also Engels (2012: 172).

- (22) a. **In the box, with a flashlight John found a hammer.*
 b. **[In the box/With a flashlight], a hammer *[in the box/with a flashlight], John found.*

If we look at frame/event external adjuncts from these perspectives, we get a robustly different picture. In (23) we find no principle C violation (20b and 23 are from Frey 2003).

- (23) *On Beni's birthday, he_i took it easy.*

Unlike an event-internal CIADJ, these adjuncts may co-occur with TOP, either preceding or following it (24).

- (24) a. *Probably, Linda, you have met.*
 b. *Linda, probably, you have met.*

Regarding information structure, we first should note that any clause-initial adjunct may, but does not have to be interpreted contrastively. Apart from some cases, CIADJs are not necessarily interpreted as contrastive. Accordingly, sentences in (18) and (19) may be uttered without evoking alternatives in the discourse. So (18a) does not necessarily suggest that there are other places where one cannot do anything and (19a) does not necessarily implicate that there is another container where John found something other than a hammer. So CIADJs may function as neutral topics, and this is a critical difference in comparison with TOP.

It should be added however that there are CIADJs which only allow for the contrastive reading. There are some adjunct-types that highly disprefer a neutral topic interpretation. For instance, manner or measure adjuncts are often ungrammatical clause-initially. Example (25a) is from Ernst (2002). The sentence improves if the adverb receives “strong contrastive stress” (Ernst 2002: 470), that is, if it is an instance of TOP with an adjunct (25b).

- (25) a. **Tightly, she must hold on to the railing.*
 b. *TIGHTLY she must hold on to the railing.*

I think the reason for this must be semantico-pragmatic rather than syntactic in nature. In particular, following Shaer (2004: 388), we can assume that “the degraded acceptability of a sentence containing a fronted adverbial may have its source in the difficulty of inferring the relation of the adverbial not only to its host sentence but also to previous discourse.” That is to say, although the CIADJ-construction would assign a topic interpretation to these adjuncts, they cannot assume this role because they are non-referential. The +CONTRASTIVE feature eases the referentiality requirement, as noted earlier.

Returning to referential CIADJs like (18) and (19), I follow Maienborn (2001) and Ernst (2002: 399-402) in that these cases of CIADJs are topic-like entities. In (18a) *in New York*, as a frame-setter introduces a discourse topic in a general discussion about cities. Similarly, *these days* in (18b) introduces a temporal frame. (19a-b) could be imagined in the context of a story, so the box or the metal detector are most likely the links to the previous discourse. Accordingly, it is plausibly viewed as an entity introducing a new subtopic. It follows that the context should be one in which the hearer can accommodate such a new subtopic, like (26a). (26a) also

illustrates that the CIADJ is not necessarily contrastive, as no tension can be detected when no alternative is provided. If the context is not one in which the *box* is a plausible link to the discourse, the sentence is weird. In (26b), even though it is plausible that one may find a box in a cellar, it is still hard to establish the *box* as a plausible subtopic (even if the definite article is replaced with an indefinite one to neutralize the identifiability-effect of *the*).

- (26) a. *John found several containers in the cellar: a box, a bag and a jar. He wanted to see what's inside. In the box, he found a hammer. (...) In the bag, there was another hammer.*
- b. *John went into the cellar. #In a box, he found a hammer.*

Additionally, just like the Hungarian “topic-position” (as in Figure 1 in section 2.1), sentential adverbs may also occupy this slot, as the sentential adverb *probably* does so in (24). Following the remarks made in 2.1, I classify such cases as complete or background information.

3.2.2 Analysis of clause-initial adjuncts

As the contrastive uses of clause-initial adjuncts are covered by the analysis provided for TOP, this section is only concerned with non-contrastive CIADJs. It has been established earlier that CIADJs may be divided into at least two subcategories as far as their syntax is concerned. Frame-setters/event external adjuncts fully belong to the left periphery, while event internal adjuncts maintain ties to their clause-internal interpretational site.

Let us first approach the event-internal type. An alternative IP-adjoined position with annotations is provided for the event-internal type in (27) and (28).

- (27) IP → AdvP/PP IP
 $\downarrow \in (\uparrow \text{UDF})$
 $\downarrow \in (\uparrow \text{ADJPATH})$
 $\downarrow_i [\text{NEW} = -]$
 $\downarrow_i [\text{D-STR} = +]$

- (28) ADJPATH $\equiv \{ \text{XCOMP} \mid \text{COMP} \mid \text{OBJ} \}^* \text{ADJ}$
 $(\rightarrow \text{LDD} \neq -) \quad (\rightarrow \text{TENSE})$

Here the phrase-structural category is restricted to AdvP and PP, in order to prevent inappropriate categories from being parsed as adjuncts (e.g. nominals in LD). The ADJPATH is a modified version of the TOPPATH: it encodes similar restrictions (to capture the similar syntactic behavior of TOP and event-internal CIADJs) but is restricted to adjuncts. So the fronted adjunct may be embedded at an arbitrary depth, through XCOMPs, COMPs and tensed OBJ functions.

The one major modification is that there is no disjunction in the second half of the notation. This is because although extracting an argument out of an adjunct is possible (as in 6), extracting an adjunct from an adjunct is ungrammatical. (29 is to be construed as *on the third floor* being an adjunct for *that room*.)

(29) **On the third floor, Chris teaches his classes in that room.*

If this configuration is not possible for the adjunct for distance or semantico-pragmatic reasons (as in 25a), they may be analyzed using the TOP annotation, introduced earlier, where the +CONTRASTIVE feature makes non-referential entities acceptable.

Another difference is that these CIADJs are specified at information structure as having the feature-specification of topics: –NEW, +D-STRUCTURING.

Frame setter/event external adjuncts behave differently: as we saw earlier, they do not “reconstruct”, see the data in (23). This suggests that the integration of such adjuncts into the sentence is generally looser. According to Minimalist approaches, such entities are “base-generated” in some high position of the syntactic tree. This high position could be some CP- or IP-adjoined position, or in a designated TopicP, depending on the details of the analysis.

There is also an alternative approach. Shaer (2004), building on the ideas of Haegeman (1991), claims that such clause-initial adjuncts are “syntactic orphans”: they are not integrated into the phrase structure. In other words, they are not proper (syntactic) parts of the sentence.

One of the main arguments of Shaer (2004) is that certain polarity items like *ever* in (29a) are not permitted even if a supposedly base-generated CIADJ would in principle license them (*ever* needs to be under the scope of *only*). If the CIADJ is undoubtedly part of the sentence, as in (29b), the polarity item is grammatical. One might conclude that inversion itself is the critical factor in the licensing of polarity items like *ever*, but heavy contrastive stress (as in Topicalization) may also mitigate ungrammaticality to a certain extent, see (30c).

- (30) a. **Only in New York, John could ever have fun.*
 b. *Only in New York could John ever have fun.*
 c. ?*ONLY IN NEW YORK, John could ever have fun.*

However, Engels (2012), referencing Haegeman (1995) points out that the facts in (30) may have alternative explanations, e.g. the non-projection of the relevant feature of *only* outside the PP. Whatever the exact formulation of the constraints is, the complete exclusion of these adjuncts from the f-structure of the sentences would make their semantic association with the sentence problematic. Additionally, we shall see in the next section, that LD-Eng has certain properties which make the orphan-analysis plausible for it but these properties are crucially different with respect to CIADJs. Thus, it is motivated to side with the more conservative approach where the CIADJs are proper parts of the sentence.

In light of the discussion, I propose that the following annotation is possible for such CIADJs.

- (31) IP → AdvP/PP IP
 ↓ ∈ (↑ ADJ)
 { ↓_i [NEW=–]
 ↓_i [D-STR=+] |
 (↓ ADV-TYPE)=_c SENT
 ↓_i [D-STR=–] }

These adjuncts are not extracted: they modify at the sentence-level so they are not identified with any clause-internal function. At information structure, they are either topics (in the case of referential expressions) or completive/background information, in the case of sentence-adverbs (*probably, luckily, etc.*).

The analysis of a sample sentence with a frame-setting and an event-internal adjunct is shown below.⁴³

(32) *Yesterday, for an hour I was looking for a taxi.*

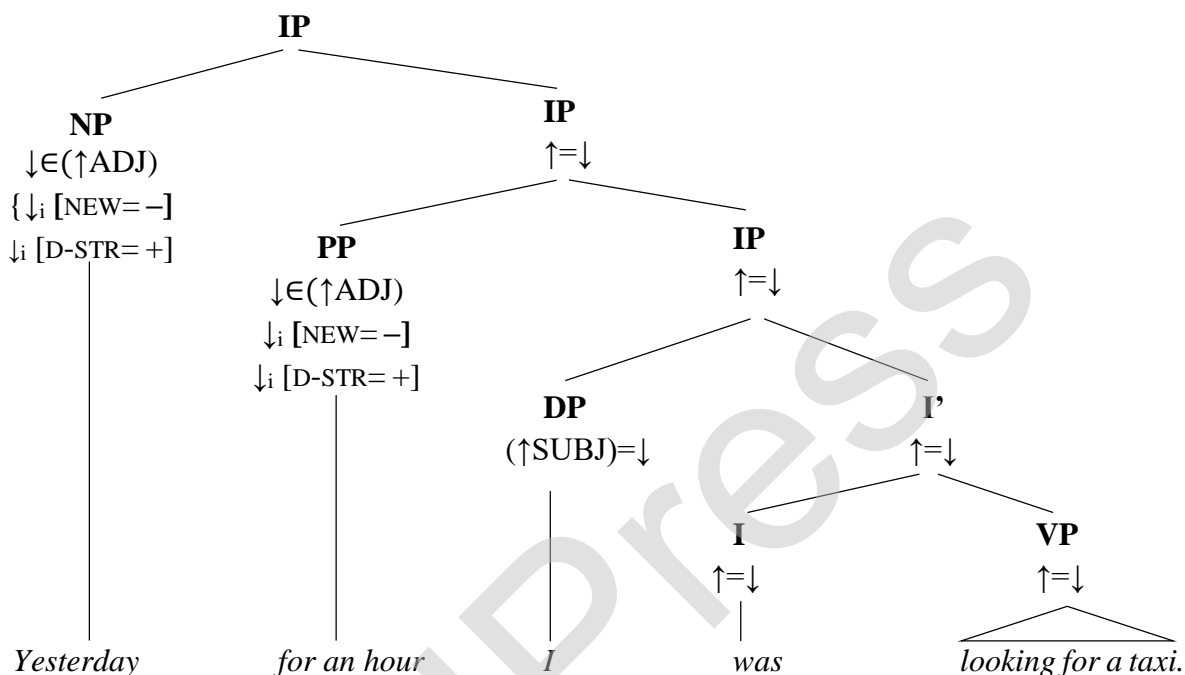


Figure 2a.

Constituent-structure of (32)

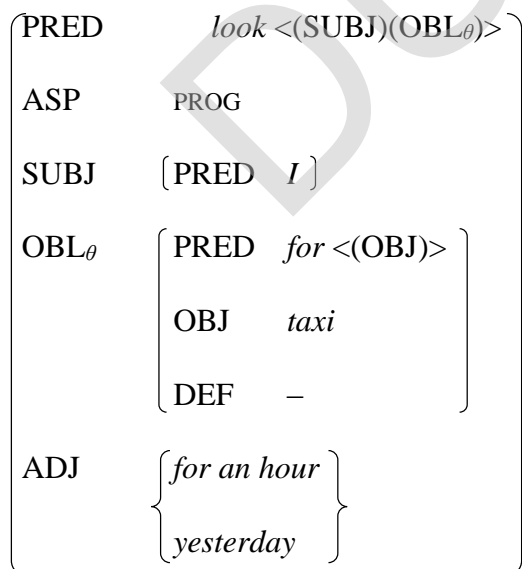


Figure 2b. Functional-structure of (32)

⁴³ Although I take the ordering of adjuncts to be free as far as syntax is concerned, other factors such as semantic or scopal considerations do impose constraints, see Ernst (2002), Engels (2012: 177-188).

3.3 Left dislocation in English

The term “left dislocation” (LD) was first used by Ross (1967) to label English sentences where there is some discourse-prominent element at the left periphery of the sentence and there is a co-referent resumptive pronoun later in the clause. I will use the label “host” for the prominent entity itself (*John*), repeated here as (33) and “associated pronoun” or “pronominal associate” for the co-referential pronoun (*him*).

(33) *John_i, I like him_i.*

As we can see, left dislocation minimally differs from TOP in the fact that here we usually find a pronoun in the canonical position of the fronted constituent. However, just like in the case of CIADJs, if we look deeper, a wide range of other differences surface.

3.3.1 Properties of left dislocation in English

First, in contrast with TOP, the fronted element in left dislocation in English (LD-Eng) must be a referential expression. For this reason, the examples with a fronted VP, AP and CP in (34) are only marginally acceptable, as opposed to the natural-sounding examples in (2).

- (34) a. ???*Surrender_i, we will never [do so]_i.*
b. ???*Happy_i, Tom will never be [like that]_i.*
c. ???*[That Tom was a movie star]_i, we would have never guessed that_i.*

From this it also follows that quantified or indefinite nominals are quite degraded in LD-Eng, as already noted in Rodman (1974). Rodman actually marks the examples in (35) with a *, but Shaer (2009) shows that some examples of this sort may be found in naturally occurring texts, e.g. see (36). Still, the phenomenon seems to be quite restricted.

- (35) a. ???*[A boy]_i, I saw him_i.*
b. ???*Someone_i, he_i is coming.*
c. ???*Everybody_i, they_i are doing it_i.*
(36) *And [most folks]_i, they_i don't seem like they notice that you've answered that same question 400 times that hour.*

Secondly, the effects related to island-constraints that were observed with TOP-Eng are absent in the case of LD-Eng. It is clear that syntactic factors are irrelevant for the dependency between the host and its pronominal associate.

- (37) a. *John_i, I doubt the claim that you like him_i.* (CNPC)
b. *John_i, I can't guess why you like him_i.* (wh-island)
c. *John_i, we think that David laughed when we selected him_i.* (adjunct-island)

- d. *John_i, that you saw him_i surprised me.* (subject island)
- e. *John_i, I whispered that I saw him_i.* (nonbridge verb)

Thirdly, the binding-patterns are also different. The fronted reflexive (38a) and the pronoun in (38b) remain unbound, which leads to ungrammaticality for the former case. This pattern is the opposite of what was seen in (7) earlier with TOP-Eng. This suggests that there are no “reconstruction”-effects in LD-Eng.

- (38) a. *?Pictures of himself_i, John_i likes them.* (Principle A)
- b. *Pictures of him_i, John_i likes them.* (Principle B)

That sentences like (38a) are not entirely unacceptable could be a result of some poorly-understood processes that make variable binding possible even when the necessary syntactic configurations do not hold. Actually, such claims have also been made in connection with LD-Eng, see e.g. Vat (1981), who reports that (39a) is not entirely ruled out. See also (39b), a sentence without any left dislocation, from Shaer & Frey (2003). Based on a purely syntactic view of binding, the sentences in (39) should be ungrammatical.

- (39) a. *?His_i first article, I think [every linguist]_i would consider it a failure.*
- b. *?When his_i boss is happy, [every office-worker]_i is happy too.*

Lastly, while in TOP the case of the fronted constituent is always the one that it would get in its canonical position (40a), in LD-Eng the two may be different (40b).

- (40) a. *Me/*I, John likes.*
- b. *Me_i, I_i like beer.*

In principle, the connection between the left-dislocated host element and the pronoun can be thought of as one involving syntactic or merely a pragmatic relationship. I think that the latter option seems more likely though, since one can find examples where, given the proper context, there is no pronoun present in the construction (these are “unlinked topics”). The following examples are from Lambrecht (2001).

- (41) a. *Tulips, do you have to plant new bulbs every year?*
- b. *Austin, at least you can sit near the AC.*

The relation between LD-Eng and other fronting constructions in a single sentence can not be easily assessed. According to Grohmann (2003: 139) LD-Eng always has to precede TOP, as in (42). Shaer (2009: 389), on the basis of examples like (43), debates this.

- (42) a. *Mary_i, John_j, she_i likes_j.*
- b. **Mary_i, John_j, he_j likes.*

- (43) a. *Now [this junk]_i, [my father]_j, he_j was always collecting.*
 b. *Now [my father]_i, [this junk]_j he_i was always collecting.*

Another contended issue is to what extent LD-Eng can be regarded as a root-clause phenomenon. According to most standard accounts, LD-Eng is ungrammatical in subordinate clauses but once again, the opposite claim is also present in the literature. Note that there is no debate about the embeddability of TOP or CIADJ (44b-c).

- (44) a. *?John_i said that Mary_j, he_i likes her_j.*
 b. *John_i said that Mary_j, he_i likes.*
 c. *John said that in New York, there is always something to do.*

Related to this is the fact that while TOP and CIADJ allow syntactic dependencies to arch over them (45c-d),⁴⁴ LD-Eng blocks such phenomena (45a-b). This is probably partly because of the fact mentioned above that LD-Eng is somewhat marked in embedded positions. The long-distance dependency exacerbates the already dispreferred configuration.

- (45) a. **How would you say that Robin_i, he_i would solve this problem?*
 b. **This is a problem that Robin_i, he_i wouldn't solve.*
 c. *How do you think that this problem, Robin would solve?*
 d. *How do you think that in New York, Robin would behave?*

These last three properties (ordering, embeddability, intervening) all seem to point in one direction: LD-Eng is a more peripheral structure in some sense than TOP and the CIADJ construction. As such, it always has reduced grammaticality when the structure occupies a sentence-internal position.

The main syntactic properties having been surveyed, let us move to the information structural properties of left dislocation. Prince (1998) claims that there are three basic functions for LD-Eng: island-amnesty, simplifying discourse processing, signaling a “poset-inference”

In the first use, it is actually applied as covert topicalization. The speaker would like to use a topicalization, but faces a syntactic obstacle, e.g. an island, and thus is forced to put a resumptive pronoun in the canonical position of the initial element. One such example is shown in (46).

- (46) *Tom_i, the story about *(him_i) was funny.*

As such uses are clearly forced by core syntax and have nothing to do with information structure, I exclude them from the scope of this discussion.

The second function of LD-Eng is “simplifying discourse processing.” According to Prince (1998) this means that by using LD-Eng, people remove discourse-new entities from positions that are dispreferred for them. Prince’s (1998) example of this is the following segment:

⁴⁴ This is not entirely unrestricted though, see Browning (1996: 252-253).

- (47) *My sister got stabbed. She died. Two of my sisters were living together on 18th Street. They had gone to bed, and this man, their girlfriend's husband, came in. He started fussing with my sister and she started to scream. The landlady_i, she_i went up, and he laid her out. So sister went to get a wash cloth to put on her, he stabbed her in the back.*

According to Prince (1998), *the landlady* in its original position would be a subject and subjects are generally dispreferred as discourse-new entities. One can also approach this from the perspective of Lambrecht's (1994: 185) "Principle of the separation of reference and role": do not introduce a referent and talk about it in the same clause.

On the axis of referential givenness, it could also be added that such left dislocated elements tend to occur if the topic is shifted, that is, when a not yet topical entity is promoted to the topic-status. This is also the conclusion of Frey (2005), who shows that LD-Eng cannot be used to simply to continue a given topic, it must "break the discourse continuity and start a new discourse unit" (Frey 2005: 23). See (48):

- (48) *I heard some news about John. #John, Mary kissed him.*

That LD-Eng can be used to introduce a brand-new entity can also be shown from the fact that it is conceivable that someone, looking for a particular *Tom*, enters a room and utters the following sentence, containing an LD-Eng:

- (49) *Tom_i, where is he_i?*

The same could hardly be conceivable with TOP as the discourse-conditions for evoking contrast are much stricter. There must be a discourse question already under discussion to which the contrast and the set of alternatives can be related.

The third use of LD-Eng according to Prince (1998) is to trigger an inference on the part of the hearer that the entity represented by the initial NP stands in a salient partially-ordered set in relation to some entity or entities already evoked in the discourse-model. Partially ordered sets ("posets") are "defined by a partial ordering R on some set of entities, e, such that, for all e-1, e-2, and e-3 that are elements of e, R is either reflexive, transitive, and antisymmetric or, alternatively, irreflexive, transitive, and asymmetric" (Prince 1998). In essence, this means that the left-dislocated entity has some set relation with other elements.

Prince (1998) sees these functions as separate entities. However, subsequent research suggests that there may be a way to have a unified view of functions 2 and 3 (as was stated, the first function is set aside).

Gregory & Michaelis (2001), based on a corpus study on TOP and LD-Eng, suggest that the overarching function of LD is that of "topic promotion", that is, to bring entities into the discourse. They have compared all the LD-Eng tokens with all the TOP tokens and have found three factors that back this claim up.

First, they examined the givenness of LD-Engs, compared to TOPs. They used Gundel, Hedberg & Zacharski's (1993) cognitive statuses to determine the referential givenness of an element in the discourse. The authors found that LD-Eng has relatively low givenness in the discourse, as opposed to TOP-Engs. The latter had higher activation status, which is expected

if they are always contrasted to some discourse elements, as was established in the previous section. LD-Eng was used to bring new topics into the discourse, which is in accordance with its function of topic-promotion.

Gregory & Michaelis's (2001) second target for investigation was the anaphoricity of left-dislocated and topicalized entities. They categorized tokens according to the type of the anaphoric link that the fronted element had to the discourse (from highest to lowest): directly mentioned, the entity is member of a set that has been mentioned, none. They found that LD-Engs tended to have low anaphoricity, which is to be expected of a topic-promoter.

Gregory & Michaelis's (2001) final factor was topic persistence. They measured to what extent the fronted elements in LD-Eng and TOP tend to remain topics of the subsequent discourse. They found that LD-Eng has a high topic persistence, as opposed to TOP. This is in line with what we have discussed in connection with these structures: LD-Eng is a topic promoter, so one expects that the entity introduced by it is going to be talked about. We do not necessarily have such expectations for contrasted elements introduced by TOP.

Thus, it is plausible to claim that LD-Eng is a construction that is linked to the topic information structural category. It may have contrastive uses, as in Prince's (1998) poset use, but unlike TOP-Eng, this contrastivity is not compulsory, as can be clearly seen in (46).

The link of LD-Eng and topics in general is further supported by a corpus study by Snider & Zaenen (2006) who found that there is a positive correlation between LD-Eng and animacy. This is expected if LD-Eng is topic-related, since cross-linguistic investigation often observes a connection between animacy and topichood (and topic-marking).⁴⁵

3.3.2 Analysis of left dislocation in English

Left dislocation in English (LD-Eng) resembles frame-setting CIADJ in that it seems to be loosely integrated into the sentence. As we have seen in the previous section, it is unaffected by islands and it induces no Principle C effects. Based on this one might assume that an analysis like the one posited for CIADJ would be suitable for LD as well, where they are "base-generated" in a left-peripheral node, and are not functionally plugged into the internal part of the sentence. However, I think it is undesirable to collapse the analysis of CIADJ and LD-Eng. The reasons are the following.

While CIADJ are easily assigned the grammatical function ADJUNCT, it is not at all clear what GF should be assigned to the nominals that occur in LD-Eng. This makes a situation that is difficult to handle with LFG's functionally annotated phrase-structure rules. Assigning no GF is not a particularly appealing option given the central role of f-structure in LFG.

LD-Eng also differs from CIADJ in a number of ways in terms of their syntax. While CIADJs are flexibly placed around TOP, LD-Eng seems to be marked in post-TOP positions. Although the data is not entirely straightforward, it does contrast with CIADJ, where there is no doubt about the positional flexibility. The situation is similar with respect to embeddability. The embeddability of LD is questionable, while CIADJ (and TOP) are fully licensed in subordinate clauses. It appears that LD is "radically" left-peripheral. From this perspective it seems plausible that LD-Eng may block or disrupt syntactic dependencies reaching over it.

⁴⁵ See Dalrymple & Nikolaeva (2011, chapter 1 and references therein) for an outline of correlation between animacy and topichood in the context of object-marking in a number of languages.

What this suggests is that LD-Eng is even less integrated than CIADJ. In my view, this is because the fronted elements in LD-Eng are loose “attachments” to the phonetic string of the sentence. They are unattached to the syntax of the containing sentence at any syntactic level, so they function as a fully independent entity in relation to the host. As such, placed sentence-internally they disrupt the inner structure of the sentence that they are phonetically part of. Thus, they are essentially syntactic orphans in the sense of Haegeman (1991) and Shaer (2009). The relation between the host (the left-peripheral element) and the pronominal associate is like the relation between entities in two different utterances. Then the degradation in (42)-(45) arguably comes from processing factors, more precisely, from the difficulty of parsing phonologically intermingled independent utterances.

This approach is feasible in LFG as the phonetic string in itself could be regarded as a representational level in LFG. It is the acoustic data, which the function π maps to the constituent-structure (the level where constituency is represented) and the function β maps to the prosodic structure (where intonation, stress, etc. are represented, see Mycock & Lowe 2013). As LD-Eng-constituents are independent entities in the host string, they and the matrix sentence are mapped from the phonetic string to entirely different c- and f-structures.

It is true that the f-structure of left-dislocated element would be rather fragmentary. *John* in (1c) (repeated here as 50a) would most probably only have a bare PRED feature, but some sort of analysis for such fragments is independently needed for constructions like vocatives or elliptic answers.

- (50) a. *John_i, I like him_i.*
 b. *Hey, John, what do you see?*
 c. A: *What may I serve you?*
 B: *Tea.*

Any agreement-like phenomena like number matching in (43a) follow from general discourse-considerations, like pronoun-antecedent relationship. Also, since there is no syntactic link between the fronted element and any sentence-internal correlate, case-mismatches of the sort illustrated in (40) are not unexpected.

Reconstruction effects are also expected to be absent, as in (38) and (39). That some reconstruction may still be possible could be the result of some poorly-understood processes at play. Shaer & Frey (2003: 476) speculate that what happens is “semantic subordination of the expression containing the variable to the expression containing the quantificational expression”.

As these independent elements may project to information structure in their own right, they may still assume the topic role that they usually have in discourse. Also, they may be referred back to by a pronoun in the host sentence, to form an explicit link, but this is not mandatory: in the case of unlinked topics (40), the relationship is mere pragmatic inference. Consequently, no syntactic formalization is necessary for the analysis of LD-Eng.

3.4 Conclusion to chapter 3

In this chapter I presented an LFG-based account of the three clause-initial, discourse-related constructions of English: topicalization, clause-initial adjuncts and left dislocation. Of the three, TOP-Eng is the most integrated structure: it is a functional dependency, marking contrastive topics/foci at information structure. Event-internal CIADJs are integrated like TOP, but these are not necessarily contrastive. Other, frame setting CIADJs are not tied to sentence-internal positions, they fully belong to the left-periphery as neutral/contrastive topics. Finally, LD-Eng, which is also a topic-marking device at information structure is a “syntactic orphan”, it is not parsed as a part of the host-sentence. Consequently, it is radically extra-sentential, barred from clause-internal positions.

As result of the discussion in this chapter, the following phrase structure of the English left-periphery emerges. As indicated throughout the chapter, I take the positions of these fronted elements to be IP-adjunctions. Frame-setting/event external adjuncts may occur either before or after topicalized phrases. TOP shares a node with event-internal CIADJs. They are integrated via versions of a functional identificational equation. Also, their i-structural specification differs, requiring only TOP to be contrastive.

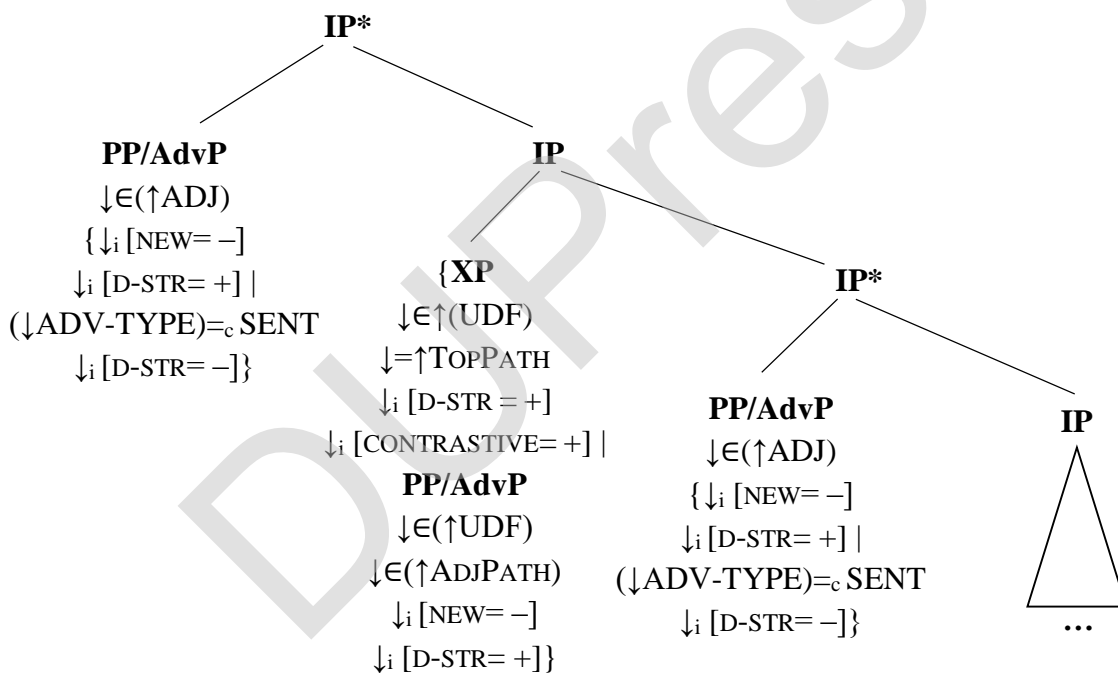


Figure 3.
The English left periphery

CHAPTER 4

CONSTRUCTIONS ON THE HUNGARIAN LEFT PERIPHERY

In this section we turn our attention to Hungarian and investigate some left-peripheral constructions in this language. In harmony with the goals of the monograph, each construction has repercussions for the information structure of the sentence.

The first part of this chapter is about configurations that are reminiscent of left dislocation structures in English so I follow the literature in calling these structures “left dislocation” (LD). However, it should be born in mind that I do not want to suggest that the analysis of these Hungarian sentence-types is necessarily parallel to that of English left dislocation. In fact, as we will see, the Hungarian LD actually has two subtypes, each of which should receive its own analysis and only one of these analyses is akin to the analysis of LD-Eng as presented in 3.3.

The second part of the chapter focuses on a construction which has already been mentioned briefly in 2.2, operator fronting (OF). OF also falls into distinct subtypes, with one that resembles topicalization in English in that it behaves as a long-distance, “movement”-like dependency. The other pattern has no counterpart in the English constructions that have been surveyed. That version of OF is also a long-distance dependency, but one that is anaphoric in nature, with semantic co-reference rather than functional identification being the underlying mechanism. This OF is going to be referred to as “prolepsis” and it will be argued to be sharing properties with what is called “control” in syntactic theory.

4.1 Left dislocation in Hungarian

In the previous section I have used the term “left dislocation” for those English constructions where there is some clause-initial nominal entity that is often co-referent with some clause-internal pronoun.

The Hungarian constructions that we are to investigate also show this pattern, which justifies the label for it. However, it should be noted that while the clause-internal pronominal counterpart was argued to be optional in LD-Eng, the Hungarian constructions (LD-Hun) are defined on the basis of the pronoun, so we will not have optionality here in this sense.

We will see that the phenomenon of “LD-Hun” as covers two distinct structures. These, as we will see, display different syntactic and information structural properties. They are going to be labelled as follows. These examples are formally quite similar but the subsequent discussion will reveal a number of contrasts. Following the terminology introduced in the previous chapter, I will use the label “host” for the prominent entity itself (*Jánost* ‘John.ACC’ in (1)) and “associated pronoun” or “pronominal associate” for the co-referential pronoun (*őt/azt*). For expository purposes, I coindex the host and the associate in the LD-Hun examples below.

- Topic left dislocation (TLD)
 - (1) *Jánost_i, azt_i meghívtuk.*
John.ACC that.ACC invited.1PL
'John, we invited him.'
- Free left dislocation (FLD)
 - (2) *Jánost_i, őt_i meghívtuk .*
John.ACC him.ACC invited.1PL
'John, we invited him.'

Left-dislocation in Hungarian has been the subject matter of a number of papers. The most notable references are Kenesei et al. (1998), Lipták & Vicente (2009), Lipták (2010, 2012), Baloghné Nagy (2013) and den Dikken & Surányi (2017). In my discussion, I will build on these sources in terms of empirical background. However, as none of these are LFG-papers, my theoretical perspective will be different. Also, none of these papers give a systematic survey of LD-Hun, so our discussion can also be regarded as a complementation of the aforementioned sources.

It is a common property of the two LD-Hun types illustrated in (1)-(2) that the pronominal associate can only occur at the left periphery. The presence of the resumptive pronoun at the right edge of the clause causes ungrammaticality.

- (3) **Meghívtuk Jánost_i, azt_i / őt_i.*
invited.1PL that.ACC that.ACC him

In the following section, I survey the properties of the subtypes of LD-Hun and then I present my analysis for them.

4.1.1 Properties of topic left dislocation

In (1), the left-dislocated phrase is *János* 'John' the subject matter of *Kati*'s love, and the associated pronoun is the demonstrative *azt* 'that.ACC'. The left-peripheral host and the pronominal associate are phonologically integrated into the sentence, there is no major intonational break between them and the clause.

Topic left dislocation (TLD) occurs in the "topic field" of the main- or subordinate clauses, intermingled with other topics.

- (4) *János Marit_i, azt_i Tamással reggel látta.*
John Mary.ACC that.ACC Thomas.with morning saw.3SG
'As for Mary, John, with Thomas saw her in the morning.'
- (5) *Erika azt mondta, hogy Júliával_i azzal_i gyakran találkozik.*
Erika that.ACC said.3SG that(C) Julia.with that.with often meets.3SG
'Erika said that as for Julia, she meets her often.'

Although stylistically marked, it is grammatically acceptable to have multiple instances of TLD-Hun in a sentence. The following example is from Den Dikken and Surányi (2017). (Stylistic or processing factors may put a limit to the process, but not syntax *per se*.)

- (6) *Jánost_i azt_i Péterrel_j azzal_j még nem kevertem össze.*
 John.ACC that.ACC Peter.with that.with yet not mixed.up.1SG
 ‘John, with Peter I have never mixed him up.’

If the host-elements and the pronouns are separated in a configuration like (7), the order of the pronouns preferably follow the order of the hosts, but the alternative order is not ruled out either.

- (7) a. *Jánost_i Péterrel_j azt_i azzal_j még nem kevertem össze.*
 John.ACC Peter.with that.ACC that.with yet not mixed.up.1SG
 b. *Jánost_i Péterrel_j azzal_j azt_i még nem kevertem össze.*
 Both: ‘John, with Peter I have never mixed him up.’

As regards positioning the host nominal and the associate pronoun, Lipták (2011) claims that a strict adjacency is mandatory (so 8 ought to be ungrammatical), while Balogné Nagy (2013) finds (8) acceptable. Interestingly, Lipták (2012) contradicts her earlier account, marking such structures grammatical. I take the position that while adjacency is a stylistic preference, it is not part of the syntactic description of TLD.

- (8) *?Jánost_i, tegnap azt_i mindenki látta a koncerten.*
 John.ACC yesterday that.ACC everyone saw.3SG the concert.on
 ‘John, everyone saw him at the concert.’

Various lexical classes and grammatical functions may be included in TLD. (9a) illustrates this with an oblique complement, (9b) with an infinitive and (9c) with a predicative adjective.

- (9) a. *A házban_i, [abban_i/ott_i] nincs senki.*
 the house.to that.in there not.be nobody
 ‘The house, nobody is there.’
 b. *Enni_i, azt_i szeretek.*
 eat.INF that.ACC like.1SG
 ‘To eat, I like doing that.’
 c. *Gazdag_i, az_i nem vagyok.*
 rich that not am
 ‘Rich, I am not that.’

(9a) also shows that sometimes there is a choice with regards the used demonstrative. The case-marked form of the basic demonstrative *az* ‘that’ is the standard option but if there is semantically matching specialized pronoun like the locative *oda* ‘there’ in the lexical inventory of the language, that may also be used. Thus *onnan* ‘from.there’, *oda* ‘to there’, etc. are also available in the appropriate contexts.

It can be said that the choice of the demonstrative basically follows the pattern of general pronoun selection of Hungarian: whatever demonstratives would be selected in non-LD contexts, such pronouns are also utilized in Hungarian LD.

However, there are some peculiarities. As shown in (1) above, personal names may be associated with a demonstrative pronoun in Hungarian LD. However, in non-LD contexts, such a reference would be considered pragmatically infelicitous, or at least impolite (regarding *John* not a person but a thing), and a personal pronoun would be the default choice.

- (10) A: *Jánost hívtad meg?*
 John.ACC invited.2SG PV
 ‘Is it John that you invited?’
 B: *Igen, [#azt/ őt].*
 yes, that.ACC him
 ‘Yes, #that/ him.’

However, this pragmatic infelicity cannot be felt in the TLD-examples brought up so far, which indicates that the semantics/pragmatics of this LD-demonstrative is not completely identical to regular demonstratives.

Another point of divergence between regular demonstratives and the ones used in the LD-construction has to do with number agreement. Interestingly, a plural host may be also associated with a singular LD-pronoun. Such a pattern would not be possible in regular discourse using demonstrative pronouns. (11a) is a constructed example, while (11b-c) are from the Hungarian National Corpus.

- (11) a. *Ezután a könyveket_i azt/ azokat_i visszavitettem a könyvtárba,*
 then the books.ACC that.ACC those.ACC back.took.1SG the library.to
majd hazamentem.
 then home.went.1SG
 ‘So I took the books back to the library and went home.’
- b. *...jól van, az adatokat_i azt_i főlírom rendeléskor.*
 right be.3SG the data.ACC that.ACC up.write.1SG order.at
 ‘Right, I take a note of the data when ordering.’
- c. *Ha már okozott esetleg következményeket is, az ilyen eseteket_i*
 if already caused.3SG perhaps consequences.ACC too the such cases.ACC
azt_i már mindig nehezebb kikezelni...
 that.ACC already always difficult.COMPARATIVE treat.INF
 ‘If it perhaps has some consequences, these cases are more difficult to treat.’

The third interesting divergence from the standard usage of demonstrative pronouns is that a seemingly accusative-marked LD-pronoun may be associated with a host that does not bear the OBJ grammatical function, as shown in (12), from Lipták & Vicente (2009: 661). (12a) is the LD-structure, and (12b) shows that the infinitival phrase must be the subject of the main predicate *jó* ‘good’.

- (12) a. *Úszni_i, azt_i jó volt.*
 swim.INF that.ACC good was
 ‘To swim, that was good to do.’
- b. *Jó volt [úszni/ az úzás/ *az úzást].*
 good was swim.INF the swimming.NOM the swimming.ACC
 ‘Swimming was good.’

Apart from such special cases, the host and the pronominal associate show case-matching. (13) is the minimally modified version of (1).

- (13) **János_i, azt_i meghívtuk.*
 John.NOM that.ACC invited.1PL
 Intended: ‘John, we invited him.’

The pronominal associate shows distal deixis by default. Proximal deixis is only possible if the host explicitly contains a proximal element. This is not a special property of Hungarian LD, the pattern shows up in other parts of Hungarian too, e.g. the pronominal associate of subordinate clauses is also distal by default (as mentioned in section 2.2.1).

- (14) *#Jánost_i, ezt_i meghívtuk.*
 John.ACC this.ACC invited.1PL
 Intended, approx.: ‘John, we invited this one.’
- (15) *[Ezt a fiút]_i, ezt_i meghívtuk.*
 this.ACC the boy this.ACC invited.1PL
 ‘This boy, we invited him.’
- (16) *Azt/ #ezt mondtam, hogy Jánost meghívtuk.*
 that.ACC this.ACC said.1SG that(c) John.ACC invited.1PL
 ‘I said that we had invited John.’

Otherwise can be said that general rules of Hungarian apply when it comes to the selection of the pronoun. In other words, TLD-Hun involves an anaphoric dependency. The syntactic constraints are imposed on top of these semantic/pragmatic considerations.

This may be clearly exposed if the host is a place-name. As was shown in (9a) earlier, there are two ways to refer to such entities: the appropriately case-marked form of *az* ‘that’, or a semantically appropriate spatial pro-form. Accordingly, we can use either of these forms in TLD-Hun, see (17).

- (17) a. *A házban_i, abban/ ott_i senki nem volt.*
 the house.in that.in there nobody not was
 ‘As for the house, there was nobody.’
- b. *A kertbe_i, abba/ oda_i gyakran kiment.*
 the garden.to that.to to.there often out.went.3SG
 ‘As for the garden, he often went out there.’

- c. *A konyhából_i, abból_i/ onnan_i gyakran jöttek jó illatok.*
 the kitchen.from that.from from.there often came good scents
 ‘As for the kitchen, good scents often came from there.’

However, when referring to certain types of places (e.g. cities), the case-marked forms cannot be used, only the spatial pronoun is suitable. An example can be seen in (18), where although the city itself bears the *-re* ‘onto’ ending in the meaning of *to*, when referring back to the city, only *oda* ‘there’ is appropriate.

- (18) A: *Szeged-re mész?*
 Szeged-onto go.2SG
 ‘Are you going to Szeged?’
 B: *Igen, *arra/ oda megyek.*
 yes that.onto there go.1SG
 ‘Yes, I’m going there.’

This pattern is preserved at TLD-Hun, as attested by the following examples: it is always the generally used pronoun that occurs. If a case-marking is featured in several uses (for example, in 19e-f), the functionally appropriate form is used.

- (19) a. *Szegedre_i, *arra_i/ oda_i szívesen megyünk.*
 Szeged.onto that.onto there gladly go.1PL
 ‘As for Szeged, we gladly go there.’
- b. *Debrecenben_i, *abban_i/ ott_i sokan élnek.*
 Debrecen.in that.in there many live.3PL
 ‘As for Debrecen, many people live there.’
- c. *Izlandon_i, *azon_i/ ott_i szívesen laknék.*
 Iceland.on that.on there gladly live.COND.1SG
 ‘As for Iceland, I would gladly live there.’
- d. *Vidékre_i, *arra_i/ oda_i nem költöznék.*
 countryside.onto that.onto there not move.COND.1SG
 ‘As for the countryside, I wouldn’t move there.’
- e. *Az egyetemen_i, *azon_i/ ott_i sokan dolgoznak.*
 the university.on that.on there many work.3PL
 ‘As for the university, many people work there.’
- f. *Az egyetemen_i, azon_i/ *ott_i sokan dolgoznak.*
 the university.on that.on there many work.3PL
 ‘As for the university, many people work on it.’ (many people work on the plans, construction, etc. of the university)

An interesting consequence of this is that in certain cases, the semantic and formal requirements are in conflict so there is no perfect form for the resumptive pronoun. One such case is (20).

- (20) *Olyan helyekre_i, ?oda_i/ ?olyanokra_i/ ?azokra_i nem megyünk.*
 such place.PL.onto there such.PL.onto those.onto not go.1PL
 ‘As for such places, we won’t go there.’

The problem in (20) is that the grammatically most appropriate form, which semantically and formally matches the antecedent, would be **odákra* ‘there.PL.onto’, but such a form does not exist, since *oda* (‘there’) is an adverb, which does not take plural and adpositional morphology. The other possible forms are either semantically or formally in conflict with the left dislocated constituent. What is evident from our discussion regarding the form of the associate is that one cannot impose rigid syntactic criteria on it; semantic considerations play the crucial role.

Dislocated adverbials also fall under such considerations. In (21) only the adverbial pronoun may be used, the case-marked nominal is ungrammatical (even though the case-marking itself is appropriate).

- (21) *Szép-en_i, *az-on_i/ úgy_i csak Kati dolgozik.*
 nice-on that-on so only Kate works
 ‘(As for) nicely, only Kate works like that.’

One may wonder at this point why is it possible for a non-referential element like *szépen* ‘nicely’ to participate in a TLD, which (as the label suggests) is supposed to be a topic-marking device. The answer is that TLD is neutral with respect to contrastivity, that is, both standard and contrastive topics may be marked with it. As noted in section 1.3.2, the referentiality-requirement typically associated with neutral topics is absent in the case of contrastive ones. This is the case in for example in (21), which can only be interpreted contrastively. Accordingly, the intonation must be a fall-rise one. Some more examples for non-referential, contrastive left dislocated elements are shown in (22).

- (22) a. *Állatot_i, azt_i nem tartok.*
 animal.ACC that.ACC not keep.1SG
 ‘As for animals, I don’t keep one.’
- b. *[Kevés fiút]_i, azt_i Mari hívott meg.*
 few boys.ACC that.ACC Mary invited.3SG PV
 Mary invited few boys (contrasted to e.g. John, who invited many).

Neutral topics may not be non-referential, which can also be shown by investigation of the interaction of TLD-Hun with copular sentences. Traditionally (e.g. Declerck 1989), these are classified into five subtypes:⁴⁶

- (23) a. *János okos (volt).*⁴⁷ → Predicational
 John smart was
 ‘John is/(was) smart.’

⁴⁶ For an overview of the state of affairs pertaining to copular clauses, see Mikkelsen (2011).

⁴⁷ In Hungarian, the copula is zero in present tense. For more on Hungarian copular clauses, Laczkó (2012) in LFG and Hegedűs (2013) in minimalist framework.

- b. *A hajnalcsillag az esthajnalcsillag (volt).* → Equative
 the morning.star the evening.star was
 ‘The morning star is/(was) the evening star.’
- c. *Az a férfi (ott) János.* → Identificational
 that the man there John
 ‘That man (there) in John.’ (In a deictic context.)
- d. *A demokrácia egy olyan rendszer (volt), ahol a népakarat*
 the democracy one such system was where the will.of.the.people
érvényesül. → Definitional
 reign.3SG
 ‘Democracy is a system where the will of the people rules.’
- e. *A legjobb jelölt János (volt).* → Specificational
 the best candidate John was
 ‘The best candidate is/was John.’

A non-contrastively interpreted TLD may occur with the subject of each of these types, as seen in (24), with the added pronouns *az* ‘that’, except for the specificational subtype. (24e) is only grammatical on the contrastive reading.

- (24) a. *János_i az_i okos (volt).*
 b. *A hajnalcsillag_i az_i az esthajnalcsillag (volt).*
 c. *[Az a férfi]_i az_i (ott) János.*
 d. *A demokrácia_i az_i egy olyan rendszer (volt), ahol a népakarat érvényesül.*
 e. *A legjobb jelölt_i az_i János volt.* – grammatical only on the contrastive reading

The problem with (24e) lies in the semantics of specificational copular sentences. The initial constituents in (24a-d) are regular referential aboutness topics, the sentence says something about them. This is not the case in the specificational sentences (24e). The constituent *a legjobb jelölt* ‘the best candidate’ is not an entity that is the subject of predication, but it is a predicate itself, introducing a variable (“the best candidate is X”). It is not a referential entity, so it is not a topic. Not being topic-entity, it is not a suitable host for the neutral topic use of TLD. The other nominative entity is also not suitable for the pronoun, as it identifies the variable introduced by the predicate (“X=János”). As such, its interpretation is that of an information focus.

If the TLD-d element is contrastive, it may participate in “reconstruction”: being systematically interpreted as if it occupied a lower position in a syntactic tree. So TLD may avoid Principle A-violations as in (25a), and TLD-d elements are capable of taking narrow scope (25b).

- (25) a. *Magát_i azt_i szereti János.*
 himself.ACC that.ACC likes John
 ‘(as for) Himself, John likes.’

- b. [Az *pro*_i⁴⁸ *anyját*]_j *azt*_j *mindenki*_i *szereti*.
 the mother.POSS.3SG.ACC that.ACC everyone likes
 ‘Everyone_i likes his_i mother.’

4.1.2 Analysis of topic left dislocation

First let consider the “signature-elements” of TLD, the pronouns themselves: what is their role in the structure? This is not evident at first sight, as they are optional in all LD-Hun constructions. Nothing happens if they are omitted, as in (26), which is the pronoun-less version of (1).

- (26) *Jánost* *meghívtuk*.
 John.ACC invited.1PL
 ‘We invited John.’

However, we saw in the previous that the selection of the pronouns is partially based on semantics. That is, these pronouns are not expletives/dummy elements, they do contribute semantically and pragmatically to sentences. In LFG-theoretic terms, this means that they should have a PRED feature.

It follows that there should be contexts where their presence/absence makes a difference. As it turns out, there are such contexts and the TLD pronoun basically requires the host to be a referentially anchored.

Consider (27) first, taken from Gécseg & Kiefer (2009). They use it to illustrate that topics in Hungarian are not necessarily specific, contra É. Kiss (2002: 10), as *valaki* ‘somebody’ in (27) may have a non-specific reading (see also footnote 33 in section 3.1.). However, if an LD-Hun pronoun is added to such a sentence, it becomes infelicitous, if uttered out of the blue.

- (27) *Elmondom mi történt tegnap. Valaki (#az_i) bekopogott az*
 tell.1SG what happened.3SG yesterday somebody that in.knocked the
ajtómon.
 door.POSS.1SG.on
 ‘I tell you what happened yesterday. Somebody knocked at my door.’

Another instance where the pronoun makes a difference is when it is attached to an indefinite nonpronominal expression, like the example in footnote 23, repeated here as (28). Here the pronoun may not be added as the indefinite is not specific.

⁴⁸ The “pro” here represents the unpronounced possessor. I use it in the string for clarity purposes, not because I posit a zero c-structural element. Such unpronounced entities are only present in f-structure within the LFG-framework (due to Economy of Expression).

- (28) *Elmondom mi történt tegnap.* ('I tell you what happened yesterday.)
 [Egy gyerek]_i (#azi) leesett a villamosról, de szerencsére nem sérült meg.
 one child that fell.off.3SG the tram.from but luckily not injured.3SG
 PV
 'A child fell off the tram but luckily, he wasn't injured.'

Nonetheless, both *valaki* 'somebody' and *egy gyerek* 'a child' with TLD may be acceptable in contexts where the referents are members of contextually given sets.

- (29) a. *Sok embert meghívtam. Valaki_i az_i eljött, valaki nem.*
 lot person.ACC invited.1SG somebody that came.3SG somebody not
 'I invited a lot of people. Some came, some didn't.'
- b. *Több gyerek játszott a homokozóban. [Egy gyerek]_i az_i kitalálta, hogy épít egy homokvárat.*
 several child played.3SG the sandpit.in one child that
 came.up.with.3SG that(C) build.3SG one sandcastle.ACC
 'Several children were playing in the sandpit. One of them came up with the idea of building a sandcastle.'

A similar contrast may be construed with *bárki* 'anyone'. In (30a) may be interpreted in a way that *bárki* 'anyone' is restricted to a certain group of people. (30b), where this anchored interpretation is not available, as the meaning unrestrictedly refers to people in general, is infelicitous.

- (30) a. *Bárki_i (az_i) nem jöhet be.*
 anyone that not come.POT.3SG in
 Intended: 'Not just anyone may come in.' (Lit.: 'Anyone, they may not come in.')
- b. *Ha bárki_i (*az_i) bejött, adtunk neki enni.*
 if anyone that in.came.3SG gave.1PL him.DAT eat.INF
 Intended: 'If anyone came in, we gave them food.' (Lit.: 'If anyone, they came in, we gave them food.')

Another indication of the semantic nature of this pronoun is its incompatibility with idiom-chunks. Consider (31).

- (31) *A fene_i (#az_i) megette ezt az egész ügyet.*
 the heck that ate.3SG this the whole issue.ACC
 'This whole issue is screwed.' (Lit.: 'The heck, that ate this issue.')

(31) is an intriguing sentence, as there is an idiom chunk in the topic field, which in itself should make the sentence anomalous, in theory. (Compare: #*The beans, John spilled (them).*) For some reason which is not really clear at this point, the pronoun-less version of the sentence is acceptable, even on the idiomatic reading. Several examples of this sort may be found via internet search. Nevertheless, adding the LD-pronoun makes the sentence semantically

anomalous by forcing a degree referentiality on the subject phrase *a fene* ‘the heck’, which it is not compatible with.

It has to be noted that the force of this argument is somewhat diminished by the fact that splitting the idiom up by any means reduces the grammaticality of the sentence.

- (32) *A fene (?már) megette ezt az egész ügyet.*
 the heck already ate.3SG this the whole issue.ACC
 ‘The whole issue was already screwed.’ (Lit.: ‘The heck already ate this issue.’)

However, while (32) with the interjecting *már* ‘already’ sounds marked, it is still not totally unacceptable, in contrast with the LD-version of (31). I take this as an indication that apart from the syntactic problem of breaking the continuity of an idiom, the semantics of the pronoun is also behind the problem in (31).

At this point it should be restated that TLD is not necessarily contrastive. This is important because otherwise one could argue that the explanation behind the data in (27)-(31) is simply the difficulty of construing contrastive readings for the sentences.

Additionally, I would like to call attention to Arregi (2003: 40), who describes similar effects in Spanish CLLD. In (33), *algo* ‘something’ may not be associated with the pronominal clitic *lo* ‘it’.

- (33) *Algo_i Juan si (*lo_i) comió.*
 something Juan yes it ate.3SG
 ‘Something, Juan did eat.’

Arregi (2003: 40) argues that “the distribution of the clitic is determined by the interpretation of the clitic itself... In left dislocation, the clitic is interpreted as an individual variable”. While the proper semantic/pragmatic characterization of the LD-pronominal is yet to be worked out, it seems to be clear that it has to be interpreted some way, which precludes an analysis where it is an empty formative.

From this it might seem that the importance of the pronoun lies in imposing a specificity restriction on the host element. This conclusion, however, is premature, as generics may readily be supplemented with an LD-pronoun. (34), with generic subjects, is grammatical and felicitous, regardless of contrastivity.

- (34) a. *[A delfin]_i az_i egy okos állat.*
 the dolphin that one smart animal
 ‘The dolphin is a smart animal.’
 b. *[Egy gyerek]_i az_i nem tudhat mindent.*
 one child that not know.COND.3SG everything.ACC
 ‘A child cannot know everything.’

While the full semantics analysis of the TLD-pronouns is beyond the scope of this investigation, I would like to propose that we can gain some valuable insight by looking at the referential givenness-conditions on the usage of the pronominals.

As was discussed in section 1.3.1, referential givenness is about “a relation between a linguistic expression and a corresponding non-linguistic entity in the speaker/hearer’s mind” (Gundel 1999). One approach about the characterization of referential givenness is the Givenness Hierarchy of Gundel, Hedberg and Zacharski (1993). They distinguish between six cognitive statuses that a referring expression may have. The statuses are the following, from lowest to highest. The definitions are from Gundel et al. (2010), the examples are from Gundel, Hedberg and Zacharski (1993):⁴⁹

- (35) a. Type identifiable: associate type representation.
I couldn't sleep last night. A dog (next door) kept me awake.
- b. Referential: associate unique representation by end of sentence.
I couldn't sleep last night. This dog (next door) kept me awake.
- c. Uniquely identifiable: associate unique representation by end of NP.
I couldn't sleep last night. The dog (next door) kept me awake.
- d. Familiar: associate representation in memory.
I couldn't sleep last night. That dog (next door) kept me awake.
- e. Activated: associate representation in working memory.
I couldn't sleep last night. That kept me awake.
- f. In focus: associate representation in focus of attention.
I couldn't sleep last night because of your dog. It kept barking.

I propose that the distal demonstrative that is used in LD-Hun by default may be felicitously used if the referent is at least “referential” in this sense: a unique entity may be associated with the referent of the pronoun in the mind of the interlocutors. This subsumes the notion of specificity, which is a favorable aspect as von Heusinger (2011) argues that specificity should be understood as “referential anchoring”. For example, when using a specific indefinite (as in 38) the hearer must establish a permanent representation of the referent. What I would like to add to this is that in the case of generics, the referent may be construed as a unique class of entities. That is, the concept of being “referential” in the sense of the Givenness Hierarchy generalizes over specific indefinites and generics. However, nonspecific indefinites are only type-identifiable, as is clear from (35a). This explains the infelicitous examples in (27)-(31) as the fronted elements in these cases fail to single out an entity.

Another question about LD is the nature of the relationship between the host and the pronoun: which of them is the dominant participant in the sentence? Here I agree with Zaenen (1997), who argues for an analysis of Icelandic left dislocation where the pronoun is an adjunct of the host. This is the most plausible analysis for Hungarian as well. The alternative is the reversed constellation, whereby the pronoun is the argument of the main predicate and the host is an adjunct, like an appositive construction. While such an analysis might be plausible for some Germanic LD-constructions,⁵⁰ it is definitely not for Hungarian. To prove this, first recall

⁴⁹ Note that the higher statuses entail the lower ones, e.g. anything that is “referential” is also “type identifiable”.

⁵⁰ Frey (2004) and Alexiadou (2006) propose analyses along this path. Whether Zaenen’s (1997) analysis should be revised too is matter of further inquiry. I will explore some of the cross-linguistic and theoretical landscape of LD in section 5.2.

the data from (12) where it is an infinitival complement that satisfies the subcategorization requirement of the main predicate and not an accusative element like the LD-pronoun. Moreover, evidence for the primary status of the host over the pronoun can also be seen from object definiteness agreement patterns.

As illustrated in (36), finite verbs in Hungarian show definiteness agreement with their objects. Demonstrative pronouns count as definite objects, evidenced by (37).

- (36) a. *Egy fiút lát-tál.*
 one boy.ACC see-PAST.2SG.INDEF
 ‘You saw a boy.’
- b. *A fiút lát-tad.*
 the boy.ACC see-PAST.2SG.DEF
 ‘You saw the boy.’
- (37) *Azt lát-tad/ *lát-tál.*
 that.ACC saw-2SG.DEF see-PAST.2SG.INDEF
 ‘You saw that.’

In an LD-sentence, it is always the host and not the pronominal associate that determines the definiteness agreement of the verb. Hence in (38a) the verb shows indefinite agreement, triggered by *egy almát* ‘an apple.ACC’, even though there is the demonstrative LD-pronoun in the sentence, which in principle could trigger definite conjugation. (See also den Dikken & Surányi 2017: 571-572).

- (38) a. [*Egy fiút*]_i *azt*_i *lát-tál.*
 one boy.ACC that see-PAST.2SG.INDEF
 ‘A boy, you ate it.’
- b. [*A fiút*]_i *azt*_i *lát-tad.*
 the apple.ACC that see-PAST.2SG.DEF
 ‘The apple, you ate it.’

As noted, Zaenen (1997) proposes an analysis for Icelandic LD whereby the pronoun is regarded as an adjunct of this topical host, as shown in (39).

- (39) S → XP XP V NP
 (↑TOP)=↓ (↑TOP-ADJ)=↓ ↑=↓ (↑SUBJ)=↓

Based on the considerations outlined in this discussion, I propose an analysis in a similar spirit. This is shown in Figure 1 for topic left dislocation in Hungarian, exposed via annotated phrase structure. In (40) we see the proposed annotation for the pronominal associate.

egyedül ‘alone’ does not vary depending on the subject), that is why the matching has to be stated separately.

(41) *Ég* *geri* *petta* *einn.*
 I MASC.SG.NOM will.do this alone MASC.SG.NOM.
 ‘I will do this alone.’

(42) a. *Én* *ezt* *egyedül* *fogom* *csinálni.*
 I this.ACC alone will.1SG do.INF
 ‘I will do this alone’

b. *Ők* *ezt* *egyedül* *fogják* *csinálni.*
 they this.ACC alone will.3PL do.INF
 ‘They will do this alone.’

The second point is that I propose to handle case-discrepancies with alternate lexical entries for the respective pronouns. This differs from the approach of Lipták & Vicente (2009) and Lipták (2012), where predicate left dislocation (e.g. (12a)) is analyzed as being the result of a process that is distinct from other instances of LD. Lipták & Vicente (2009) propose that the accusative case on the pronoun in (12a) is the manifestation of default case in Hungarian. In my approach, the accusative case is just apparent, this alternative lexical entry of the pronoun is caseless. I consider this to be a better approach as a unitary underlying mechanism is posited for all LD-structures in Hungarian. Moreover, it is not evident that accusative is the default case in Hungarian, see e.g. (9c), where the adjective is associated with a nominative pronoun. Also, in contrast to English, left-peripheral, hanging pronouns are not in the accusative case, which argues against accusative being the default in Hungarian.⁵¹

(43) *Mei, I* like beer.

(44) [*Éni*/ **Engemi*], *éni* *szeretem* *a* *sört.*
 I me I like.1SG the beer.ACC
 ‘Me, I like beer.’

This latter construction is distinct from TLD, it is an instance of free left dislocation (FLD), to which we in the next section.

A final point to make is that I assume that the LD-pronoun is specified for the person feature (3rd person), but the apparently singular one is actually underspecified with respect to number, which enables it to appear in sentences like (11).

⁵¹ Bartos (2002, footnote 5) notes that the dative may surface in imperative root infinitives. This may also be regarded as some sort of a default case, crucially non-identical to the accusative.

(i) *?A* *füknak* *leülni!*
 the boys.DAT sit.INF
 ‘Boys, sit!’

4.1.3 Properties of free left dislocation

Free left dislocation (FLD) in Hungarian represents a relatively unmapped territory of the Hungarian language. It can be traced back to Kenesei et al. (1998), who first observed that left dislocation structures can be associated with nontopic functions as well. This is true for FLD, as we shall see. Let us repeat our example from (2) as (45).

- (45) *János_i, ő_{ti} meghívtuk .*
John.ACC him.ACC invited.3PL
'John, we invited him.'

In contrast to TLD, which prosodically forms a unit with the rest of the sentence, the left-peripheral element in FLD is set apart by a noticeable intonational break.

Another immediately striking difference is that personal names are associated with personal pronouns, as one would expect in standard discourse. This feature of FLD can be most clearly explicated in conjunction with another property of the construction, the wider range of information structural categories that can be involved. In addition to the topic discourse function, the FLD pronoun can also be a focus of the main clause (as noted by Kenesei et al. 1998). This is seen in (46), where the focussed pronoun in the preverbal position pushes the preverb *meg* (contributing to the perfective interpretation of the sentence) to the postverbal field. In such cases, using a demonstrative like the ones in TLD triggers the sort of pragmatic infelicity demonstrated in (10) above.

- (46) *János_i, [ő_{ti focus} / #az_{ti focus}] hívtuk meg.*
John.ACC him that.ACC invited.1PL PV
'John, we invited [HIM/#THAT].'

Another difference compared to TLD is that there is no obligatory case-matching in FLD. (47) forms a contrast with (13) above.

- (47) a. *János_i, ő_{ti} meghívtuk.*
John.NOM him invited.1PL
'John, we invited him.'
- b. *János_i, ő_{ti} hívtuk meg.*
John.NOM him invited.1PL PV
'John, we invited him.'

FLD contrasts with TLD in that it becomes marked if the host element is not string-initial. This obviously happens in subordinate clauses, but the same effect may appear in main clauses as well. Consider the FLD (a)- and TLD (b)-examples below.

- (48) a. *?Mari Jánosnak_i, neki_i adott ajándékot.* FLD
Mary John.DAT him.DAT gave.3SG gift

- b. *Mari Jánosnak_i, annak_i adott ajándékot.* TLD
 Mary John.DAT that.DAT gave.3SG gift
 ‘John, Mary gave him a present.’
- (49) a. *?Mondtam, hogy Jánost_i, őt_i meghívtuk.* FLD
 said.1SG COMP John him invited.1PL
- b. *Mondtam, hogy Jánost_i, azt_i meghívtuk.* TLD
 said.1SG COMP John that.ACC invited.1PL
 ‘I said that John, we invited him.’

Finally, the FLD-Hun also differs from TLD-Hun in terms of binding and scope. Compared to the fully grammatical forms in (25), the parallel structures with FLD-Hun are definitely marked, even though contrastivity obtains. In other words, no “reconstruction”-effects can be observed. (Although the effects discussed in connection with LD-Eng and binding may blur the picture a bit, see example 38 in Chapter 3).

- (50) a. *??Magát_i, Őt_i szereti János.*
 himself him.ACC loves.3SG John
- b. *[??Az pro_i anyját]_j, Őt_j szereti mindenki_i.*
 the mother.POSS.3SG.ACC her.ACC likes everyone

Considering all these, it seems highly plausible that FLD-Hun constitute a structure distinct from other versions of LD-Hun. The exact nature of the distinction will be fleshed out in the next section.

4.1.4 Analysis of free left dislocation

Based on the considerations outlined the previous section, I argue that the most plausible analysis for FLD is one where the left-peripheral entity is syntactically independent from the rest of the sentence. In other words, just like LD-Eng, it is regarded as a “syntactic orphan”, using the terminology of Haegeman (1991) and Shaer (2009). The relation between the host (the left-peripheral element) and the pronominal associate is like the relation between entities in two different utterances. To back up the claim of FLD-Hun being extra-syntactic, let us review its properties. Many of these should sound familiar from our discussion of LD-Eng from previous sections.

The connectivity-effects are weak between the fronted element and the resumptive pronoun – there need not be number- or case-identity, just like in LD-Eng. Since the fronted element bears no syntactic connection with the pronoun, strict featural matching is not required. However, general discourse considerations are still operational, so wildly inconsistent specifications will be infelicitous (e.g. referring to a distal expression with a proximal pronoun, as in (51b)). However, this is no more surprising than the strangeness of (52), for which obviously no one would think that a syntactic characterization would be necessary.

- (51) a. *A könyv_i, János AZT_i vitte vissza a könyvtárba.*
 the book.SG.NOM John that.ACC took.3SG back the library.to
 ‘The book, John took THAT back to the library.’
- b. *#AZt_i a könyvet, ezt_i vittem vissza a könyvtárba.*
 that.ACC the book.ACC this.ACC took.1SG back the library.to
 ‘That book, I took THIS back to the library.’
- (52) *Look at [that car]_i! #This_i/That is so beautiful!*

FLD’s occurrence is preferred in a sentence-initial position. While the sentences in (48a) and (49a) cannot be starred, there is some degradation compared to a sentence-initial version. The degree of this markedness could be subject to inter-speaker variation. The mild ungrammaticality such sentences arguably comes from processing factors, more precisely, from the difficulty of parsing phonologically intermingled independent utterances. This is exactly the situation with LD-Eng, for which I’ve also subscribe to an “orphan”-analysis.

When referring to people, a personal pronoun must be used. Since personal pronouns are the default choice for referring to personal names in Hungarian and demonstratives are pragmatically infelicitous, the data in (46) is unsurprising

Also, since reconstruction is a sentence-internal process, it is expected not to save sentences like (50). (However, remember from section 4.3 that variable binding may still be possible via some extrasyntactic process (e.g. Shaer & Frey’s (2003) suggestion of “semantic subordination”).

Thus, the properties of FLD-Hun all make sense if we assume that the fronted element is extra-sentential, as in LD-Eng. We could think of it in the following way: the speaker first makes an announcement of a salient entity and then makes some assertion about this entity. The real structure behind it could be more faithfully indicated with a punctuation similar to that in (53).

- (53) *János(t)_i – tegnap őt_i hívtam fel.*
 John(.ACC) yesterday him called.1SG up
 ‘John – yesterday I called HIM up.’

As stated in section 3.3.2, From an LFG-perspective, this can be conceptualized along the line of Mycocks & Lowe (2013: 446), who regard the phonetic string itself as a representational level, which the function π maps to the constituent-structure and the function β maps to the prosodic structure.

4.2 Operator fronting in Hungarian

The other discourse-related long-distance dependency under scrutiny in Hungarian is what I label “operator fronting” (OF). The name reflects my aim to provide a concise theory- and analysis-neutral description of the phenomenon. In the previous literature, the structure has also been called “focus raising” (Kenesei 1992, Lipták 1998, Gervain 2004), “operator raising” (Gervain 2002) and “long operator movement” (É. Kiss 2002). These names for the phenomenon mirror the Chomskyan frameworks that these analyses are couched in. The

framework of this monograph is Lexical-Functional Grammar, so I have tried to find a label which has the least amount of implicature about the analysis. “Movement” and “raising” explicitly suggest that the analysis employs some kind of dislocation and are obviously Chomskyan terms. “Fronting” still has some derivational flavor to it, but it is at least not an exclusively Chomskyan term, it is more theory-neutral.

A basic example for operator fronting is shown in (54a), where the locative phrase *Párizsba* ‘to Paris’ is a complement of the embedded verb, yet it occurs in the main clause, sentence-initially, in a fronted position. It also bears some discourse function (hence the “operator” in the name)⁵². For contrast, (54b) shows a standard subordinate clause, with no fronted material.

- (54) a. *Párizsba mondtad, hogy méysz.*
 Paris.into said.2SG that(c) go.2SG
 ‘To Paris you said that you will go.’
- b. *Azt mondtad, hogy Párizsba méysz.*
 that said.2SG that(c) Paris.into go.2SG
 ‘You said that you will go to Paris.’

4.2.1 Properties of operator fronting

As was shown in (54), the essence of OF is that some element that would normally be assumed to belong to the finite subordinate clause surfaces in the matrix clause. In (46), the affected element is an oblique function (a prepositional phrase, in c-structural terms), but there is a range of other options as well. In (55), it is demonstrated that the fronted element could also be a subject (55a), an object (55b) or an adjunct (55c) of an embedded clause.

- (55) a. *János mondtad, hogy jön a partira.*
 John said.2SG that(c) comes the party.onto
 ‘(Of) John you said that he will come to the party.’
- b. *Egy kutyát mondtad, hogy láttál.*
 one dog.ACC said.2SG that(c) saw.2SG.INDEF
 ‘A dog you said that you saw.’
- c. *Tegnap mondtad, hogy sokat ettél.*
 yesterday said.2SG that(c) lot.ACC ate.2SG
 ‘Yesterday you said that you ate a lot.’ (intended: the eating was yesterday)

A note is in order here. Fronting subjects, obliques or adjuncts may result in ambiguous structures, as sometimes these elements may also be interpreted in such a way that they belong to the matrix clause. This is clearly observable in (55c), where in principle *yesterday* could also refer to the time of saying. In (55a), if *say* had 3rd person singular conjugation, the sentence would be ambiguous between the meaning “(someone) said that John will come to the party”

⁵² The information-structural import of the fronting is ambiguous. That is why I use topicalization in some of the translations, which also has several interpretational options, as has been discussed in 3.1.

analogous to the given meaning in (55a), or the unremarkable interpretation “John said that he is coming to the party”, as the conjugation could be equally triggered by the *pro*-dropped 3SG subject or *János*. If there is an overt subject plus a nominative fronted element as in (56), the sentence’s judgement is degraded (Szűcs 2010), because of the ambiguity and processing difficulty caused by the two initial nominative constituents.

- (56) *?János Kati mondta, hogy jön.*
 John Kate said.3SG that(c) comes
 ‘John (of) Kate said that he will come. / (Of) John, Kate said that he will come.’

The case is similar with obliques. In (57), *Párizsban* (‘in Paris’) could refer to the location of the running (the OF interpretation) or to the location of saying.

- (57) *Párizsban mondtad, hogy futottál.*
 Paris.in said.2SG that(c) ran.2SG
 ‘Of Paris you said that you ran there.’ / ‘You said in Paris that you ran.’

Although such cases may be interesting from a language processing point of view, they will not be addressed in this monograph. I just note that this complication exists, so one must be careful about the construction of examples and try to avoid problematic ones if possible (e.g. by choosing an oblique that cannot be interpreted in the matrix clause, as in 54a).

The pronoun associated with subordinate clauses in Hungarian (see 54b) cannot be present when OF takes place, as evidenced by (58).

- (58) *(*Azt) Párizsba mondtad (*azt), hogy méysz.*
 that.ACC Paris.to said.2SG that.ACC that(c) go.2SG

The distance between the fronted element and its standard position is not limited to the immediate subordinate clause; it can span across several clauses.

- (59) *Párizsba mondtad, hogy hallottad, hogy megy János.*
 Paris.to said.2SG that(c) heard.2SG.DEF that(c) goes John
 ‘To Paris you said that you heard that John will go.’

There are a couple of interesting variations to be observed about OF. The first one is a possible “case-switch”: the fronted constituent may bear the case assigned by the matrix predicate, even if that case does not correspond to the function of the fronted element in the embedded clause. For instance, (60b) shows a fronted embedded subject, (60c) shows an embedded oblique and (60d) shows an embedded possessor. These may all bear the accusative case assigned to the OBJ argument of the main verb. In (60c-d), a resumptive pronoun must surface in the embedded clause. As noted in den Dikken (2010), such resumptive pronouns are also marginally possible

with fronted accusative subjects (60b)⁵³. The resumptive pronoun is impossible if the fronted element bears its original case (e.g. 60a).

- (60) a. *János mondtad, hogy (*ő) jön a partira.*
 John.NOM said.2SG that(c) he comes the party.to
 ‘(Of) John you said that he will come to the party.’
- b. *János-t mondtad, hogy (?ő) jön a partira.*
 John-ACC said.2SG that(c) he comes the party.to
- c. *Párizs-t mondtad, hogy oda utazol.*
 John-ACC said.2SG that(c) there travel.2SG.
 ‘(About) Paris you said that you will go there.’
- d. *János-t mondtad, hogy az (?ő) órája elveszett.*
 John-ACC said.2SG that(c) the his watch.POSS.3SG lost
 ‘(Of) John you said that his watch was lost.’

If an embedded indefinite object is fronted to the matrix object position, the main verb may show definite conjugation, triggered by the embedded clause (which counts as a definite object), or it may show indefinite conjugation corresponding to the fronted object. In this case, the conjugation distinguishes between two possible configurations.

- (61) *Két almát mondtad/ mondtál, hogy vettél.*
 two apple.ACC said.2SG.DEF said.SG.INDEF that(c) bought.2SG.INDEF
 ‘(Of) two apples you said that you bought.’

It has not been noted in the literature that the case-mismatch presented in (60) is actually not restricted to main clause objects and accusative-marking. As discussed in section 3.2, subordinate clauses may also be associated with inherently case-marked pronouns. The fronted element may also replace such a pronoun, assuming its case.

- (62) a. *Arról gondolkodtam, hogy János jön a partira.*
 that.from thought.1SG that(c) John comes the party.to
 ‘I was thinking about John coming to the party.’
- b. *Jánosról gondolkodtam, hogy jön a partira.*
 John.from thought.1SG that(c) comes the party.to
 About John I was thinking that he is coming.
- c. *Attól tartok, hogy János jön a partira.*
 that.from be.afraid.1SG that(c) John comes the party.to
 ‘I’m afraid of John coming to the party.’

⁵³ Since Hungarian is a pro-drop language, personal pronouns are avoided if independent factors do not necessitate them, hence the marked nature of (60b-d). Adding emphatic stress to the resumptive improves such sentences.

- d. *Jánostól tartok, hogy jön a partira.*
 John.from be.afraid.1SG that(c) comes the party.to
 ‘Of John I was afraid that he is coming.’

The second interesting variation is about agreement matching/mismatching, which was first observed by Gervain (2002): if the fronted constituent that has been case-switched contains a quantifier or a numeral, the embedded verb may show either singular or plural agreement. This is unexpected, since such elements in a simple (non-subordinated) sentence invariably trigger singular agreement, as in (63d). This is shown with a fronted embedded subject bearing accusative case in (63a-b). Oblique-marked case-switched elements pattern similarly as shown by (63c) below. If the fronted element bears its original nominative case, only singular agreement is permitted, see (63e). From this it seems that it is the case-switching that enables the number-mismatch.

- (63) a. *Az összes lány-t mondtad, hogy jön.*
 the each girl-ACC said.2SG that(c) come.3SG
 ‘(Of) each girl you said that they are coming.’
- b. *Az összes lány-t mondtad, hogy jönnek.*
 the each girl-ACC said.2SG that(c) come.3PL
- c. *Az összes lányról gondolkodtam, hogy jön/ jönnek.*
 the each girl-from thought.1SG that(c) come.3SG come.3PL
 ‘(Of) every girl I was thinking that they come.’
- d. *Az összes lány jön / *jönnek.*
 the each girl come.3SG come.3PL
- e. *Az összes lány mondtad, hogy jön / *jönnek.*
 the each girl.NOM said.2SG that(c) come.3SG come.3PL

Thus, taking the two axes of variation with fronted elements (original/switched case, matching/mismatching agreement), there are four possible combinations, out of which the one with original case and mismatching agreement is impossible. As for the three remaining combinations, the question is whether people judge them uniformly or if there is some dialectal/idiolectal variation.

According to Gervain (2002), there are two groups of speakers. One group accepts both matching (singular in 63a) and mismatching (plural in 63b) agreement on the embedded verb, but only switched case (accusative in 63a) on the fronted element. That is, they accept sentences like (63a) and (63b). The other group accepts both original (nominative) and switched (accusative) case, but only with matching (singular) agreement, so they accept sentences like (63a) and (63c). Jánosi (2013) debates this and claims that while there could be some speaker variation, no consistent speaker groups can be distinguished. My general outlook will be more along the lines of Jánosi, so OF will be mostly treated as a unitary phenomenon with regard to acceptability patterns. A few remarks with respect to an account for potential dialects will be made nevertheless in section 4.2.2.

Let us examine a number of other differences between “case-switched” fronted elements and one bearing their original case. Gervain (2002) shows that the case-switched version may bypass island-violations, for example the complex-NP-constraint:

- (64) a. *Az elnök-öt mondtad, hogy hallottad a hírt, hogy jön.*
 the president-ACC said.2SG that(c) heard.2SG the news.ACC that(c)
 comes
 ‘(Of) the president you said that you heard the news that he will come.’
- b. **Az elnök mondtad, hogy hallottad a hírt, hogy jön.*
 the president.NOM said.2SG that(c) heard.2SG the news.ACC that(c)
 comes

Another example for this is shown in (65). This is based on the analysis of subordinate clauses where in presence of an associate pronoun, the CP itself is an adjunct and thus an island for extraction.

- (65) a. *János-t mondtad, hogy azt hallottad, hogy jön.*
 John-ACC said.2SG that(c) that.ACC heard.2SG that(c) comes
 ‘(Of) John you said that you heard that he is coming.’
- b. **János mondtad, hogy azt hallottad, hogy jön.*
 John.NOM said.2SG that(c) that.ACC heard.2SG that(c) comes
 ‘(Of) John you said that you heard that he he is coming.’

Also, Gervain (2009) shows that no strong crossover effects⁵⁴ are present in the case-switched version, in contrast to the case-retaining one.

- (66) a. *Két szomszéd-ot panaszoltál egymásnak, hogy zajos/ zajosak.*
 two neighbor-ACC complained.2SG each other.to that(c) noisy.SG noisy.PL
 ‘(Of) Two neighbor you complained about each to each other that they are noisy.’
- b. **Két szomszéd panaszoltad egymásnak, hogy zajos.*
 two neighbor.NOM complained.2SG each other.to that(c) noisy.SG

Furthermore, Gervain (2009) observes that the opposite pattern may be observed with reconstruction-effects: it yields ungrammaticality with the (switched) accusative version, but does not do so if the case is retained.

⁵⁴ Strong crossover (SCO) is the label for the phenomenon where a movement of a wh-phrase “crosses over” a pronoun that binds it.

(i) **Who_i does he_i think you saw t_i?*

- (67) a. **Két rossz hír-t egymásról mondtál, hogy letaglózta a fiúkat.*
 two bad news-ACC each other.about said.2SG that(c) saddened.3SG the boys.
 ‘Two pieces of bad news about each other you said that saddened the boys.’
- b. *Két rossz hír egymásról mondtad, hogy letaglózta a fiúkat.*
 two bad news.NOM each other.about said.2SG that(c) saddened.3SG the boys.
 ‘(Of) two pieces of bad news about each other you said that saddened the boys.’

Fronting idiom-chunks is impossible with the accusative version, and marginally possible with the nominative, as per (68)

- (68) a. **A gépszíj-at mondtad, hogy elkapta Jánost.*
 the driving.belt-ACC said.2SG that(c) caught.3SG John.ACC
 ‘(Of) the driving belt you said that it had caught John./ *John has to work a lot lately.’
- b. ??*A gépszíj mondtad, hogy elkapta Jánost.*
 the driving.belt.NOM said.2SG that(c) caught.3SG John.ACC
 ‘(Of) the driving belt you said that it caught John./ John has to work a lot lately.’

In all the examples considered so far, the fronted element was preverbal. However, it is also possible to place a case-switched fronted element postverbally, as in (69b-c). Then there is a verum focus on the verb. As discussed in 3.2, there is a similar pattern with the associated pronouns of standard subordinate clauses, with which fronted elements are in complementary distribution, see (69a). Such variation is impossible with non-case-switched fronted constituents, shown in (69d).

- (69) a. ?*De hiszen te MONDTAD azt, hogy jön János.*
 but you said.2SG that that(c) comes John
 ‘But you DID say that John is coming.’
- b. ?*De hiszen te MONDTAD János-t, hogy jön.*
 but you said.2SG John-ACC that(c) comes
 ‘But (about) John, you DID say that he is coming.’
- c. ?*De hiszen te tartasz Jánostól, hogy jön.*
 but you be.afraid.2SG John.from that(c) comes
 ‘But you ARE afraid of John, that he is coming.’
- d. **De hiszen te MONDTAD János, hogy jön.*
 but you said.2SG John.NOM that(c) comes

In the cases where a fronted subject bears a switched accusative case in the matrix clause, the main verb was *mond* ‘say’. This was done for expository reasons, as this is the most commonly used verb in this construction. But this type of OF is obviously not restricted to *mond*. At the

very least, it is assumed to work with so-called bridge verbs. According to Kenesei (1992), the licensing verbs are *akar* ‘want’, *szeretne* ‘would like’, *mond* ‘say’, *hisz* ‘believe’, *képzél* ‘imagine’, etc. In É. Kiss (2002: 253) we find “*akar* ‘want’, *szeretne* ‘would like’, *kell* ‘need’, *szabad* ‘may’, *lehet* ‘is possible’, *nyilvánvaló* ‘is obvious’, *valószínű* ‘is likely’, as well as verbs of saying and verbs denoting mental activities, among them *mond* ‘say’, *ígér* ‘promise’, *állít* ‘claim’, *gondol* ‘think’, *hisz* ‘believe’, etc.”

Before investigating the list of verbs more extensively, a baseline generalization can be made: the verb must be one that occurs with a subordinate clause, associated with a pronoun.

- (70) a. **Azt futottuk, hogy János megerősödjön.*
 that ran.1PL that(c) John become.stronger.SBJV.3SG
 ‘We ran every day so that John would become stonger.’
- b. **János(t) futottuk, hogy megerősödjön.*
 John(.ACC) ran.1PL that(c) become.stronger.SBJV.3SG

What the main accounts in the literature (Kenesei 1992, Lipták 1998, É. Kiss 2002, Gervain 2002) tend to agree on is that it is the set of bridge verbs that licenses OF, only differing in which particular examples they highlight.

Let us explore the landscape in more detail. As a starting point, let us take a look at the set of bridge verbs as listed by Kálmán (2001), which is the largest list in the literature for Hungarian. Kálmán’s (2001) criterion is that a verb is a bridge verb if it can occur in a construction labelled “fake question-word question” by Kálmán (2001). It is a construction where two question words are present: *mit* ‘what.ACC’ in the main clause and another one in the subordinate clause. According to Kálmán (2001), the construction is triggered by a locality condition: in Hungarian, the question word in front of a main verb is always a local complement of the main verb. From this it follows that extraction, should be impossible in Hungarian, as illustrated in (71a). The “fake question-word question”-construction is thus a way to bypass this restriction (71b-c). Although I do not think either the locality condition or the subsequent analytical idea of “fake-question words” is a proper characterization of the state of affairs, let us look at the data nevertheless. I restrict the set of verbs to those relevant for OF, that is, to those which could occur with an accusative associate pronoun. All the grammaticality judgments in (71) are from Kálmán (2001).

- (71) a. **Ki mondtad, hogy jön a partira?*
 who said.2SG that(c) comes the party.to
 ‘Who did you say will come to the party?’
- b. *Mit mondtál/ gondolsz/ állítasz/ álmodtál/ ?sejtesz/*
 what.ACC said.2SG think.2SG claim.2SG dreamed.2SG suspect.2SG
 **fontolgatsz/ ígértél, hogy ki jön a partira?*
 contemplate.2SG promised.2SG that(c) who comes the party.to
 ‘What did you say/think/claim/dream/suspect/contemplate/promise, who is coming to the party?’

- c. *Mit szeretnél/ akarsz/ javasolsz/ tanácsolsz/*
 what.ACC would.like.to.2SG want.2SG suggest.2SG advise.2SG
parancsolsz, hogy ki jöjjön a partira?
 order.2SG that(c) who come. SBJV.3SG the party.to
 ‘What do you like/want/suggest/advise/order, who should come to the party?’

So according to Kálmán (2001), the set of bridge verbs in Hungarian includes: *mond* ‘say’, *gondol* ‘think’, *állít* ‘claim’, *álmodik* ‘dream’, *sejt* ‘suspect’, *szeretne* ‘would like to’, *akar* ‘want’, *javasol* ‘suggest’, *tanácsol* ‘advise’ and *parancsol* ‘command’.

At first sight, there are some unexpected items on this list. While *mond* ‘say’, *gondol* ‘think’ and *állít* ‘claim’ are core members of the set of bridge-verbs cross-linguistically, *álmodik* ‘dream’, *tanácsol* ‘advise’ and *parancsol* ‘command’ are not.

- (72) a. **Who did you dream that you gave a flower to?*
 b. **Who did you advise/command that we should invite?*

Also, I would like to express doubts about Kálmán’s (2001) locality condition: it seems to me that (71a) is a grammatical OF-sentence. The contrast is particularly stark if it is compared to (73):

- (73) **Ki álmodtad, hogy jön a partira?*
 who dreamed.2SG that(c) comes the party.to
 ‘Who did you dream will come to the party?’

As for the other judgments in (71b-c), I would like to express doubt for these also, for all of the verbs seem to me to be grammatical in the sentences. I think the reason for this is that the “fake question-word question” is an alternative strategy which circumvents extraction, so it is possible even with nonbridge verbs. Since nothing is extracted in a “fake question-word question”, it is hard to see what the relevance of the bridge quality could be in those cases. In such sentences, the OBJ argument of the predicate is the question word *mit* ‘what.ACC’.

What happens if we investigate the verbs along the lines of proper extraction? It turns out that the already introduced distinction between elements bearing their original case and “case-switched” elements surfaces again. When we look at nominative fronted subjects, we get the cross-linguistically standard set of bridge verbs, as shown in (74a-b). However, accusative-marked fronted subjects occur with a much wider range of verbs, as per (74c-d). In other words, the “case-switched” OF is less lexically restricted.

- (74) a. *Ki mondtad/ gondoltad/ állítottad/ remélted/ *álmodtad/*
 who said.2SG thought.2SG claimed.2SG hoped.2SG dreamed.2SG
**fontolgattad/ *furcsállottad/ *kézelted/ *sérelezted, hogy*
 contemplated.2SG found.strange.2SG doubted.2SG resented.2SG that(c)
jön a partira?
 comes the party.to
 ‘Who did you say/think... will come to the party?’

- b. *Ki akarod/ szeretnéd/ javaslod/ *parancsolod/ *tanácsolod,*
 who want.2SG would.like.2SG suggest.2SG order.2SG advise.2SG
hogy jöjjön a partira?
 that(c) come..SBJV.3SG the party.to
 ‘Who do you want/would you like... that (s)he should come to the party?’
- c. *Kit mondtál/ gondoltad/ állítottad/ reméltél/ *álmodtál/*
 who.ACC said.2SG thought.2SG claimed.2SG hoped.2SG dreamed.2SG
fontolgattál/ furcsállottál/ ?kételtél/ ?sérelemezteél, hogy
 contemplated.2SG found.strange.2SG doubted.2SG resented.2SG that(c)
jön a partira?
 comes the party.to
 ‘Who did you say/think... will come to the party?’
- d. *Kit akarsz/ szeretnél/ javasolsz/ *parancsolsz/*
 who.ACC want.2SG would.like.2SG suggest.2SG order.2SG
?tanácsolsz, hogy jöjjön a partira?
 advise.2SG that(c) come. SBJV.3SG the party.to

A similar distribution may be obtained if we use indicative OF examples. That is, OF with nominative-marked fronted subjects (original case) is grammatical with the set of cross-linguistically attested bridge verbs (see 75a-b), while the accusative version (switched case) shows greater flexibility (see -d).

- (75) a. *János mondtad/ gondoltad/ ??állítottad/ remélted/ *álmodtad/*
 John said.2SG thought.2SG claimed.2SG hoped.2SG dreamed.2SG
**utáltad *fontolgattad/ ?furcsállottad/ *kételted/ *sérelemezteél,*
 hated.2SG contemplated.2SG found.strange.2SG doubted.2SG resented.2SG
hogy jön.
 that(c) comes
 ‘(Of) John you said/thought/... that he is coming.’
- b. *János akarod/ szeretnéd/ javaslod/ *parancsolod/ *tanácsolod,*
 John want.2SG would.like.2SG suggest.2SG order.2SG advise.2SG
hogy jöjjön.
 that(c) come. SBJV.3SG
 ‘(Of) John you want/would like/... that he should come.’
- c. *Jánost mondtad/ gondoltad/ ?állítottad/ remélted/ *álmodtad/*
 John.ACC said.2SG thought.2SG claimed.2SG hoped.2SG dreamed.2SG
??utáltad fontolgattad/ furcsállottad/ ?kételted/ ?sérelemezteél,
 hated.2SG contemplated.2SG found.strange.2SG doubted.2SG resented.2SG
hogy jön.
 that(c) comes
 ‘(Of) John you said/contemplated/found it strange/doubted/resented that he is coming.’

- d. *Jánost akarod/ szeretnéd/ javaslod/ *parancsolod/ ??tanácsolod,*
 John.ACC want.2SG would.like.2SG suggest.2SG order.2SG advise.2SG
hogy jöjjön.
 that(c) come. SBJV.3SG
 ‘(Of) John you want/would like/suggest/order/advise that he should come.’

This is not to say that the accusative version is completely free in its distribution. For instance, in the examples above, *álmodik* ‘dream’ or *parancsol* ‘order’ rejects it even when the fronted element bears a switched case.

It seems then that the case-retaining version of OF (where the fronted element bears the case that it would also have in its original position, e.g. 74a-b, 75a-b, but also oblique elements in earlier examples like 62) generally patterns with the prototypical bridge verbs.

With case switching accusative OF (where the fronted element becomes the OBJ/OBL_θ of the main clause, regardless of its original function), the state of affairs is more delicate. There seems to be a core set of verbs, with which case-switching OF is flawless. These are *mond* ‘say’, *gondol* ‘think’, *ígér* ‘promise’ *kérdez* ‘ask’, *remél* ‘hope’, also *beszél* ‘speak’ and *mesél* ‘tell’. Even this is larger than the set of bridge verbs, licensing the nominative version OF.

After these comes a big “grey zone”, i.e. verbs that are marked but acceptable to a considerable extent. Examples for this are *fontolgat* ‘contemplate’, *sejt* ‘surmise’, *tud* ‘know’, *sajnál* ‘pity’, *tanácsol* ‘advise’, *gyanít* ‘suspect’, etc. Although there could be differences in acceptability (also between speakers), these are more or less acceptable in accusative OF.

Finally, some verbs which otherwise occur with a subordinate clause are poor in OF. Examples for this group are *álmodik* ‘dream’, *hazudik* ‘lie’, *válaszol* ‘answer’, *felel* ‘reply’, *terjeszt* ‘spread (the news)’, *utál* ‘hate’, *parancsol* ‘order’, *érez* ‘feel’, *nyilatkozik* ‘speak to the press’, as well as manner of speaking verbs, like *suttog* ‘whisper’ or *ordít* ‘shout’.

Instances where the case-switch happens with an inherent oblique case seem fairly unrestricted.

- (76) a. *Abban bízok, hogy János jön.*
 that.in trust.1SG that(c) John comes
 ‘I trust that John will come.’
 b. *Jánosban bízok, hogy jön.*
 John.in trust.1SG that(c) comes
 ‘(Of) John I trust that he will come.’
- (77) a. *Attól tartok, hogy János jön.*
 that.from be.afraid.1SG that(c) John comes
 ‘I am afraid that John will come.’
 b. *Jánostól tartok, hogy jön.*
 John.from be.afraid.1SG that(c) comes
 ‘(Of) John I’m afraid that he will come.’
- (78) a. *Gratulálok ahhoz, hogy megszületett a gyereked.*
 congratulate.1SG that.to that(c) was.born.3SG the child.POSS.2SG
 ‘I congratulate you on the birth of your child.’

- b. *Gratulálok a gyerekedhez, hogy megszületett.*
 congratulate.1SG the child.POSS.2SG.to that(c) was.born.3SG
 I congratulate you on your child, on it being born.
- (79) a. *Azzal számolok, hogy János nyer.*
 that.with count.1SG that(c) John wins
 ‘I expect that John will win.’
- b. *Jánossal számolok, hogy nyer.*
 John.with count.1SG that(c) wins
 ‘Of John I expect that he will win.’

What I would like to conclude from this discussion is that the literature’s original claim that OF is restricted to bridge verbs, cannot be maintained. There is a much wider range of verbs which participate in OF, although the distribution is not completely free, especially with regards to switching to accusative case.

A final syntactic point to be discussed is the nature of the complementizer in OF. Remember from section 2.2.1 that in standard subordinate clauses, if the focus of the sentence is a pronoun, *hogy* ‘that(C)’ may be absent, as in (80). An interrogative subordinate clause makes the complementizer optional even in the case of nonpronominal foci, as in (81).

When OF takes place, the complementizer is always compulsory if the embedded clause is indicative as per (82)-(83). When case-switching takes place, the complementizer seems obligatory even if the subordinate clause is interrogative (84a), in contrast with (81). Omission of the complementizer seems better if the fronted element retains its original case (84b-c).

- (80) *Azt mondtad, Párizsba mész.*
 that said.2SG Paris.into go.2SG
 ‘You said that you will go to Paris.’
- (81) *[Jánostól]_{IF/CF} kérdeztem, hova megy.*
 John.from asked.1SG where goes.
 ‘I asked John where he is going.’
- (82) **Párizsba mondtad, mész.*
 Paris.into said.2SG go.2SG
 ‘To Paris you said that you will go.’
- (83) **János-(t) mondtad, jön.*
 John-ACC said.2SG. comes
 ‘(Of) John you said that he will come.’
- (84) a. **Jánost kérdezted, jön-e.*
 John.ACC asked.2SG. comes-QUESTION PRT
 Intended: ‘(Of) John you asked if he will come.’
- b. *?János kérdezted, jön-e.*
 John.ACC asked.2SG. comes-QUESTION PRT
 Intended: ‘(Of) John you asked if he will come.’

- c. *?Párizsba kérdezted, megyek-e.*
 Paris.into asked.2SG go.1SG-QUESTION PRT
 Intended: ‘To Paris you asked me if I go come.’

However, remember that the fronted element may be associated with something embedded at a distance, not just with a grammatical function in an immediately subordinated clause (see 85). When that happens, it is only the first complementizer that is obligatory, the others are droppable along the general restrictions. This has gone unnoticed in the previous literature.

- (85) a. *János(t) mondtad, *(hogy) gondolod, (hogy) jön.*
 John(.ACC) said.2SG that(C) think.2SG that(c) comes
 ‘(Of) John you said that you think that he will come.’
- b. *János(t) mondtad, *(hogy) Maritól kérdezték, (hogy) jön-e.*
 John(.ACC) said.2SG that(C) Mary.from asked.3PL that(c) comes-Q-PRT
 ‘(Of) John you said that they asked Mary whether he will come.’

As regards information structure, it seems clear that the fronted elements bear some kind of discourse prominence, regardless of the potential case-switch. This translates as them being interpreted as either information focus (IF), Contrastive focus (CF) or contrastive topic (CT). One can rightly observe that these are exactly the discourse functions that the pronouns associated with subordinate clauses bear.

While this basic picture is true regardless of the type of the OF, there are certain tendencies with regards to the preferred discourse functions. If a subject is fronted and it retains its original (nominative) case, speakers judge the sentence to be more felicitous if the fronted subject is a contrastive topic. That is, everything else being equal, (86a) is preferred over (86b). In fact, (86b) is often rejected by native speakers. However, the same speakers judge (86a) much more favorably (for discussion, see, Szűcs 2013), which suggests that it is not extraction *per se* that degrades (86b), but the information structure. Here capitalization and the translations reflect the different information structures.

- (86) a. *[János]_{CT} MONDTAD, hogy jön.*
 John said.2SG that(c) comes
 ‘As for John, you DID say that he is coming.’
- b. *[JÁNOS]_{IF/CF} mondtad, hogy jön.*
 ‘It was John that you said that he is coming.’

No such difference can be observed in the case of fronted nonsubjects: (87a-b) are equally felicitous in the appropriate context.

- (87) a. *[Párizsba]_{CT} MONDTAD, hogy mész.*
 Paris-to said.2SG that(c) go.2SG
 ‘To London, you did say that you will go.’
- b. *[PÁRIZSBA]_{IF/CF} mondtad, hogy mész.*
 ‘It is to Paris that you said that you will go.’

Also no preference can be observed in the case-switched version of OF. (88), which minimally differs from (86b), is fully felicitous.

- (88) *[JÁNOST]_{IF/CF} mondtad, hogy jön.*
 John.ACC said.2SG that(c) comes
 ‘OF JOHN you said that he comes.’

In sum, the main information structural requirement about OF is that some discourse prominence should be attributed to the fronted element. This is plausible given that these structures deviate from canonical sentences, where no such dependencies are established. Accordingly, there should be some motivation for the additional syntactic processes. As for the CT-preference in the case of fronted subjects in nominal case, I am going to argue that independent factors are responsible for it.

4.2.2 Analysis of operator fronting in Hungarian

Let us repeat the basic examples for OF, for which an analysis is to be presented in this section.

- (89) a. *János(-t) mondtad, hogy jön.*
 John(-ACC) said.2SG that(c) comes
 ‘(Of) John you said that he will come.’
- b. *János-ban bízok, hogy jön.*
 John.in trust .1SG that(c) comes
 ‘In John I trust that he will come.’
- c. *Párizsba mondtad, hogy mész.*
 Paris.into said.2SG that(c) go.2SG
 ‘To Paris you said that you will go.’

Recent analyses of Operator Fronting (Gervain 2002, Coppock 2003, Jánosi 2014) recognize that the data presented in the previous section lend themselves for two distinct analytical strategies, of which I give a brief overview here, laying more emphasis on Coppock’s (2003) LFG-analysis.

The first strategy is for constellations where no case-switching happens, like the nominative version of (89a), where the subject of the embedded clause is fronted or (89c), where an oblique argument occupies the fronted position. In such scenarios, the fronted element actually belongs to the embedded clause so the matrix position is the result of some syntactic process. In Chomskyan terms, this means that the fronted element is “generated” in the subordinate clause and then moves into the matrix position. Since agreement between the fronted element and the embedded predicate is established locally before the movement, no number mismatch of the sort illustrated earlier may occur. The relevant data are repeated here for convenience. As the morphosyntactic rules of Hungarian clearly establish singular agreement between a quantified phrase and a predicate, as in (90a), this is carried over to the OF-sentence of (90b).

- (90) a. *Az összes lány jön / *jönnek.*
 the each girl come.3SG come.3PL
 ‘Each of the girls are coming.’
- b. *[Az összes lány]_i mondtad, hogy t_i jön / *jönnek.*
 the each girl.NOM said.2SG that(c) come.3SG come.3PL
 ‘(Of) each of the girls you said that they are coming.’

The second strategy that the fronted element is “base-generated” in the main clause. Here, since the fronted constituent is directly related to the main predicate, it may assume the case that the relevant predicate is able to assign. Association with the embedded predicate is established via some anaphoric process, e.g. resumption. Number variation may occur because of the nonsyntactic nature of the association process (more on this below). This results are sentences like (63a-b), repeated here as (91).

- (91) *Az összes lány-t mondtad, hogy jön/ jönnek.*
 the each girl-ACC said.2SG that(c) come.3SG come.3PL
 ‘(Of) each of the girls you said that they are coming.’

Coppock (2003) is an adaptation of Gervain’s (2002) minimalist account to LFG. The primary concerns for her account are fronted subjects. She subscribes to the expletive-associate-chain analysis of subordinate clauses (Kenesei 1992/1994) and consequently she analyzes fronted accusative-marked subjects as athematic arguments of the main verb, in the same way that an expletive pronoun is athematic. Therefore, she makes the claim that the lexical entry of a verb participating in OF looks like (92), where the object function is located outside the angle brackets, indicating it being a non-thematic argument.

- (92) *mond* <(SUBJ)(COMP)> (OBJ)

The strategies are modelled by positing that the groups use different mechanisms to unify the athematic matrix object with the appropriate thematic function in the embedded clause: the movement-group uses functional identification, whereas the resumption-group utilizes anaphoric identification. Recall from section 1.2 that the essential difference between the two mechanisms is that functional identification means strict f-structural identity, so a single element is shared by two sub-f-structures, while anaphoric identification is a looser, semantic kind of relationship, a coreference of two distinct elements. On the one hand, functional identification requires fully identical grammatical features, so no mismatching agreement is allowed downstairs. On the other hand, anaphoric identification may allow for a number mismatch. The two approaches are schematically illustrated in Figure 2a and 2b (for simplicity, the discourse functions have been omitted).

- (93) *Az összes lányt mondtad, hogy jön/ jönnek.*
 the each girl.ACC said.2SG that(c) come.3SG come.3PL
 ‘Every girl you said that they will come.’

- (95) *Egy kutyát mondtad, hogy láttál.*
 one dog.ACC said.2SG.DEF that(c) saw.2SG.INDEF
 ‘A dog you said that you saw.’

Here *egy kutyát* ‘a dog.ACC’ would be in Spec/FP in Coppock’s (2003) terminology but it would not be the object of the main verb, contrary to the annotation in (94). We may know this because *egy kutyát* ‘a dog’ is indefinite, but *mondtad* ‘said.2SG.DEF’ is in definite conjugation. This definite conjugation is triggered by the clause, which acts as the object, as outlined in section 2.2.

There are some empirical problems as well. The account also leaves the phenomenon of idiom chunks not being able to participate in case-switched OF unexplained. In comparable cases of English raising structures, idiom chunks are possible, since the idiomatic meaning is kept intact via functional identification.

- (96) *I believe the cat to be out of the bag.* (“I believe that the secret is revealed.”)

One might argue that it is only because the fronted element in OF gets a focus discourse function that idiom chunks are ungrammatical. However, if the fronted subject is postverbal, according to standard assumptions about Hungarian, the fronted constituent would not get a topic or focus discourse-function. Still, case-switched OF is unacceptable.

- (97) *#De hiszen te mondtad a gépszíjat, hogy elkapta*
 but you said.2SG the driving belt.ACC that(c) caught.3SG
Jánost.
 John.ACC
 ‘But you said of the driving belt that it has caught John. / *But you said that John has had to work a lot lately.’

Also, it must be noted that this account is built on the idea that the associate pronoun of subordinate clauses in Hungarian is a non-thematic element, an expletive. However, as I have demonstrated in 2.2.2, a number of factors argue against this view.

Nevertheless, the fundamental objection that may be raised against Coppock (2003) is the downgrading of Semantic Coherence. The well-formedness conditions are fundamental in LFG, so modifying them to be violable seems a daring move, especially considering that this is done on the basis of a single construction as no other phenomena has been shown to violate the (Semantic) Coherence condition. If an alternative solution exists, it should be preferred. The alternative is regarding the object as thematic in Figure 5, and this is exactly the analytical idea that is going to be adopted.⁵⁵

I accept the basic premise that two analytical avenues are required for the analysis of OF:

⁵⁵ From my perspective, there is also a c-structure related issue with Coppock (2003). Although this is not directly related to the analysis of OF, it might be worth mentioning that Coppock (2003) posits an FP (Focus Phrase) functional category. However, this is not justified in LFG for Hungarian, as there is no lexical item invariably associated with the F head, so the Economy of Expression is violated. As was shown in 3.1, a satisfactory account can be achieved without such projections if focused elements are assumed to be in Spec/VP, following (Laczkó) 2014.

- An approach with a direct long-distance dependency between the fronted element and the embedded predicate. Let us use the label “LDD-OF” for this.
- An approach where the fronted element is primarily an associate of the main predicate and there is some anaphoric link to the embedded clause. Let us use the label “proleptic OF” for this.

Of these two cases, the LDD-OF has the more straightforward analysis. These cases are “fronting” constructions in a more traditional sense, so *János* ‘John.NOM’ and *Párizsba* ‘to Paris’ in (89) functionally and thematically belong to the embedded clause, the only trait that ties them to the main clause is their position. Such configurations are sensitive to factors that are linked to movement-like dependencies in Chomskyan frameworks: they are ungrammatical if the embedded function associated with the fronted element is in an island, they are also bad in crossover configurations, they show reconstruction effects and the main verb must be one that has the bridge property. Although there is no such thing as movement in LFG, the framework of this monograph, functional identification is a tool that is a close nonderivational parallel to it, as already seen with TOP.

Proleptic OF, where some element associated with an embedded function actually becomes the object or an oblique of the main verb, is more complex from a theoretical point of view. In many respects its properties are the direct opposites of the case-retaining version: it is unaffected by islands, crossover and by the non-bridge quality of the main verb and it shows no reconstruction effects. In Chomskyan frameworks, such properties have warranted a base-generation analysis (as in Gervain 2002). Here, in an LFG framework, I argue that this version of OF should be viewed as “prolepsis” – a type of dependency posited for a number of similar phenomena in the literature. I will give a typological and a theoretical overview of proleptic and related structures in Chapter 5, here let us just state the basic definition as worded by Salzmann (2017: 1): prolepsis is a construction where a structural complement of the matrix verb is semantically related to the predicate of a finite embedded clause. This definition leaves the details of the analysis unspecified, and rightly does so, since beyond crucial similarities there are also important differences in the cross-linguistic realizations of prolepsis. In the case of Hungarian, my position is that the proleptic constituent should be analyzed as a thematic argument of the main verb, which is associated with the embedded function via obligatory anaphoric identification.

Let us now go into the analysis of OF. As discussed earlier, fronted elements in OF may bear a range of discourse functions: information/contrastive focus and contrastive topic, although there are preferences for certain discourse functions depending on the function of the fronted constituent. Nominative-marked phrases tend to be contrastive topics, while accusative and oblique-marked elements prefer a focus interpretation. What this means for is that in OF, the fronted elements may occur under the S node in the topic field or in Spec/VP.

As claimed in the previous section, in cases where no “case-switch” takes place, the dependency between the fronted element and the function associated with it in the embedded clause should be analyzed as functional identification, much like in the case of TOP-Eng. These are the cases where a subject, an object or an oblique argument is fronted and the original case is retained. To model this, we must add the following optional specifications to the relevant nodes:

These annotations encode that a constituent occupying these positions can be identical with some grammatical function in an embedded clause. Following the convention adopted in 3.1, I label the path to the embedded function as OFPATH and its details are shown in (99).

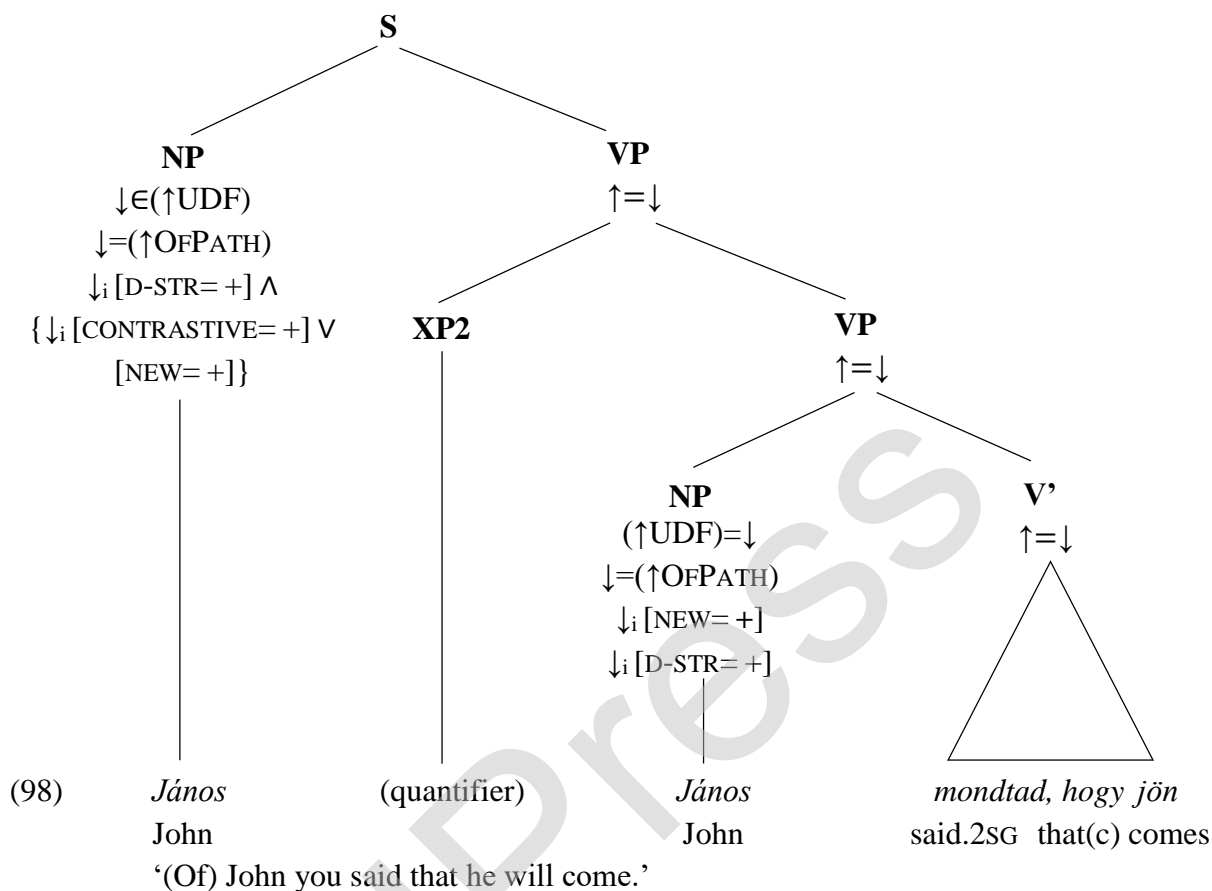


Figure 3.

C-structure for OF in case-retaining scenario

- (99) OFPATH \equiv {OBJ | SUBJ}⁺ GF
(→TENSE)
(→LDD \neq -)

(99) indicates that the fronted element may have any grammatical function in an arbitrarily deeply embedded complement clause. Recall from 3.2 that complement clauses are subject or object arguments of the matrix predicates if there is no associate pronoun. The ⁺ notation is called a “Kleene-plus” meaning “at least one”, so it ensures that embedding does take place. As they are finite clauses, these objects will have tense features, as opposed to scenarios where the SUBJ or OBJ is a lexical noun or pronoun associated with an embedded clause (where the clause itself is an ADJ). In such cases, OF will not take place so CNPC- and adjunct-islands will be ruled out. The annotation also requires every predicate along the path to have a positively specified LDD feature (that is, the predicates must have bridge-quality).

In accordance with the interpretation of OF, the appropriate features are contributed to these elements at information structure. This ensures that, depending on their position, the Operator fronted constituents will be interpreted as IF/CF/CT.

As noted in the previous section, it is possible that there is a group of people who do not accept fronted nominative constituents (Gervain 2002 vs. Jánosi 2013). If this is indeed the case, these speakers' mental grammar lack the equation in (99), so they derive OF with another strategy (proleptic OF), which is to be discussed below.

As discussed, speakers may vary according to their preferences of these positions. The general pattern is that nominative fronted embedded subjects are strongly preferred as contrastive topics while nonsubject functions may be any D-STRUCTURING category except for neutral Topics. I think that rather than positing syntactic restrictions, invoking general mapping principles between grammatical function and IS-function could give us a simpler explanation for this pattern. There is a strong cross-linguistic tendency for the association of subjects and topicality, since subjects are default topics. As such, it is no surprise that even when fronted they prefer configurations in which they are interpreted as topic-like entities. Since the neutral topic category is unavailable for fronted elements, the contrastive topic discourse function is the closest match. Other grammatical functions do not possess this association, so they can easily accommodate either discourse function. I offer no formal treatment of these preferences here but combining LFG with Optimality Theory could be a fruitful research avenue for this topic. Such a treatment of the issue would allow for the specification of constraints expressing the preference patterns, so a mapping like SUBJ-CT would be more optimal than SUBJ-IF/CF. Taking the focus-interpretation as an example, Figure 4 shows the f-structure corresponding to (98).

In accordance with our discussion in 3.2, I take *mond* 'say' and similar verbs to subcategorize for a SUBJ and an OBJ grammatical function. The OBJ can be a lexical DP/NP, a pronoun, a bare S or a CP-clause, as is the case is in this particular example.

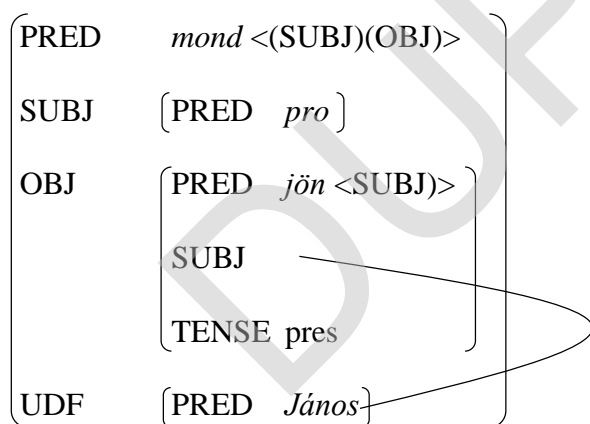


Figure 4.

F-structure corresponding to (98)

However, if the object argument is expressed as a pronoun, the clause itself is an adjunct. As such, it is an island, which means that functional identification into it is not licensed by (94). Accordingly, a pronoun cannot surface here, as (100) shows.

- (100) **János mondtad azt, hogy jön.*
 John said.2SG that.ACC that(c) comes

Similarly, non-bridge verbs, having a negatively specified LDD-feature, block functional identification, as in (101).

- (101) **János álmodtam, hogy jön.*
 John dreamed.1SG that(c) comes
 Intended: ‘(Of) John I dreamt that he will come.’

Since the fronted element is fully present via functional identification in both the matrix and the embedded f-structure, reconstruction effects are explained. Also, since functional identification means complete syntactic identity, the lack of potential number-variation in the embedded clause is also expected. An example is repeated here for convenience.

- (102) *Az összes lány mondtam, hogy jön/ *jönnek.*
 the each girl said.1SG that(c) come.3SG come.3PL
 ‘(Of) each girl I said that they will come.’

Proleptic OF, in which the fronted element actually becomes the argument of the main predicate, requires more consideration, especially from a thematic, argument-structural point of view. This is because, in this case, the lexical entry for the relevant main predicate is augmented. Consider (103).

- (103) a. *Én Jánost mondtam, *(hogy jön).*
 I John.ACC said.1SG that(c) comes
 Intended: ‘(Of) John I said that he will come.’
 b. *Én Jánosban bizok, *(hogy jön).⁵⁶*
 I John.in trust.1SG that(c) comes
 Intended: ‘About John I believe that he will come.’

Without context, the *that*-clause is not omissible. However, so far, I have treated *mond* ‘say’ and similar verbs as subcategorizing for two grammatical functions, a SUBJ and an OBJ. If this is the case, it seems strange that the *that*-clause is compulsory, given that *I* and *John* satisfy these requirements already. In order to account for this, I posit that via an argument-structural operation, the lexical entry of certain predicates (i.e. those that participate in this pattern) is altered, resulting in the modification of the aforementioned <(SUBJ)(OBJ)> frame to a frame like the one in (104).

⁵⁶ Since *bizik* ‘trust’ may easily take a bare OBL_θ argument (98b) is grammatical without the *that*-clause, but not in the relevant sense. Without the *that*-clause (98b) means that “I trust in John” (he is a trustworthy person), while the intended reading is that I have a trust in connection with him, that he will come. The distinction is easier to see if the OBL_θ is replaced with something that is not trustable in the first sense, like an inanimate object. In that case, the irrelevant reading is semantically anomalous, while the intended one is not.

- (i) *Én az új virágokban bizok, #(hogy megszépítik a kertet).*
 I the new flowers.in trust.1SG that(c) make.nice.3SG the garden.ACC
 ‘I have a trust regarding the new flowers, that they will make the garden nice.’

- (104) a. *mond* ‘say’, *gondol* ‘think’, *hisz* ‘believe’, etc. <(SUBJ)(OBJ)(COMP)>
 b. *bízik* ‘trust’, *tart* ‘be.afraid’, *számol* ‘expect’, etc. <(SUBJ)(OBL_θ)(COMP)>

In order to be able to analyze the modification of the argument structure of these predicates, we first have to make some notes on the initial mapping. I have adopted Kibort’s (2007) version of Lexical Mapping Theory (LMT) in section 1.2, so that is going to be the frame of the discussion here. For convenience, I repeat the featural decomposition of grammatical functions, the universal valency frame and the mapping principles here.

Table 1.

	-o	+o
-r	SUBJ	OBJ
+r	OBL _θ	OBJ _θ

Grammatical functions in LMT

- (105) Universal valency frame: -o/-r -r +o -o ... -o
 arg1 arg2 arg3 arg4 arg_n

- (106) **Mapping Principle:** The ordered arguments are mapped onto the highest (i.e. least marked) compatible function on the markedness hierarchy. (markedness: having positive feature-specification)

Markedness hierarchy: SUBJ (-o/-r) > OBJ (-r/+o), OBL_θ (-o/+r) > OBJ_θ (+o/+r)

Let us take a look at predicates requiring an object complement (89a). These verbs semantically subcategorize for an agent and theme/propositional argument (represented in 107a-a’). In principle, the <(SUBJ)(OBJ)> subcategorization can follow from two mappings of these arguments onto the valency template (107b-b’):

- | | | | | | | | | |
|----------|-----------|------|------|--|-----|-----------|------|------|
| (107) a. | predicate | ag | prop | | a’. | predicate | ag | prop |
| b. | | -r | +o | | b’. | | -o | -r |
| | | arg1 | arg3 | | | | arg1 | arg2 |
| c. | | SUBJ | OBJ | | c’. | | SUBJ | OBJ |

Either of these configurations is in harmony with the mapping principles and yields the desired outcome. However, there is evidence from nominalization that the mapping actually taking is the second one. The +o specification in (107a) would make it impossible to nominalize the object, as the resulting nominal would have the POSS function, which is -o, according to Laczkó (1995). However, this prediction is not borne out, as (108)-(109) show.

- (108) a. *Azt* *mondta*, *hogy* *tökéletes* *vagy*.
 that.ACC said.2SG that(c) perfect be.2SG
 ‘You said that you are perfect.’

- b. *Annak az állandó mondása, hogy tökéletes vagy, nem szép.*
 that.POS the constant say.NOUN that(c) perfect be.2SG not nice
 approx. ‘The constant saying of you being perfect is not nice.’
- c. *Azt bizonyították, hogy bűntény történt.*
 that.ACC proved.3PL that(c) crime happened
 ‘They proved that a crime had happened.’
- d. *Annak a bizonyítása, hogy bűntény történt, könnyű volt.*
 that.POS the prove.NOUN that(c) crime happened.3SG easy was
 ‘Proving that a crime had happened was easy.’
- (109) a. *Azt firtatták, hogy milyen a magánéletem.*
 that.ACC asked.3SG that(c) how the personal.life.POSS.1SG
 ‘They asked me about my personal life.’
- b. *Annak a firtatása, hogy milyen a magánéletem, bosszant.*
 that.POSS the ask.NOUN that(c) how the personal.life.POSS.1SG annoys
 ‘Getting asked about my personal life annoys me.’

The question is what happens that results in the modification of a subcategorization frame like (107a’) to something like (104). The fronted element becomes the OBJ of the main verb, but its behavior is different from the OBJ of the basic template. Unlike standard objects, it cannot be nominalized, compare (110-112) with (108-109).

- (110) a. *Jánost mondtad, hogy jön.*
 John.ACC said.2SG that(c) comes
 ‘(Of) John you said that he will come.’
- b. **János állandó mondása, hogy jön, nem szép dolog.*
 John constant say.NOUN that(c) comes not nice thing
 intended, approx.: ‘The constant saying of John that he will come is not a nice thing.’
- (111) a. *?Jánost bizonyították, hogy jön.*
 John.ACC proved.3PL that(c) comes
 ‘(Of) John they proved that he will come’
- b. **János bizonyítása, hogy jön, könnyű volt.*
 John prove.NOUN that(c) comes easy was
 intended, approx.: ‘Proving of John that he will come was easy.’
- (112) a. *Jánost firtatták, hogy jön-e.*
 John.ACC asked.3PL that(c) comes-QUESTION PRT
 ‘(Of) John they asked if he will come.’
- b. **János firtatása, hogy jön-e, bosszant engem.*
 John ask.NOUN that(c) comes-QUESTION PRT annoys me
 intended, approx.: ‘The asking of John if he comes annoys me.’

This suggests that the newly added object is specified as +o, which means that it cannot map to the POSS function. Now if the extra argument (tentatively labelled as having a subject matter

thematic role) is simply added to a base template as a +o argument, in accordance with the Mapping Principle, the following would be the outcome.

(113) a.	predicate	agent	proposition	subject matter
b.		–o	–r	+o
		arg1	arg2	arg3
c.		SUBJ	OBJ	OBJ _θ

However, (113) cannot be right, as it would predict the possibility of the simultaneous occurrence of a normal object and this added object. But this is not the way OF works. The newly added object can only co-occur with a clause, never with a pronoun or other nominal object.

(114)	<i>*Jánost</i>	<i>mondta</i>	<i>[a hírt]/</i>	<i>[azt, hogy jön].</i>
	John.ACC	said.2SG	the news.ACC	that.ACC that(c) comes

It follows that the added object is the sole OBJ argument of the predicate. If this is the case, the propositional argument cannot be an OBJ anymore. This is ensured if we posit that when the new argument is added, the proposition is not mapped to the second position of the universal valency frame, but instead moves to the fourth, –o position. This is possible in the Kibort (2007)-version of LMT, as morphosemantic alternations are modelled as semantic arguments realigning along the universal valency frame (see the discussion about the dative shift in section 2.4). The result is (115).

(115) a.	predicate	ag	subject matter	proposition
b.		–o	+o	–o
		arg1	arg3	arg4
c.		SUBJ	OBJ	COMP

It is interesting to note that Kálmán (2001:115) mentions some rare cases where such objects appear with standard oblique dependents.

(116) a.	<i>Merre</i>	<i>mondták</i>	<i>a</i>	<i>kocsmát?</i>
	what.direction	said.3PL	the	pub.ACC
	approx.: ‘What did they say of the pub, where it is?’, ‘Which way did they say the pub was?’			
b.	<i>Hova</i>	<i>gondolod</i>	<i>a</i>	<i>szekrényt?</i>
	to.where	think.2SG	the	wardrobe.ACC
	approx.: ‘What do you think of the wardrobe, where should it be put?’, ‘Where do you think the wardrobe should be put?’			

This might lend support for COMPs being analyzed as kinds of OBLs. Proleptic operator fronting standardly realizes this OBL function as a *that*-clause, but as (116) shows, standard oblique dependents are also possible sometimes.

Now let us take a look at verbs with oblique complements as in (89b). The initial mapping is shown in (117). Then a similar argument structural process takes place (118). A subject

matter argument is newly added to the argument position of the original oblique complement, and the propositional argument simultaneously moves to the fourth (–o) argument position.

- (117) a. predicate ag proposition
 b. –o +r
 arg1 arg2
 c. SUBJ OBL_θ
- (118) a. predicate ag subject matter proposition
 b. –o +r –o
 arg1 arg2 arg4
 c. SUBJ OBL_θ COMP

Another change that happens along with the addition of the new argument is that the lexical entry specifies these entities as obligatorily co-referent with some element in the embedded clause. Taking *mond* ‘say’ a prototypical verb of this class, its lexical entry is the following.

- (119) mond ‘say’ <(SUBJ)(OBJ)(COMP)>
 OBJ INDEX={COMP⁺ GF⁺} GF INDEX

As (119) indicates, the coreferent element may have any GF in the immediately following COMP, or it may be embedded somewhere at depth. Co-reference is indicated by the identity of the indices.

What has not been addressed so far is the nature of the extra argument. I propose that the phenomenon that has just been described is an instance of “prolepsis”: a construction where “a structural complement of the matrix verb is semantically related to the predicate of a finite embedded clause” (Salzmann 2017:1). The extra argument is referred to as a “proleptic” one. Theoretical and cross-linguistic perspectives on it will be provided this in the next chapter. Generally, it should be viewed as a lexical augmentation process, whereby the adicity of certain predicates is expanded with a new argument bearing a subject matter theta-role. I agree with Kotzoglou and Papangeli (2007), who, building on Pesetsky (1995) and Reinhart (2002), suggest that the proleptic argument is properly characterized as bearing the “subject matter” theta-role. (Kotzoglou and Panangeli 2007 concern themselves with a Greek proleptic construction. I will introduce that and other proleptic structures in the next chapter.)

As discussed in section 4.2.1, predicates can accept this extra proleptic argument to different degrees. Verbs subcategorizing for OBL_{θs} are generally free. This is probably because obliques are generally semantically closely related to the governing predicates. A reflex of this is their “+restricted” specification in LMT. The semantic connection with objects is weaker, so their occurrence in proleptic OF is more restricted. As a reminder, here is the basic picture:

- (i) Verbs that are entirely acceptable in prolepsis: *mond* ‘say’, *gondol* ‘think’, *ígér* ‘promise’, *kérdez* ‘ask’, *remél* ‘hope’, *beszél* ‘speak’ and *mesél* ‘tell’.
- (ii) Verbs that are moderately acceptable in prolepsis: *fontolgat* ‘contemplate’, *sejt* ‘surmise’, *tud* ‘know’, *sajnál* ‘pity’, *tanácsol* ‘advise’, *gyanít* ‘suspect’.

- (iii) Verbs that are ungrammatical in prolepsis: *álmodik* ‘dream’, *hazudik* ‘lie’, *válaszol* ‘answer’, *felel* ‘reply’, *terjeszt* ‘spread the news’, *parancsol* ‘order’, *érez* ‘feel’, *suttog* ‘whisper’, *ordít* ‘shout’.

Here, the question of whether we can find an explanation or at least motivation for this pattern arises. Although I think it ultimately boils down to idiosyncratic lexical properties and individual differences, there are some factors that can be identified as affecting the possibility of proleptic OF.

As we will see in the next chapter, in languages with proleptic structures, the proleptic element is often not an argument, but a thematic adjunct bearing some oblique case-marking. This pattern can be found in Hungarian, too. Verbs that occur with prolepsis almost always also occur with a relative dependent, which intuitively has the same subject matter thematic relationship to the main verb that the proleptic argument has. Verbs that reject OF often do not occur with such an oblique.

- (120) a. *Azt mondtam/ gondoltam/ ígértem/ meséltem/ gyanítottam*
 that.ACC said.1SG thought.1SG promised.1SG told.1SG suspected.1SG
János-ról hogy jön.
 John-DEL that(C) comes
 ‘I said/thought/promised/told/suspect about John that he will come.’
- b. *Azt éreztem/ utáltam/ tanácsoltam/ parancsoltam/ (*János-ról),*
 that.ACC felt.1SG hated.1SG advised.1SG ordered.1SG John-DEL
hogy jön/ jöjjön.
 that(C) comes/ come.SBJV.3SG
 ‘I felt/hated/advised/ordered (*of John) that he (should) come.’

However, the correspondence is far from perfect. Although it seems to be a solid generalization that verbs occurring with direct object prolepsis can alternatively take a relative dependent, the reverse is not true: there are predicates that can take a relative dependent but are dispreferred in proleptic object OF. The relative dependent is completely optional, so it is in all probability a thematic adjunct of these predicates.

- (121) a. *Azt terjesztettem/ nyilatkoztam/ válaszoltam/ ?suttogtam*
 that.ACC spread.1SG spoke.to.the.press.1SG answered.1SG whispered.1SG
Jánosról, hogy jön.
 John.DEL t hat(C) comes
 approx.: ‘I spread (the news)/ spoke to the press/ answered about John that he will come.’
- b. *???Jánost terjesztettem/ nyilatkoztam/ válaszoltam/ suttogtam,*
 John.ACC spread.1SG spoke.to.the.press.1SG answered.1SG whispered.1SG
hogy jön.
 that(C) comes
 Intended: ‘I spread (the news)/spoke to the press/ replied/ whispered of John that he will come.’

So the possible oblique dependent can be seen as a prerequisite for proleptic OF, but other factors are also at play. Another such factor seems to be that the verb should be obligatorily associated with a proposition. Verbs that reject the direct object OF are often such that they have an intransitive counterpart. Verbs that are good with OF tend not to have intransitive versions. For these verbs, the proposition itself seems to be an additional argument not required by the basic semantics of the verb. Prolepsis is obligatorily associated with a proposition, so it is expected that these predicates do not constitute a proper basis for the argument-structure of prolepsis.

- (122) a. **János mondott/ gondolt/ ígért/ gyanított.*
 John said.3SG thought.3SG promised.3SG suspected.3SG
 b. *János hazudott/ nyilatkozott/ válaszolt/ suttogott.*
 John lied.3SG spoke.to.the.press.3SG answered.3SG whispered.3SG
 ‘John lied/spoke to the press/answered/ whispered.’

A final factor that I would like to mention is that it seems that the verb’s flexibility in realizing its propositional argument seems to correlate with its ability to participate in proleptic OF. It was discussed in 3.2 that in Hungarian subordinate clauses, neither the object pronoun nor the complementizer is obligatory by default.

- (123) (*Azt*) *mondom/ gondolom/ ígérem,* (*hogy*) *jövök.*
 that.ACC say.1SG think.1SG promise.1SG that(c) come.1SG
 ‘I say/think/promise (that) I will come.’

This means that the OBJ argument may be either a lexical noun, a pronoun, a *that*-clause or a *that*-less, bare clause. However, not every verb allows this flexibility. Although almost every verb at issue allows the complementizer to be deleted (with the restrictions discussed earlier), not all of them allow the pronoun to be absent (124a) if there is a propositional CP. Furthermore, only a small subset of this latter group allows the deletion of both the complementizer and the pronoun (124b).⁵⁷ As it happens, this last group (*mond* ‘say’, *gondol* ‘think’, *ígér* ‘promise’, *remél* ‘hope’, *gyanít* ‘suspect’) coincides with the best direct object proleptic verbs. (The tense varies in (124) to create pragmatically more plausible sentences, e.g. one cannot dream while uttering a sentence).

- (124) a. *Mondom/ gondolom/ sérelmezem/ sejtem/ *feleltem/ *képzelem/*
 say.1SG think.1SG resent.1SG surmise.1SG replied.1SG imagine.1SG
**álmodtam, hogy János jön.*
 dreamed.1SG that(C) John comes.
 ‘I say/think/ resent/ surmis/replied/imagine/dreamed that John will come.’

⁵⁷ While the complementizer can be freely deleted if the pronoun is present, once it is absent, the complementizer becomes obligatory for many verbs. The reasons for this are unknown to me. It is possible that investigating this along the line of E. Kiss’s (2005) theory of event structure would provide some insights.

- b. *Mondom/ gondolom/ ígérem/ remélem/ gyanítom/ *sérelnem/*
 say.1SG think.1SG promise.1SG hope.1SG suspect.1SG resent.1SG
**sejtem János jön.*
 surmise.1SG John comes.

Ultimately, I claim it is lexical variation that decides whether the proleptic argument can be added to the lexical entry of a predicate. At least three factors contribute to this potential: the ability to occur with a relative dependent, the obligatory nature of the propositional argument and the categorial flexibility with which the proposition is realized. The more of these factors a verb is positive about, the more likely it is that it will accept the addition of an extra subject matter argument to its lexical frame, which results in proleptic direct object Operator fronting. From this perspective, the proleptic object is a “derived argument” in Toivonen’s (2013) sense: an argument that is not part of the initial argument list, but added through an argument structure operation.

This time no additional phrase-structural rules are necessary, since the positions needed independently exist in Hungarian grammar. The proleptic element is the object/oblique argument of the predicate (*mond* in this case), and it can sit in the positions that may be occupied by the respective arguments of these verbs.

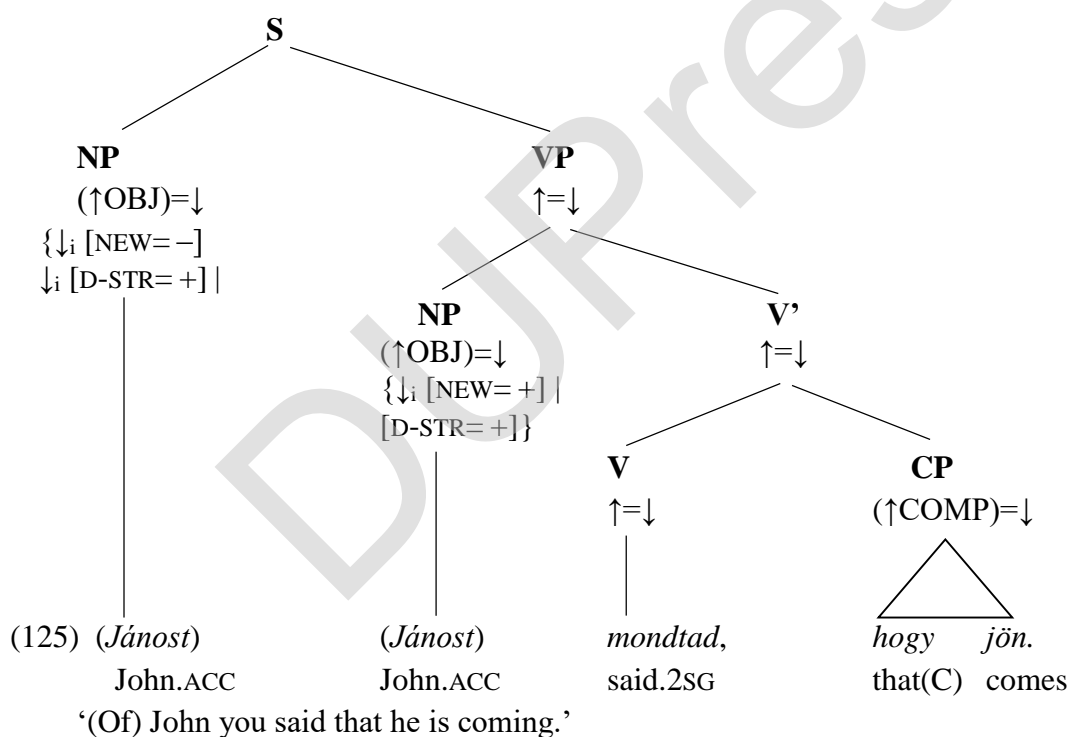


Figure 5.

C-structure for OF in the proleptic scenario

The corresponding f-structure is shown in Figure 6. Crucially, the proleptic object is the thematic argument of the verb, *mond* ‘say’ in (125). Let us summarize the arguments for this. First, let us recall that Coppock’s (2003) analysis posited that the proleptic element is non-thematic, but this leads to a violation of Semantic Coherence, as a meaningful element (the proleptic direct object) would not be associated with a thematic argument slot of any predicate. An alternative analysis, if possible, is preferable to the suspension of a basic LFG principle.

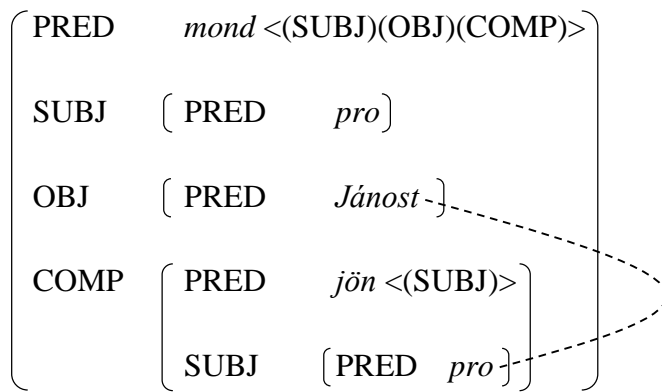


Figure 6.
F-structure for (125)

Second, the proleptic element is in complementary distribution with the associate pronoun of demonstrative clause and I have argued extensively in section 2.2.2 that the pronoun is also a thematic argument and not an expletive. While a proleptic account could in principle be posited even under the expletive-analysis, the plausibility of such an account is greater if the thematic status of the OBJ-argument is uniform in the argument-structural variations of the predicates. Also, this makes the analysis of object and oblique proleptic elements uniform, as the latter are to be analyzed as arguments anyway.

Third, OF is incompatible with idiom chunks, even when they are postverbal, as per (126).

- (126) #(*De hiszen*) *te mondtad a gépszíjat, hogy elkapta Jánost.*
 but you said.2SG the driving belt.ACC that(c) caught John.ACC
 ‘But you said about the driving belt that it had caught John. / *But you said that John has had to work a lot lately.’

Finally, in certain contexts, the complement clause may be elided (127a). This is comparable to the English examples in (128), based on Bresnan (1982). *Persuade* takes a thematic object, while *believe* is a raising verb, taking a nonthematic object. It is then the case that *John* in (128a) is thematically independently licensed in the main clause, while in (128b) it is left without a thematic integration, causing ungrammaticality. The fact that proleptic OF patterns with *persuade* lends support for the thematic object analysis of the proleptic element. We will further explore the connection between prolepsis and control in 5.3.

- (127) a. A: *Végül Péter jött.*
 finally Peter came.3SG
 ‘Finally, Peter came.’
 B: *De hiszen te Jánost mondtad!*
 but you John.ACC said.2SG
 Intended: ‘But it was (of) John you said (that he would come)!’
- (128) a. *Someone had to wash my car. I persuaded John (to wash my car).*
 b. *Someone stole my car. I believed John *(to have stolen my car).*

Since the proleptic element is linked to the embedded grammatical function via an anaphoric link and not functional identity, a compulsory matching of grammatical features is not expected, so there can be space for the number variation of the kind shown earlier. We will see in section 5.3 that anaphoric identification often leads to number mismatches in similar setups.

- (129) *Az összes lányt mondtad, hogy jön/ jönnek.*
 the each girl.ACC said.2SG that(c) come.3SG come.3PL
 ‘(Of) each girl you said that they are coming.’

This may be elucidated with reference to an ambiguity between the notions of “INDEX-agreement” and pragmatic agreement, following the terminology of Wechsler & Zlatić (2003). In their theory, agreement may be of several types: CONCORD (the syntactic side, usually participating in NP/DP-internal agreement), INDEX (based on the semantic content of the nominal, manifest in e.g. subject-verb agreement, anaphora) and pragmatic agreement (based on conceptual and discourse factors). As we have seen in (63d), the default INDEX value for quantified noun-phrases in Hungarian is singular (in a simple sentence: ‘each girl come.3SG/*3PL’). However, there is an anaphoric binding relationship between the proleptic element and the embedded *pro*. As Wechsler & Zlatić (2003: 84) explain, “this invites the possibility of pragmatic/semantic agreement. Hence the range of options for personal pronoun-antecedent agreement are INDEX agreement and pragmatic/semantic agreement”. This is exactly what we see in (129). Despite the singular index, these quantified nominals are conceptually plural, enabling the possible plural agreement. Such phenomena may be a matter of dialectal/idiolectal variation, so if Gervain (2002) is right and there is a group of speakers who do not accept plural agreement in similar cases, they may be accounted for by positing that they exclusively resort to INDEX-agreement and disallow the pragmatic strategy.

Since the proleptic element is a genuine main clause entity only semantically plugged into the embedded clause, the lack of reconstruction and island effects follow from the analysis.

Finally, some notes have to be made about the complementizer in OF. The basic pattern is the one observed in standard subordinate clauses: the complementizer is obligatory if the main clause contains a focus or contrastive topic (apart from the associate demonstrative pronoun). All instances of OF are tied to such discourse functions, so *hogy* ‘that(C)’ is expected to be necessary.

- (130) *János(t) mondtad, *(hogy) jön.*
 John(ACC) said.2SG that(c) comes
 ‘(Of) John you said that he will come.’

A difference appears if interrogative subordinate clauses are considered. In standard subordinate sentences, this enables complementizer-drop, even in the cases of focussed main-clause elements. Proleptic OF does not entirely behave according to expectations. (131), with an interrogative main verb, is degraded under the proleptic interpretation (“of John you asked whether he would come”).

- (131) *??Jánost kérdezted, jön-e.*
 John.ACC asked.2SG comes-Q-PRT

I see two avenues for explanation. The first is that one may consider the proleptic lexical entry of the relevant predicates. The grammatical function of the embedded clause is a COMP and the most canonical realization of the COMP function is a CP. Then the obligatory nature of the complementizer could be a reflex of this. The second possibility is to argue that the fact that (131) is not fully ungrammatical indicates that the general rules of Hungarian still apply and the degradation is the result of some yet to be identified independent factors.

If the fronted element retains its original case (LDD-OF), the original lexical entry is involved, so complementizer-drop should be an option, which is the case: (132) is much better than (131).

- (132) *?János kérdezted, jön-e.*
 John.ACC said.2SG comes-Q-PRT
 ‘(Of) John you asked if he will come.’

To conclude this section, let us summarize the analysis of Operator fronting. There are two distinct ways in which OF may happen.

In the first case, some element of the subordinate clause appears in the preverbal field of the matrix clause and gets some information structural specification. There is a strict syntactic dependency, functional identification between the fronted element and its embedded position.

In the second case, via an idiosyncratic lexical process the basic entry of certain predicates is altered and an extra argument, bearing a “subject matter” theta role is added. This is called prolepsis. Hence the originally two-place predicate becomes a three-place one, subcategorizing for a subject, an object and complement clause. The extra argument is realized as an oblique or an object (depending on the original subcategorization of the base predicate) and is obligatorily co-indexed with an embedded grammatical function.

4.2.3 On split operator fronting

A variation of OF recently received some attention in Jánosi (2013). What she calls “long split focalization” involves some nominal that is modified by an adjective. It is possible to front the adjective-noun complex in a way that the adjective remains in the embedded clause, as illustrated in (133).

- (133) *Autó mondtad, hogy új áll a garázsban.*
 car said.DEF.2SG that(c) new stand.3SG the garage.in
 ‘(Of) a car you said that a new one is standing in the garage.’

In (133) the fronted element is *autó* ‘car’ and there is the adjective *új* ‘new’ associated with it in the subordinate clause. The main verb shows definite conjugation, which indicates that it agrees with the subordinate clause in this respect (subordinate clauses count as definite objects). This construction is in parallel with the “movement” version of OF, where the fronted element is functionally identified with its embedded canonical position and it is not thematically related to the main verb.

Split fronting can also behave like the proleptic version of OF. An example for this is shown in (134), where the fronted element bears accusative case and the main verb is in indefinite conjugation, matching the fronted phrase.

- (134) *Autót mondtál, hogy új áll a garázsban.*
 car.ACC said.INDEF.2SG that(c) new stands the garage.in
 ‘(Of) a car you said that a new one is standing in the garage.’

The split construction also parallels the unsplit version in other respects: as opposed to the functionally identified version, the proleptic pattern is insensitive to islands (135a) and can show number-mismatch (135b). Also, just like in unsplit OF, other grammatical functions can also be fronted (135c).

- (135) a. *Autót mondtál/*mondta, hogy hallottad a hírt, hogy új áll a garázsban.*
 car.ACC said.INDEF.2SG/ DEF that(c) heard.2SG the news.ACC that(c) new stands the garage.in
 Intended: ‘(Of) a car you said that you have heard the news that a new one is standing in the garage.’
- b. *Autót mondtál/*mondta, hogy újak állnak a garázsban.*
 car.ACC said.INDEF.2SG/ DEF that(c) new.PL stand the garage.in
 Intended: ‘(Of) a car you said that new ones are standing in the garage.’
- c. *Autót mondtál/mondta, hogy újat vettél.*
 car.ACC said.INDEF.2SG/ DEF that(c) new.ACC bought.2SG
 ‘(Of) a car you said that you had bought a new one.’

What this suggests is that the underlying mechanisms of the split and unsplit versions of OF are the same. Split OF is possible as there is a general process by which adjectives and nouns may be separated in Hungarian (for details, see Jánosi 2014, chapter 2). When this happens, the adjective receives the case associated with the nominal (136b).

- (136) a. *Péter új autót vett.*
 Peter new car.ACC bought.3SG
 ‘Peter bought a new car.’
- b. *Autót vett Péter újat.*
 car.ACC bought.3SG Peter new.ACC
 ‘Peter bought a new car.’

Given the independent motivation for the possibility of this noun phrase split in (131b) and the analytical parallel of split and unsplit OF, split OF does not constitute an independent issue for the purposes of this monograph. What seems worth noting is that in the proleptic version, the role of the resumptive element (the co-indexed element in the COMP) is fulfilled by the case-marked adjective.

If future research explores the LFG-analysis of split adjectival constructions in general, I think the details of the analysis should follow from the general approach adopted in this chapter.

4.2.4 On long topicalization

There is a fronting construction in Hungarian that has been mentioned in the literature and resembles OF in some respects but has not been discussed so far. É. Kiss (2002) calls it “long topicalization” (LT) and I adopt this terminology. LT looks like OF in that there is a fronted element which is related to some embedded grammatical function. However, there are several crucial differences between OF and LT. First of all, in the case of LT the discourse function of the fronted element is a neutral Topic. In (137) there is no sign of contrastivity or other discourse-prominence associated with *János* ‘John’.

- (137) *János azt gondolom, hogy jön.*
John that.ACC think.1SG that(c) comes.
‘John, I think, will come.’

Also observe in (137) that the demonstrative pronoun is present, in contrast with OF structures. This means that the CP is an adjunct, from which extraction should be impossible. Other seemingly island-constraint violating examples are also shown, taken from É. Kiss (2002:258). (138a) contains a temporal adjunct and (138b) a complex noun phrase.

- (138) a. *János már dél felé járt az idő, [amikor felébredt].*
John already noon towards went.3SG the time when awoke.3SG
‘John, it was already about noon when (he) woke up.’
b. *Jánost nincs [az az ember aki fel tudná bosszantani].*
John.ACC isn’t that the man who up could make.angry.INF
‘John, there is no man who could make (him) angry.’

Moreover, the complementizer is completely optional in such sentences.

- (139) *János azt gondolom, jön.*
John that.ACC think.1SG comes
‘John, I think will come.’

From this it is clear that LT is a distinct structure. I agree with É. Kiss (2002:259) in that such elements are left-peripheral hanging topics. From the perspective of this monograph, the construction is akin to LD-Eng or FLD-Hun: a “syntactic orphan”, unintegrated into the host sentence. Thus, it is expected that the extra-sententiality effects associated with these structures are to be observed with LT as well. They indeed are: non string-initial LTs are dispreferred in (140) in the same way as LD-Eng and FLD-Hun are. These effects are not observable with OF (141).

- (140) a. *?Szerintem János azt mondtad, hogy jön.*
in.my.opinion John that.ACC said.2SG that(c) comes
‘In my opinion, John, you said that he will come.’

- b. *?Úgy hallottam, hogy János azt gondolod, hogy jön.*
 so.DIST heard.1SG that(c) John that.ACC think.2SG that(c) comes
 ‘I’ve heard that John, you think that he will come.’
- c. **Te vagy az az ember, aki János azt gondolja, hogy jön?*
 you are that the person who John that.ACC think.3SG that(c) comes
 Intended: ‘Are you the person who (of) John_i thinks that he_i comes?’
- (141) a. *Szerintem János(-t) mondtad, hogy jön.*
 in.my.opinion John-(ACC) said.2SG that(c) comes
 ‘In my opinion (of) John you said that he will come.’
- b. *Úgy hallottam, hogy János(-t) gondolod, hogy jön.*
 so heard.1SG that(c) John-(ACC) think.2SG that(c) comes
 ‘I’ve heard that (of John) you think that he will come.’
- c. *Te vagy az az ember, aki János(-t) gondolja, hogy jön?*
 you are that the person who John-(ACC) think.3SG that(c) comes
 ‘Are you the person who (of) John thinks that he will come?’

Thus, long topicalization is to be treated as an independent syntactic structure, set apart from OF. It could be viewed as a topic-marking counterpart of FLD-Hun. Actually, such a state of affairs is not unexpected, as nothing prescribes that left-peripheral syntactic orphans in Hungarian are to be associated only with a Focus discourse-function.

4.3 Conclusion to chapter 4

In this chapter I investigated Hungarian left-peripheral constructions. The two main areas of investigation were left dislocation (LD-Hun) and operator fronting (OF). Both of them turn out to be comprised of several distinct configurations.

In LD-Hun, an element in the topic field of the Hungarian sentence (a neutral or a contrastive topic) is associated with a pronominal. There is an anaphoric dependency between them and some degree of syntactic feature-matching is also required.

In operator fronting, an element that is associated with an embedded grammatical function appears in the matrix clause and bears some prominent discourse function (contrastive topic, contrastive/information focus). This may happen in two ways: either along the lines of conventional “fronting”, or as prolepsis. In the first case, the fronted element is related to its embedded position via functional identification. This means strict syntactic matching, including case, person and number features. In prolepsis, the fronted element becomes the argument of the main predicate and an anaphoric link is established with the embedded function. The dependency is weaker here, allowing for case and number mismatches and the circumvention of island-constraints.

Apart from these, I also discussed two configurations where a fronted element is “radically” left-peripheral. In free left dislocation and long topicalization, the sentence-initial constituent is a “syntactic orphan”, an element not integrated into the host sentence.

CHAPTER 5

CROSSLINGUISTIC AND THEORETICAL PERSPECTIVES

Throughout the monograph we have discussed various left-peripheral (“fronted”) discourse-related constructions in English and Hungarian. The aim of this chapter is to put them into perspective by looking at similar structures in other languages, making comparisons and also to investigate how the proposed analyses fit into the existing theoretical space of fronting constructions.

The configurations discussed in the monograph may be put into four categories:⁵⁸

- i. **Fronting:** a left-peripheral element is “extracted” from a sentence-internal position, without a subsequent associated pronoun. Although the c-structural position of the fronted element is left-peripheral, functionally it is plugged into its canonical position. This manifests in strict agreement-properties and syntactic reconstruction effects. English topicalization, certain clause-initial adjuncts (event-internal ones) and the long-distance dependency version of Hungarian operator fronting belong here.
- ii. **Integrated left dislocation:** a discourse-prominent element of the left-periphery is associated with a subsequent pronoun. The construction is syntactically integrated, with formal correspondences between the left-dislocated element and the associate pronoun. Among the discussed structures, Hungarian topic left dislocation exemplifies this.
- iii. **Non-integrated (“hanging”) left dislocation:** the left peripheral element is loosely attached to the sentence. The host may contain a coreferent pronominal. No formal dependencies are established, the connection between the initial element and the rest of the utterance (and a possible coreferent pronominal) is pragmatic in nature. English (hanging) left dislocation and free left dislocation Hungarian were analyzed as such constructions.
- iv. **Prolepsis:** an internal argument of the main clause predicate is anaphorically identified with an embedded grammatical function. This is the analysis proposed for proleptic operator fronting.

In this chapter I give an overview of the theoretical and cross-linguistic landscape of these constructions. The literature standardly discusses two versions of left dislocation together and I will follow this pattern, but their distinct nature will be emphasized. As prolepsis has been given relatively little attention in the literature, I will give the most exposition to this construction here.

⁵⁸ A fifth category could be the set of base-generated structures, e.g. event-external CIADJs. These are not particularly remarkable from our perspective, as both their position and integration into the sentence falls out from basic phrase-structural rules of the given language. In other words, these count as canonical constructions.

5.1 Perspectives on fronting

Both English topicalization (TOP) and case retaining operator fronting (OF) involve the extraction of an argument from a sentence-internal position. In both structures, the fronted element maintains strong ties with its original position, which results in connectivity-effects: the case, the number and the person features of the fronted element must match those that it would receive in its original position.

The main syntactic difference between TOP and OF is that the former is licensed in a specialized left-peripheral IP-adjoined position, while OF is part of the standard Hungarian sentence-structure as an element in the regular topic- or focus-field of the Hungarian sentences.

This difference translates into a difference in information structure: while TOP is linked to contrastive i-structural categories (contrastive topic and contrastive focus), OF may also receive an information focus interpretation. The contrastiveness of a fronted element in Hungarian is only preferred in the case of fronted subjects. This, however, is probably not related to the structural properties of the construction, but has to do with the individual languages themselves. Hungarian is a discourse-configurational language with a relatively flexible word-order, so constituents occupying various left-peripheral positions with the appropriate discourse functions is the default scenario. That is, various permutations of the word order (SVO, VSO, OSV, etc.) are routinely produced by this language.

In contrast, English has a much more rigid constituent structure, with SVO being the default word order. Thus, deviations from this standard pattern have more radical consequences. In other words, there has to be substantial motivation for the establishment of long-distance dependencies like TOP. Contrastiveness seems to function as such a motivation.⁵⁹ This seems to be corroborated by the fact that even in English, obligatory contrastiveness is only observed in the case of fronted arguments. As pointed out in section 3.2, clause initial adjuncts are not necessarily contrastive (though they can be). This makes sense given that the positions of adjuncts are much freer even in a configurational language like English, see for example (1), from Frey (2003: 170). As the various adjunct-positions are parts of the core patterns of the language, no special considerations (contrastiveness) are required for their clause-initial placement either.

- (1) *(Unfortunately) she (unfortunately) will (unfortunately) be talking about this subject.*

This is partially in line with López's (2016) argumentation, who, referring to Baker (1996), points out that it might be argued that in polysynthetic languages, all arguments are compulsorily in a left-peripheral position. As this is the standard setup for speakers of these languages, no special information structural properties are associated with these left-peripheral elements. The connection to my proposal is that special IS-status is evoked only if the base position of the relevant elements in the given language is moderately fixed. If the relevant grammatical entities are flexibly placed by default, or, on the other end, such entities are obligatorily "dislocated" in some sense (depending on the syntactic framework one assumes), the association of the left-peripheral element with i-structure will be weaker.

⁵⁹ It is interesting to note it at this point that contrastive topics, as opposed to neutral topics, may be nonreferential. From this it seems that contrastiveness can generally function as a licenser for otherwise dispreferred patterns of language.

It has to be noted here though that a lot depends on what one means by contrastivity.⁶⁰ According to López (2016), all frontings (in his terminology, “d-type dislocations”) are contrastive. Given our discussion of CIADJs in English and LDD-operator fronting, this is not the case (see also the next section for related discussion about left dislocation (i.e. constructions with a pronoun associate)). As discussed in 1.3.1, “contrast” may be interpreted at the semantic level. In this case all foci (i.e. answers to questions) will be “contrastive”, so the LDD-OF will have this feature even when it is associated with the noncontrastive information focus discourse function. Probably some modification of the CIADJ-examples could make them fit this weaker notion of “contrast”. For instance, the event-internal adjunct *in the box* in (2) (this is example (19a) from section 3.2.1) might be argued to be related to some contextually relevant set of containers.

(2) *In the box, John found a hammer.*

However, in my “pragmatic” view of contrast explicated in section 1.3.1, the alternatives have to be really evoked in the minds of the interlocutors. That is, the alternatives have to be not only contextually relevant, but also salient. This is not the case in (2), and also not the case in LDD-OF examples which might contain an information focus (that is, any LDD-OF with a fronted nonsubject). In comparison, the alternatives are much more salient in the topicalization example in (3) where an argument is fronted.

(3) *Into the box, John put a hammer. (Into the bag, he put a screwdriver.)*

On the empirical side, the role of contrastiveness has been noted with regards to quite a few constructions in other languages as well. For example, Vilkuna (1995) suggests that Finnish has a left-peripheral position which may be occupied either by a contrastive focus or a contrastive topic. The following examples are from Molnár & Winkler (2010: 1399).

- (4) a. A: *Pekka lensi Tukholmaan.*
 Pekka flew.3SG to.Stockholm
 ‘Pekka flew to Stockholm.’
 B: *Eihän, vaan [Reykjavikiin]_{CF} Pekka lensi.*
 no, but to.Reykjavik Pekka flew
 ‘That’s not true, to Reykjavik Pekka flew.’
- b. *[Tukholmaan]_{CT} Pekka lensi Finnairilla. [Reykjavikiin]_{CT}*
 to Stockholm Pekka flew.3SG by.Finnair to Reykjavik
Icelandairilla.
 by.Icelandair
 ‘To Stockholm, Pekka flew by Finnair. To Reykjavik, by Icelandair.’

As contrastive topics need not be referential, even a verb may occupy this position as in (5), from Vilkuna (1995: 252).

⁶⁰ For more on this topic, see Repp (2016).

- (5) *Mikko on hyvä ruoanlaitossa mutta [leipoa]_{CT} hän ei osaa.*
 Mikko is good cooking.in but bake.INF he not.3SG can
 ‘Mikko is good at cooking, but bake, he can’t.’

A similar example was mentioned in Dutch, in section 1.3.1, repeated here as (6).

- (6) a. *Ik geloof dat [alleen DIT boek]_{CF} Jan Marie gegeven heft.*
 I believe that only this book John Mary given has
 ‘I believe that John has given ONLY THIS BOOK to Mary.’ (and not another book)
- b. *Ik geloof dat [zo’n boek]_{CT} alleen JAN Marie gegeven heeft.*
 I believe that such-a book only John Mary given has
 ‘I believe that such a book, only John has given to Mary.’ (others may have given other books)

German offers another variation on this theme. Frey (2010) calls attention to the fact that certain sentence-initial elements in German must receive heavy stress and should be interpreted contrastively. For instance, consider the pair in (7).

- (7) a. *Einen Kollegen hat heute ein Polizist verhört.*
 a.ACC colleague has today a policeman interrogated.3SG
 ‘A policeman interrogated a colleague today.’
- b. *GRÜN will Maria die Tür streichen.*
 green wants Mary the.ACC door paint.INF
 ‘Mary wants to paint the door GREEN (and not some other color)’, ‘GREEN Mary wants to paint the door.’

According to Frey (2010: 1418), it is only (7a) that is “acceptable even if the prefield elements are not stressed beyond the word accents”. That is, while *einen Kollegen* ‘a colleague’ is acceptable as a neutral topic or Information focus, *grün* ‘green’ must be contrastive. Although Frey (2010) is not explicit about this, CT and CF are both acceptable interpretations.

Frey’s (2010) explanation is that the fronting of the two constructions is executed by distinct mechanisms. In (6a) we see “formal movement”: “the highest maximal phrase in the middle field is moved to the adjacent prefield (...) without any intonational or interpretative effects of its own”. In light of our earlier discussion, this is a “default” fronting, which is part of the standard structure of the language.

(6b) on the other hand involves A-bar movement (in Frey’s (2010) framework). *Grün* ‘green’ would never occupy the fronted position, which motivates an additional interpretative effect, namely contrastivity. Formal movement is clause-bound, which means that any element fronted from an embedded sentence may only undergo the second mechanism, as in (8).

- (8) *Den CHEF/*chef meint Maria, dass Paul zur Party einladen sollte.*
 the.ACC boss thinks Mary that(c) Paul to.the party invite.INF
 should.3SG
 ‘The boss Mary thinks that Paul should invite to the party.’

So what Frey (2010) analyzes as A-bar movement is a close correlate of English topicalization, while “formal movement” is closer to OF. Note that neither A-bar movement nor TOP is obligatory in any context: contrast is dependent on the discourse, including the intention of the speaker. Frey (2010) demonstrates this with (9), where B1, with *einen Kuchen* ‘a cookie’ is a CF is preferred over B2. The latter with *einen Kuchen* as IF is nevertheless a possible answer.

- (9) A: *What does Paul want? Ice-cream or cookie?*
 B1: *Einen Kuchen möchte Paul.*
 a.ACC cookie wants Paul
 ‘An ice-cream Paul wants.’
 B2: *Paul möchte einen Kuchen.*
 Paul wants a.ACC cookie
 ‘Paul wants a cookie.’

This is parallel to the observation made by Krifka (2008) (mentioned in section 2.6.1) that the answer in (8) is not necessarily contrastive. Accordingly, (10/B1) may optionally be expanded into a TOP structure if B wants to evoke contrast.

- (10) A: *What do you want to drink, tea or coffee?*
 B1: *Tea.*
 B2: *Tea I want (and not coffee).*
 B3: *I want tea.*

Also, while contrast is obligatory in Frey’s (2010) A-bar movement construction and in TOP-Eng, it is not exclusively related to these configurations, as in-situ contrast is also possible.

- (11) a. *Paul möchte EINEN KUCHEN.*
 Paul wants a.ACC cookie
 ‘Paul wants a cookie.’
 b. *I want TEA, and not a coffee.*

That is, while certain configurational positions seem to be necessarily contrastive, a contrastive element does not have to move into these slots. These considerations argue against the cartographic view of contrast (e.g. Molnár & Winkler 2010), where “contrast” is a syntactic feature, triggering movement for checking in the specifier a designated position. I agree with Neeleman & van de Koot (2008) in that such fronting constructions should be viewed as interface optimizations, whereby the syntactic structure of a sentence is matched to an information structural template (the designated contrastive positions).

5.2 Perspectives on left dislocation

Since Ross (1967), who introduced the term “left dislocation” (LD) for constructions whereby some discourse-prominent entity is placed at the left periphery of the clause, with a subsequent co-referential pronoun, a large body of literature has emerged about LD. Some of the most notable instances are Cinque (1977), the edited volume of Anagnostopoulou et al. (1997) and Grohmann (2003). There seems to be a consensus that at least two subtypes of LD should be distinguished. In one type of LD, there is some syntactic dependency between the host and the associated pronoun and the construction itself is properly (syntactically) integrated into the containing sentence. This LD is commonly referred to as “i-type” left dislocation (iLD henceforth). The second type of LD is thought of as a looser kind of dependency. There, the host and the pronoun are only related pragmatically, and the host itself is also assumed to be in some sense less integrated of the core clause structure. This LD is usually called “n-type” left dislocation (for “non-integrated”, nLD).⁶¹ Frey (2004) illustrates the two LD-types with the following German sentence pair. Similar constructions have also been reported in other Germanic languages, e.g. Dutch, Icelandic.

- (12) a. *Den Hans_i, den_i mag jeder.*
the.ACC Hans d-pron.ACC likes everyone
- b. *Den Hans_i, jeder mag ihn_i.*
the.ACC Hans, everyone likes him
'Hans, everyone likes him.'

A related construction is clitic left dislocation (CLLD), which is standardly analyzed as falling into the i-type LD category. It's most obvious feature is that the pronominal associate is not a personal or a demonstrative pronoun, but a clitic. The Greek example in (12) is from Alexiadou (2006). CLLD has also been reported in Italian, Spanish, etc.

- (13) *Ton Jani_i den ton_i ksero.*
the.ACC John.ACC NEG clitic.ACC know.1SG
'John, I do not know him.'

I-type and n-type LD in sharp contrast with respect to syntactic distribution and connectivity effects. For example, left-dislocated constituents containing a reciprocal pronoun are only grammatical in the i-type construction. A Principle A violation causes ungrammaticality in the Dutch example in (14b), as *elkaars* ('each other') is not integrated into the sentence. In (13a), even though the reciprocal is string-initial, it is still a proper part of the sentence, so binding may obtain via syntactic reconstruction. These data are from Vat (1981).

- (14) a. [*Elkaars jassen*]_i, *die_i dragen ze niet graag.*
each other's coats d-pron wear they not willingly
'Each other's coats they don't like to wear.'

⁶¹ López (2016) refers to the same theoretical distinction as “d-type” and “h-type” dislocations. In the cited sources, various labels (contrastive left dislocation, German left dislocation, hanging topic left dislocation) are used for the particular constructions.

- b. **[Elkaars jassen]_i, zij_i dragen ze niet graag.*
 each other's coats them wear they not willingly

Just like in German, different pronouns are used in Dutch iLD (a d-pronoun, *die*) and nLD (a personal pronoun, *zij*).

Another example for the connectedness in iLD and the lack of connectedness in nLD is constituted by the case-marking patterns. In the iLD example of (15a), the left peripheral *þessum hring* ‘this ring.DAT’ is in dative case, matching the coreferent pronoun. In contrast, in (14b), which is an nLD example, the dislocated entity is in the default, nominative case. This is illustrated by the Icelandic data in (15) from Zaenen (1997), citing Thráinsson (1979).

- (15) a. *[þessum hring]_i, honum_i hefur Ólafur lofað Mariú.*
 this ring.DAT it.DAT has Olaf promised Mary.DAT
- b. *[þessi hringur]_i, Ólafur hefur lofað Mariú honum_i.*
 this ring.NOM, Olaf has promised Mary.DAT it.DAT
 ‘This ring, Olaf has promised it to Mary.’

In Zaenen’s LFG-analysis for Icelandic, the left-dislocated element in the nLD structure is introduced under a special, designated non-iterative “E-node” in the constituent structure. This category is assumed to be on the left periphery, above the proper sentential domain. This serves to explain the fact that according to Zaenen (1997), Icelandic nLD, in contrast to Icelandic iLD, is not possible in embedded clauses, see the contrast between (16a) and (16b).

- (16) a. *Jón segir að [þessum hring]_i, honum_i hafi Ólafur lofað Mariú.*
 John said that this ring.DAT it.DAT has Olaf promised Mary.DAT
- b. **Jón segir að [þessi hringur]_i, Ólafur hafi lofað Mariú honum_i.*
 John said that this ring.NOM Olaf has promised Mary.DAT it.DAT
 John said that this ring, Olaf has promised it to Mary.’

As already noted in section 4.1.2, for the iLD-structure, Zaenen (1997: 133) proposes that the “simplest way left in LFG to treat the anaphoric copy is as an adjunct”. As for her analysis of nLD, I think that the “orphan-approach” (which I advocate) can capture the essence of the analysis equally well, but it does not bear the burden of having exotic nodes in the c-structure.

German also offers good grounds for the investigation of LD. Similarly to the situation in Icelandic and Dutch, Grohmann (2003) distinguishes two main types. The first one is shown in 16a, and it involves a loosely integrated left peripheral element paired with a sentence internal demonstrative/personal pronoun, with a topical interpretation. The second, (17b) invariably involves a “d-pronoun”, which is a kind of a demonstrative pronoun. As we can see, case matching is compulsory in the i-type construction.

- (17) a. *Den/ Der Hans_i, jeder mag ihn_i.*
 the.ACC the.NOM Hans everyone likes him
 ‘Hans, everyone likes him.’

- b. *Den / *Der Hans_i, den_i mag jeder.*
 the.ACC the.NOM Hans d-pron likes everyone
 ‘As for Hans, everyone likes him.’

Prosodically, the fronted element is separated from the host sentence by a noticeable intonational break in the n-type LD. As is the situation with in English left dislocation and Hungarian free left dislocation, reconstruction effects are weak, see (18) from Grohmann (2003) and Shaer & Frey (2004) (though as noted in connection with example 38 in chapter 3, there could be some not yet understood semantic reconstruction processes at play).

- (18) a. *?[Sein_i erster Artikel]_j, ich glaube, dass [jeder Linguist]_i ihn_j als*
 his first article I believe.1SG that(c) every linguist it as
Mißerfolg betrachten würde.
 failure consider would.3SG
 ‘His first article, I think every linguist would consider it a failure.’
- b. *[Freunde von einander]_i, Herforder erzählen ihnen_i selten Lügen.*
 friends of each.other Herfordians tell.3PL them rarerly lies
 ‘Friends of each other, Herfordians rarely tell them lies.’

Furthermore, nLD in German seems to be marked if placed sentence internally. (19) is from Grohmann (2003: 148), where *einen Arschtritt* ‘a kick in the ass’ is a topicalized entity and *dieser Kandidat* ‘this candidate’ is the constituent in nLD-Ger.

- (19) **Einen Arschtritt [dieser Kandidat]_i, sollte man ihm_i geben.*
 a.ACC kick-in-the-ass this.NOM candidate should one him give
 Intended: ‘A kick in the ass, this candidate, one should give him.’

The i-type construction in German displays the opposite behavior regarding these phenomena. There are reconstruction effects and sentence-internal positions are grammatical.

- (20) *[Seine Mutter]_j, die_j verehrt [jeder Junge]_i.*
 his mother DEMONSTRATIVE admires every boy
 ‘His mother, every boy admires.’
- (21) *Ich meine, dass Hans_i, den_i jeder mag.*
 I think.1SG that(c) Hans DEMONSTRATIVE everyone likes
 ‘I think that Hans, everyone likes.’

These data suggest that iLD-Ger forms a natural pair with TLD and nLD-Ger is related to FLD. Nevertheless, there are certain differences between the Germanic type of LD and what we find in Hungarian, especially in the case of the i-type structure.

As noted, my analysis for TLD is similar to that of the analysis of Zaenen (1997) for Icelandic LD. Frey (2004), in a Minimalist framework, also argues for an analysis of this sort, where the left-dislocated phrase and the pronoun are independently “base-generated” and co-indexed. However, in his account, the pronoun is in a theta-position and the left-peripheral phrase is a CP-adjunct, so the functional hierarchy is the opposite of Zaenen’s (1997) and mine:

the left-peripheral element is the adjunct and the pronoun is the “host” element. As already mentioned in section 4.1.2, while this might be the right approach for Germanic LD, it is definitely not the one for Hungarian. Apart from the arguments already mentioned (subcategorization, object agreement), let us also note that the LD-pronoun by itself may be fully felicitous in German given the appropriate context (as in 22a), this is not the case in Hungarian, as the demonstrative cannot refer to a person outside this construction, see (22b). Without the host *Jánost* ‘John.ACC’, the pronoun could only refer to some nonhuman entity. Thus an analysis where the host is an adjunct is more plausible in Germanic LD than in Hungarian.

- (22) a. (*Den Hans_i*), *den_i* *mag* *jeder*.
 the.ACC Hans d-pron.ACC likes everyone
- b. #(*Jánost_i*), *azt_i* *mindenki* *kedveli*.
 John.ACC that.ACC everyone likes
 ‘John, everyone likes him.’

The CP-adjoined position of the left-peripheral element in German is supported by the fact that it can occur in a pre-complementizer position in a subordinate clause, as in (23a), from Frey (2004, footnote 14). This configuration is sharply ungrammatical in Hungarian, see (23b).

- (23) a. *Maria* *glaubt*, *den* *Hans_i*, *dass* *den_i* *jeder* *mag*.
 Mary believes the.ACC Hans that(c) d-pron everyone likes
 ‘Maria believes that Hans, everyone likes him.’
- b. **Mari* *hiszi*, *Jánost_i*, *hogy* *azt_i* *mindenki* *kedveli*.
 Mary believes John.ACC that(c) that.ACC everyone likes

Another divergence from the Germanic pattern is that in these languages, LD is restricted to root clauses and subordinate clauses introduced by bridge verbs Frey 2004: 226). This is not the case in Hungarian, where TLD is freer in its distribution, as the contrast between the German data in (24a) and the Hungarian data in (24b) attests.

- (24) a. **Maria* *bezweifelt*, *den* *Hans_i*, *dass* *den_i* *jeder* *mag*.
 Mary doubts the.ACC Hans that(c) d-pron everyone likes
- b. *Mari* *kétli*, *hogy* *Jánost_i*, *azt_i* *mindenki* *szereti*.
 Mary doubts that(c) John.ACC that.ACC everyone likes
 ‘Maria doubts that Hans, everyone likes him.’

These data about subordinate clauses suggest that TLD in Hungarian is closer to the core sentential domain than the Germanic LD type. This likens the Hungarian construction to clitic left dislocation (CLLD) structures (see (12) above), which are analyzed as being IP-adjuncts by Alexiadou (2006). While on our framework, there is no IP in Hungarian, the parallel is that the construction is located in the standard sentential domain, which is IP in configurational languages and S in a language like Hungarian. This gives a straightforward explanation for the contrasts in (23)-(24).

It may be added here that since the topic field is inherently iterative in Hungarian, there is no point of talking about the host being an adjunct, in contrast to other instances of CLLD, noted above. This difference in phrase-structural configuration may be one of the reasons why the host is able to dominate the pronominal associate in terms of functional structure.

Another CLLD-like property of TLD is that it allows for stacking, which is not possible in the Germanic type of LD. Consider the data in (25)-(27), where the non-Hungarian examples are from Alexiadou (2006). (25) is an Italian sentence which shows multiple instances of CLLD. (26) is Dutch LD demonstrating the ungrammaticality of multiple LDs. The Hungarian equivalent in (27) is possible.

- (25) *Di vestiti_i a me Gianni_j in quel negozio_k non mi_j ce_k ne_i ha mai comprati.*
 DET clothes to me Gianni in that shop not to.me there of.them has
 ever bought
 ‘As for clothes, for me, Gianni has never bought them in that shop.’
- (26) **Jan_i op school_j die_i daar_j zag ik niet.*
 John at school that there saw I not
- (27) *Jánost_i az iskolában_j azt_i ott_j nem láttam.*
 John.ACC the school.in that.ACC there not saw.1SG
 ‘John, in the school, I didn’t see him there.’

Thus it seems that TLD is closer to CLLD constructions than Germanic LD, as far as the syntactic distribution is concerned. However, the LD-pronoun in Hungarian is not a clitic, but a demonstrative-like element, like in the Germanic type.

Finally, for information structure, it is to be noted that n-type LDs are commonly associated with hanging topics. However, nothing conceptually excludes other DFs, so potential association with focus in Hungarian FLD just fills a typologically available but unattested scenario.

TLD is tied to the topic discourse function, but again this is not a necessity for i-type dislocations cross-linguistically. Both Grohmann (2003: 145) and Frey (2004: 213) assert that German left dislocation may be used as a contrastive focus. This is shown in (28). Thus, the inventory of information structural categories for LD constructions has to be established on the basis of individual languages.

- (28) A: *Have you met Anna yesterday?*
 a. B1: *Nein. Den Martin_i, den_i habe ich gestern getroffen.*
 no the.ACC Martin d-pron.ACC have I yesterday met
 b. B2: *Nem. Martinnal_i, azzal_i találkoztam.*
 no Martin.with that.with met.1SG
 ‘No. I met Martin yesterday.’

In sum, left dislocation structures cross-linguistically divide into a type where the left-peripheral element is a loosely attached syntactic orphan and into another type, which is integrated. In the latter, from an LFG-perspective the pronoun is plausibly analyzed as an adjunct to the host. From this distinction, a number of syntactic contrasts follow, such as

connectivity effects (formal feature matching, binding phenomena) and the ability to occur in non-peripheral positions (e.g. in an embedded clause).

Hungarian free left dislocation is an nLD, while topic left dislocation falls into the iLD-category. Within the iLD-category, TLD resembles clitic left dislocation as far as its syntactic distribution is concerned, but the form of the associate is pronominal and not a clitic, which is closer to the Germanic type.

5.3 Perspectives on prolepsis

In my discussion of Operator fronting in Hungarian, I have argued for an analysis where the case-switched element in the main clause is a thematic argument of the main verb and it is co-indexed with an argument of the embedded clause. I labelled this configuration “proleptic OF”. In this part of the monograph I explore how the purported analysis fits into the cross-linguistic picture both in terms of descriptive data and from the perspective of theoretical syntax in general.

The term “prolepsis” originates in rhetoric, where it is used to describe a figure of speech, in which the speaker anticipates and answers possible objections to their argument. The element of anticipation or foreshadowing is carried over to linguistics, where an argument also related to an embedded verb already manifests in the matrix clause.

As a working definition, let us refer to Salzmann (2017: 1), who characterizes prolepsis as “a construction where a structural complement of the matrix verb is semantically related to the predicate of a finite embedded clause”.

In this section, I will first explore the cross-linguistic landscape and show that similar accounts have been given for phenomena in various languages (English, German, Madurese, Greek, Korean, Japanese).

After the crosslinguistic survey, I will turn to the theoretical repercussions of the analysis. I will argue that prolepsis can be regarded as a kind of control, though this requires widening the concept of control to a certain extent. However, the parallels are clearly visible and we will see that from an LFG perspective it is actually not unexpected that such structures may exist.

This will lead us into section 5.4, where I will outline what kind of typological picture of control may be established in LFG. I will also survey other occupants of the typology.

5.3.1 Proleptic structures cross-linguistically

“Prolepsis” was introduced into the modern linguistic literature by Higgins (1981) for the characterization of sentences like (29), from Zacapoaxtla Nahuat (a Mexican language).

- (29) *Nimickaki ke tiwa:la.*
hear.1SG.2SG that(c) came.2SG.
'I hear that you came.'

Although it is not evident from the English translation, the object of the main verb is an incorporated 2nd person singular object (“you”) and it is co-referent with the subject of the subordinate clause.

Some discussion of the structure and its possible English parallels takes place in Massam (1985). Following this, prolepsis was out of the theoretical limelight for a long time. The notion was resurrected by the work of Davies (e.g. 2001, 2005) and Salzmann (2006, 2017). In what follows, I survey the literature on proleptic constructions in different languages.

Reflecting on Higgins (1981), Massam (1985) discusses two types of sentences in English for which a proleptic analysis might be entertained. All the examples in this section come from these works. The first is a pattern with perception verbs:

- (30) a. *Catherine saw the nuns doing yoga.*
b. *Barbara saw Geoffrey and Anvil singing.*

According to our characterization, (30) is not a good example for prolepsis even at first glance. This is because the complement clause is non-finite. Disregarding this, the reason why Higgins (1981) and Massam (1985) entertain the possibility of a proleptic analysis is that *the nuns* and *Geoffrey and Anvil* seem to fulfil two roles in the sentences: they are the objects of perception and the doers of the action.

However, even in this sense the proleptic account cannot be maintained. This is because it can be shown that there is no thematic relationship between the main verbs and the objects. Although the first intuition about (30) may be that the objects act as targets of perception (that is, it is entailed that *Catherine* saw *the nuns* and *Barbara* saw *Geoffrey and Anvil*), this may be an illusory. Massam (1985) gives the following examples to dispel the illusion. In none of these do we obtain the reading that the object by itself is a target for perception. Instead, it is the entire event that is heard/smelled/seen.

- (31) a. *I heard George start up a chainsaw.*
b. *I could smell my neighbors cooking dinner.*
c. *I saw him crush the Huey.*

Based on this, Massam (1985) argues for an analysis involving Exceptional Case Marking (ECM), where the matrix verb only subcategorizes for a clause. The expressions *George/my neighbors/him* are in the embedded small clause, receiving “exceptional case marking”, similarly to the Government-Binding Theory analysis of nonthematic objects of *believe*, *consider*, etc. (e.g. *I believe him to be happy*). Massam (1985: 167) notes that the illusion of direct perception arises because “when the event determined by the clause is seen, it is normally the case (though not always and not necessarily) that the participants in the event are also seen”. The nonthematic relationship between the main verb and the object is also reinforced by the availability of idioms in the construction. This is impossible in genuine cases of prolepsis (like Hungarian), as discussed in the previous chapter.

- (32) a. *One could hear the shit hitting the fan in the next room.*
b. *We used to see the fur fly when the phone bill came due.*

On the basis of these considerations, we can discard a proleptic analysis for perception verbs in English. The other construction that Higgins (1981) and Massam (1985) discuss is when a propositional verb takes a bare or a prepositional object and a clausal complement containing an element co-referent with the main clause object. The bare object version is not grammatical in contemporary English. In (33) we see some examples from “Biblical English” (in 33b-c, the objects have undergone passivization).

- (33) a. *And God saw the light, that it was good.*
 b. *Then shall the prophet be known that the Lord hath truly sent him.*
 c. *Now if Christ be preached that he rose from the dead.*

Contemporary English may use the preposition *of* in the relevant structures.

- (34) a. *I know of Mrs. Dalloway that the light gave her a headache.*
 b. *I read of Carrol that she was awfully shy.*
 c. *I've heard of him that he didn't realize he was oppressed.*

Apart from the preposition, this construction closely parallels the state of affairs discussed with regards to Hungarian Operator fronting. The (prepositional) objects bear the same kind of “aboutness”-relationship with the main verb that I have characterized as a subject matter theta role. (34a) also shows that the co-referent pronoun in the embedded clause may bear grammatical functions other than that of the subject (the possessor function in 34a).

Although Higgins (1981) and Massam (1985) do not mention this fact, idioms are excluded from the construction. (35a) can only receive a literal interpretation. Note the contrast with the standard raising to object example in (35b)

- (35) a. *#I believe of the shit that it hit the fan.*
 b. *I believe the shit to have hit the fan.*

According to Massam (1985), the thematic status of these *of*-complements is unclear. While (35) would suggest a thematic relationship, the optionality of the *of*-PP makes an argumental analysis unlikely. Massam (1985) leaves this question open. In my analysis of Hungarian OF, I argued that the verbs licensing OF generally allow an oblique complement, which can be considered a thematic adjunct. The oblique (delative) complement (*-ról/-ről*) is quite similar in meaning to these *of*-PPs. Therefore, it seems plausible to argue for a thematic adjunct classification of these constituents, reconciling their ambivalent properties. While the lexical process that was argued to turn these thematic adjuncts into real (derived) arguments seems not to be operational in current English, the biblical examples indicate that this process was part of the language. It should be noted at this point that a remnant of this lexical process might be that certain British English dialects allow sentences like (27) (É. Kiss 2002: 255, citing Chomsky 1981).

- (36) *Whom do you suggest should be the president?*

The accusative case of the question word must have been assigned by the matrix verb. According to É. Kiss (2002: 255), “Chomsky (1981: 174) suggests that in such constructions the raised constituent receives accusative case from the intermediate verb when passing through the specifier of the lowest CP. It is only a nominative-marked constituent, bearing a phonologically null case marker, whose case can be superseded by an accusative assigned by the matrix verb”. While this is a possible explanation (one has to assume multiple case-assignment for it to work), an alternative could be an analysis where *whom* is base-generated in the matrix-clause as the argument of *suggest*. This would be very similar indeed to a proleptic analysis. Further data could reveal the behavior of idioms, the possibility of resumptive pronouns or the effect of island-constraints on this construction. As such data are not discussed by Chomsky (1981) or É. Kiss (2002), I must leave this question open.

German has a similar proleptic construction, as is discussed in detail in Salzmann (2006, 2017). The examples here are from these works. In German, the proleptic construction involves the preposition *von*. Like in English, there is no equivalent in German for the version where the proleptic object is the direct object or an oblique argument of the matrix verb. (Salzmann notes that prolepsis is much more common in combination with relative clauses like in (37a) but it is also possible to use it with *wh*-clauses as and topicalizations, as per (37b-c).

- (37) a. *ein Maler, von dem ich glaube, dass Maria ihn mag*
 a painter of who.DAT I think.1SG that(c) Mary him likes
 ‘a painter whom I think that Mary likes’
- b. *Von welchem Maler glaubst du, dass Maria ihn mag?*
 of which.DAT painter think.2SG you that(c) Mary him likes
 ‘Which painter do you think that Mary likes?’
- c. *Von dem Maler glaube ich, dass Maria ihn mag.*
 of the.DAT painter think.1SG I that(c) Mary him likes
 ‘Of the painter I think that Mary likes.’

Unlike Hungarian proleptic OF, which is lexically restricted to a considerable extent, German prolepsis seems to work with basically every predicate that takes a complement clause. (38a) illustrates it with the German equivalent of *whisper*, a verb that would not work in Hungarian as per (38b).

- (38) a. *beim al-Qaida-Terrornetzwerk, von dem gleichzeitig sämtliche*
 at.the al-Qaida-terror.network of which simultaneously all
Experten augenzwinkernd flüstern, dass es überhaupt nicht
 experts with.a.wink whisper.3PL that(c) it at.all not
mehr existiert
 anymore exists
 ‘with the Al-Qaida terror network that all experts simultaneously whisper with a wink that it does not exist anymore’
- b. **Jánost suttogtam, hogy jön a partira.*
 John.ACC whispered.1SG that(c) comes the party.onto
 Intended: “Of John I whispered that he is coming to the party.”

Because of this free distribution, Salzmann (2006, 2017) discards an analysis where the proleptic constituent is an argument of the main verb. In the same way as in English, the optionality of the PP also argues against such an analysis. Based on various data involving binding, superiority and restrictions on extraction, Salzmann posits that the proleptic constituent is base-generated in a main clause argument position (despite it not being an argument), and is treated as the subject of a predication constituted by a null operator and the subordinate clause. For details see Salzmann (2006:193-208), here I present an example that shows that the construction is not sensitive to island-constraints. This directly parallels the state of affairs observed concerning proleptic OF in Hungarian. (39a) shows a CNPC-island, while (39b) is a *wh*-island.

- (39) a. *der Mann, von dem ich denke, dass Marie jedes Buch liest,*
the man of who.DAT I think.1SG that(c) Mary every book reads
das er schreibt
which he writes
‘the man of whom I think Mary reads every book that he writes’
- b. *der Mann, von dem ich glaube, dass niemand weiß,*
the man of who.DAT I believe.1SG that(c) nobody knows
wie er heißt
how he is.called
‘the man of whom I think nobody knows what he is called’

Salzmann (2006, 2017) also notes that an analysis where the proleptic element is a simple adjunct is also undesirable, as the *von*-constituent is licensed only if there is a co-referent element in the subordinate clause. As (40) shows, a clause that is simply “about” the proleptic PP is not enough to license it.

- (40) **Von Computern glaube ich, dass jeder einen PC kaufen*
of computers.DAT believe.1SG I that(c) everyone a.ACC PC buy.INF
sollte.
should.3SG
‘I believe of computers that everyone should buy a PC.’

In the two languages surveyed so far, the proleptic entity is introduced by a prepositional phrase as a thematic adjunct. This is in contrast with the Hungarian situation, where the proleptic entity could also be a direct object or oblique argument. However, one can find data and analyzes that bear a closer resemblance to Hungarian OF in this respect. Some languages where this is the case are Madurese, Greek, Korean and Japanese, to which languages we know turn.

Madurese is a language related to Indonesian, spoken in Southeast-Asia. It is a close relative of Balinese and Javanese. Davies (2005) provides a detailed description of its proleptic construction. (40) is an example for Madurese prolepsis. This and other Madurese examples in this section are taken from Davies (2005).

- (41) *Siti ngera Hasan bari' melle motor.*
 Siti think Hasan yesterday buy car
 'Yesterday Siti thought about Hasan_i that he_i bought a car.'

According to Davies (2005) there is no case- or tense-marking in Madurese. Nevertheless, the positioning of *bari'* ('yesterday') is a fair indication that *Hasan* is in the main clause.⁶² The question is how it gets there. Davies (2005) considers two options: the movement of *Hasan* from the embedded clause (a raising/movement analysis) or base-generation there (the proleptic analysis). Based on various data, which are to be briefly discussed below, he opts for the second possibility. These properties bear similarity to the ones that also characterize Hungarian OF, or the English and German cases discussed in the previous sections.

First, it is possible to pronounce a resumptive pronoun in the embedded clause. This is possible even though Madurese is (similarly to Hungarian) a pro-drop language (so pronouns are generally dropped if independent factors (e.g. information structure) do not require their presence). Hence, (42) is a grammatical sentence.

- (42) *Siti ngera Hasan bari' aba'eng melle motor.*
 Siti think Hasan yesterday he buy car
 'Yesterday Siti thought about Hasan_i that he_i bought a car.'

Second, Davies (2005) observes that as the argument structure of the predicate and the projected syntax is altered in prolepsis, one can detect a shift in meaning compared to a version of the sentence where no prolepsis takes place. While (43a) describes a situation, (43b) puts the emphasis on an individual.

- (43) a. *Ita a-bukteagi ja' Hasan ngeco' sapedha.motor*
 Ita prove that(c) Hasan stole motorcycle
 'Ita proved that Hasan stole the motorcycle.'
 b. *Ita a-bukteagi Hasan ja ngeco' sapedha.motor*
 Ita prove Hasan that(c) stole motorcycle
 'Ita proved about Hasan that he stole a motorcycle.'

Third, it is possible for the proleptic object to refer to nonsubjects in the embedded clause. (44a) illustrates this with an embedded object, while (44b) shows an embedded possessor.

- (44) a. *Siti ngera Hasan ja dokter juwa mareksa aba'eng.*
 Siti think Hasan that(c) doctor DEMONSTRATIVE examine he
 'Siti thinks about Hasan that the doctor examined him.'
 b. *Marlena a-bala-agi Hasan ja embi'-eng ngekke Ali.*
 Marlena say Hasan that(c) goat bite Ali
 'Marlena said about Hasan_i that his_i goat bit Ali.'

Fourth, nonbridge verbs also participate in Madurese prolepsis, e.g. *kabarragi* ('spread the news'), which is unexpected under a proper fronting ('movement') approach.

⁶² Some details about Madurese morphosyntax are to be found in Davies (2005: 646-649). These details are set aside here.

- (45) *Terdokter juwa ekabarragi Ina mon aba'engi ngobad ana'eng Marlana*
 doctors DEM spread.the.news Ina if they cure child Marlana
 'Ina spread the news about the doctors that they cured Marlana's child.'

In fact, Davies (2005) claims that the construction is possible with any verb capable of taking a clausal argument, which is like the situation in German.

Fifth, idiom chunks are not interpretable in Madurese prolepsis.

- (46) a. *Nase'la daddi tajjin.*
 rice.already become porridge
 'It is too late to do anything about it.' (lit.: 'The rice has become porridge.')
- b. *Siti ngera nase bari ja la daddi tajjin.*
 Siti think rice yesterday that(c) already become porridge
 Only lit.: 'Siti thought about the rice that it had become porridge.'

Sixth, island-constraints do not affect Madurese prolepsis, as the following sentence with CNPC shows.

- (47) *Wati ng-enga'-e Atin careta-na ja aba'eng ng-angkep maleng.*
 Wati remember Atin story that(c) she capture thief
 'Wati remembered about Atin the story that she captured the thief.'

Finally, Davies (2005) notes that alternatively, the proleptic construction may involve a PP. This also likens Madurese to the Hungarian, German and English structures.

- (48) *Siti ngera parkara Hasan ja epareksa dokter juwa.*
 Siti think about Hasan that(c) examine doctor DEMONSTRATIVE
 'Siti thinks about Hasan that that doctor examined him.'

The data show that the properties of Hungarian OF show considerable parallels with those of Madurese prolepsis as described by Davies. The difference between Hungarian OF and Madurese prolepsis seems to lie in the structures' productivity. In Hungarian, only a subset of the verbs that occur with a thematic adjunct PP can participate in the direct object version of prolepsis. However, according to Davies's (2005) assessment, all such verbs are suitable for the direct object version of prolepsis.

In Greek, the construction that most closely resembles prolepsis is referred to as "quasi-ECM" by Kotzoglou & Papangeli (2007). In (49) it can be seen that the matrix verb takes a direct object and this direct object is co-referent with the (null) subject of the embedded clause. In (49a) the embedded clause is subjunctive, but in Ingria (1981) and Philippaki-Warburton (1987) we find examples with indicative clauses as well. (49b) is an example from Ingria (1981).

- (49) a. *I epivates perimenan ton kapetanio na ferthi me*
 the passengers expected.3PL the captain.ACC SBJV behave.3SG with
aksioprepia.
 dignity
 ‘The passengers expected of the captain that he should behave with dignity.’
- b. *Theoro ton jani pos ine eksipnos.*
 consider the John.ACC that(c) he smart
 ‘I consider John to be smart. / I believe of John that he is smart.’

Kotzoglou & Papangeli (2007) report that the construction occurs with a few matrix verbs, including, including the ones in (33) and also *ipologhizo* ‘estimate’, *pistevo* ‘believe’, *ksero* ‘know’, *thelo* ‘want’. The words in this set express some mental state or propositional belief, and as such, they are comparable to the Hungarian data.

Although the subject of the embedded clause is null in (49a), Kotzoglou and Papangeli (2007) show by various means that the embedded clause does in fact contain an independent subject. Here, only one piece of evidence for this is reproduced. Emphatic modifiers in Greek attach to and agree in case with a host. It is possible to put such an emphatic modifier in the embedded clause in the “quasi-ECM” construction, which indicates that there must be a host for it. Moreover, it bears nominative case, which suggests that the host is not the direct object of the distant main clause.

- (50) *I epitheorites ithelan to jani na lisi monos tu/*
 the inspectors wanted.3PL the John SBJV solve.3SG alone.NOM
**mono tu to provlima.*
 alone.ACC his the problem
 ‘The inspectors wanted of John that he should solve the problem on his own.’

As in the case of in proleptic constructions in the previously discussed languages, the Greek “quasi-ECM” is not compatible with idiom chunks, see (51).

- (51) **Perimena psilus na tu bun st’ aftia.*
 expected.1SG fleas.ACC SBJV his get.3PL in.the ears
 Only lit.: ‘I expect of the fleas that they will get into his ears.’ (*idiomatic: I expect him to become suspicious.)

Another similarity with the previously discussed languages is the possibility of a paraphrase with a PP proleptic element, as shown in (52).

- (52) *I epivates perimenan apo ton kapetanio na ferthi me*
 the passengers expected.3PL from the captain SBJV behave.3SG with
aksioprepia.
 dignity
 ‘The passengers expected from the captain that he should behave with dignity.’

Although Kotzoglou and Papangeli (2007) seem to argue for an argumental analysis for both the PP and the direct object version of the structure, given the analyses discussed for the

proleptic constructions so far it seems more likely that an adjunct status should be assigned to the PP instead, while the argumental analysis could be maintained for the direct object version.

All these properties liken the “quasi-ECM” constructions to the proleptic constructions discussed so far. There are a number of differences, however. The first is that the construction is not unbounded, that is, the proleptic element must be co-indexed with an argument in the immediately embedded clause. Also, this embedded argument must be the subject of the embedded clause (Salzmann 2017: 25).

These properties make the Greek construction bear closer resemblance to traditional object-control structures like the English example in (53).

(53) *I persuaded John to leave.*

Indeed, this is the kind of analysis that is suggested by Kotzoglou & Papangeli (2007). They argue that the “quasi-ECM” construction is a case of “semantic control”, whereby the proleptic element gets a “weak theta role” from the main verb and then it gets co-indexed with the embedded subject. This assessment is very close to what I argue to be the case.

Finally, let us turn to Korean and Japanese, which possess the following constructions that could be regarded as cases of prolepsis. (54) is a Korean example from Yoon (2007), while (55) is a Japanese one, from Hoji (2005).

(54) *Cheli-nun Yenghi-lul yenglihaysstako mitnunta.*
 Cheli-TOP Yenghi-ACC smart believed
 ‘Cheli believed Yenghi to be smart.’

(55) *John-wa Mary-o Itariazin da to omotteita.*
 John-TOP Mary-ACC Italian be that(c) thought
 ‘John believed Mary to be Italian.’

There has been a debate about the proper analysis of such sentences. Three analyses have been put forward. In the first one (proposed for example by Hong S. M. 2005 for Korean and Tanaka 2002 for Japanese), (54)-(55) is analyzed as “raising to object”, where the accusative marked element starts out as the subject of the embedded clause and then moves to the main clause. This is what the standard analysis for the English glosses of the sentences is. Another kind of raising analysis is offered by Yoon (2007), who proposed that what moves is the “major subject” (a topic constituent) of the embedded clause. In this analysis the accusative-marked elements do not start out as the objects of the embedded clauses, but are instead base-generated at the left periphery of the embedded clause and bind a zero pronominal variable in the object position. The third analysis is basically a prolepsis-analysis. In this analysis, proposed for example by Hong K. S. (1997) for Korean and Hoji (2005) and Miura (2008) for Japanese, the accusative-marked constituent is base-generated as the direct object of the main predicate.

Yoon (2007) explains in detail that the first, object-raising analysis can be readily ruled out. To start with, the embedded clauses are finite and the construction can target nonsubject functions of the embedded clause. Such facts, while not entirely fatal, are unexpected under the standard raising analysis. (56a) is a Korean example involving a possessor and (56b) is from Japanese.

- (56) a. *Na-nun Cheli-lul meli-ka cohtako mitnunta.*
 I-TOP Cheli-ACC head good believe
 ‘I believe of Cheli that his head is good.’/ ‘I believe Cheli to be smart.’
- b. *Taroo-wa Hanako-o atama-ga ii to shinjiteiru.*
 Taroo-TOP Hanako-ACC head-NOM good COMP believes
 ‘Taroo believes of Hanako that his head is good.’/ ‘Taroo believes Hanako to be smart.’

Moreover, like the structures in the previous section, the Japanese construction at hand is not affected by island-constraints. (57) is a Korean example.

- (57) *Na-nun Yenghi-lul kunye-ka hanun ili mopemcekilako sayngkakhanta.*
 I-TOP Yenghi-ACC she-NOM do work exemplary think
 ‘I think of Yenghi that the things she does are exemplary.’

Finally, a number of interpretative properties militate against the standard raising analysis. For example, the construction fails on the familiar idiom chunk test.

- (58) *Hankwuksalam-un cakun kochwu-lul mayptako sayngkakhanta.*
 Korean-TOP small pepper-ACC hot think
 Only lit.: ‘Koreans believe small pepper to be hot.’ (*Idiomatic: ‘Koreans believe that size does not matter.’)

The distinction between the “major subject”-analysis and the prolepsis analysis is actually quite delicate. Since the “major subject” is a kind of topic that sits at an edge-position of the lower clause, the island and idiom facts follow. In (57) what is inside the island is the zero pronominal co-indexed with the major subject. Also, the major subject being a topic, incompatibility with idioms is expected.

What makes Yoon (2007) side with the “major subject”-analysis is that there are certain properties of the accusative element that suggest that it originates in the lower clause. For example, sometimes an oblique case assigned by the embedded predicate is retained and the accusative is stacked on top of it, demonstrated in (59).

- (59) a. *Cheli-hanthey-(man)-i mwuncey-ka issta.*
 Cheli-DAT-(only)-NOM problem-NOM exist
 ‘(Only) Cheli has problems.’
- b. *Na-nun Cheli-hanthey-(man)-ul mwuncey-ka isstako mitnunta.*
 I-TOP Cheli-DAT-only-ACC problem-NOM exist think
 ‘I believe only Cheli to have problems.’

Interestingly, although Yoon (2007) subscribes to the “major subject” analysis, he admits that in certain cases, an alternative parse is possible for such accusative-marked entities, as “many of the matrix verbs that govern subject-to-object raising have a usage where they take the ACC-marked DP as an argument” (Yoon 2007: 99). These may be genuine proleptic structures. Furthermore, these argumental uses alternate with an adpositional version. A Korean example is shown in (60).

- (60) *Na-nun Cheli-lul/ Cheli-eytayhay kunyese*k*-i tollassta-ko sayngkakhay.*
 I-TOP Cheli-ACC/ Cheli-regarding that.guy-NOM crazy-that(c) think.1SG
 ‘I think of Cheli that the guy is crazy.’

Only when special care (like case-stacking in 59) is taken to rule out the proleptic parse can the “major subject”-analysis be firmly established.

It might be added that, according to Miura (2008), pronouncing the embedded subject is generally dispreferred in Japanese, but this effect can be ameliorated by adding a focus particle to the resumptive pronoun. This is comparable to the Hungarian situation, compare (61) below with (60b) and footnote 53 in section 4.2.1.

- (61) *Taroo-ga Ken-o kare-koso-ga tensaida to omotteita.*
 Taroo-NOM Ken-ACC 3SG.MASC-FOC-NOM genius that(c) thought.3SG
 ‘Taroo thought of Ken_i that HE_i was a genius.’

Here I cannot do justice to the Korean and Japanese situation. While Yoon (2007) may be right that (54)-(59) are best analyzed not as prolepsis but as raising, (60) suggests that prolepsis may play a role in the syntax of these languages as well. Further research is needed for the to establish a precise differentiation between the two structures.

Summary of the crosslinguistic landscape

Having surveyed proleptic constructions in a number of languages, it is time to take stock. The following properties connect all these constructions:

- i. There is a matrix verb that has 3 dependents: the subject, the proleptic element and a finite complement clause.
- ii. The proleptic element bears a subject matter semantic relationship with the main verb.
- iii. Because of this subject matter semantic relationship, the proleptic element must be a referential entity.
- iv. There is a resumptive pronoun in the complement clause, co-indexed with the proleptic element.

The dimensions of divergence among the constructions are the following:

- i. How productive is the proleptic pattern lexically? How many matrix verbs allow prolepsis? (fully productive, e.g. German, Madurese or more restricted, e.g. Hungarian, Greek)
- ii. Functionally, is the proleptic element a (thematic) adjunct (as in German, English) bearing oblique case, or is it a direct object/oblique argument (as in Hungarian, Greek, Madurese)?
- iii. Are there restrictions on the function and the distance of the embedded resumptive? (Greek: yes, other languages: no.)

It seems that every language surveyed has a version of prolepsis where the proleptic element is definitely not an argument of the main verb, but a thematic adjunct (elements marked with adpositions like *-róll/-ről* in Hungarian, *of* in English, *von* in German, *apo* in Greek, *eytayhay* in Korean).

The version with thematic adjuncts and the argumental oblique version in Hungarian seem to be more productive than the direct object version of prolepsis. In Hungarian we observed that virtually all verbs allowing prolepsis with a direct object have an adpositional variant. The reverse is not true: there are verbs that allow the proleptic element to be realized as an adpositional phrase, but they do not allow the accusative version. Also, in German and English the construction seems to be quite free, as Salzmann (2017) also notes. The two versions are equally free in Madurese as discernible from Davies's (2005) account, despite an absence of a detailed account of this freeness. In Greek and Japanese, the accusative version is possible with a restricted set of verbs.

All this suggests that the adpositional thematic adjunct version could be a base structure for the accusative version. If the predicate is able to take an adpositional dependent as a thematic adjunct, then an additional argument-structural mechanism of the kind proposed in section 5.2.3 can advance this relationship into a true argumental one, realizing the proleptic element as a direct object in the process and thus turning it into a "derived argument" in Toivonen's (2013) sense. The productivity of the conversion seems to be a language-specific property. In Hungarian, Greek and Japanese it is fairly restricted, but for example "in Madurese (...), there is a very productive applicative-type process that creates core (non-prepositional) arguments out of oblique arguments" (William D. Davies, p.c.). This seems to be on the right track, though the obliques should be thought of as thematic adjuncts, not arguments. The function of the proleptic element and its resumptive dependent in the embedded clause is quite free, with the exception of Greek where it is restricted to the immediate subject.

Something along these lines is also suggested in den Dikken (2010: 1) in a Minimalist framework. In this account the distinction between the delative and accusative version as prolepsis in Hungarian translates as the proleptic XP originating in either inside or outside the the matrix clause VP. I will discover such theoretical repercussions of prolepsis in the next section.

5.3.2 Theoretical perspectives on prolepsis

Let us now change perspective and try to find a place for prolepsis in the general space of grammar. The Minimalist analysis offered by Salzmann (2006, 2017) is that prolepsis involves predication, with the aid of a null operator. In his approach, verbs taking a CP-complement are specified by a lexical redundancy rule that they can "take a CP whose head is specified for requiring a silent operator in its spec" (Salzmann 2017: 26). This silent operator turns the CP into a predicate, which then licenses the proleptic object as the subject of this predication. The predicate and the proleptic object together constitute a proposition that satisfies the semantic requirements of the main verb.

From an LFG-perspective, the null operator analysis is obviously problematic, since it involves a zero c-structural entity and such elements are not part of the architecture. Apart from this theory-specific objection, Salzmann does provide a detailed and principled analysis for prolepsis in German which possibly extends to other languages as well. The compulsory nature

of the co-referent entity constitutes a problem for an LFG account if it is accepted that the proleptic PP is an adjunct. It seems that distinguishing this adjunct from normal adjuncts by assuming it to be a thematic adjunct is motivated. However, it is not clear to me at this point how this translates to the obligatory presence of the co-referent entity.

However, it has to be added though that Salzmann's (2017) claim about obligatoriness may be overstated. As he himself notes in his footnote 24, a statement about X quite naturally includes X. Since the semantics of proleptic adjunct is quite close to an aboutness relation, it is no surprise that there is a tendency to include the relevant entity in the complement clause itself. I cannot make definitive claims about the German datum in (40), repeated here as (62), but the Hungarian equivalent in (63a) does not seem to be that bad, especially compared to the argumentalized proleptic version (63b).

- (62) **Von Computern glaube ich, dass jeder einen PC kaufen sollte.*
of computers.DAT believe.1SG I that(c) everyone a.ACC PC
buy.INF should.3SG
'I believe of computers that everyone should buy a PC.'
- (63) a. ?*A számítógépek-ről azt hiszem, hogy mindenkinek venni kéne egy PC-t.*
the computers-DEL that.ACC believe.1SG that(c) everyone.DAT buy.INF
should one PC-ACC
'I believe of computers that everyone should buy a PC.'
- b. **A számítógépek-et hiszem, hogy mindenkinek venni kéne [*egy PC-t/ egyet].*
the computers-ACC believe.1SG that(c) everyone.DAT buy.INF should
one PC-ACC one
'(Of) computers I believe that everyone should buy one.'

A similar contrast is obtained in (64). The version with the delative adjunct is less than perfect, but the argumental version is plainly ungrammatical. If a proper context is set up like in (65a), the oddness of not having a coreferential element in the complement clause may be diminished. A proper context is one where the aboutness relation may be plausibly established, e.g. via the inclusiveness relation between the garden and the trees in (65a). The same context does not help the argumental proleptic version.

- (64) a. ?*Azt gondolom János-ról, hogy a lakás unalmas.*
that.ACC think John-DEL that(c) the flat boring
'I think of John that the flat is boring.'
- b. **János-t gondolom, hogy a lakás unalmas.*
John-ACC think.1SG that(c) the flat boring
'(Of) John I think that the flat is boring.'
- (65) a. *Azt gondolom a kert-ről, hogy gyönyörűek a fák.*
that think.1SG the garden-DEL that(c) beautiful the trees
'I think of the garden that the trees are beautiful.'

- b. *A kert-et gondolom, hogy gyönyörűek a fák.
 the garden-ACC think.1SG that(c) beautiful the trees
 ‘I think of the garden that the trees are beautiful.’

What this suggests is that the requirement of having an obligatory coreferent element in the case of proleptic adjuncts may be illusory, simply being a result of pragmatic constraints on communication. The English translations in (62)-(65) seem to be amenable to such an approach.

In our version of Hungarian proleptic OF, the linking issue does not arise, since the lexical entry participating in OF was argued to be an added argument as a result of a lexical process, which may include the addition of such specifications. In what follows, I will argue that there is theoretical and empirical motivation for an approach where argumental prolepsis is taken to be a subspecies of control-constructions.

To investigate prolepsis from the perspective of control, let us first present a broad theoretical overview of control. “Control” in the Chomskyan tradition is a term for constructions where the referent unpronounced subject of an embedded nonfinite clause is determined (“controlled”) by a main clause subject or object. The unpronounced subject is called a PRO.

- (66) a. *John_i tried [PRO_{i/*j} to go].*
 b. *I persuaded John_i [PRO_{i/*j} to go].*

The overt subject (*John* in 66) is called the controller, while the PRO is the controllee. The controllee is traditionally assumed to be restricted to the immediately embedded subject.

Control (66) is often contrasted with “raising”, shown in (67). The latter superficially mirrors the control-sentences in (66), but it is argued to have a distinct analysis, whereby the controller is originally an element of the embedded clause which subsequently moves to the main clause.

- (67) a. *John_i seemed [~~John_{i/*j}~~ to laugh].*
 b. *I believe John_i [~~John_{i/*j}~~ to be laughing].*

It is a commonplace observation that *John* is semantically related to *try* and *persuade* but not to *seem* and *believe*. That is, while *John* is a “trier” in (66a) and a “persuadee” in (66b), he is not a “seemer” in (67a) or a “believer” in (67b).

Table 1 shows a collection of control and raising verbs from Landau 2013:10 (for more predicates, see Davies and Dubinsky 2004: 11-12).

	Subject	Object
Control	<i>try, condescend, promise, decide, plan, agree, hope, prefer, wonder, refrain</i>	<i>persuade, encourage, recommend, appeal, force, plead, order, urge, dissuade</i>
Raising	<i>seem, appear, turn out, happen, begin, continue, stop, likely, certain, sure</i>	<i>believe, consider, prove, show, take, expect</i>

Table 1.

Control and raising predicates

Beneath the surface similarity, there is a wide range of properties that distinguishes the two constructions. Without the goal of completeness, we will survey here the most important distinctions.⁶³

At least two features follow from the fact that the subject of *try* in (66a) is thematic, while the subject of *seem* in (67a) is non-thematic. First, expletive subjects can be associated with *seem* but not with *want*. (68b') shows that the problem is not simply that *want* cannot take a *that*-clause complement.

- (68) a. *It seems that John is happy.*
 b. **It wants that John is happy.*
 b'. **It wants for John to be happy.*

Second, since they lack individual meaning, idiom chunks can only appear in raising constructions, not in control ones.

- (69) a. *The cat seems to be out of the bag.* (idiomatic or literal)
 b. *The cat wants to be out of the bag.* (only literal)
 (70) a. *I believe the cat to be out of the bag.* (idiomatic or literal)
 b. *I persuaded the cat to be out of the bag.* (only literal)

The basic analytical distinction between equi and raising is that while in the case of control, the controller is entirely interpreted in the main clause (in GB/MP terms one may say that it is base-generated there), while in raising the controller also sustains strong ties with the controllee-position in the subordinate clause.

In GB/MP, the analysis of raising is relatively uncontroversial. Putting the technical details aside, the standard analysis posits that the subsequent matrix subject is base-generated in the subject position of the embedded predicate and is moved ("raised") into the matrix clause subject/object nonthematic subject or object position.⁶⁴ For the analysis of control, there are

⁶³ An extended survey is to be found in Landau (2013: 8-28).

⁶⁴ Raising to subject is entirely uncontroversial, while Exceptional Case Marking (ECM) is sometimes offered as an alternative to raising to object. This debate largely stems from the theoretical layout of GB/MP, so I will not go into details here.

two main approaches: a movement-based approach (Boeckx, Hornstein & Nunes 2010) and Landau’s (2015) two-tiered account.

According to the movement-theory of control, the analysis parallels that of raising. That is, the controller also starts out in the embedded clause and then moves into its final, main-clause position. The only difference is that in the case of control, both the embedded and the main clause positions are thematic (or A-positions, in Chomskyan terminology). This means that the constituent in question ends up with two theta-roles, so the original formulation of the Theta-criterion has to be loosened.

The other main approach is Landau’s (2015) “two-tiered” model. In this model, control may be established in two ways: either by predication (tier 1) or by variable binding (tier 2), where the second tier may be superimposed on the first one in attitudinal contexts. As we will see soon, Landau’s view is actually very close to the LFG’s conception of control, where there are also two basic ways to establish a control relation.

In LFG, the term “control” is used more inclusively, for all constructions in which there is a lexically specified identification (functional/anaphoric) between two f-structure entities, so both raising and “GB/MP’s control” are covered by the term. What the GB/MP tradition calls “control” runs by the name “equi⁶⁵” in LFG-literature. In other words, the Chomskyan tradition makes a control-raising distinction and no general term covering both, LFG has an equi-raising distinction and “control” is a general term for this range of phenomena. Since the chosen framework of this monograph is LFG, I use the latter terminology.

LFG utilizes functional or anaphoric identification in the analysis of control constructions. The identification is specified in the lexical entries of the relevant predicates. Usually, only one of these entities is phonetically realized (in other words, only one of them has a c-structural exponent).

As mentioned in section 1.2, functional identification means strict f-structural identity, with one element providing value for two f-structural attributes. In the case of raising predicates, the subject or the object argument is nonthematic. The raised constituent is functionally identified with the embedded subject. This identification is specified in the lexical entry of *seem*, shown in (61).

- (71) *seem* V <(↑XCOMP)> (↑SUBJ)
 (↑SUBJ)=(↑XCOMP SUBJ)

The propositional argument is encoded as an XCOMP, the open complement function. This grammatical function allows its SUBJ to be provided from the outside. The f-structure for (57a) is provided in Figure 1.

⁶⁵ An abbreviation for “equivalent NP-deletion”, a terminology from early Chomskyan analyses.

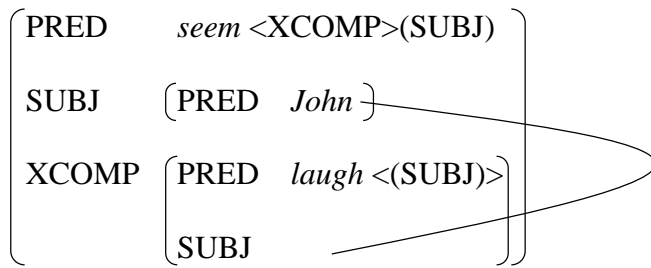


Figure 1.

F-structure of (67a) (*John seemed to laugh.*)

According to Falk (2001), equi-constructions can involve either functional or anaphoric identification, depending on the specification of the individual predicates. This is conceptually quite close to Landau's (2015) dichotomy of predicative/variable-binding control: XCOMP is a predicative complement in LFG, while anaphoric identification actually involves binding a *pro* variable in the f-structure.

Falk (2001: 137) proposes syntactic tests to determine which analysis is the correct one in individual cases. These are based on the difference between functional and anaphoric identification. Recall that functional identification is a strict syntactic identity, while anaphoric identification is more like a semantic connection.

(...) in functional control, the controller must be present (in the f-structure) and it must be a core function. If it were absent, the subordinate clause would be incomplete because it would lack a SUBJ, and (...) only core functions can be specified by a control equation. On the other hand, neither of these is necessary for anaphoric control: pronouns need not have antecedents and there is no restriction on the grammatical function of an antecedent of a pronoun. On the other hand, anaphoric control should allow split controllers, because pronouns can take split antecedents; while a functional controller is the single element specified by the control equation. (Falk 2001: 137)

Based on these, Falk (2001: 138) shows that *try* involves functional identification, while *agree* utilizes anaphoric identification.

- (72) a. *It was tried (by the geneticist) to clone dinosaurs.*
 b. *It was agreed (by the geneticist) to clone dinosaurs.*
- (73) a. *The geneticist tried to clone dinosaurs.* (SUBJ of clone: the geneticist only)
 b. *The geneticist agreed to clone dinosaurs.* (SUBJ of clone: the geneticist + possibly other people)

What (73b) also shows is a phenomenon called "partial control" (named by Landau 2000), where in addition to the matrix controller, other understood controllers are also possible. Similar predicates are e.g. *want, prefer, yearn, arrange, hope*, etc. Verbs that pattern with *try* are referred to as "exhaustive control" verbs. Other examples are *avoid, forget, fail, refrain, decline, neglect*, etc. The most satisfactory analysis of this phenomenon is under discussion in the linguistic literature, see Boeckx, Hornstein & Nunes (2010:182-190), Landau (2016), Pitteroff,

Alexiadou & Fischer (2017); also see Haug (2013) in LFG framework. Without going into the details of the problem of partial control, it can be stated that any LFG analysis for it must be based on anaphoric identification (semantic coreference), as the strict syntactic identity of functional identification would make any semantic mismatch impossible.

However, while the discussion in the previous paragraph does have the logical consequence that functional identification always means exhaustive control, it does not entail that all cases of exhaustive control should be viewed as involving functional identification. That is to say, while cases like partial control must be analyzed in terms of anaphoric identification, I would not like to suggest that anaphoric control can never have strict semantics regarding the control relationship. In fact, for cases of equi for which Falk (2001) proposes functional identification, Dalrymple (2001) puts forward an analysis in terms of “obligatory anaphoric identification”. While I agree with Falk (2001) in that functional identification is more appropriate for the English examples at hand, I also agree with him in that obligatory anaphoric control is a possibility in LFG (see Falk 2001: 138). So while functional identification is a unitary phenomenon in terms of the semantics of the control relationship, anaphoric identification may have different types, ranging from “obligatory” through “quasi-obligatory” (Haug 2013) to arbitrary. The analysis of specific constructions in specific languages should boil down to careful investigation of these particular constructions and languages.

Let us make an interim summary at this point. So far, we have discussed two aspects of variation in control constructions from an LFG-perspective. On the one hand, a control construction may be equi involving a thematic controller or it can be an instance of raising involving a nonthematic controller. On the other hand, the control relationship may be either anaphoric or functional. Since functional identification can be thought of as an analogue for movement, while anaphoric identification can be likened to PRO-based GB/MP approaches, the LFG architecture provides theoretical space for both main Chomskyan approaches to equi, without having to choose between them. Such a flexibility is a long-standing merit of LFG and has been noted to be in line with linguistic diversity (Levinson & Evans 2010). As a result, the following taxonomy emerges.

The reason for anaphoric raising being unexpected is that such a configuration would lead to a violation of Semantic Coherence, thus the LFG architecture excludes such a construction. Actually such an analysis was argued by Coppock (2003) for Hungarian OF, which I have analyzed as prolepsis. I have argued extensively that a prolepsis-account is on both theoretical and empirical grounds. Since anaphoric raising has not been put forward for other constructions in any language, the taxonomy is legitimate at this point.

Here, the following question arises: how does prolepsis fit into this picture? I have argued that the proleptic constituent is a thematic argument of the main verb, so it is an equi-type construction. Also, in order to account for the possible variation of the number feature of the embedded predicate, I analyzed it as involving anaphoric identification. Consequently, prolepsis is related to the *agree*-type constructions. From this perspective, prolepsis is a kind of object-equi (like sentences with *persuade*, *encourage*, *force*, etc.), utilizing an obligatory anaphoric identification.

CONTROL-TYPE		Example
Thematicity of controller	Nature of identification	
Equi	anaphoric identification	canonical control, <i>agree</i> -type
	functional identification	canonical control, <i>try</i> -type
Raising	anaphoric identification	NOT EXPECTED
	functional identification	canonical raising

Table 2.
LFG's taxonomy of control (first version)

The primary feature in which prolepsis differs from canonical control-constructions is that it involves a finite complement clause. In earlier stages of generative grammar, it was thought that control involves non-finite complements only. However, this has changed and several constructions in various languages have been described that necessitate positing control into finite complements. If this is so, then the taxonomy presented in Table 2 can be expanded so that it encompasses finite versions of control.

CONTROL-TYPE			Example
Thematicity of controller	Nature of identification	Finiteness	
equi	anaphoric	finite complement	PROLEPSIS
		non-finite complement	canonical control, <i>agree</i> -type
	functional	finite complement	
		non-finite complement	canonical control, <i>try</i> -type
raising	anaphoric	finite complement	not expected
		non-finite complement	not expected
	functional	finite complement	
		non-finite complement	canonical raising

Table 3.
An LFG-taxonomy of control (expanded version)

The question here is whether there is justification for making a parallel between prolepsis and equi-type control relations. I consider such a parallel to be justified.

Let us see here a prototypical Hungarian proleptic example in (74) and its f-structure in Figure 2. The proleptic object is a matrix clause argument and it is co-referent with the embedded clause subject. The co-reference is lexically induced by the lexical entry of the

proleptic verb containing the equation $OBJ\ INDEX = \{COMP^+ GF^+\} GF\ INDEX$. Compare this with a garden-variety object-equi structure from English: (75) and Figure 3.

Apart from the information structural side (which is a by-product of Hungarian sentence-structure) the similarity is evident. In both cases, the main clause thematic object is anaphorically identified with the subject of the embedded clause. Although there are some important differences between the standard case of control and prolepsis, there are crucial parallels to be drawn. Because of these similarities I think it is warranted to treat prolepsis as a special subtype of control.

- (74) *Jánost mondtad, hogy jön.*
 John.ACC said.2SG that(C) comes
 ‘(Of) John you said that he will come.’

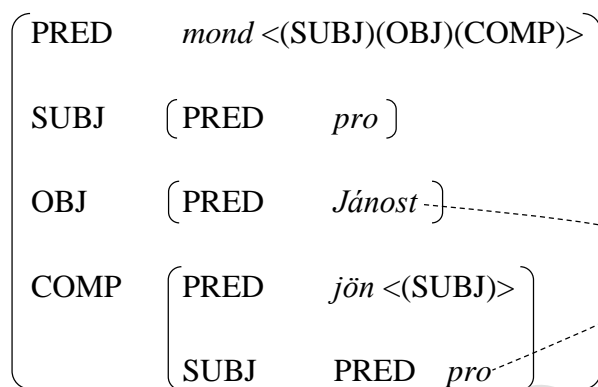


Figure 2.
 F-structure of (74)

- (75) *John persuaded Mary to go.*

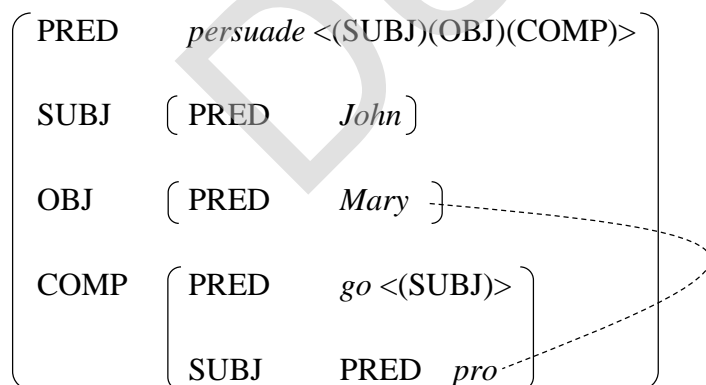


Figure 3.
 F-structure of (75)

First let us see the differences. One difference is that equi is traditionally seen as involving non-finite complement clauses, while prolepsis involves a finite one. However, this standard view is false, as control can involve certain finite embedded clauses (Landau 2013), so this is not a real difference, but only a typological variation. We will discover such cases in section 5.4.2 (and we will also clarify what is meant by a clause being “(non-)finite”).

Another difference is that standard control is a local dependency between a main clause argument and an immediately embedded clause. Prolepsis, on the other hand, can be long-distance.

- (76) a. *John persuaded Mary to kiss herself.*
 b. **John persuaded Mary to say that he should kiss herself.*
 c. *Jánost mondtad, hogy hallottad, hogy jön a partira.*
 John.ACC said.2SG that(C) heard.2SG that comes the party.to
 ‘(Of) John you said that you heard that he is coming to the party.’

A third difference is that in standard equi, the controlled element is always the embedded subject (67a), while prolepsis can involve a variety of grammatical functions (67b-c).

- (77) a. **I persuaded Mary for John to like.*
 b. *Két almát mondtál, hogy vettél.* (embedded OBJ)
 two apple.ACC said.2SG.INDEF that(C) bought.2SG.INDEF
 ‘(Of) two apples you said that you had bought.’
 c. *Párizst mondtad, hogy oda utazol.* (embedded OBL_θ)
 Paris.ACC said.2SG.DEF that(C) there travel.2SG.
 ‘(About) Paris you said that you will go there.’

The fact that prolepsis involves an anaphoric and not a functional link is probably relevant for these two differences. Anaphoric dependencies involve an f-structural *pro* in LFG, which is no different in terms of the PRED feature from ordinary pronouns. That is, the inherent distinction between PRO and *pro* in GB/MP is absent from LFG. As pronouns are not restricted to find their antecedents in the immediately containing clause, a long-distance anaphoric dependency is actually not that unexpected.

As for the grammatical function of the controllee, in the theory of subjects in LFG framework in Falk (2006b), it is established that lexically specified functional control equations can only refer to subjects, because of a restriction by made Universal Grammar. However, no such constraint has been put forward in connection with anaphoric dependencies. This can provide space for the conceptual possibility of prolepsis as anaphoric identification with a range of grammatical functions.

Now let us take stock of the similarities, apart from the analytical similarity already shown in connection with Figure 2 and 3 and the facts about the non-availability of idiom-chunks, which was demonstrated in 4.2.1. Landau (2013), surveying the properties of equi, shows that in control, the controllee (the embedded subject) behaves as a bound variable. Consider the pair in (68).

- (78) a. *Only John claimed to have won.*
 b. *Only John claimed that he won.*

(78a) only has the so-called “bound variable reading” (also called “sloppy reading”): only John claimed about himself that he won. That is, if another person, e.g. Bill also claimed that he

(=Bill) won, (78a) is false. In the same scenario (John claims that John won, Bill claims that Bill won), (78b) is not necessarily false. This is because it may also have the “strict reading”, i.e. that John is the only person who claims that John won. In other words, the identity of the embedded subject is tied to the main clause subject in (78a) and thus it varies according to who is understood as such (John, Bill). In (78b) it can also be strictly anchored to one referent. To clarify this further, see the paraphrases below.

- (79) a. Bound variable reading: John = Only x [x claimed x won].
 b. Strict reading: John = Only x [x claimed John won].

The same lack of ambiguity may be observed in object-control sentences, e.g. in (80).

(80) *Only John forced himself to eat the tasteless food.*

(80) only has the sloppy reading: no other person “self-forced” to eat the food, though it may be possible that several other people also forced John to do so.

Prolepsis behaves like this, as attested by (81a), which can have the meaning that *János* was the only person who claimed self-victory, not that nobody else claimed *János* to be the winner. Compare this with (81b), which is ambiguous (*he* may refer to whoever the matrix subject is (the sloppy reading) or invariantly to *János* (the strict reading) or to somebody else (this is also the strict reading, with an outside referent)).

- (81) a. *Csak János mondta magát, hogy (ő) nyert.*
 only John said.2SG himself.ACC that(c) he won.3SG
 ‘Only John said (of) himself that he had won.’
 b. *Csak János mondta azt, hogy (ő) nyert.*
 only John said.2SG that.ACC that(c) he won.3SG
 ‘Only John said that he had won.’

A similar phenomenon can be observed in elliptical contexts. In equi, the elided part is interpreted sloppily.

- (82) a. *Mary expected to attend the ceremony, and Sue did too (expect to attend the ceremony).*
 b. *Mary encouraged Paul to attend the ceremony, but not David (encourage to attend the ceremony).*

The unexpressed subjects of the elided parts are understood as *Sue* and *David*. So (82a) does not mean that *Sue* expected *Mary* to attend the ceremony, and (82b) cannot mean that *Mary* encouraged *David* in order that *Paul* attends the ceremony.

Again, prolepsis shows this pattern: (83) cannot mean that “you said of Péter that *János* comes”, the agent of *jön* ‘come’ must be *Peter*.

- (83) *Én Jánost mondtam, hogy jön, te pedig Pétert (mondta),*
 I John.ACC said.1SG that(C) comes you but Péter.ACC said.2SG
hogy jön).
 that(c) comes
 ‘(Of) John I said that he comes and you did too Peter.’

The similarities are rooted in the fact that some obligatory equi constructions use the same kind of anaphoric dependency that is utilized in prolepsis.

Consequently, I deem it motivated to locate the typological place of prolepsis in the realm of control, as an instance of obligatory anaphoric finite equi. Given Landau’s (2015) analysis, where predication is involved in control, it may turn out in the long run that a convergence with Salzmann’s (2006, 2017) predicative approach to prolepsis is more than a remote possibility.

In the remaining part of the monograph, I will explore the taxonomy in Table 3 and present examples from various languages to fill it with. These sections do not intend to provide an exhaustive theoretical account for control in general. That would be far outside of the scope of this monograph. Rather, I set out to provide an empirical overview from the perspective of the monograph, though some analytical suggestions will be made nevertheless.

5.4 Control in the finite domain

In order to discuss control in the finite domain, first we have to clarify what is meant by the term “finite”. Surprisingly, although the term has been used for a long time in the study of language and linguists in general seem to have an intuitive understanding of what it is, the exact meaning is elusive and quite difficult to pin down.

There are at least four dimensions that have been evoked in the literature for the characterization of finiteness: tense, agreement, the ability to host its own subject and the capability to assign nominative case to this subject. Much of the difficulty of defining the term arises because these dimensions are largely independent and it is not clear whether a specific property on each of these dimensions is required for a clause to be classified as “finite”, or partially aligned specifications are enough. Also, it is not settled whether any of these dimensions should be regarded as a primary trait of finiteness. Even if a primary trait is delimited, there remains the question of how the other dimensions are related to this primary trait and to each other. Some progress has been made in this respect, but no consensus has been reached and no answers for these pressing questions will be provided in this monograph either. What will be provided is an overview of structures that can be classified as finite according to some of these criteria and some remarks from the theoretical perspective of LFG.

Before briefly describing the aforementioned dimensions, it must be indicated at this point that although in some versions of Chomskyan syntax the option of finite control is a difficulty,⁶⁶ this is not the case in LFG. As far as no other principle of LFG is violated, a clause having tense, agreement, etc. specifications has nothing to do with individual predicates

⁶⁶ For raising, the main clause being able to check case, tense and agreement features is assumed to block the movement of elements out of these clauses, unless other unchecked features (*wh*, focus, etc.) motivate the operation.

For equi, the controllee in standard GB/MP is PRO, an intrinsically zero nominal element. According to the “PRO theorem”, PRO must be ungoverned, but the subject position of any finite clause is governed.

instantiating control-equations in the f-structures. So from an LFG-perspective, finite control constructions should definitely be an option in the linguistic landscape.

Let us briefly introduce the ingredients of finiteness (for in depth discussion, see Nikolaeva ed. 2007). “Tense” is one such ingredient. Having a tense-specification on a predicate is usually taken as an indication of its finiteness. A complicating issue is that tense can be defined either as a morphosyntactic property (overt tense-marking morphemes) or as “semantic tense”, that is, being able to establish independent time reference. The first case is self-explanatory. As for the second, consider English infinitives. These are not marked for tense morphologically but since Stowell (1982) it is known that they can be divided into two groups along their ability to host independent temporal reference. Consider (74).

- (84) a. *#Yesterday, John tried/dared/managed/forgot/avoided/failed/got to solve the problem tomorrow.*
b. *Yesterday, John wanted/agreed/preferred/arranged/hoped/refused to solve the problem tomorrow.*

The second property on our list is agreement, by which we mean morphological person/number marking on the predicate. According to the theory of Landau (2013), tense and agreement exhaust the list of finiteness ingredients.

The property of hosting a subject is sometimes also mentioned as a defining trait of finite clauses (Subirats-Rüggeberg 1990). To make this dimension fully operational, one has to factor different conceptions of subjects into it. In GB/MP, raising sentences do have a trace/copy in the subject position of infinitival clauses, while equi sentences are commonly analyzed as a PRO occupying these positions.⁶⁷ Also, some analyses regard the *for*-phrases associated with infinitivals (*I hope for Bill to win*) and the objects of “Exceptional Case Marking” sentences (*I believe John to be happy*) as some sort of subjects.⁶⁸ In LFG, a further question that arises regarding this criterion is whether the relevant level for checking the presence of a subject is the c- or the f-structure. The subjects of XCOMPs are “empty” only in the sense that they are predicated from the outside via functional identification. Otherwise, the subject is f-structurally “fully there”. (This is paralleled by the copy-theory of movement in Minimalism.) In anaphoric identification, the complement clause (bearing a COMP grammatical function) has its independent “pro” subject, which in LFG theory is no more special than the zero subjects of pro-drop languages, but this “pro” is absent in the c-structure. The reverse situation is less likely, since any c-structural element would contribute some attribute-value pair to the f-structure, unless it is excluded by some special mechanism.

Finally, in the theory of Cowper (2016) a clause hosting a subject is not sufficient for a clause to be finite: it is the subject bearing structural case that is the crucial diagnostics of finiteness. It is not entirely clear whether structural case only involves the nominative or subjects with accusative, dative or oblique cases are also relevant. For Cowper, finiteness has no semantic content in itself so features like “tense” can only be indirectly associated with it.

After this familiarization with the concept of finiteness, let us investigate some finite control structures in various languages. At any rate, prolepsis is unquestionably finite since the

⁶⁷ Also, one must reckon with the “pro” subjects of pro-drop languages.

⁶⁸ Of course, such cases may be argued not to be proper subjects. My aim here is to highlight surrounding issues rather than adjudicating them.

complement clause has independent tense, agreement and can host its subject both in f- and c-structure.

The schedule is the following. First, we will look at English and some Bantu languages which have raising constructions out of finite clauses. Actually this is a small sample of languages with such constructions, for an extensive discussion, see Ademola-Adeoye (2010). The following step will be a discussion of finite equi-constructions in Hungarian, Greek, Serbo-Croatian and Turkish. Finally, before concluding the chapter, a summary of these finite control constructions will be outlined.

5.4.1 Finite raising

The basic f-structural setup for a raising construction needs the following ingredients: a predicate with a non-thematic argument and an XCOMP. The athematic argument is functionally identified with the empty subject function of the XCOMP. The identification is encoded in the lexical entry of the predicate in the form of an equation like (SUBJ)/(OBJ)=(XCOMP SUBJ).⁶⁹ Since the defining property of XCOMPs is this functional identification and not their c-structural categorial expression (Asudeh 2002), the c-structural categorial status of the XCOMP is irrelevant: it can equally be a VP, an IP or a CP. In these latter cases, the clause may bear a full range of tense and agreement features. If the embedded subject is independently expressed, we have what the literature calls a “copy raising” construction. If it is not, we have “hyperraising”. We will see that e.g. English instantiates the first type (85), while Bantu languages the second one (86). (86) is from Carstens & Diercks (2013).

(85) *Richard_i seems like he_i smokes.*

(86) *Chisaang'i chi-lolekhana mbo chi-kona*
 animal SM-seem that(c) SM-sleep.PRES
 ‘The animals seem to be sleeping.’

Let us take a closer look at these constructions. copy raising (CR) in English can occur with verbs related to perception. In CR, these take a finite complement clause (with agreement and independent tense) headed by *like/as if* and there is an obligatory co-referent pronoun (the “copy pronoun”) in this complement clause.

(87) *John_i seems/appears/looks/feels/sounds like he_{i/*j} is going to win the race.*

A complicating factor in CR is that the main clause subject might be regarded as a thematic argument of the main verb in certain cases. Although the most recent LFG-analysis of CR by Asudeh & Toivonen (2009) explicitly denies this possibility, claiming that the CR-subject can only bear a “semantic role” (a weaker relation than thematic roles), Landau (2011) convincingly argues that an argumental analysis is simpler and more satisfactory for these cases of CR. In

⁶⁹ In the languages surveyed here, it was the matrix SUBJ that undergo identification with the embedded subject. However, Ademola-Adeoye (2010) provides a wide range of examples for finite object-hyperraising structures.

fact, if these predicates assign a “perceptual source” thematic role to their subjects, a copy in the complement clause is no longer necessary. In this case, the concept of control is unnecessary. The construction in (88) is not CR then, by definition.

(88) *John seems/appears/looks/feels/sounds like Mary is going to win the race.*

Of course, pragmatic difficulties may arise when the perception of the main clause subject cannot be plausibly used for inference about the event of the embedded clause, but these restrictions are not syntactic. The main predicates in such cases have dyadic lexical entries similar to (89), where the SUBJ is associated with a perceptual source thematic role and the COMP is a proposition.

(89) *seem/appear/...* <(↑SUBJ)(↑COMP)>

Hence, these cases must be disregarded as no-control scenarios. However, there are cases where the matrix clause subject is non-thematic. This should come as no surprise since the predicates under consideration can all occur with expletive *it* subjects.

(90) *It seems/appears/looks/feels/sounds like/that John is going to win the race.*

Landau (2011) shows that in cases where the main clause subject cannot possibly be construed as a perceptual source due to world knowledge or contextual reasons, a different scenario emerges: the resumptive element becomes obligatory.⁷⁰

(91) *The problem_i sounds like it_{i/*j} is difficult to solve.*

So the following discussion pertains to only those cases of copy raising that are of this second type, i.e. where an obligatory control relationship is involved. As the name suggests, this a raising-type control relationship, the main clause subject does not bear any thematic relationship to the main verb. This can be seen from the fact that expletive-like elements and idiom chunks may participate in CR (Potsdam & Runner 2001).⁷¹

- (92) a. *It seems like it's raining harder than it is.*
b. *There seemed like there was a problem.*
c. *The shit appears as though it's going to hit the fan very soon.*

Based on these examples, it could be plausible to represent CR as the following (simplified) f-structure. Figure 4 essentially resembles a standard raising f-structure.

(93) *Richard_i seems like he_i smokes.*

⁷⁰ This Landau's (2011) “Psource-Copy Generalization”: a copy is not necessary only if the subject is a Psource.

⁷¹ There could be some dialectal variation in the judgment about these sentences.

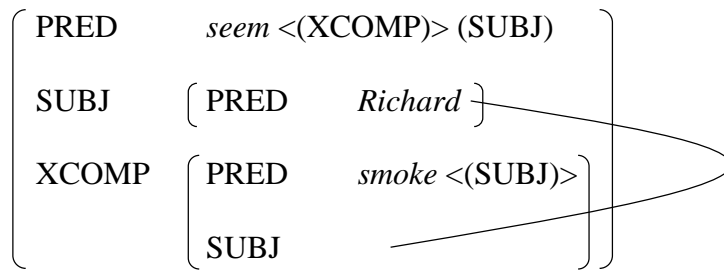


Figure 4.

A potential f-structure of (93)

This intuitive approach is not without problems though. Let us begin with the lesser problem. In the original formulation of LFG’s control-theory (Bresnan 1982), XCOMPs are associated with lexical categories (V, N, P, A) at c-structure. They lack a structural subject-position, which enables them to have their subject specified from the outside via functional identification. However, in this representation there is a subject position in the complement clause, standardly assumed to be Spec/IP or TP in English. However, if one follows Asudeh (2002) in abandoning the XCOMP-VP association and assuming that “the defining property of XCOMP is not its c-structural category, but rather whether it contains a grammatical function that is the target of a functional control equation”, this problem can be bypassed.

A more serious problem is that although I have conveniently left the XCOMP SUBJ empty in Figure 4, allowing it to be targeted by functional identification, the attribute-value pairs contributed to by the pronoun in the complement clause should in fact be there, including a PRED feature, which is barred from functional identification. Thus, in its current form, an analysis like Figure 4 results in violation of Semantic Consistency.

As an alternative, one may adopt the analysis provided by Asudeh (2002). In that analysis, the main clause subject is not directly associated with the embedded copy, but the prepositional connective elements *like/as* function as predicates that can take a thematic subject. This thematic subject is then anaphorically identified with the copy pronoun. Thus, the f-structure would be the following (Figure 5).

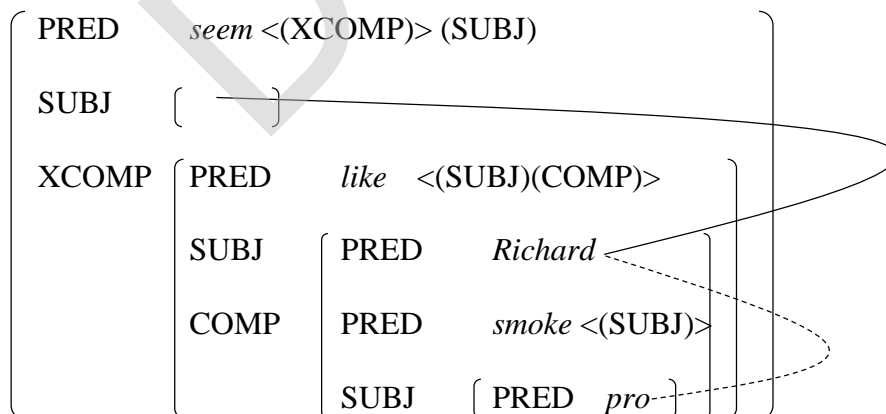


Figure 5.

Asudeh’s (2002) f-structure for (93)

However, the analysis of Figure 5 does not contain finite raising anymore, since functional identification (indicated by the solid line) only happens between the main clause and the

intermediate predicate (*like*), which itself counts as non-finite. Nevertheless, the analysis cannot handle sentences like (92), where a nonthematic element (expletive, idiom chunk) is “raised”, since this element would be properly associated with the predicate in the embedded clause.

To handle cases like (92), an alternative lexical entry for *like* is posited by Asudeh (2002), which takes a nonthematic subject and an XCOMP complement. Figure 6 is Asudeh’s f-structural representation of (92b) (*There seemed like there was a problem*).

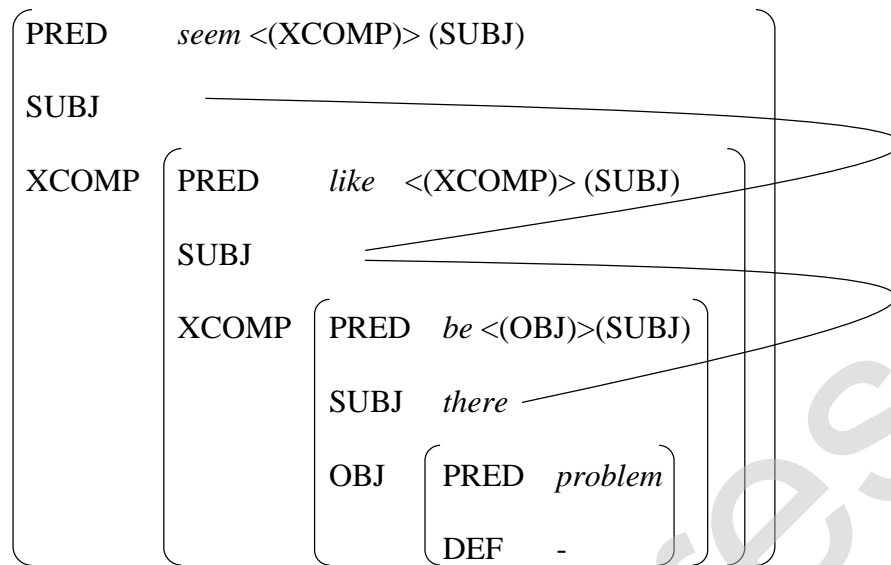


Figure 6.
Asudeh’s (2002) f-structure for (92b).

Independent evidence for the necessity of some proper “raising” mechanism is provided by Fujii (2005), who observes that data from binding argues for reconstruction in CR, see (94). From his theoretical perspective, this argues for A-movement. In LFG, this is a hallmark of functional identification.

- (94) *[Stories about each other]_i seem like they_j have frightened [John and Mary]_i.*

Let us continue this section by showcasing the typological counterpart of copy raising: hyperraising, where some constituent is genuinely “raised” out of a finite complement clause, without intermediaries such as *like* in CR and there is no resumptive element in the embedded clause.

As Carstens & Diercks (2013) note, many Bantu languages exhibit such constructions. The finiteness of these clauses is manifested by the possible agreement and tense features inside them. Zeller (2006) shows one such construction in Nguni, while Carstens & Diercks (2013) demonstrate these with Lubukusu and Lusaamia. Nguni hyperraising is restricted to one predicate, the modal-like *fanele* (‘ought to’), while Lubukusu and Lusaamia involve standard raising predicates like *lolekhana* (‘seem’) and *bonekhana* (‘appear’). Another difference between these languages is that Nguni hyperraising takes place from a subjunctive complement clause, while Lubukusu and Lusaamia raise out of an indicative clause. Traditionally subjunctives are regarded as “less finite” than indicative clauses, since subjunctives often have limited tense-assigning capabilities (they are often restricted to future-orientation or a tense

dependent on the main verb). Although according to Zeller (2006), the tense of the subjunctive clause in the Nguni hyperraising construction is independent, it is safe to say that the structures in Lubukusu and Lusaamia are even “more finite” than the Nguni ones.

Apart from these differences, a similar set of data from these languages supports the hypothesis that these languages do have finite raising constructions. The hyperraising sentences (the “b” examples below in 95-97) have non-raising alternates (the “a” examples in 95-97) with truth-conditional equivalence, so that in the “b” sentences, the main verb subject can be “reconstructed” into the complement clause. (The glosses are somewhat simplified for more focus on substantial issues. E.g. representation of the complex agreement system is omitted. The reader is referred to the cited works for full glosses.)

- (95) a. *Ku-fanele ukuthi uMdu a-khulume isiZulu name.* (Nguni)
 LOC-ought that(C) Mdu SM-speak.SBJV Zulu with.1SG
 ‘Mdu should speak Zulu with me.’
- b. *UMdu u-fanele ukuthi a-khulume isiZulu na-mi.*
 Mdu SM-ought that(that) SM-speak.SBJV Zulu with.1SG
 ‘Mdu should speak Zulu with me.’
- (96) a. *Ka-lolekhana mbo chisaang’i chi-kona.* (Lubukusu)
 SM-seem that animal SM-sleep.PRES
 ‘It seems that the animals are sleeping.’
- b. *Chisaang’i chi-lolekhana mbo chi-kona*
 animal SM-seem that(C) SM-sleep.PRES
 ‘The animals seem to be sleeping.’
- (97) a. *Bi-bonekhana koti eng’ombe chi-ng’were amachi.* (Lusaamia)
 SM-appear that(C) cow SM-drink water
 ‘It appears that the cows drank the water.’
- b. *Eng’ombe chi-bonekhana chi-ng’were amachi.*
 cow SM-appear SM-drink water
 ‘The cows appear to have drunk the water.’

Both Zeller (2006) and Carstens & Diercks (2013) show data from passivization in support of the claim that these constructions are parallel to traditional raising constructions. Zeller (2006) demonstrates the truth-conditional equivalence of sentence pairs where an embedded object is passivized into the main clause through a raising verb as per (98). This is a characteristic of raising constructions, see the truth-conditional equivalence of English raising sentences in (99a-b). The same is not true for equi sentences such as (99a’-b’).

- (98) a. *Udokotela u-fanele ukuthi a-bhek-e isiguli.*
 doctor SM-ought that(C) SM-examine-SBJV patient
 ‘The doctor must examine the patient.’
- b. *Isiguli si-fanele ukuthi si-bhek-w-e ng-udokotela.*
 patient SM-ought that(C) SM-examine-PASS- SBJV by-doctor
 ‘The patient must be examined by the doctor.’

- (99) a. *John seems to visit Mary.* a'. *John tries to visit Mary.*
 b. *Mary seems to be visited by John.* b'. *Mary tries to be visited by John.*

Besides, Carstens & Diercks (2013) show that under the assumption that *suubil* ('believe') in Lubukusu does not assign a thematic role to its object, the thematic role of *omukeni* ('guest') can only originate from the embedded verb, proving the ties to the embedded clause (100).

- (100) *Omukeni ka-asuubil-wa mbo k-ola*
 guest SM-believe.PAST-PASSIVE that(c) SM-PAST.arrive
 'The guest was believed to have arrived.'

Finally, Zeller (2006) shows that idiom chunks may be "raised" in Nguni. Although Carstens & Diercks (2013) do not provide such data, it can safely be assumed that the same is true for Lubukusu and Lusaamia.

- (101) *Izandla zi-fanele ukuthi zi-gez-an-e.*
 hand SM-ought that(C) SM-wash-REC-SBJV
 'It's vital that two people do something for each other.' Lit.: 'It's vital that one hand washes another.'

Earlier in this section, the embedded pronoun providing a PRED feature (which should have been identified with the main clause subject) constituted an analytical difficulty for English copy raising. Neither Zeller (2006) nor Carstens & Diercks (2013) mention the possibility of overt embedded subjects. In fact, according to personal communication with Vicky Carstens, such elements are unlikely to occur. Consequently, the consistency-problem associated with English CR does not arise in Bantu so these structures can be unproblematically represented along the lines of Figure 4 in the previous section.

To conclude this section, we make a few remarks about the lexical entries of finite raising predicates and about some of their typological consequences. As already mentioned, the lexical entries are fundamentally identical to the lexical entry for the standard raising versions of these verbs, with an <(XCOMP)>(SUBJ) grammatical function requirement. However, in these uses, the XCOMP is realized as a CP⁷², while in CR it is realized by prepositional elements (*like, as if*, etc.). This could be a problem on the view that PPs are canonically associated with OBL functions. However, as mentioned in section 1.2, there is a debate in LFG about the status of COMP and XCOMP, some arguing that they are special OBLs. If these proposals are on the right track, it should come as no surprise that these grammatical functions may have PPs as their c-structural realizations.

A complication with CR is that the PRED feature of the embedded copy pronoun clashes with the PRED feature of the matrix subject. In this sense, hyperraising is less problematic for an analysis, since the SUBJ of the complement clause is not and cannot be expressed. The question is what kind of deeper reasons lie beneath these patterns.

My proposal is that the obligatoriness of a c-structurally filled subject position is crucial in this respect. That is, in English we must have a copy pronoun simply because the language is one that requires the subject position of an indicative clause to be filled. This is dubbed as

⁷² I follow Falk (2002: 139-140) in assuming that infinitival *to* is a complementizer.

the famous EPP-requirement in GB/MP. For this requirement, let us use the name “EPP”, without commitment to the GB/MP-theory as a whole. The only way to satisfy the EPP and the lexically encoded functional identification equation in English is to insert a c-structural subject into the complement clause. Although this leads to a violation of Semantic Consistency, this solution is arguably better than the alternative, violating the EPP. In fact, Asudeh (2012) shows that it is possible to remove the extra semantic resource of the pronoun, by using “manager resources” in a “glue semantic” approach in LFG.

In languages where the EPP is not as strong as in English, the consistency-violation is not necessary. Thus, in Bantu (and in other languages with hyperraising), the embedded subject is zero: there is no element in the sentence that would provide a conflicting PRED value. Since there is no independent motivation for the presence of such embedded subjects, the only thing their presence would contribute to the construction is a violation of Semantic Consistency. Accordingly, they are ruled out.

This leads us to a typological prediction: pro-drop languages (languages with no strong EPP requirement) should instantiate hyperraising, while non-pro-drop languages (with strong EPP, like English) should have copy raising. What I am saying is that copy raising and hyperraising are essentially the same structure: finite raising with functional identification. The differences then follow from language-particular properties. This seems to be on the right track: Ademola-Adeoye (2010:vi) makes the same assertion if we look aside from the difference in theoretical persuasion.⁷³

An important empirical generalisation, first noted by Ura (1994), which is empirically supported by the data discussed in this thesis, is that if a language has Hyperraising or Hyper-ECM, it is also a pro-drop language. On the basis of this generalisation, it is argued that Hyperraising and Hyper-ECM constructions involve the use of resumptive *pro* in the embedded subject position, while languages with Copy raising and Copy ECM use overt pronouns. Apart from this difference, it is argued that these A-movement constructions are identical in all crucial respects.

While the cited works and my research radically differ in theoretical background, our conclusion is the same: finite versions of raising are possible, and the overt presence of the controlled subject is dependent on the status of the pro-drop parameter in a given language.

5.4.2 Finite equi

Let us now turn to equi-type control constructions. As a reminder, the controller is a thematic argument of the main verb in these constructions, with the controller entering a dependency with the (usually covert) subject of a complement clause. Just like raising, equi is canonically associated with non-finite complements. In spite of this widespread generalization, as Landau (2013: 88) notes, “the existence of finite control was recognized already in the 1980s as a pervasive feature of the Balkan languages”. As we will see apart from these Balkan examples, a number of other languages also display such configurations, including Hungarian, Greek and Turkish.

⁷³ I found Ademola-Adeoye’s (2010) work after reaching the conclusion myself.

Attention was first called to the possibility of finite control in Serbo-Croatian by Zec (1987)., These Serbo-Croatian constructions involve subjunctive clauses, containing agreeing verb forms. Tense is constrained to present forms.

- (102) a. *Petar_i je pokusao da PRO_{i/*j} dodje na vreme.*
 Peter AUX tried.3SG that(c) comes on time
 ‘Peter tried to come on time.’
- b. *Ana je naterala Mariju_i da PRO_{i/*j} dodje.*
 Anna AUX forced.3SG Maria.ACC that(c) comes
 ‘Anna forced Marija to come.’

The fact that these kind of sentences involve anaphoric rather than functional identification is shown by two pieces of evidence. First, an overt complement subject is possible in the form of an emphatic pronoun as per (103). This is comparable to the case of overt infinitival subjects in Hungarian.

- (103) a. *U takvim situacijama se odman pokusa da se pobegne.*
 in such situations one immediately tries that(c) one escapes
 ‘In such situations one immediately tries to escape.’
- b. *Ana_i je naterala Mariju_j da one_{*i/j} dodje.*
 Anna AUX forced.3SG Maria.ACC that(c) she comes
 ‘Anna forced Marija that she should come.’

Second, split control (where the controllee is controlled by two matrix clause elements) is possible as in (104).

- (104) *Petar je naterao Mariju da zajedno pobjegnu.*
 Peter AUX advised.3SG Maria.ACC that(c) together escape.3PL
 ‘Peter advised Mary to escape together.’

This is not the same as the previously mentioned partial control, since the matrix arguments now jointly specify the controllee exhaustively. Also, the licensing of syntactic plurality is the standard, rather than an exception (Landau 2013: 172-174). As in the case of partial control, this is unexpected under functional identification, which states the full identity of strictly two f-structural elements. From an LFG perspective, this can be seen as another type of the possible manifestations of anaphoric control.

These properties argue for an independent, controlled anaphor in a COMP grammatical function required by these control predicates.

An overview of Greek finite equi is to be found in Spyropoulos (2007), on which this section is based. Greek allows equi with subjunctive complements, which may count as (semi-)finite because of the full range of agreement marking in them. However, their temporal properties are restricted: according to Spyropoulos (2007), they are either fixed in temporal reference (a future-oriented reading) or have “anaphoric tense”, which is determined by the temporal characteristics of the main clause. I adopt the terminology of Spyropoulos (2007) where the

first type bears the name of “dependents subjunctives” (DS) and the second type is labelled “anaphoric subjunctives” (AS).

AS shows what one would expect in the case of obligatory anaphoric control. The anaphoric nature of this configuration is reflected in the possibility of an overt subject in certain contexts as per (105a), while the obligatory nature is manifest in the impossibility of partial control, see (105b).

- (105) a. *o janis_i kseri na xorevi ki aftos_{i/*j} kalo tsamiko*
 the John.NOM know SBJV dance.3SG and he good tsamiko
 ‘John knows how to dance tsamiko well too.’
- b. **i zoi ema_{θe} na kolimbane*
 the Zoe learned.3SG SBJV swim.3PL
 Intended, approx.: ‘Zoe learned to swim with others.’

DS displays some interesting differences. Apart from licensing the kind of overt subjects that can be seen in (105a), it also allows an unusual version of partial control. Normally, in partial control, only semantic, but not syntactic plurality is licensed, see (106a), where the reciprocal cannot appear. Unlike this, in Greek DS, syntactic plurality is also licensed (96b), as can be seen from the plural agreement on *pane* ‘go.3PL’.

- (106) a. *John told Mary that he preferred to meet (*each other) at 6.*
- b. *episa ti maria na pane a psonia tin triti*
 persuaded.1SG the Maria.ACC SBJV go.3PL for shopping the Tuesday
 Intended, approx.: ‘I persuaded Maria that they should go for shopping on Tuesday.’

This argues for a looser version of anaphoric control. I think the reason for this possibility is that anaphoric control can come in various degrees of strictness. The number feature of the complement subject seems to be a point of variation. In light of the fact that control can be suspended by certain DS predicates, it is very plausible that Greek DS is on the lenient side of anaphoric control.⁷⁴

- (107) *o janis prospathise na er_{θun}, ala afti δ en ta kataferan*
 the John tried.3SG SBJV come.3PL but they NEG them manage.3PL
 ‘John tried for them to come, but they didn’t make it’

On the basis of (107), Spyropoulos (2007) notes that it is not without merit to argue that Greek DS are simply a no-control scenario. However, he maintains the control-based analysis for two reasons. One, not every DS predicate allows control suspension, as the possibility of control suspension is dependent on the semantics of the individual predicates. Two, if the feature specification of the matrix controller matches that of the controllee, independent reference is not possible (as in 108a), unlike genuine cases of no-control, exemplified by (108b). I follow his conclusion and regard Greek DS as a “liberal” type of anaphoric control.

⁷⁴ According to Spyropoulos (2007: 166), “there is a variation with respect to the availability of control suspension, which depends on the semantics of the matrix predicate.”

- (108) a. *o janisi prospaθise na PRO_{i/*j} ftasi noris*
 the John tried.3SG SBJV arrives early
 ‘John tried to arrive early.’
- b. *o janisi elpizi na PRO_{i/j} ftasi noris*
 the John hopes SBJV arrives early
 ‘John hopes that s/he will arrive early.’

Based on the flexible nature of Greek subjunctive control, the availability of this special type of partial control and the possibility of overt subjects and control into arguably indicative clauses, Spyropoulos (2007: 174) suggests that control in Greek involves a “pro” instead of a “PRO”. Hence, he proposes that “the hypothesis that the controlled null subject of Greek (...) subjunctives is a controlled “pro” breaks the strict association between control and PRO. It implies that control is not a property of PRO only and that other categories can also be controlled”. As the “PRO” vs. “pro” distinction is nonexistent in LFG to begin with, such a proposal is certainly a promising step towards theoretical convergence.

Turkish presents a particularly interesting case of finite equi. There is a set of verbs which occur with a direct object and a complement clause, the subject of which is controlled by the direct object. This is the description of a standard object-control structure. The only peculiarity of the Turkish construction is that the complement clause has full tense and agreement specifications. In other words, this is a finite object-control structure.

- (109) *Ben Ali-yi yarın balığı yiyecek sanıyordum.*
 I Ali-ACC tomorrow fish eat.FUT.3SG assumed.1SG
 ‘I thought that Ali would eat the fish tomorrow.’

Sentences like (109) had previously been analyzed either as Exceptional Case Marking-constructions, with *Ali-yi* ‘Ali-ACC’ being part of the complement clause (e.g. Aygen 2002), or as *Ali-yi* having moved from the embedded clause into the matrix clause (e.g. Özsoy 2001). However, Ince (2006) shows that neither analysis is correct: in Minimalist framework, the proposed analysis is that the accusative element is base-generated as the main clause object. I just cite one piece of evidence only for each claim, the reader is referred to Ince’s (2006) work for a full discussion.

As (110) shows “Turkish does not permit scrambling of any element out of embedded finite clauses to the right periphery of matrix clauses” (Ince 2006: 210), but the accusative element can be scrambled to the right periphery, so it must be part of the main clause (110b).

- (110) a. **Dinleyiciler biz içtik sanıyorlar viski-yi.*
 auditors we drank.1PL assuming.3PL whisky-ACC
 Intended: ‘The auditors believe we drank the whisky.’
- b. *Hasan Ayşeden nefret ediyoruz sanıyor biz-i.*
 Hasan Aysen.ABL hatred do.3PL assuming.1PL we-ACC
 ‘Hasan thinks that we hate Ayşe.’

It can be shown that the accusative element is base-generated in the matrix clause by using the familiar test of idiom chunks, as they are anomalous in this construction. Thus it can be assumed that the object is a thematic argument of the main verb.

- (111) a. *Alinin anası bellendi.*
 Ali.GEN his.mother was.screwed.3SG
 Idiomatic: ‘They really messed up Ali.’ Literal: ‘They raped Ali’s mother.’
- b. *Alinin anası-nı bellendi sanıyordum.*
 Ali.GEN his.mother-ACC was.screwed.3SG assumed.1SG
 Only literal: ‘I thought that Ali’s mother was raped.’

Ince (2006) does consider a proleptic analysis for this construction, but rejects it on the grounds that unlike prolepsis, it is restricted to embedded subjects, so it must involve a “PRO” rather than a “*pro*”. From my perspective, it is true that prolepsis is much more flexible than standard object control in that it allows for nonsubjects to be controlled and for a greater distance between the controller and the controllee. However, the fundamental analytical idea is the same, i.e. that a thematic matrix object is identified with an embedded argument. Since Ince (2006) mentions no possibility of overt embedded subjects or partial control, I assume this to be a case of functional identification.⁷⁵ That is, the Turkish construction represents a typological variation of equi that we have not encountered yet: it is finite, but instantiates functional, rather than anaphoric control (just as Hungarian prolepsis, Serbo-Croatian and Greek do). In other words, it can be considered a thematic (equi) counterpart of copy raising/ hyperraising structures.

Finally, let us survey the Hungarian landscape, apart from prolepsis. Here, there are two phenomena that are relevant for the topic of finite equi (apart from prolepsis). The first one is that Hungarian also possesses inflected infinitives; and the second one being the possibility of overt subjects in certain infinitival clauses.

Inflected infinitives only occur if the controller bears dative case and has an experiencer thematic relation to the predicate. According to Rákosi (2006b), inflection on the infinitive is possible if the infinitival clause hosts its own subject, which can be translated as anaphoric identification with a controlled *pro* subject. This is in line with Falk (2001: 138) who states that since “anaphoric control involves an ordinary COMP function, an overt SUBJ should be an option.”

However, there is another logical possibility: the case of no-control. That is, it could happen that the dative dependent fully belongs to the complement clause, serving as its subject and therefore no control relation is present in the sentence. To illuminate the issue let us take the predicate *fontos* ‘important’. The two conceptually possible analyses for a sentence with this predicate and an inflected infinitive are shown in (112). (The “PROs” in the string are only for expository purposes.)

- (112) a. *Fontos Jánosnak_i [PRO_i hazamen-ni-e].*
 important John.DAT home.go-INF-3SG

⁷⁵ This in turn may give some explanation for its restrictedness, given Falk’s (2006) approach to functional identification, which was mentioned earlier.

- b. *Fontos [Jánosnak hazamennie].*
 ‘It is important for John to go home.’

Rákosi (2006b), agreeing with É. Kiss (2002), shows that (82b) is the proper analysis. This is supported by the observation that (112) does not necessarily mean that it is *János* ‘John’ and not somebody else that judges going home to be important. This stems from the fact that the dative is not an obligatory dependent of *fontos* ‘important’. *Fontos* (and many other predicates with experiencer semantics) are essentially monadic in nature. This is reflected in the bracketing of (112b). However, they may license a dative dependent as a thematic adjunct. The thematic adjunct may then control the embedded subject, but in this case the infinitival agreement is not possible.

So if one wants a construal where it is necessarily *János* ‘John’ who judges his going home as important, one can only get this interpretation with a non-inflecting infinitive (this crucial interpretational difference is mirrored by the translation). Note that since *János* is not an argument, the morphology of such a dependent is not fixed, a PP-alternative with *számára* ‘for’ is possible (Rákosi 2006b).

- (113) *Fontos Jánosnak_i/ János_i számára [PRO_i hazamen-ni-(*e).*
 important John.DAT John for.3SG home.go-INF-3SG
 ‘John judges it to be important for him to go home.’

Because of this, this class of predicates is not really suitable for investigating finite control constructions in Hungarian. However, there exist a number of predicates which do take a dative constituent as an argument which establishes a control relationship with the infinitival clause. Such predicates are *számít* ‘matter’, *derogál* ‘feels derogatory to do’ or *nem akarózik* ‘does not feel like’ and *sikerül* ‘succeed’. According to Rákosi (2006b), the datives associated with these predicates are arguments. This for example can be seen from the fact that such datives are not replaceable with *számára* (‘for’)-PPs (*számít/derogál/nem akarózik Jánosnak/ *János számára*). With these, the dative is always a matrix clause element, so the structural representation should always be along the lines of (112a). This in turn predicts that no inflected infinitive is expected for these predicates under Rákosi’s (2006b) assumptions. On the other hand, since these are subject clauses, anaphoric control is expected for them, since functional identification as a control relation is uniquely associated with the XCOMP grammatical function. Based on this observation, there might be grounds for expecting inflection. It turns out that infinitival inflection is possible with these predicates. Although Rákosi (2006b:219) marks a sentence like (114a) ungrammatical, I do not share this intuition. Also, the examples in (114b-c) are entirely acceptable.

- (114) a. *Jánosnak derogál velem mutatkoz-ni-(a).*
 János.DAT feels.derogatory with.me show.up-INF-3SG
 ‘It feels derogatory for John to show up with me.’
 b. *Jánosnak nem akarózik hazamen-ni(-e).*
 János.DAT not want.MIDDLE.3SG home.go-INF-3SG
 ‘John does not feel like going home.’

- c. *Jánosnak sikerült hazamen-ni(-e).*
 János.DAT succeeded home.go-INF-3SG
 ‘John succeeded in going home.’ (Lit.: ‘Going home was successful for John.’)

It must be added that even Rákosi (2006b: 218, footnote 13) admits that if the dative argument is implicit, such sentences are quite acceptable. While such an observation may have some merit to it, I doubt that it would boil down to any fundamental configurational difference. In my view, inflected infinitives like (114) are acceptable, independently from the controller being implicit or explicit.

Recall that the anaphoric dependency in proleptic OF enabled an agreement-mismatch between the matrix proleptic argument and the embedded predicate. As it turns out, a similar phenomenon may be observed with the constructions at hand. Note however that in contrast to OF, quantification in this case actually bars agreement-mismatch.

- (115) a. *A lányoknak derogált hazamen-ni-e / hazamenniük.*
 the girls.DAT felt.derogatory home.go-INF-3SG home.go-INF-3PL
 ‘It felt derogatory to the girls to go home.’
- b. *A összes lánynak derogált hazamen-ni-e / *hazamenniük.*
 the every girl.DAT felt.derogatory home.go-INF-3SG home.go-INF-3PL
 ‘It felt derogatory to every girl to go home.’

There is an open question concerning the optionality of infinitival agreement. Cases like (114) suggest that it is entirely optional, without noticeable difference in meaning. On the other hand, in the case of thematic adjunct controllers, there is a noticeable semantic difference between the sentences with and without agreement. Although the two cases are differentiated by the argument or adjunct nature of the controller, the control-mechanism itself is the same, so the cause of the difference remains obscure. I leave the answer to this question to further research.

The other constellation in which finiteness of a control clause is relevant is the possibility of having overt subjects in infinitival clauses. It was Szabolcsi (2009) who observed that if the infinitival subject is under the scope of some operator, it is possible to pronounce the subject. Note that in the absence of the focus operator, this is ungrammatical. (116a) shows this with a raising structure (Szabolcsi convincingly shows that there is a raising construal for *(el)kezd* ‘begin’), while (116b) is an example with equi.

- (116) a. *János elkezdett [*(csak) ő kapni kitüntetések].*
 John began.3SG only he get.INF awards.ACC
 ‘It began to be the case that only John got awards.’
- b. *János szeretne [*(csak) ő énekelni].*
 John would.like.3SG only he sing.INF
 ‘John would like it to be the case that only he sings.’

Interesting data may be gained if one puts together the two Hungarian constructions under discussion. One observation is that these overt subjects cannot license infinitival agreement.

- (117) a. **János elkezdett csak ő kap-ni-a kitüntetések.*
 John began.3SG only he get-INF-3SG awards.ACC
- b. **János szeretne csak ő énekel-ni-e.*
 John would.like.3SG only he sing-INF-3SG

If Rákosi (2006b) is right in that a proper subject in the complement clause enables the agreement to appear, this might suggest that these subjects are somehow deficient. Note also that in the case of *sikerül* ‘succeed’ or *fontos* ‘important’, where the controller is in the dative, the overt infinitival subject also bears this case. There, the possibility to license infinitival agreement is not impaired by the overt subject.

- (118) *Jánosnak sikerült/ fontos volt csak neki hazamenni(e).*
 János.DAT succeeded.3SG important was only him.DAT home.go.INF(.3SG)
 ‘John succeeded in that only he went home.’/ ‘It was important for John that only he goes home.’

Szabolcsi (2009) notes that it is quite cumbersome to construe sentences with the meanings of (116) and (118) without the overt nominative in the infinitival clause, so it might be the case that these overt subjects are last-resort elements. This could somehow lead to their less than full f-structural presence. There are technical ways in LFG to remove unwanted f-structural elements, such as the use of the restriction operator (Kaplan & Wedekind 1993) or the use of “manager resources” (Asudeh 2012), but working out an analysis is left to future research.

I conclude this section with a few remarks about the interaction of overt infinitival subjects with predicates that take only an optional dative thematic adjunct. This produces some interesting data since the change in word order triggered by the focus-operator exposes some configurational issues. Consider the data in (119).

- (119) a. *Fontos [csak Jánosnak hazamen-ni-*(e).]*
 important only John.DAT home.go-INF-3SG
 ‘It is important that only John goes home.’
- b. *Csak Jánosnak fontos hazamen-ni-(e).*
 ‘It is important that only John goes home.’/ It is important only for John that he goes home.

Given the compulsory pre-predicate position of focussed elements in Hungarian, *csak Jánosnak* ‘only John.DAT’ can only belong to the complement clause in (119a), which in turns triggers the compulsory agreement. In (119b), we can have two parses. One is similar to the one already mentioned with regards (112) above, the difference being that the dative is scrambled to the beginning of the sentence. Also, *csak Jánosnak* ‘only John.DAT’ could also be a thematic adjunct associated with *fontos* ‘important’ and it controls into the complement clause. In this case, no agreement is necessary.

Also, if the dative is a thematic adjunct belonging to the main clause, an additional dative can appear in the infinitival clause as its overt subject. In (120) the agreement would indicate the presence of *János* in the complement clause, leaving no position for the other dative subject *csak neki* ‘only for him’.

- (120) *Fontos Jánosnak csak neki hazamen-ni-(*e).*
important John.DAT only for.him home.go-INF-3SG
'It is important for John that only he should go home.'

5.4.3 Summary of finite control

Now that we have surveyed a number of languages with finite control constructions it is time to put the pieces together and in doing so, assemble the bigger picture of control-typology from an LFG-perspective.

The starting point has been that the LFG architecture provides a solid framework and points of references for building a typology like this. From the earliest days onwards, two fundamental axes of variation have been recognized: the thematic relationship of the controller to the main verb, and the nature of the identification process. If the controller element is a thematic argument, we have “equi” structures, if it is athematic, we have “raising”. The nature of the identification may be either functional, which is the strict identity of two f-structural elements or anaphoric, which is a referential dependency. The former is instantiated into the open complement function XCOMP, while the latter involves the closed complement function COMP. Furthermore, while functional identity always requires full syntactic and semantic matching, the anaphoric link may be of various strength, ranging from strict obligatory control through “quasi-obligatory” control (in English, Haug 2013) to weak/no-control (as in Greek DS). What I have done is supplement this taxonomy with another fundamental concept of grammar: finiteness.

Now we are able to present a populated version of Table 3 from section 5.3.1. The starting point was that although prolepsis differs from standard cases of equi in a number of important properties (distance, grammatical functions), they share the same fundamental analytical idea: a thematic grammatical function in the main clause is anaphorically identified with an embedded one. Although the strictness of the anaphoric link varies in the constructions in the slot, it transpires in the variations of number features of the controllee (singular/plural variation in certain cases of proleptic OF, partial control in Greek, split control in Serbo-Croatian). Also, the complement clause is a full, independent clause (COMP), hosting its own subject. I have also identified a finite version of functionally identified equi in Turkish. These are straightforward extensions of the already established cases of canonical equi into the finite domain.

Raising has also been shown to cut across the distinction of finiteness. Finite raising structures exist: not just in lesser-studied Bantu languages, but arguably in English too, in the form of copy raising, which is actually hyperraising for languages with compulsory overt subjects. The LFG architecture readily accommodates this typology. Moreover, the slot that is unexpected given the theoretical assumptions of LFG (raising with anaphoric identification) is still empty, which is a welcome result for this research paradigm.

CONTROL-TYPE			Example
Thematicity of controller	Nature of identification	Finiteness	
equi	anaphoric identification	finite complement	argumental prolepsis, inflected infinitives in Hungarian, Greek and Serbo-Croatian control
		non-finite complement	“agree-type” canonical control
	functional identification	finite complement	Turkish object control
		non-finite complement	“try-type” canonical control
raising	anaphoric identification	finite complement	not expected
		non-finite complement	not expected
	functional identification	finite complement	copy raising in English, Bantu hyperraising
		non-finite complement	canonical raising

Table 4.
An LFG-taxonomy of control (final version)

5.5 Conclusion to chapter 5

This chapter has provided some theoretical and cross-linguistic perspective of the left-peripheral discourse-related constructions that have been the topic of the monograph. Three main configurations were distinguished: fronting proper, left dislocations and prolepsis.

In fronting constructions, a left-peripheral element is “extracted” from a sentence-internal position, but it maintains strong syntactic ties to its original position via functional identification. The configurations in this category are English Topicalization and case-retaining Hungarian Operator fronting (LDD-OF). Additionally, certain Finnish and German fronting constructions were also argued to be manifestations of fronting. There is a tendency for such structures to be interpreted contrastively, but this is not strictly necessary.

In left dislocation constructions, the peripheral discourse-prominent element may be related to a pronominal associate. There are two subtypes of such constructions. One type is syntactically integrated left dislocations (iLD), where the link between the fronted element and the pronoun is grammatically encoded and enforced: these are topic left dislocation in

Hungarian, and certain types of Icelandic, Dutch and German left dislocations may also be located here. The other left dislocation type is comprised of “radically” left peripheral elements, that are generated as a “syntactic orphans” and are only related to the host sentence via pragmatic inference. English (hanging topic) left dislocation and Hungarian free left dislocation are the examples for this, but other Germanic languages (Dutch, Icelandic) also possess such nLD-structures.

In López (2016), fronting and the integrated type of LD are seen as a unitary category, as “d-type dislocations” are contrasted with the non-integrated LD, which is called “h-type dislocation”. From this perspective, English not having the appropriate type of pronouns (d-type pronouns) is the reason for this language not having a pronoun-assisted integrated left dislocation. While this might be true, it should not be seen as a cross-linguistic generalization as a language may have integrated (“d-type”) left dislocation and fronting at the same time, as in the case in German, Dutch, Icelandic and Hungarian. In other words, (integrated) LD and fronting are not in contemporary distribution.

Finally, in the case of anaphoric equi, an element is inserted as a matrix-clause argument and is related to an embedded grammatical function via obligatory anaphoric/functional identification. Various nonfinite and finite equi constructions belong here, prolepsis being in the latter category, with an anaphoric link.

As was seen, both anaphoric and functional links may be established either by the specification of a relevant lexical entry (in the case of control-constuctions) or by general phrase-structure rules (as in fronting and left dislocation).

In sum, it seems that a left-peripheral discourse-related constituent may be related to the clause in the following ways:

- There is a syntactic dependency between the fronted element and the rest of the sentence
 - functional identification (structurally or lexically encoded)
 - anaphoric identification (structurally or lexically encoded)
- There is a loose pragmatic dependency between the fronted element and the rest of the sentence (syntactic orphans)

Table 5 summarizes the picture that emerges if we consider the structures that we have discussed throughout the book.

CONSTRUCTION	INTEGRATION
TOP-Eng LDD-OF CIADJ (English) – event internal Finnish contrastive fronting German fronting (formal/ A-bar)	functional identification – structurally encoded
Functionally identified equi structures (English, Turkish, Greek)	functional identification – lexically encoded
Anaphorically identified equi structures (Greek, Serbo-Croatian) Prolepsis	anaphoric – lexically encoded
iLD structures: TLD, iLD-German/Icelandic/Dutch	anaphoric – structurally encoded
LD-Eng FLD nLD in German/Icelandic/Dutch	pragmatic
CIADJ (English) – event external	base-generation, no identification

Table 5.
Clause-initial discourse-related constructions

The emerging picture is not unlike what den Dikken (2010) presents as the “typology of A’-dependencies”. Disregarding left dislocations (as they lie outside of his concern), he claims that A’-dependencies may be established in three ways: successive-cyclic movement via ν P-edges, resumptive prolepsis, and scope marking. While in this monograph there is no correlate for scope-marking constructions, the first two may easily be related to the LFG-view presented here. Though the technical details are obviously different, functional identification is conceptually the LFG-correlate of successive-cyclic movement. Den Dikken (2010) is not particularly explicit about the details of prolepsis, he largely follows Salzmann (2006), but adds that the proleptic argument may be generated either as an adjunct or as an argument (“the proleptic XP originates in the matrix clause, either inside or outside the VP”). For prolepsis, Salzmann’s (2006, 2017) Minimalist and my LFG-based implementation differ substantially (predication vs. anaphoric control), though convergence of the approaches in the future is possible, given Landau’s (2015) predicative theory of control.

What is not discussed in den Dikken (2010) are the extrasentential syntactic orphans at the left-periphery, which are only related to the host sentence via pragmatic inference.

CHAPTER 6

CONCLUSION

In this monograph I have investigated several left peripheral, clause-initial (“fronting”) discourse-related sentence types in English (topicalization, clause-initial adjuncts, left dislocation) and Hungarian (left dislocation, operator fronting) within the framework of Lexical-Functional Grammar.

After a general outline of the aims and scope of the monograph in **chapter 1**, I presented the theoretical framework, i.e. Lexical-Functional Grammar. The most attention was given to information structure, for which a new feature-based taxonomy was offered. The aim of this new system is a fruitful amalgamation of earlier theoretical taxonomical work. It is based on the notions of relational newness, discourse-structuring capacity and contrastiveness (as a subcase of the discourse-structuring feature). These features classify the following i-structural categories: topic, contrastive topic, information focus, contrastive focus, completive information and background information.

In **chapter 2**, I outlined my view of the relevant Hungarian grammar. The basis for it was the recent work of Laczkó (2014a, 2014b, 2015), which in turn is an adaptation of É. Kiss (1992). Hungarian sentence-structure is seen as being comprised of a hierarchical preverbal field and a flat postverbal area. The preverbal field is the discourse-structural prominent part of the sentence. Under an exocentric S-node we find an iterative “topic-field”, quantifiers, and a Spec/VP “focus-position”. I supplemented Laczkó’s framework with my proposals about information structure.

As for subordinate clauses, I argued contra the standard view (put forward in Kenesei (1992/1994) by making the claim that the optional associate pronouns of subordinate clauses are not expletives, but contentful demonstratives, as Tóth (2000) first argued. I supported my view with theoretical, empirical and typological arguments.

In **chapter 3**, I investigated the left-peripheral discourse-related constructions in English: topicalization, clause-initial adjuncts and left dislocation. The first was argued to be a genuine fronting construction, involving functional identification and the connectivity-effects that follow from this. Information structurally, it marks +CONTRASTIVE categories. The analysis of Clause-initial adjuncts was split between event-internal adjuncts, which behave like topicalized elements and frame-setting ones, which are “base-generated” on the left periphery and do not engage in clause-internal dependencies. Unlike TOP, CIADJs are not necessarily contrastive. The third structure, English left dislocation was argued to be an extrasentential syntactic “orphan”, functionally unintegrated into the host sentence.

Chapter 4 was about the analysis of the Hungarian constructions: various Left dislocations and Operator fronting. Both turned out to involve various subtypes. LD-Hun was split between the syntactically integrated topic-related structure and the focus related version, which should receive an extra-clausal analysis similarly to LD-Eng.

Operator fronting constructions involve a dichotomy: it has an LDD-version, which is a proper fronting construction, involving functional identification. It is thus a syntactic correlate of TOP. Also, there is proleptic OF, in which a thematic matrix argument is anaphorically identified with some embedded function. Information structurally, OF is more uniform: it is involved in the marking of +D-STRUCTURING categories, except for neutral topics. A detailed

analysis, including various levels of linguistic representation was provided for each of the constructions discussed.

Finally, **chapter 5** explored the cross-linguistic and theoretical perspectives of the constructions discussed throughout the monograph. First, I examined some fronting constructions and showed that it is not unusual for languages to dislocate constituents to the left periphery of the clause, in order to give them discourse prominence. Such processes are often accompanied by contrastive interpretations (as in Finnish and certain German configurations) but this is not necessary (as in Hungarian OF or other structures in German). German was also invoked as a basis for comparison of constructions with a pronominal associate (“left dislocation”), along with Dutch and Icelandic. It was shown that a binary split of left dislocation is common: the left-peripheral element may either be a “hanging topic” or it may be syntactically plugged into the host sentence.

Proleptic structures were also surveyed in a number of languages followed by an investigation of how prolepsis fits into the space of theoretical syntax. I concluded that from an LFG-perspective, prolepsis may be looked at as a specific subtype of “control”: finite, long-distance obligatory anaphoric equi. From this stance, I explored the landscape of control constructions in general, putting the focus on less-studied, finite control constructions (including both raising- and equi-type structures).

A number of issues remain undiscovered and thus potential targets for future research. Apart from the specific issues that were mentioned throughout the monograph, a number of additional points are to be added here.

For instance, it has yet to be seen how the different taxonomies that I have presented (information structure, control) can withstand the test of time. As I have identified Topicalization as a construction that makes explicit use of the feature +CONTRASTIVE and the Hungarian preverbal position (the “focus-position”) as a locus for [+NEW V +D-STRUCTURING], hopefully other constructions may be identified which refer to other possibilities e.g. +D-STRUCTURING or –NEW.

As for control, it is predicted that certain slots of the taxonomy (anaphoric raising) should remain empty. Time will tell whether this is the case. Also, more research on the relationship of control and prolepsis is needed (especially in an LFG framework) to consolidate my proposal of a unified approach for them. Furthermore, anaphoric identification has been argued to have different subtypes, based on the strictness of the connection between the controller and controllee. The precise characterization of these subtypes, their relatedness and the overall ramifications of this for the LFG-theoretic anaphoric control are yet to be investigated.

I have argued extensively for the plausibility for a proleptic account of the kind presented here for certain cases of operator fronting, and I gave an outline of similar constructions in a number of languages. However, an essential part of the analysis is clouded at this point. It was claimed that an argument-structural process can turn certain thematic adjuncts into derived arguments (as a relative dependent turns into a proleptic object, or perceptual source adjuncts turn into real subject arguments in copy raising), but exactly why, when and with what constraints such argument structural processes may happen is opaque.

It is also not clear at this point what the precise analysis of non-integrated syntactic orphans should be. LFG’s multi-level correspondence architecture offers a conceptually appealing framework for such an analysis. To do so, detailed work on the formal representation of discourse structure should be carried out.

Finally, the lineup of clause-initial discourse-related constructions in the target-languages is not complete, especially regarding English. I have not discussed Inversion (*rarely do I see John*), *it*-clefts (*it was John that I invited*) or *wh*-clefts (*what I want is a nice cup of tea*). All of these have interesting syntactic and information structural properties which have yet to be explored.

DUPress

REFERENCES

- Ademola-Adeoye, Feyisayo Fehintola. 2010. *A cross-linguistic analysis of finite raising constructions*. University of KwaZulu-Natal dissertation.
- Alba-Salas, Josep. 2004. Lexically selected expletives: evidence from Basque and Romance. *SKY Journal of Linguistics* 17. 35-100.
- Alexiadou, Artemis. Left dislocation (including CLLD). 2006. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax*, vol. 2, 668-699. Blackwell Publishing.
- Alsina, Alex. 2008. A theory of structure-sharing: focusing on long-distance dependencies and parasitic gaps. In Miriam Butt and Tracy H. King (eds.), *Proceedings of the LFG08 Conference*, 5-25. Stanford, CA: CSLI Publications.
- Alsina, Alex, KP Mohanan & Tara Mohanan. 2005. How to get rid of the COMP. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG05 Conference*, 21-41. Stanford, CA: CSLI Publications.
- Alrenga, Peter. 2005. A sentential subject asymmetry in English and its implications for complement selection. *Syntax* 8(3). 175-207.
- Anagnostopoulou, Elena, Henk van Riemsdijk & Frans Zwarts (eds.). 1997. *Materials on left dislocation*. Amsterdam: John Benjamins.
- Ariel, Mira. 1988. Referring and accessibility. *Journal of Linguistics* 24. 67-87.
- Arregi, Karlos. 2003. Clitic left dislocation is contrastive topicalization. In *Proceedings of the 26th Annual Penn Linguistics Colloquium* (Penn Working Papers in Linguistics 9(1). 31-44.) Philadelphia: Penn Linguistics Club.
- Asher, Nicholas & Alex Lascarides. 2003. *Logics of conversation*. Cambridge: Cambridge University Press.
- Asudeh, Ash. 2002. Richard III. In Andronis, Mary, Erin Debenport, Anne Pycha & Keiko Yoshimura, eds. *CLS 38: the main session*, vol. 1. Chicago, IL: Chicago Linguistic Society.
- Asudeh, Ash. 2012. *The logic of pronominal resumption*. Oxford: Oxford University Press.
- Asudeh, Ash & Ida Toivonen. 2012. Copy raising and perception. *Natural Language and Linguistic Theory* 30(2). 321-380.
- Aygen, Gülşat. 2003. Extractibility and the nominative case feature on tense. In A. Sumru Özsoy, Didar Akar, Mine Nakipoglu-Demiralp, Eser E. Taylan & Aksu Koc (eds.), *Studies in Turkish linguistics*, 81-94. Istanbul: Bogazici University Press.
- Baker, Mark. 1996. *The polysynthesis parameter*. Oxford: Oxford University Press.
- Baloghné Nagy Gizella. 2013. *Left dislocation in Optimality Theory*. Eötvös Loránd University dissertation.
- Bartos Huba. 2002. Root infinitives. In Kenesei István & Siptár Péter (eds.), *Approaches to Hungarian 8: papers from the Budapest conference*, 13-38. Budapest: Akadémiai Kiadó.
- Bennis, Hans. 1986. *Gaps and dummies*. Dordrecht: Foris.
- Berman, Judith. 2001. On the cooccurrence of *es* with finite clause in German: an LFG analysis. In Walt D. Meurers & Kiss Tibor (eds.), *Constraint-based approaches to Germanic syntax*, 7-30. Stanford, CA: CSLI Publications.

- Biberauer, Theresa. 2008. Semi null-subject languages, expletives and expletive *pro* reconsidered. In Elliott Lash, Yi An Li & Thomas Rainsford (eds.), *Cambridge occasional papers in linguistics 4*, 24-43. Cambridge: University of Cambridge.
- Boeckx, Cedric, Norbert Hornstein & Jairo Nunes. 2010. *Control as movement*. Cambridge: Cambridge University Press.
- Börjars, Kersti, Rachel Norldinger & Louisa Sadler. 2019. *Lexical-Functional Grammar: an introduction*. Cambridge: Cambridge University Press.
- Bresnan, Joan. 1978. A realistic transformational grammar. In Morris Halle, Joan Bresnan and George A. Miller (eds.), *Linguistic theory and psychological reality*. Cambridge, MA: MIT Press.
- Bresnan, Joan (ed.). 1982. *The mental representation of grammatical relations*. Cambridge, MA: MIT Press.
- Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Wechsler. 2016. *Lexical-functional syntax*, 2nd edition. Wiley Blackwell.
- Bresnan, Joan & Jonni M. Kanerva. 1989. Locative inversion in Chicheŵa: a case study of factorization in grammar. *Linguistic Inquiry* 20(1). 1-50.
- Breul, Carsten. 2004. *Focus structure in generative grammar*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Brody, Michael. 1990. Remarks on the order of elements in the Hungarian focus field. In Kenesei István (ed.), *Approaches to Hungarian*, vol. 3, 95-122. Szeged: JATE.
- Browning, Maggie A. 1996. CP recursion and that-t effects. *Linguistic Inquiry* 27(2). 237-255.
- Butt, Miriam & Tracy H. King. 1996. Structural topic and focus without movement. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG96 Conference*. Stanford, CA: CSLI Publications.
- Büring, Daniel. 1997. *The meaning of topic and focus: the 59th street bridge accent*. London: Routledge.
- Büring, Daniel. 2003. On d-trees, beans, and b-accent. *Linguistics & Philosophy* 26(5). 511–545.
- Carstens, Vicki & Michael Diercks. 2013. Parameterizing case and activity: hyper-raising in Bantu. In Seda Kan, Claire Moore-Cantwell, Robert Staubs (eds.), *Proceedings of the 40th annual meeting of the North East Linguistic Society*, 99-118. Amherst: University of Massachusetts Graduate Linguistic Student Association.
- Chafe, Wallace R. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Charles Li (ed.), *Subject and topic*, 27-55. New York: Academic Press.
- Choi, Hye-Won. 1996. *Optimizing structure in context*. Stanford University dissertation.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, Noam. 1986. *Knowledge of language, its nature, origin, and use*. New York: Praeger.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Massachusetts: The MIT Press.
- Cinque, Guglielmo. 1977. The movement nature of left dislocation. *Linguistic Inquiry* 8(2), 397-412.
- Cook, Philippa & John Payne. 2006. Information structure and scope in German. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG06 Conference*, 124-144. Stanford, CA: CSLI Publications.

- Coppock, Elizabeth. 2003). Sometimes it's hard to be coherent. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG03 Conference*, 126-143. Stanford, CA: CSLI Publications.
- Cowper, Elizabeth. 2016. Finiteness and pseudofiniteness. In Kristin M. Eide (ed.), *Finiteness matters: On finiteness related phenomena in natural languages*, 47-77. Amsterdam: John Benjamins.
- Dalrymple, Mary. 1993. *The syntax of anaphoric binding*. Stanford, CA: CSLI Publications.
- Dalrymple, Mary. 2001. *Lexical Functional Grammar*. San Diego: Academic Press.
- Dalrymple, Mary & Irina Nikolaeva. 2011. *Objects and information structure*. Cambridge: Cambridge University Press.
- Dalrymple, Mary & Helge Lødrup. 2000. The grammatical functions of complement clauses. In Butt, Miriam & Tracy H. King (eds.), *Proceedings of the LFG00 Conference*, 104-121. Stanford, CA: CSLI Publications.
- Dalrymple, Mary, John J. Lowe & Louise Mycock. 2019. *The Oxford reference guide to Lexical Functional Grammar*. Oxford: Oxford University Press.
- Davies, William D. 2001. Against raising in Madurese (and other Javanic languages). *Proceedings of the Chicago Linguistics Society* 36. 57-69.
- Davies, William D. 2005. Madurese prolepsis and its implications for a typology of raising. *Language* 81. 645-665.
- Davies, William D. & Stanley Dubinsky. 2004. *The grammar of raising and control – A course in syntactic argumentation*. Blackwell Publishing.
- Davies, William D. & Stanley Dubinsky. 2009. On the existence (and distribution) of sentential subjects. In Donna B. Gerds, John C. Moore & Maria Polinsky (eds.), *Hypothesis A/Hypothesis B: linguistic explorations in honor of David M. Perlmutter*, 111-128. Cambridge, MA: MIT Press.
- Declerck, Renaat. 1988. *Studies on copular sentences, clefts and pseudoclefts*. Dordrecht: Foris.
- den Dikken, Marcel. 2010. *On the strategies for forming long A'-dependencies: Evidence from Hungarian*. Unpublished manuscript, Ms. CUNY Graduate Center.
- den Dikken, Marcel & Surányi Balázs. 2017. Contrasting contrastive left dislocation explications. *Linguistic Inquiry* 48(4). 543-584.
- É. Kiss Katalin. 1992. Az egyszerű mondat szerkezete. In Kiefer Ferenc (ed.), *Strukturális magyar nyelvtan 1: Mondattan*, 79-177. Budapest: Akadémiai Kiadó.
- É. Kiss Katalin (ed.). 1995. *Discourse configurational languages*. New York: Oxford University Press.
- É. Kiss Katalin. 1998. Identificational focus versus information focus. *Language* 74. 245-273.
- É. Kiss Katalin. 2002. *The syntax of Hungarian*. Cambridge: Cambridge University Press.
- É. Kiss, Katalin. 2005. Event types and discourse linking in Hungarian. *Linguistics* 43(1). 131-154.
- É. Kiss Katalin & Pintér Lilla. 2014. Identificational focus revisited. Paper presented at CL50 (the 50th anniversary meeting of the Chicago Linguistic Society).
- Enc, Mürvet. 1991. The semantics of specificity. *Linguistic Inquiry* 22(1). 1-25.
- Ernst, Thomas. 2002. *The syntax of adjuncts*. Cambridge: Cambridge University Press.
- Falk, Yehuda N. 2001. *Lexical-Functional Grammar*. Stanford, CA: CSLI Publications.
- Falk, Yehuda N. 2006a. On the representation of case and agreement. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG06 Conference*, 184-201. Stanford, CA: CSLI Publications.

- Falk, Yehuda N. 2006b. *Subjects and Universal Grammar: an explanatory theory*. Cambridge: Cambridge University Press.
- Falk, Yehuda N. 2007. Do we wanna (or hafta) have empty categories? In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG07 Conference*, 184-197. Stanford, CA: CSLI Publications.
- Frascarelli, Mara. 2007. Types of topics in German and Italian. In Susanne Winkler & Kerstin Schwabe (eds.), *On information structure, meaning and form*, 87-116. Amsterdam, Philadelphia: John Benjamins.
- Frey, Werner. 2003. Syntactic conditions on adjunct classes. In Ewald Lang, Claudia Maienborn & Catherine Fabricus-Hansen (eds.), *Modifying adjuncts*, 163-209. Berlin, New York: Mouton de Gruyter.
- Frey, Werner. 2004. Notes on the syntax and pragmatics of German left dislocation. In Horst Lohnstein & Susanne Trissler (eds.), *The syntax and semantics of the left periphery*, 203-233. Berlin: Mouton de Gruyter.
- Frey, Werner. 2005. Pragmatic properties of certain German and English left peripheral constructions. *Linguistics* 43. 89-129.
- Frey, Werner. 2010. \bar{A} -movement and conventional implicatures: About the grammatical encoding of emphasis in German. *Lingua* 120(6). 1416-1435.
- Gazdik Anna. 2011. *Multiple questions in French and Hungarian*. Eötvös Loránd University dissertation.
- Gécseg Zsuzsanna. 2001. A kontrasztív topik szintaxisáról és szemantikájáról. *Magyar Nyelv*, September, December.
- Gécseg Zsuzsanna. 2013. Adalékok az információs szerkezet és mondat szerkezet összefüggéseinek kérdéséhez. *Magyar Nyelv* 109(1). 32-49.
- Gécseg Zsuzsanna & Kiefer Ferenc. 2009. A new look at information structure in Hungarian. *Natural Language and Linguistic Theory* 27. 583-622.
- Gerőcs Mátyás, Babarczy Anna & Surányi Balázs. 2014. Exhaustivity in focus: experimental evidence from Hungarian. In Joseph Emonds & Markéta Janebová (eds.), *Language use and linguistic structure* (Olomouc Modern Language Series 3), 181-195. Palacky University Press.
- Gervain Judit. 2002. *Linguistic methodology and microvariation in language: The case of operator-raising in Hungarian*. MA Thesis, University of Szeged.
- Gervain Judit. 2004. The syntax and semantics of the resumptive dependency in Hungarian focus-raising constructions. In Benjamin Shaer, Werner Frey & Claudia Maienborn (eds.), *Proceedings of the Dislocated Elements Workshop* (ZAS Papers in Linguistics, 35(1)), 111-131. Berlin: ZAS.
- Gervain Judit. 2009. Resumption in focus(-raising). *Lingua* 119. 687-707.
- Gregory, Michelle L. & Laura A. Michaelis. 2001. Topicalization and left-dislocation: a functional opposition revisited. *Journal of Pragmatics* 33. 1665-1706.
- Grohmann, Kleanthes K. 2003. *Prolific domains: on the anti-locality of movement dependencies*. Amsterdam: John Benjamins.
- Gundel, Jeanette K. 1974. *The role of topic and comment in linguistic theory*. Doctoral dissertation, University of Texas.
- Gundel, Jeanette K. 1985. 'Shared knowledge' and topicality. *Journal of Pragmatics* 9(1). 83-107.

- Gundel, Jeanette K. 1999. Topic, focus, and the grammar-pragmatics interface. In Jim Alexander, Na-Rae Han & Michelle Minnick Fox (eds.), *Proceedings of the 23rd Annual Penn Linguistics Colloquium* (University of Pennsylvania Working Papers in Linguistics 6.1), 185-200. University of Pennsylvania.
- Gundel, Jeanette K., Mamadou Bassene, Bryan Gordon, Linda Humnick & Amel Khalfaoui (2010). Testing predictions of the Givenness Hierarchy framework: A crosslinguistic investigation. *Journal of Pragmatics* 42. 1770-1785.
- Gundel, Jeanette K., Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69. 274-307.
- Gutzmann, Daniel, Katharina Hartmann & Lisa Matthewson. 2020. Verum focus is verum, not focus: Cross-linguistic evidence. *Glossa: a journal of general linguistics* 5(1). 1-48.
- Haegeman, Liliane. 1991. Parenthetical adverbials: the radical orphan approach. In Shuji Chiba et al. (eds.), *Aspects of modern English linguistics: papers presented to Masatomo Ukaji on his 60th birthday*, 232-254. Tokyo: Kaitakusha.
- Haegeman, Liliane. 2003. Notes on long adverbial fronting in English and the left periphery. *Linguistic Inquiry* 34. 640-649.
- Haug, Dag. 2013. Partial control and the semantics of anaphoric control in LFG. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG13 Conference*, 274-294. Stanford, CA: CSLI Publications.
- Hedberg, Nancy. 2000. The referential status of clefts. *Language* 76. 891-920.
- Hegedüs Veronika. 2013. *Non-verbal predicates and predicate movement in Hungarian.*, Tilburg University dissertation. Utrecht: LOT Publications.
- Higgins, Roger. 1981. Proleptic objects and verbs of perception in Zacapoaxtla Nahuatl. *Texas Linguistic Forum* 18. 69-88.
- Hoekstra, Teun. 1983. The distribution of sentential complements. In Hans Bennis & Wus van Lessen Kloeke (eds.), *Linguistics in the Netherlands*, 93-103. Dordrecht: Foris.
- Hoji, Hajime. 2005. Major object analysis of the so-called raising-to-object constructions in Japanese. Paper presented at New Horizons in the Grammar of Raising and Control Workshop. LSA Summer Institute, Harvard University.
- Hong, Ki-Sun. 1997. Yenge-wa kwuke-uy insang kwumwnun pikyo pwunsek (Subject-to-object raising constructions in English and Korean). *Language Research* 33. 409-434.
- Hong, Soo-Min. 2005. "Exceptional" case-marking and resultative constructions. University of Maryland dissertation.
- Horváth Júlia. 2010. 'Discourse features', syntactic displacement and the status of contrast. *Lingua* 120. 1346-1369.
- Ince, Atakan. 2006. Direct complement clauses as object control structures in Turkish. In Erin Bainbridge & Brian Agbayani (eds.), *Proceedings of the thirty-fourth Western Conference On Linguistics*, 208-221. Department of Linguistics, California State University, Fresno.
- Ingria, Robert J. P. 1981. *Sentential complementation in Modern Greek*. Massachusetts Institute of Technology dissertation.
- Jánosi Adrienn. 2013. *Long split focus constructions in Hungarian with a view on speaker variation*. KU Leuven dissertation. CRISP.
- Kálmán László. 1985. Word order in neutral sentences. In Kenesei István (ed.), *Approaches to Hungarian* 1, 13-23. Szeged: JATE.
- Kálmán László (ed.). 2001. *Leíró magyar nyelvtan (Mondattan I.)*. Budapest: MTA-ELTE.

- Kaplan, Ronald M. & Jürgen Wedekind. 1993. Restriction and correspondence-based translation. In *Proceedings of the 6th Meeting of the EACL*. European Chapter of the Association of Computational Linguistics. University of Utrecht.
- Kenesei István. 1992. Az alárendelt mondatok szerkezete. In Kiefer Ferenc (ed.), *Strukturális magyar nyelvtan II: Mondattan*, 529-714. Budapest: Akadémiai Kiadó.
- Kenesei István. 1994. Subordinate clauses. In Kiefer Ferenc & É. Kiss Katalin (eds.), *The syntactic structure of Hungarian*, 141-165. San Diego: Academic Press.
- Kenesei István. 2006. Focus as identification. In Molnár Valéria & Susanne Winkler (eds.), *The architecture of focus* (Studies in generative grammar 82), 137-168. Berlin, New York: Mouton de Gruyter.
- Kenesei István, Robert M. Vago & Fenyvesi Anna. 1998. *Hungarian*. Taylor & Francis: Routledge, London.
- Kibort, Anna. 2007. Extending the applicability of Lexical Mapping Theory. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG07 Conference*, 250-270. Stanford, CA: CSLI Publications.
- Kibort, Anna. 2008. On the syntax of ditransitive constructions. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG08 Conference*, 312-332. Stanford, CA: CSLI Publications.
- King, Tracy H. 1997. Focus domains and information structure. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG97 Conference*. Stanford, CA: CSLI Publications.
- Kotzoglou, George & Dimitra Papangeli. 2007. Not really ECM, not exactly control: the “quasi-ECM” construction in Greek. In William D. Davies & Stanley Dubinsky (eds.), *New horizons in the analysis of control and raising*, 111-131. Dordrecht: Springer.
- Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55. 243-276.
- Krifka, Manfred & Renate Musan. 2012. Information structure: overview and linguistic issues. In Manfred Krifka & Renate Musan (eds.), *The expression of information structure*, 1-43. Berlin: De Gruyter Mouton.
- Laczkó Tibor. 1995. *The syntax of Hungarian noun phrases: A Lexical-Functional approach*. (Metalinguistica 2). Frankfurt am Main: Peter Lang.
- Laczkó Tibor. 2012. On the (un)bearable lightness of being an LFG style copula in Hungarian. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG12 Conference*, 341-361. Stanford, CA: CSLI Publications.
- Laczkó Tibor. 2014a. The essentials of an LFG-analysis of Hungarian finite sentences. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG14 Conference*, 325-345. Stanford, CA: CSLI Publications.
- Laczkó Tibor. 2014b. An LFG analysis of preverbs in Hungarian. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG14 Conference*, 346-366. Stanford, CA: CSLI Publications.
- Laczkó Tibor. 2015. On operators in the preverbal domain of Hungarian finite sentences. In Miriam Butt & Tracy H. King (eds.) *Proceedings of the LFG15 Conference*, 187-207. Stanford, CA: CSLI Publications.
- Laczkó, Tibor. 2017. *Aspects of Hungarian syntax from a lexicalist perspective*. Doctor of the Hungarian Academy of Sciences thesis. University of Debrecen.
- Lambrecht, Knud. 1994. *Information structure and sentence form: topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

- Lambrecht, Knud. 2001. Dislocation. In Martin Haspelmath et al. (eds.), *Language typology and language universals*, 1050-1078. Berlin/New York: Walter de Gruyter.
- Landau, Idan. 2011. Predication vs. aboutness in copy raising. *Natural Language and Linguistic Theory* 29. 779-813.
- Landau, Idan. 2013. *Control in generative grammar: a research companion*. Cambridge: Cambridge University Press.
- Landau, Idan. 2015. *A two-tiered theory of control*. Cambridge, MA: MIT Press.
- Landau, Idan. 2016. Agreement at PF: an argument from partial control. *Syntax* 19(1). 79-109.
- Lang, Ewald, Claudia Maienborn & Catherine Fabricus-Hansen (eds.). 2003. *Modifying adjuncts*. Berlin, New York: Mouton de Gruyter.
- Levinson, Stephen C. & Nicholas Evans. 2010. Time for a sea-change in linguistics: Response to comments on 'The Myth of Language Universals'. *Lingua* 120. 2733-2758.
- Lødrup, Helge. 2012. In search of a nominal COMP. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG12 Conference*, 383-404. Stanford, CA: CSLI Publications.
- Lipták Anikó. 1998. A magyar fókuszemelések egy minimalista elemzése. In Büky László and Maleczki Márta (eds.), *A mai magyar nyelv leírásának újabb módszerei* 3, 93-116. Szeged: University of Szeged.
- Lipták Anikó. 2011. The structure of the topic field in Hungarian. In Paola Beninca & Nicola Munaro (eds.), *Mapping the left periphery*, 163-198. Oxford: Oxford University Press.
- Lipták Anikó. 2012. Correlative topicalization. *Acta Linguistica Hungarica* 59(3). 245-302.
- Lipták Anikó & Luis Vicente. 2009. Pronominal doubling under topicalization. *Lingua* 119. 650-686.
- López, Luis 2009. *A derivational syntax for information structure*. Oxford: Oxford University Press.
- López, Luis. Dislocations and information structure. 2016. In Féry, Caroline & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*. Oxford: Oxford University Press.
- Maienborn, Claudia. 2001. On the position and interpretation of locative modifiers. *Natural Language Semantics* 9(2). 191-240.
- Massam, Diane. 1985. *Case theory and the projection principle*. Massachusetts Institute of Technology dissertation.
- Mikkelsen, Line. 2011. Copular clauses. In Claudia Maienborn, Klaus von Stechow & Paul Portner (eds.), *Semantics: an international handbook of natural language meaning*, vol. 2, 1805-1829. Berlin: Mouton de Gruyter.
- Molnár Valéria. 2002. Contrast – from a contrastive perspective. *Language and Computers* 39. 147-162.
- Molnár Valéria. 2006. On different kinds of contrast. In Molnár Valéria and Susanne Winkler (eds.), *The architecture of focus* (Studies in generative grammar 82), 197-233. Berlin, New York: Mouton de Gruyter.
- Molnár Valéria & Susanne Winkler. 2010. Edges and gaps: contrast at the interfaces. *Lingua* 120. 1392-1415.
- Moro, Andrea. 1997. *The raising of predicates*. Cambridge: Cambridge University Press.
- Mycock, Louise. 2006. *The typology of constituent questions: a Lexical-Functional Grammar analysis of 'wh'-questions*. University of Manchester dissertation.

- Mycock, Louise. 2013. Discourse functions of question words. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG13 Conference*, 419–439. Stanford, CA: CSLI Publications.
- Mycock, Louise & John Lowe. 2013. The prosodic marking of discourse functions. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG13 conference*, 440–460. Stanford, CA: CSLI Publications.
- Needham, Stephanie & Ida Toivonen. 2011. Derived arguments. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG11 Conference*, 401–421. Standford: CSLI Publications.
- Neeleman, Ad & Hans van de Koot. 2008. Dutch scrambling and the nature of discourse templates. *Journal of Comparative Germanic Linguistics* 11. 137–189.
- Neeleman, Ad, Elena Titov, Hans van de Koot & Reiko Vermeulen. 2009. A syntactic typology of topic, focus and contrast. In van Jeroen Craenenbroeck (ed.), *Alternatives to cartography*, 15–52. Berlin: Mouton de Gruyter.
- Nikolaeva, Irina (ed.). 2007. *Finiteness*. Oxford: Oxford University Press.
- Onea, Edgar. 2009. Exhaustiveness of Hungarian focus. Evidence from Hungarian and German. In Arndt Riester & Edgar Onea (eds.), *Focus at the syntax-semantics interface* (Working Papers of the SFB 732: Incremental Specification in Context), vol. 3, 53–68. University of Stuttgart.
- Oravecz Csaba, Váradi Tamás & Sass Bálint. 2014. The Hungarian gigaword corpus. In *Proceedings of LREC 2014*.
- Özsoy, A. Sumru. 2001. On small clauses, other bare verbal complements and feature checking in Turkish. In Eser Erguvanlı & Eser E. Taylan (eds.), *The verb in Turkish*, 213–237. Amsterdam: John Benjamins Publishing Company.
- Philippaki-Warbuton, Irene. 1987. The theory of empty categories and the pro-drop parameter in Modern Greek. *Journal of Linguistics* 23(2). 289–318.
- Pesetsky, David 1987. Binding problems with experiencer verbs. *Linguistic Inquiry* 18(1). 126–140.
- Pesetsky, David. 1995. *Zero syntax: experiencers and cascades*. Cambridge, Massachusetts/London, England: The MIT Press.
- Pintér Lilla. 2016. A magyar szerkezeti fókusz kimerítő értelmezésének kísérletes vizsgálata. In Balázs József, Bojtos Anita, Paár Tamás, Tompa Zsófia, Turi Gergő and Vadász Noémi (eds.), *Studia Varia*, 191–212. Budapest: Pázmány Péter Katolikus Egyetem Bölcsészeti és Társadalomtudományi Kar.
- Pitteroff, Marcel, Artemis Alexiadou & Silke Fischer. 2017. On partial control in German. *Journal of Comparative Germanic Linguistics* 20. 139–185.
- Prince, Ellen F. 1981. Towards a taxonomy of given-new information. In Peter Cole (ed.), *Radical pragmatics*. New York: Academic Press.
- Prince, Ellen F. 1998. On the limits of syntax, with reference to left-dislocation and topicalization. In Peter Cullicover & Louise McNally (eds.), *Syntax and semantics 29: the limits of syntax*. New York: Academic Press.
- Prince, Ellen F. 1999. How not to mark topics: ‘Topicalization’ in English and Yiddish. In *Texas Linguistics Forum* (chapter 8). Austin: University of Texas.
- Prince, Alan & Paul Smolensky. 1993. *Optimality theory: constraint interaction in generative grammar (RuCCS Technical Report #2)*. Rutgers University Center for Cognitive Science.

- Postal, Paul M. & Geoffrey K. Pullum 1988. Expletive noun phrases in subcategorized positions. *Linguistic Inquiry* 19. 635-670.
- Rákosi György. 2006a. On the need for a more refined approach to the argument-adjunct distinction: the case of dative experiencers in Hungarian. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG06 Conference*, 416-436. Stanford, CA: CSLI Publications.
- Rákosi György. 2006b. *Dative experiencer predicates in Hungarian*. University of Utrecht dissertation.
- Rákosi György. 2012. Non-core participant PPs are adjuncts. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG12 Conference*, 524-543. Stanford, CA: CSLI Publications.
- Rákosi György & Laczkó Tibor. 2005. Verbal category and nominal function – Evidence from Hungarian Subject Clauses. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG05 Conference*, 353-370. Stanford, CA: CSLI Publications.
- Reinhart, Tanya. 1981. Pragmatics and linguistics: an analysis of sentence topics. *Philosophica* 27. 53-94.
- Reinhart, Tanya. 2002. The theta system – an overview. *Theoretical Linguistics* 28. 229-290.
- Repp, Sophie. 2010. Defining ‘contrast’ as an information-structural notion in grammar. *Lingua* 120. 1333-1345.
- Repp, Sophie. 2016. Contrast: dissecting an elusive information-structural notion and its role in grammar. In Caroline Féry & Shinichiro Ishihara (eds.), *Handbook of information structure*. Oxford University Press.
- Roberts, Craige. 1996. Information structure in discourse: towards an integrated formal theory of pragmatics. In Jae-Hak Yoon, Andreas Kathol and Craige Roberts (eds.), *Papers in semantics* (Working Papers in Linguistics 49). The Ohio State University.
- Rodman, Robert. 1974. On left dislocation. *Papers in Linguistics* 7. 437-466.
- Rooth, Mats E. 1985. Association with focus. University of Massachusetts, Amherst, MA dissertation.
- Ross, John R. 1967. *Constraints on variables in syntax*. Massachusetts Institute of Technology dissertation.
- Rothstein, Susan D. 1995. Pleonastics and the interpretation of pronouns. *Linguistic Inquiry* 26. 499-529.
- Salzmann, Martin. 2006. *Resumptive prolepsis: a study in indirect A'-dependencies*. Utrecht: LOT (LOT Dissertation Series 136).
- Salzmann, Martin. 2017. Prolepsis. In Martin Everaert & Henk van Riemsdijk (eds.), *The Wiley-Blackwell companion to syntax*, 2nd edition, vol. 5, 3203-3245. Wiley-Blackwell.
- Shaer, Benjamin. 2004. On the syntactic status of certain fronted adverbials in English. In Pawel Nowak, Corey Yoquelet & David Mortensen (eds.), *Proceedings of the Berkeley Linguistics Society* 29, 379–390. Berkeley, CA: BLS.
- Shaer, Benjamin. 2009. German and English left-peripheral elements and the orphan analysis of non-integration. In Benjamin Shaer, Philippa Cook, Werner Frey & Claudia Maienborn (eds.), *Dislocated elements in discourse: syntactic, semantic and pragmatic perspectives*, 366-397. New York: Routledge.
- Shaer, Benjamin & Werner Frey. 2004. Integrated and non-integrated left-peripheral elements in German and English. In Benjamin Shaer, Werner Frey & Claudia Maienborn (eds.), *Proceedings of the Workshop on Dislocated Elements* (pp. 465-502). Berlin: Zentrum für Allgemeine Sprachwissenschaft.

- Snider, Neal & Annie Zaenen. 2006. Animacy and syntactic structure: Fronted NPs in English. In Butt, Miriam, Mary Dalrymple & Tracy H. King (eds.), *Intelligent linguistic architectures: Variations on themes by Ronald M. Kaplan*. Stanford, CA: CSLI Publications.
- Spyropoulos, Vassilios. 2007. Finiteness and control in Greek. In William D. Davies & Stanley Dubinsky (eds.), *New horizons in the analysis of control and raising*, 159-183. Dordrecht: Springer.
- Subirats-Rüggeberg, Carlos. 1990. The relation between infinitival and sentential complements in traditional grammar. *Lingvisticae Investigationes* 14. 81-93.
- Stowell, Tim. 1981. *Origins of phrase structure*. Massachusetts Institute of Technology dissertation.
- Stowell, Tim. 1982. The tense of infinitives. *Linguistic Inquiry* 13. 561-570.
- Surányi Balázs. 2006a. Mechanisms of wh-saturation and interpretation in multiple wh-movement. In Lisa Cheng & Norbert Corver (eds.), *Wh-movement: moving on*, 289-318. Cambridge, Massachusetts: MIT Press.
- Surányi Balázs. 2006b. Scrambling in Hungarian. *Acta Linguistica Hungarica* 53(4). 393-432.
- Szabolcsi Anna. 2009. Overt nominative subjects in infinitival complements: data, Diagnostics, and preliminary analyses. In Patricia Irwin & Violeta Vasquez Rojas Maldonado (eds.), *NYU working papers in linguistics* (vol. 2: Papers in Syntax). New York University.
- Szűcs Péter. 2010. *On operator raising in Hungarian*. University of Debrecen MA Thesis.
- Titov, Elena. 2013. Do contrastive topics exist? *Journal of Linguistics* 49(2). 413-454.
- Toivonen, Ida. 2013. English benefactive NPs. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG13 Conference*, 503-523. Stanford, CA: CSLI Publications.
- Tortora, Christina M. 1997. *The syntax and semantics of the weak locative*. University of Delaware dissertation.
- Tóth Ildikó. 2000. *Inflected infinitives in Hungarian*. Tilburg: TILDIL Dissertation Series.
- Ura, Hiroyuki. 1994. Varieties of raising and their implications in the theory of case and agreement. In *MIT occasional papers in linguistics*, vol. 7. Massachusetts: MIT.
- Vat, Jan. 1981. Left dislocation, connectedness, and reconstruction. *Groninger Arbeiten zur Germanistischen Linguistik* 20. 80-103.
- von Stechow, Klaus. 2011. Specificity. In Klaus von Stechow, Claudia Maienborn & Paul Portner (eds.), *Semantics: an international handbook of natural language meaning*, vol. 2, 1025-1058. Berlin/Boston: De Gruyter Mouton.
- Ward, Gregory. 1988. *The semantics and pragmatics of preposing*. New York/London: Garland Publishing.
- Wechsler, Stephen & Larisa Zlatić. 2003. *The many faces of agreement: morphology, syntax, semantics, and discourse factors in Serbo-Croatian agreement*. Stanford, CA: CSLI Publications.
- Wedgwood, Daniel. 2009. Variation in focus. In Arndt Riester & Edgar Onea (eds.), *Focus at the syntax-semantics interface* (Working Papers of the SFB 732: Incremental Specification in Context, vol. 3), 101-120. University of Stuttgart.
- Vilkuna, Maria. 1995. Discourse configurability in Finnish. In É. Kiss Katalin (ed.), *Discourse configurational languages*, 244-268. Oxford, New York: Oxford University Press.

- Yoon, James H-S. 2007. Raising of major arguments in Korean and Japanese. In William D. Davies & Stanley Dubinsky (eds.), *New horizons in the analysis of control and raising*, 71-107. Dordrecht: Springer.
- Zaenen, Annie. 1997. Contrastive dislocation in Dutch and Icelandic. In Elena Anagnostopoulou, Henk van Riemsdijk & Frans Zwarts (eds.), *Materials on left dislocation*, 119-150. Amsterdam: John Benjamins.
- Zec, Draga. 1987. On obligatory control in clausal complements. In Iida Masayo, Steven Wechsler and Draga Zec (eds.), *Working papers in grammatical theory and discourse structure*, 139-168. Stanford, CA: CSLI Publications.
- Zeller, Jochen. 2006. Raising out of finite CP in Nguni: the case of *fanele*. *Southern African Linguistics and Applied Language Studies* 24(3). 255-275.
- Zimmermann, Malte. 2008. Contrastive focus and emphasis. *Acta Linguistica Hungarica* 55. 347-360.

DUPRESS

This monograph investigates discourse-related clause-initial, left peripheral constructions within the framework of Lexical-Functional Grammar. The following structures are to be scrutinized: English topicalization, clause-initial adjuncts and left dislocation; Hungarian left dislocation and operator fronting. The overall goal is to provide an account for these structures, especially with regard to their syntactic and information structural properties and to put them into cross-linguistic and theoretical perspective.

DUPress

ISBN 978-963-318-885-9



9 789633 188859