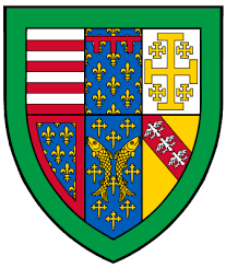
UNIVERSITY OF CAMBRIDGE

WORD ORDER AND INFORMATION STRUCTURE IN ROMEYKA: A SYNTAX AND SEMANTICS INTERFACE ACCOUNT OF ORDER IN A MINIMALIST SYSTEM

Nicolaos Neocleous

Queens' College



This dissertation is submitted for the degree of

Doctor of Philosophy

Cambridge

December 2017

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

It does not exceed the prescribed word limit for the Modern and Medieval Languages Degree Committee, that is it does not exceed 80.000 words, including footnotes, references and appendices, but excluding bibliographies.

December 2017

Nicolaos Neocleous

Abstract

Word order and information structure in Romeyka: A syntax and semantics interface account of order in a minimalist system

Nicolaos Neocleous

In this dissertation, I investigate word order and information structure in the light of recent developments within the minimalist program. I specifically pursue a principled explanation of word order within the biolinguistic perspective. In that sense, I entertain the thesis that all properties of the faculty of language contribute to a computationally efficient satisfaction of interface conditions. The language examined is Romeyka, the only Asia Minor Greek variety still spoken in the area historically known as Asia Minor (present-day Anatolia, Turkey). The objective of this study is therefore twofold: (a) descriptively, to examine word order variation in Romeyka and (b) theoretically, to investigate whether such word order variation could be a language specific property or, rather, could be accommodated in a minimalist system. Descriptively, I aim (a) to determine the pragmatically unmarked and marked word orders in Romeyka, (b) to examine their typological classification and (c) to investigate their evolution. Theoretically, this dissertation is fundamentally about the role that order plays in the efficient computation of interface conditions, mainly in regard to the syntax and semantics interface. Generative Grammar is the study of linguistic capacity as a component of human cognition. As such, Generative Grammar has made significant progress in identifying some of the computational mechanisms that distinguish man from animals; the basic tenet is that only humans appear to possess a mental (universal) grammar that permits the composition of infinitely many meaningful expressions from a finite stock of discrete units. The basic compositional operation of grammar is said to be Merge. Merge can create a set K of two linguistic objects x and y. Set K can be merged with another object z, or with another set of objects L and so on. Merge imposes a hierarchical structure, i.e. x and y are elements of K, but not vice versa. Merge is assumed not to impose order, i.e. {x, y} = {y, x}. As such, order is structure-dependent, i.e. no syntactic operation can make reference to it. It has also been claimed that hypothetical languages, in which syntactic operations are defined in linear terms, such that Merge creates an ordered pair <x, y>, are outside of the spectrum of variation defined by universal grammar. The question I am asking is whether the order of the constituents of a clause plays a role (a) in the computation from narrow syntax to the semantics interface and

(b) in the semantic component. I pursue an approach where the constituents of a clause do play such a role and ask what the implications are for the syntax and semantics interface. The findings of the dissertation show that order plays a role in the semantic component and in the computation from narrow syntax to semantic interface in Romeyka.

Acknowledgements

It is customary for PhD candidates to express their thanks to their supervisors for advice and guidance rendered throughout the course of their study. In this case, this expression of gratitude is, I think, in order, as Dr Ioanna Sitaridou gave me not only support and guidance, but also an opportunity without which the completion of this dissertation would not have been possible.

Throughout the course of this study, many individuals encouraged, supported and tolerated me. Among them, Professor Ianthi Maria Tsimpli, Professor Dr Stavros Skopeteas, Dr Victor Acedo-Matellán, Dr Theodora Alexopoulou, Dr Theresa Biberauer, Dr Bert Vaux, Dr Jenekke van der Wal and Dr Elena Ioannidou were most helpful in pointing out to me considerations and sources that had escaped my attention.

To Professor Noam Chomsky, for the interest he has shown to my research —although the final remarks of this study oppose his— should be ascribed in large part such merits as this study may claim.

My thanks also go to the Aridemi Stiftung and the A. G. Leventis Foundation for financially supporting part of this study. Special thanks are due to the late Mr Sotirios Hadjimina and Mr Alex Hadjimina from the Aridemi Stiftung for their personal interest in my studies.

I would also like to acknowledge the hospitality provided by the Romeyka speakers of the village of 'Anasta', who furnished me with invaluable language material during my fieldtrip there in 2015.

As I look back over and beyond the years of study which this work represents, I am constantly conscious of the ones whose love was an unfailing source of encouragement and support: the monastic community of the Holy Monastery of All Saints, in Imathia, Greece and especially Abbess Christodouli. I remain more deeply grateful to them than words can express.

My gratitude goes over and above all to my parents, my late father Mr Marios Neocleous and my mother Mrs Christina Pentara and my older brother, Dr Charalambos Neocleous; without their financial and emotional support provided to me throughout the course of this study, this dissertation would not have been completed. This dissertation is dedicated to the memory of my father. I miss him every day, but I am glad to know he saw this process through its completion.

Finally, my best memory from Cambridge was meeting Dimitra. And, you know, people believe that graduating from Cambridge leads you to a prestigious (academic) career. In my case, it did not. However, without studying in Cambridge, I would not have met Dimitra and

she is the most important person in my life, so you could say that it was the most important thing I built in my time in Cambridge. We have all started lifelong friendships in Cambridge and some of us even families. That is why I am so grateful to that place. Thanks, Cambridge.

All inadequacies are my own.

To the memory of my father,

Marios Neocleous

(†13 August 2018)

"Πέρα ἀπὸ τὸ αὔριο, πεπρωμένος
μᾶς περιμένει ἕνας σκελετωμένος τάφος.
Πέραν τοῦ τάφου
μᾶς καρτερᾶ τὸ μυστικό, ἡ ἀπάντηση
ποὺ θὰ βροῦμε σ' ἕνα σταυροδρόμι,
στὶς γηρασμένες νότες μιᾶς ζωῆς,
στὶς χορδὲς μιᾶς μετεμψυχωμένης Λύρας."[1]

[1] Neocleous Marios (1970). "Πέραν τοῦ τάφου." Morfosis *26 (303-304)*: p. 30.

Table of contents

List of tables

List of figures

List of maps

List of abbreviations

<…>	ordered set
≤	infix
^	linearisation feature
&	conjunction
∈	is an element of
∃	existential quantifier
Ø	zero morph
1	first person
2	second person
3	third person
α	alpha (variable)
β	beta (variable)
γ	gamma (variable)
λ	meaning
π	form
Σ	clause
ABL	ablative case
ACC	accusative case
Adv^0	adverbial head
AdvP	adverbial phrase
AG	Ancient Greek
ALP	Adverbial Licensing Principle
AMG	Asia Minor Greek
A-Top	aboutness topic
$A\text{-}Top^0$	aboutness topic head
A-TopP	aboutness topic phrase
Aux	auxiliary
A X B	Cartesian product of sets A and B
C	complementiser
C^0	complementiser head
CD	clitic doubling
cf.	confer
CFC	core functional categories
CFI	condition of full interpretation
C-Foc	contrastive focus
$C\text{-}Foc^0$	contrastive focus head
C-FocP	contrastive focus phrase
CI	condition of inclusiveness
C-I	conceptual-intentional system
CLF	classifier
CLlD	clitic left dislocation
CP	complementiser phrase
C-Top	contrastive topic
$C\text{-}Top^0$	contrastive topic head
C-TopP	contrastive topic phrase
D	definiteness
D	derivation
D	determiner
D^0	determiner head
DP	determiner phrase
DAT	dative case
D(em)	demonstrative
D(em)P	demonstrative phrase
DP	determiner phrase
DR	definite restriction
e	event variable
e	empty category
EF	edge feature
e.g.	exempli gratia
E-language	external data
EPP	extended projection principle
et al.	et alia

et seq.	et sequentes	LOC	locative case
F	female grammatical gender	MedGr	Medieval Greek
F	functional head	M	masculine grammatical gender
^{F}	focus interpretation	MF	multiple foci
F1	first factor	MG	Modern Greek
F2	second factor	MOD	modal
F3	third factor	MWQ	multiple *wh*-questions
FE	feature economy	N	narrow syntax
FI	full interpretation	N	neuter grammatical gender
FL	language faculty	NEG	negative marker
Foc	focus	NOC	non-obligatory control
FOFC	final-over-final condition	NOM	nominative case
Fn.	footnote	NP	nominal phrase
FUT	future tense	Num	numeral
GEN	genitive case	NumP	numeral phrase
GER	gerund	O	object
HelGr	Hellenistic Greek	$_{O}$	ordinary interpretation
$_{i}$	index	OPT	optative
i.a.	inter alia	P	parameter
i.e.	id est	PART	participle
iff	if only if	Past	past tense
I-Foc	information focus	PC	predicative complement
I-Foc^{0}	information focus head	PCC	person-case-constraint
I-FocP	information focus phrase	PF	phonetic form
IG	input generalisation	PHON	phonological component
IMP	imperative	PIC	phase impenetrability condition
IMPF	imperfect	PL	plural number
INF	infinitive	PLD	primary linguistic data
IPA	international phonetic alphabet	PNP	past tense, non-perfective aspect
$_{j}$	index	Pol	polarity
$_{k}$	index	PolP	polarity phrase
L1	first language	POSS	possessive
L2	second language	pro-drop	pronoun-dropping
LF	logical form	PROG	progressive aspect

PRT	particle
P	preposition
PP	prepositional phrase
Q	question
RAH	rich agreement hypothesis
REL	relativiser
S	subject
SEM	semantic component
SG	singular number
S-M	sensory-motor system
SMT	strong minimalist thesis
Spec	specifier
t	trace
T	tense
T^0	tense head
Top	topic
Top^0	topic head
TopP	topic phrase
TP	tense phrase
*u*ϕ	agree features
*u*F	uninterpretable feature
UG	universal grammar
V	verb
V^0	verb head
V2	verb second
v	value
v	little verb
v^0	little verb head
VOC	vocative case
VP	verb phrase
vP	little verb phrase
VTS	value-transfer simultaneity
x	x (variable)
X^0	X head
XP	X phrase
y	y (variable)

List of symbols

The following symbols are used in the linguistic examples to represent sounds that are not used in the International Phonetic Alphabet (IPA):

Symbol	IPA	Description
š	[ʃ]	voiceless postalveolar fricative
ž	[ʒ]	voiced palatoalveolar fricative
tš	[ʧ]	voiceless postalveolar affricate
dž	[ʤ]	voiced postalveolar affricate

Notes on intonation marking:

Lexical stress is marked by a sign assigned above the vowel of the stressed syllable. Prosodic stress is marked by capitalising the stressed syllable. Pause is marked by a comma.

1 Introduction

1.1 Scope of the dissertation

To understand how despite the *poverty of the stimulus*, nonetheless children acquire the target language is a problem inherent in every discussion of the theory of language. Modern and contemporary linguistic research centred around the names of Ferdinand de Saussure, Roman Jakobson and, more recently, Noam Chomsky has illustrated that, contrary to the 'Bloomfieldian' claim, we can speak of a human language and not only of human languages. I believe that Chomsky's argument for the principle of 'creativity' (open-endedness) of human language is decisive. This should guard us from exaggerating linguistic differences, though, of course, it does not spare us the task of interpretation of linguistic variation. Contemporary social anthropology shows us that the arrogant late nineteenth-century myth about the 'pre-logical' mentality of 'primitives' and the 'logical' mentality of Modern Man, has more chance than the primitive thought to be considered a childish superstition. Cultural evolution (after all very short in comparison with biological and inorganic evolution) has not as yet produced a new species of *homo sapiens*. Human nature is essentially one and the same. On the spatial plane now, modern science helps human to become at last more universal and ecumenical — the rest is for the anthropologists.

To this end, this dissertation is essentially an attempt to investigate how linguistic variation could be accommodated in a universal grammar; in particular, I aim to examine how word order variation could be part of such a universal grammar.

Against this background, in the generative literature, it had been argued that word order variation is out of the spectrum of universal grammar; hence it does not play any role in the efficient computation of interface conditions and mainly in relation to the syntax and semantics interface (see Chomsky et al. 2017, i.a.). The question I am asking is what if there is evidence that there is a natural language whose rules and operations are defined in linear terms? If syntactic operations could make reference to order and if order were to satisfy principles of efficient computation, then there would be evidence that variation due to order belongs to the spectrum of Universal Grammar on the one hand and as such order meets the criterion of

evolvability along with the operation Merge on the other. However, if that is the case, the emergence of nonlinear orders cross-linguistically will still remain a deep mystery.

My empirical domain is Romeyka, an Asia Minor Greek variety, for the crucial reason that linguistic phenomena that are attested in Romeyka are not attested in other Modern Greek varieties and have been phylogenetically linked directly to Hellenistic Greek. Its archaic form results from sociohistorical factors that forced it to develop in isolation to other Greek varieties for centuries (see Sitaridou 2016).

Interestingly, the empirical data from Romeyka show that the order of the constituents of a clause plays a role (a) in the computation from narrow syntax to the semantics interface and (b) in the semantic component. Thus, I develop a theoretical model to account for word order variation in Romeyka, in which the constituents of a clause do play such a role and ask what the implications are for the syntax and semantics interface.

The proposal of the study (a) provides an alternative theoretical model of information structure within the minimalist program that is principled explained, rather than the cartographic one (see Cinque 1999, Rizzi 1997), (b) offers a principled explanation of previous theories on linearisation, such as the Kayne's Linear Correspondence Axiom (LCA) (1994) and the Biberauer et al.'s (2014) Final-Over-Final Condition (FOFC), as well as (c) a principled explanation of word order change due to language contact and (d) provides evidence against the claim that word order does not play any role in the efficient computation of interface conditions and mainly in relation to the syntax and semantics interface (see Chomsky et al. 2017, i.a.).

In the remainder of this introductory chapter, I present the theoretical framework I follow in this study (§1.2.), an introduction to Romeyka (§1.3), the objectives of the study (§1.4) and the methodology of the study (§1.5). In §1.6, I present a roadmap of the dissertation.

1.2 Theoretical framework

1.2.1 The architecture of the faculty of language

The approach I pursue in this study follows the recent developments within the framework of generative grammar, i.e. the minimalist program, which was initiated by Chomsky (1995). The architecture of the faculty of language (FL) that is typically assumed within the minimalist program is given in (1):

(1) Lexicon

⇓

Narrow Syntax: External and Internal Merge

⇓

PHON ⇐ Spell-Out, Transfer ⇒ SEM

⇓

PHON ⇐ Spell-Out, Transfer ⇒ SEM

⇓

PHON ⇐ Spell-Out, Transfer ⇒ SEM

⇓

...

The basic operation in narrow syntax is Merge. There are two kinds of Merge operations, namely External Merge and Internal Merge. The former is presumably the first Merge of the derivation, while the latter results from movement. Formally, these two operations amount to the same: two objects are put together and the operation itself does not differ depending on whether the object comes from the lexicon or whether it is an object that already exists in the structure that is being re-merged. At certain points during the derivation, pieces of structure are sent off to the interfaces. These pieces are sent to both interfaces at the same time. This is done by the operation Transfer. The phonological part of Transfer is called 'Spell-Out'. In the literature, these pieces of structure are called phases.

(1) refers to the interfaces as SEM and PHON. The two interfaces are accessed by the Sensory-Motor system (*aka* the Articulatory-Perceptual system) and the Conceptual-Intentional system. This follows Chomsky (2004) who argues that: "The last line of each derivation is a pair <PHON, SEM>, where PHON is accessed by S-M and SEM by C-I. D

converges if PHON and SEM each satisfy [interface conditions]; otherwise it crashes at one or the other interface" (Chomsky 2004: 106).

An insight regarding the syntax and semantics interface is given by Chomsky et al. (2017) who postulate a quite different view of grammar if we compare it to previous stages of the generative framework:

> "Objects constructed in core syntax must be mapped onto representations that can be accessed by C-I and S-M systems: SEM and PHON, respectively. Consequently, there must be an operation TRANSFER that hands constructed objects over to the mapping components. The mapping to PHON is complex, involving the "flattening" of hierarchical structure and computation of stress, prosody etc. […] The mapping to SEM is more direct, given that hierarchical structure is the input to semantic interpretation; just how complex it is depends on the obscure question of where the boundary between the generative procedure and C-I systems is to be drawn" (Chomsky et al. 2017: 8-9).

In the aforementioned quote, no LF is conceived of as a syntactic level of representation. If there is no LF, then the syntactic structure has to be mapped directly to SEM. Semantic interpretation is therefore still derived from syntactic representations and even more directly if there is no intermediate level of representation. Thus, the view here assumes that SEM is interpretive rather than generative.

1.2.2 Variation in a minimalist system

A recent development within the minimalist program is the adoption of the biolinguistic perspective, in which, according to Chomsky (2005), three factors must be involved in the development of language in the individual, as given in (2):

(2) Three factors in language design:
   a. Genetic endowment
   b. Experience
   c. Principles not specific to the FL, the human faculty of language
   (Chomsky 2005: 6)

First, Factor I (see (2)a) is the domain of Universal Grammar (UG). Second, Factor II (see (2)b) is the external data (E-language) that constitutes the linguistic environment in which language acquisition takes place. Factor III (see (2)c) comprises "general properties of organic systems" (Chomsky 2005: 6), the results of physical constraints on the form and development of living organisms. In the case of FL, a biological organ like any other (a 'mental' organ), such third-factor constraints might include principles of efficient computation and the interface conditions imposed from outside FL by the semantic (SEM) and phonological (PHON) systems with which it interacts. Factor III is what distinguishes minimalism from other approaches to FL, offering a different benchmark for what counts as a genuine explanation (taking us "beyond explanatory adequacy", in Chomsky's words).

The goal of the minimalist program is then to move descriptive technology from Factor I (the genetic endowment, UG) to Factor III by showing that that technology is dispensable, or reducible to third-factor effects. A guiding hypothesis that we can entertain in order to pursue this aim is the Strong Minimalist Thesis (SMT), which states that no aspect of FL is without a principled, third-factor explanation —specifically, we entertain the thesis that all properties of FL contribute to a computationally efficient satisfaction of interface conditions (IC). In that sense (i.e. if SMT were true) then FL would be a "perfect" solution to IC. Such a perfect solution would comprise an empty UG: logically, if everything is Factor III, then nothing is Factor I. Clearly, this is too strong a hypothesis —UG cannot be completely empty, otherwise there would be no FL. The genetic endowment UG, then, should be maximally (but not completely) empty, consisting of a minimal unexplained residue. The question then arises as to what that minimal residue must be in order to account for the human capacity.

In such a minimalist system, variation is restricted (a) within the narrow syntax and (b) at the phonological (sensorimotor) interface. Within the narrow syntax, which is the domain of the SMT and thus a parameter-free UG, free variation is predicted to occur, with each competing option a possible choice in every derivation At the phonological (sensorimotor) interface, to which the SMT does not apply, competing options are resolved consistently in a language through parametric choices, yielding macroparametric variation at the PHON-interface.

At this point, I should address the question of the role and nature of formal hierarchies in current syntactic theory. As is well-known, Cinque (1999: 90, 106) presented an elaborate functional structure for the clause (i.e. the former IP/TP) of the following type (see (3)):

(3) $\text{Mood}_{\text{Speech Act}}$ $\text{Mood}_{\text{Evaluative}}$ $\text{Mood}_{\text{Evidential}}$ $\text{Mod}_{\text{Epistemic}}$ T(Past) T(Future) $\text{Mood}_{\text{Irrealis}}$ $\text{Mod}_{\text{Necessity}}$ $\text{Mod}_{\text{Possibility}}$ $\text{Asp}_{\text{Habitual}}$ $\text{Asp}_{\text{Repetitive(I)}}$ $\text{Asp}_{\text{Frequentative(I)}}$ $\text{Asp}_{\text{Celerative(I)}}$ $\text{Mod}_{\text{Volitional}}$ $\text{Mod}_{\text{Obligation}}$ ModAbility/Permission AspCelerative(I) T(Anterior) $\text{Asp}_{\text{Terminative}}$ $\text{Asp}_{\text{Proximative}}$ AspSgCompletive(I) AspPlCompletive Voice AspCelerative(II) AspSgCompletive(II) $\text{Asp}_{\text{Repetitive(II)}}$ $\text{Asp}_{\text{Frequentative(II)}}$ $\text{AspSg}_{\text{Completive(II)}}$

Cinque's evidence for this hierarchy came from converging facts regarding the ordering of adverbs, auxiliaries and particles in many languages and suffixes, particularly in agglutinating languages. More recently, Cinque (2006) has developed the hierarchy further.

In a similar way, Rizzi (1997) put forward an elaborated version of the left periphery (the earlier CP), splitting CP into ForceP, FocP, a possibly iterated TopP and FinP. The original proposal has been elaborated in various ways; Ledgeway (2010: 51, (80)) summarises these developments with the following cartographic structure (see (4)):

(4) DeclP FrameP1 FrameP2 ConcP HypP ExclP ThemeP1 ThemeP2 IntP C-FocP1 C-FocP2 I-FocP1 I-FocP2 FinP

Biberauer & Roberts (2015) put forth a proposal to account for the parameters in a minimalist system. According to this study, at a relatively low level of granularity, the clause consists of the core functional categories; at a higher level of granularity, there are the cartographic structures. Furthermore, these are not the only syntactically and semantically relevant levels of organisation in the clause. The phase level is higher than the core functional categories (since T^0, at least, is not inherently phasal); phases are clearly relevant to syntax and semantics and they are also computed in one way or another, at the PHON-interface. A still higher level of organisation is the Extended Projection. To summarise, Biberauer & Roberts (2015) suggest that clauses can be analysed at different levels of "magnification", as follows:

(5) Extended Projection (V^0) > phase (C^0, v^0) > CFC (C^0, T^0, v^0) > "cartographic fields" (e.g. Tense, Mood, Aspect, Topic, Focus) > semantically/lexically distinct heads.
(Biberauer & Roberts 2015: 4)

Moreover, Biberauer & Roberts (2015) develop the idea that parametric variation is an emergent property of an interaction of an underspecified UG, the PLD and third-factor computational conservativity on the part of the acquirer (see also Biberauer 2017). The two principal linguistic manifestations of the acquirer's general computational conservativity are Feature Economy (FE) and Input Generalisation (IG) (the non-language-specific third factor at work here can be thought of as the general imperative to "make maximal use of minimal means"). These can be defined as follows:

(6) a. Feature Economy (FE) (generalised from Roberts & Roussou (2003: 201):
Postulate as few formal features as possible to account for the input.

b. Input Generalisation (IG) (adapted from Roberts (2007: 275):
If a functional head F sets parameter P_j to value v_i, then there is a preference for all functional heads to set P_j to value v_i.

(Biberauer & Roberts 2015: 7)

From an acquirer's perspective, FE requires the postulation of the minimum number of formal features consistent with the input. IG embodies the logically invalid, but heuristically useful inference mechanism of learning from an existential to a universal generalisation. Like FE, it is stated as a preference, since it is always defeasible by the PLD. More precisely, Biberauer & Roberts (2015) do not see the PLD as an undifferentiated mass but take the acquirer to be sensitive to particular aspects of PLD such as movement, agreement, etc., readily encountered in simple declaratives, questions and imperatives. So, the interaction of the second (PLD) and third factor-derived (FE, IG) factors is crucial. It may seem as though IG will create superset traps for the acquirer, but this is not the case if we think of the acquirer as overgeneralising due to their ignorance of categorial distinctions. This ignorance gradually erodes through the learning process, as finer and finer distinctions are made as a consequence of the interaction of all three factors: UG leaves certain options open (essentially many aspects of the formal feature inventory), the PLD provides evidence regarding which options are needed and FE and IG ensure that the maximally conservative options are always preferred, but that the formal distinctions required to capture the observed syntactic patterns are introduced during the acquisition process (Biberauer & Roberts 2015: 7).

1.2.3 Word order and information structure in a minimalist system

In the minimalist approach, It has long been believed that ordering information is not established in narrow syntax but rather at PHON (Chomsky 1995: 334-335) (see (7)):

(7) Nonlinear syntax:
"Syntactic operations/relations make no reference to notions of linear ordering and directionality."
(Chomsky 1995: 334)

Narrow syntax only operates on hierarchical structures, as it oblivious to the processes that transform structures into linear entities. PHON is commonly assumed to contain language-variable operations of morphology-phonology (Chomsky 1965, et seq.), which implies that the mechanisms establishing linear order, i.e. linearisation, might fall within the domain of language variable operations of morphology-phonology and thus may be subject to cross-linguistic variation. This is a move that is in line with the main principles of the minimalist program, where syntax is minimalised so as to include the main generative procedures that combine elements to create larger units, such as Merge and many traditional aspects of that theory of syntax are moved to interfaces.

Besides, discourse-related elements, such as topic and focus, are purely semantic features that are only visible and accessible at the interfaces (see (8)):

(8) Configurational syntax:
"These are manifold, involving topic-focus and theme-rheme structures, figure-ground properties, effects on adjacency and linearity and many others. Prima facie, they seem to involve additional level or levels, internal to the phonological component, postmorphology but prephonetic, accessed at the interface along PF (Phonetic Form) and LF (Logical Form)."
(Chomsky 1995: 220)

This is reminiscent of previous accounts within the Government and Binding framework, in which discourse-related information is represented and interpreted at LF (Chomsky 1976, 1981, Horvath 1986, Huang 1982, Rochemont 1986) or at an abstract level of representation derived from LF, for example LF' in Huang (1984).

As a matter of principle, it is assumed that "there is no clear evidence that order plays a role at LF or in the computation from N to LF" (Chomsky 1995: 334). For instance, consider the following extract from Chomsky et al. (2017):

> "PAIR-MERGE is a formally distinct operation from Simplest MERGE, hence raises problems of evolvability. Ideally, it could be shown to be dispensable. We do not take up the challenge here; for some suggestive work on adjunction that does not invoke special operations (but at the cost of introducing other stipulations), see Hunter 2015. As for parenthesis, it seems to us that the only principled approach consistent with evolvability considerations relegates the phenomenon entirely to discourse pragmatics, obviating the need to enrich UG with special operations. That is, parenthetical expressions, which are frequently elliptical, are generated independently and interpolated or juxtaposed only in production" (Chomsky et al. 2017: 18).

According to this extract, Merge is assumed not to impose order, i.e. {x, y} = {y, x}. As such, order is structure-dependent, i.e. no syntactic operation can make reference to it. It has also been claimed that hypothetical languages, in which syntactic operations are defined in linear terms, such that Merge creates an ordered pair <x, y>, are outside of the spectrum of variation defined by UG (see Musso et al. 2003, Smith & Tsimpli 1995).

1.3 Introduction to Romeyka

1.3.1 Historical context

Romeyka[2] is the only remaining variety of Asia Minor Greek (henceforth AMG) that is still spoken in the area historically known as Asia Minor (present-day Anatolia, Turkey). AMG comprises seven varieties, namely the varieties that were spoken in the areas historically known as Bithynia, Cappadocia, Göylde, Livísi, Phárasa, Pontus and Sílli (Dawkins 1916: 5) (see Map 1):

Map 1. Asia Minor: the historical berceau (Dawkins 1916: Pl. I).

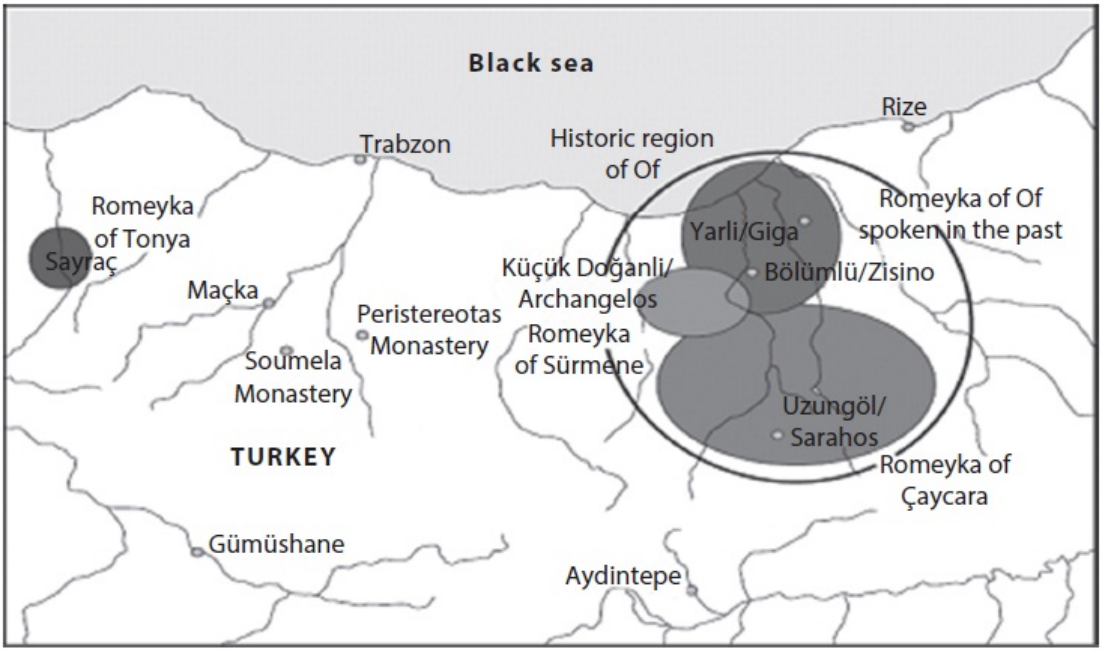
The majority of the AMG speakers were forced to relocate to Greece as refugees in the aftermath of the defeat of the Greek army in the Greek-Turkish war (1919-22) as a result of the Treaty of Lausanne for the Population Exchange between Greece and Turkey (1923). Ever since, AMG varieties have continued to be spoken in Greece and elsewhere across the world to varying extents, but because of the robust contact with Modern Greek (henceforth MG) in Greece and other languages abroad, the younger generations became attrited. In Turkey,

[2] Regarding glossonymy, I use the term 'Romeyka' following Sitaridou (2013) et seq. Schreiber & Sitaridou (2017) note on Romeyka glossonymy: "Despite the fact that Romeyka is the emic name, Mackridge (1987) uses the term 'Muslim Pontic' as a *terminus technicus* for the same variety we document (a practice followed by Brendemoen 2006, Özkan 2013 and partially by Bortone 2009). However, we prefer the term /roméika/ in line with what the majority of speakers use (some speakers may also say /romáika/ or even /rumáika/) as it would be outside current academic practice to use a term that speakers themselves do not identify with" (Schreiber & Sitaridou 2017: 2).

Romeyka is still spoken in the area historically known as Pontus (present-day, Black Sea district, Turkey) (see Sitaridou 2013).

Romeyka speakers in the Pontus area, by virtue of being Muslims, were exempt from the population exchange between Greece and Turkey following the Treaty of Lausanne in 1923. For Pontus, the result was an exodus of Greek- (and Turkish-)speaking Christians, leaving small enclaves of Greek-speaking Muslims in Turkey.

In fact, today there remain three Greek-speaking enclaves: Of/Çaykara, Sürmene and Tonya (see Deffner 1878, Mackridge 1987, 1999, Michelioudakis & Sitaridou 2012, 2013, 2016, Özkan 2013, Parcharidis 1880, Sağlam 2017, Schreiber 2018, Schreiber & Sitaridou 2017, Sitaridou 2013, 2014a, 2014b, 2016) (see Map 2):



Map 2. The historical region of Of in Pontus and current Romeyka-speaking enclaves (Sitaridou 2013: 99).

1.3.2 The sociolinguistic background

According to Mackridge (1987) and Andrews (1989), the Romeyka-speaking community in the Black Sea area consists of approximately 5.000 speakers. This population figure comes from the last available general census (Genel Nüfus Sayımı) from 1965, which records mother tongue. Based on these records, there were 4.535 Romeyka speakers. However, this number may not reflect reality due to a biased choice of Turkish as mother tongue and the exclusion of migration data (see also Brendemoen 2002, Mackridge 1987, Özkan 2013).

Crucially, Romeyka is an endangered language due to its massive contact with Turkish. In 'Anasta'[3] —the village from which I collected my data— all of the speakers are bilingual in Romeyka and Turkish, although the levels of bilingualism vary from simultaneous to additive bilingualism. Age and gender represent important sociolinguistic variables. Older speakers qualify as first language (L1) speakers, while younger speakers acquire Romeyka as second language (L2). The same holds for female speakers, the majority of whom seem to be L1 speakers, whereas male speakers are mostly L2 (see Sitaridou 2013) (see Table 1):

Table 1. Romeyka-Turkish shift (Sitaridou 2013: 104).

	Romeyka	**Turkish**
G1	L1	L2
G2	(Late) L1	(Early) L2
G3	Late L1	Early L2
G4	Early L2	L1

Vitality was found to be affected by the following speaker-related variables: (a) age; (b) language competence; and (c) gender. These variables affect language vitality in the following ways (Schreiber & Sitaridou 2017: 13):

a. The older the speaker is, the stronger vitality is.
b. The higher the linguistic competence of the speaker, the stronger the vitality is.
c. Females generally hold more positive attitudes than males.
(Schreiber & Sitaridou 2017: 13)

The language vitality of Romeyka is much more threatened than is suggested in the literature. Crucially, it will take far more for the situation to be reversed (Schreiber & Sitaridou 2017: 13).

[3] Following Sitaridou (2014a: 29, Fn. 3), I call the village where Dr Sitaridou and I conducted fieldwork 'Anasta' to preserve the anonymity of informants and the village.

1.4 Objectives

Against this background, in this dissertation I investigate word order and information structure in the light of recent developments within the minimalist program. I specifically pursue a principled explanation of word order within the biolinguistic perspective. In that sense, I entertain the thesis that all properties of the faculty of language contribute to a computationally efficient satisfaction of interface conditions. The language examined is Romeyka, that is the only Asia Minor Greek variety still spoken in the area historically known as Asia Minor (present-day Anatolia, Turkey).

The objective of this study is therefore twofold: (a) descriptively, to examine word order variation in Romeyka and (b) theoretically, to investigate whether such word order variation could be a language specific property, or rather could be accommodated in a minimalist system.

Descriptively, I aim (a) to determine the pragmatically unmarked and marked word orders in Romeyka, (b) to examine their typological classification and (c) to attenuate their evolution.

Theoretically, this dissertation is fundamentally about the role that order plays in the efficient computation of interface conditions and mainly in relation to the syntax and semantics interface. The question I am asking is whether the order of the constituents of a clause plays a role (a) in the computation from narrow syntax to the semantics interface and (b) in the semantic component. I pursue an approach where the constituents of a clause do play such a role and ask what the implications are for the syntax and semantics interface.

1.5 Methodology

1.5.1 Collecting the data

The results reported here were obtained from two corpora consisting of data collected in a remote part of the Of/Çaykara region, in a village, which will be referred to as 'Anasta', in order to preserve the anonymity of the informants and the village (Sitaridou 2014a: 29, Fn. 3): (a) a corpus consisting of data that were collected during fieldwork I carried out in July 2015 in the village of 'Anasta' (see Table 2),[4] and (b) a corpus of data that were collected during fieldwork carried out by Dr Ioanna Sitaridou during her fieldtrips to 'Anasta' in July 2012, July 2014 and July 2015 (see Table 3).[5] The two corpora comprise 14:43:51 hours of audio recordings; Corpus (a) comprises 06:48:21 hours and corpus (b) comprises 07:55:30 hours. Throughout the dissertation, I use data from both corpora. There were seven informants, whose details are shown in Table 4:

Table 2. Nicolaos Neocleous's Romeyka corpus.

Number	File	Time	Year	Speaker(s)
1	140102_0006	00:05:56	2015	S01
2	140102_0007	00:04:17	2015	S01
3	140102_0008	00:10:41	2015	S01
4	140102_0009	00:09:08	2015	S01
5	150702_0010	00:17:32	2015	S02
6	150702_0011	00:14:27	2015	S02
7	150702_0012	00:09:41	2015	S02
8	150702_0013	00:23:31	2015	S01
9	150702_0014	00:14:16	2015	S01
10	150702_0015	00:11:13	2015	S03, S04
11	150702_0016	00:23:24	2015	S03, S04
12	150702_0017	00:03:20	2015	S03, S04
13	150702_0018	00:15:13	2015	S03, S04
14	150702_0019	00:10:02	2015	S01

[4] The data on Romeyka result from fieldwork I conducted in the area in 2015 thanks to the Lister Fund granted to me by Queens' College, University of Cambridge and funding for work carried out in Cambridge granted to me by the Faculty of Modern and Medieval Languages, University of Cambridge.
[5] I am very grateful to Dr Ioanna Sitaridou for sharing with me audio recordings she collected during her fieldtrips in 'Anasta'.

15	150702_0020	00:01:43	2015	S01
16	150702_0021	00:00:25	2015	S01
17	150702_0022	00:11:06	2015	S01
18	150702_0023	00:24:58	2015	S01
19	150702_0024	00:00:22	2015	S05
20	150702_0025	00:00:32	2015	S05
21	150702_0026	00:00:02	2015	S05
22	150702_0027	00:00:03	2015	S05
23	150702_0028	00:02:55	2015	S02, S05
24	150702_0029	00:00:34	2015	S02, S05
25	150702_0030	00:00:10	2015	S01
26	150702_0031	00:10:03	2015	S01
27	150702_0032	00:07:57	2015	S01
28	150702_0033	00:06:16	2015	S01
29	150702_0034	00:09:01	2015	S01
30	150702_0035	00:00:27	2015	S01
31	150702_0036	00:03:04	2015	S01
32	150702_0037	00:02:40	2015	S01
33	150702_0038	00:11:59	2015	S01
34	150703_0039	00:13:19	2015	S02
35	150703_0040	00:23:47	2015	S01
36	150703_0041	00:07:36	2015	S01
37	150703_0042	00:16:48	2015	S01
38	150703_0043	00:05:22	2015	S01
39	150703_0044	00:01:33	2015	S01
40	150703_0045	00:01:04	2015	S01
41	150703_0046	00:17:13	2015	S01
42	150703_0047	00:00:41	2015	S01
43	150703_0048	00:44:14	2015	S06
44	150703_0049	00:08:03	2015	S06
45	150703_0050	00:01:43	2015	S06

Table 3. Ioanna Sitaridou's Romeyka corpus.

Number	File	Time	Year	Speaker(s)
1	20120713 192027	00:45:16	2012	S01, S07
2	812_0004	00:06:55	2012	S07
3	812_0006	00:01:58	2012	S07
4	812_0008	00:14:13	2012	S07
5	812_0012	00:01:15	2014	S07
6	812_0029	00:00:38	2014	S01
7	812_0044	00:27:48	2014	S07
8	812_0048	00:18:40	2014	S07
9	812_0055	00:03:52	2014	S01
10	812_0056	00:04:39	2014	S01
11	812_0057	00:04:20	2014	S01
12	812_0058	00:07:26	2014	S01
13	812_0059	00:01:46	2014	S01
14	812_0061	00:11:04	2014	S01
15	812_0062	00:16:55	2014	S01
16	812_0065	00:14:52	2014	S07
17	812_0067	00:05:29	2014	S07
18	812_0068	00:08:44	2014	S07
19	812_0069	00:03:28	2014	S07
20	812_0071	00:08:47	2014	S07
21	812_0074	00:05:36	2014	S07
22	812_0093	00:00:30	2014	S01
23	812_0103	00:28:32	2015	S01
24	812_0106	00:27:44	2015	S01
25	812_0108	00:26:57	2015	S01
26	812_0109	00:06:04	2015	S01
27	812_0110	00:00:06	2015	S01
28	812_0111	00:01:16	2015	S01
29	812_0112	00:13:01	2015	S01
30	812_0113	00:19:06	2015	S01
31	812_0114	00:03:09	2015	S01

32	812_0115	00:05:15	2015	S01
33	812_0116	00:11:25	2015	S01
34	812_0117	00:00:45	2015	S05
35	812_0118	00:02:30	2015	S05
36	812_0119	00:01:19	2015	S01
37	812_0120	00:11:11	2015	S01
38	812_0121	00:07:53	2015	S01
39	812_0122	00:10:51	2015	S01
40	812_0123	00:03:48	2015	S01
41	812_0124	00:02:00	2015	S01
42	812_0125	00:00:31	2015	S01
43	812_0126	00:16:31	2015	S01
44	812_0127	00:00:20	2015	S01
45	812_0128	00:01:23	2015	S01
46	812_0129	00:04:39	2015	S01
47	812_0130	00:02:00	2015	S01
48	812_0131	00:00:52	2015	S01
49	812_0132	00:00:09	2015	S01
50	812_0133	00:07:53	2015	S01
51	812_0134	00:44:09	2015	S06

Table 4. List of participants.

No	Participant	Gender	Year of birth	Language repertoire	Files
1	S01	Female	1969	Romeyka L1	140102_0006
				Turkish L2	140102_0007
					140102_0008
					140102_0009
					150702_0013
					150702_0014
					150702_0019
					150702_0020
					150702_0021

150702_0022
150702_0023
150702_0030
150702_0031
150702_0032
150702_0033
150702_0034
150702_0035
150702_0036
150702_0037
150702_0038
150703_0040
150703_0041
150703_0042
150703_0043
150703_0044
150703_0045
150703_0046
150703_0047
20120713
192027
812_0055
812_0056
812_0057
812_0058
812_0059
812_0061
812_0062
812_0093
812_0103
812_0106
812_0108
812_0109

					812_0110
					812_0111
					812_0112
					812_0113
					812_0114
					812_0115
					812_0116
					812_0119
					812_0120
					812_0121
					812_0122
					812_0123
					812_0124
					812_0125
					812_0126
					812_0127
					812_0128
					812_0129
					812_0130
					812_0131
					812_0132
					812_0133
2	S02	Male	1969	Romeyka L1	150702_0010
				Turkish L2	150702_0011
					150702_0012
					150702_0028
					150702_0029
					150703_0039
3	S03	Male	2001	Turkish L1	150702_0015
				Romeyka	150702_0016
				heritage	150702_0017
					150702_0018
4	S04	Male	2003	Turkish L1	150702_0015

				Romeyka	150702_0016
				heritage	150702_0017
					150702_0018
5	S05	Male	1978	Romeyka L1	150702_0024
				Turkish L2	150702_0025
					150702_0026
					150702_0027
					150702_0028
					150702_0029
					812_0117
					812_0118
6	S06	Female	1941	Romeyka L1	150703_0048
				Turkish L2	150703_0049
					150703_0050
					812_0134
7	S07	Female	1941	Romeyka L1	20120713
				Turkish L2	192027
					812_0004
					812_0006
					812_0008
					812_0012
					812_0044
					812_0048
					812_0065
					812_0067
					812_0068
					812_0069
					812_0071
					812_0074

The data collection entailed oral interviews based on structured questionnaires (see Appendix A), as well as spontaneous data. A pilot test of the questionnaires was first carried out with a Turkish native speaker (speaking the variety of Istanbul, which is said to represent the Standard

Modern Turkish variety) and a Greek native speaker (speaking the variety of Athens, which is said to present the Standard Modern Greek variety). The data were audio recorded.

Finally, I draw data on Romeyka from a body of works on Romeyka that have already been published (see Michelioudakis & Sitaridou 2012, 2013 2016, Sitaridou 2013, 2014a, 2014b, 2016, i.a.).

1.5.2 Interpreting the data

The audio recordings were transcribed in the International Phonetic Alphabet (IPA) and annotated for the purposes of the study.

This study is in essence theoretical and not experimental, in the sense that I do not provide any statistics to analyse my data. The reason is crucial and fundamental for the nature of my study, which focuses on hierarchy rather than frequency. As such, hypotheses are developed based on empirical observations on primary language data, which lead to the development and proposal of a theoretical model within the minimalist framework, which makes predictions about (a) the semantic interpretation of different word orders in Romeyka and (b) the word orders in Romeyka that are sensitive to change due to their contact with local Turkish varieties.

1.6 Roadmap

The dissertation is structured as follows:

In chapter 2, I present the basic grammatical features of Romeyka. The analysis of the Romeyka grammar is non exhaustive at all, while it is presented in a traditional way.

In chapter 3, I examine the respective position of the subject (S), verb (V) and object (O) in matrix and subordinate declarative and interrogative clauses (both direct and indirect questions) in Romeyka. The goal of this chapter is to survey word order variation in Romeyka and to identify the positions of verbs and subjects in a clause in Romeyka

In chapter 4, (a) I determine the pragmatically unmarked word order in Romeyka and (b) I examine the syntactic distribution and semantic type of the constituents in pragmatically marked word orders in Romeyka. I specifically investigate topics, foci, *wh*-questions, multiple *wh*-questions and multiple focus in Romeyka.

In chapter 5, I aim (a) to pursue a third-factor (principled) explanation for word order variation within the minimalist program and (b) to test my analysis for every potential syntactic derivation of the subject (S), verb (V) and object (O) in Romeyka and map those derivations into PF and LF rules.

In chapter 6, (a) I aim at typologically classifying Romeyka word order; (b) I compare word order in Romeyka with word order in (i) Turkish, (ii) Georgian and (iii) Pontic Greek.

In chapter 7, I examine the evolution of VO and OV alternation in matrix and subordinate clauses in Romeyka.

The dissertation concludes in chapter 8.

2 An overview of Romeyka grammar

2.0 Introduction

The aim of this chapter is to present the basic grammatical features of Romeyka spoken in 'Anasta'. The analysis of the Romeyka grammar is non exhaustive at all, while it is presented in a traditional way. The goal of this chapter is to introduce some basic grammatical aspects of Romeyka, which are expected to help the reader of this dissertation to better understand the more complex grammatical features I examine in the remainder of this dissertation.

To my knowledge, so far there has not been any grammar of contemporary Romeyka (but see Schreiber 2018 for a state-of-the-art attempt to present published and on-going grammatical research on Romeyka). However, many grammatical aspects of Romeyka have already been examined and their results are published (c.f. Michelioudakis & Sitaridou 2012, 2013, 2016, Sitaridou 2014a, 2014b, 2016, i.a.).

The grammar presented in this chapter covers aspects that are only related to the Romeyka spoken in the village of 'Anasta' as these are reflected in the corpora examined in this dissertation. As such, it does not include any microvariation reflected between the Romeyka spoken either in other villages in Çaykara or in the enclaves of Sürmene and Tonya. However, whenever this microvariation is considered vital for the explanation of specific features of the Romeyka spoken in 'Anasta', a brief comparison between them is made.

In any case, the grammar presented in this chapter is not to be taken as a comprehensive grammar of Romeyka per se; though, it depicts a sufficient picture of the basic grammatical aspects that are attested in contemporary Romeyka spoken in 'Anasta' and, to my hope, are descriptive enough to make the remainder of the dissertation easier to be followed by the reader.

The chapter is structured as follows: in §2.1, I present the basic phonological features of Romeyka; in §2.2, I present the basic morphological features of Romeyka; in §2.3, I present the basic syntactic features of Romeyka. The chapter concludes in §2.4.

2.1 Phonology

2.1.1 Phonemic inventory

2.1.1.1 Vowels

Romeyka has a nine-vowel system /i, y, ɯ, u, e, o, œ, æ, a/, shown in Figure 1 below:

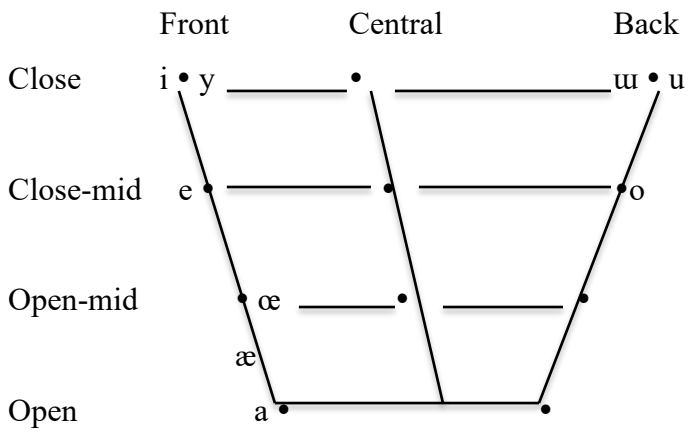


Figure 1. Romeyka vowel inventory.

In Figure 1, where symbols appear in pairs, the one to the right represents a rounded vowel.

The vowels /y/, /ɯ/ and /œ/ are only attested in Turkish loanwords (see examples of words with /y/ in (1), words with /ɯ/ in (2) and words with /œ/ in (3) below); hence, it is open to debate whether they are part of the Romeyka phonemic inventory or not. The vowel /æ/ occurs in inherited words. It is precisely the result of the reduction of the cluster /ia/ (see (4)):

(1) Romeyka:

a. her ɟyn (<Turkish *her gün* ‘every day’)

‘every day’

(S01; 150702_0013; 16:11)

b. énan c^{h}yt^{h}ýphane (< Turkish *kütüphane* ‘bookshop’)

‘a bookshop’

(S01; 812_0062; 09:11)

c. tsi birɟýlis (< Turkish *Birgül* ‘Birgül’)

‘Birgül’s’

(S01; 150702_0023; 23:23)

d. i ɟylsén (< Turkish *Gülsen* ‘Gülsen’)

‘Gülsen’

(S07; 812_0065; 05:19)

(2) Romeyka:

a. i nazlứ hanứm (< Turkish *Nazlı hanım* ‘Mrs Nazli’)

‘Mrs Nazli’

(S01; 812_0056; 02:38)

b. sɯnứf (< Turkish *sınıf* ‘grade’)

‘grade’

(S01; 150702_0015; 00:04)

c. altứ (< Turkish *altı* ‘six’)

‘six’

(S01; 150702_0015; 05:28)

d. ɟánɯ (< Turkish *yanı* ‘so’)

‘so’

(S01; 150703_0041; 00:08)

(3) Romeyka:

a. šœféris (< Turkish *şöfer* ‘driver’)
‘the driver’
(S07; 812_0044; 19:35)

b. ton œrendžín (< Turkish *öğrenci* ‘teacher’)
‘the teacher’
(S01; 812_0062; 02:42)

c. œ́ndže (< Turkish *önce* ‘before’)
‘before’
(S01; 812_0062; 16:20)

d. œméris (< Turkish *Ömer* ‘Ömeris’)
‘Ömeris’
(S01; 140102_0009; 06:13)

(4) Romeyka:

a. opsáræ (< Greek *psária* ‘fish’)
‘fish’
(S01; 150703_0041; 05:48)

b. ospítæ (< Greek *spítia* ‘houses’)
‘houses’
(S07; 812_0044; 17:46)

c. aðélfæ (< Greek *aðélfia* ‘siblings’)
‘siblings’
(S07; 812_0044; 17:21)

d. eftǽo (< Greek *ftiáo* ‘I make’)
‘I make’
(S01; 812_0110; 00:01)

2.1.1.2 Consonants

In Romeyka, 26 consonant phonemes are found, shown in Table 5 below:

Table 5. Romeyka consonants inventory.

	Bilabial	**Labiodental**	**Dental**	**Alveolar**	**Postalveolar**	**Palatal**	**Velar**
Plosive	p b		t d			c ɟ	k g
Nasal	m		n			ɲ	ŋ
Trill			r				
Fricative		f v	θ ð	s z	ʃ ʒ	ç ʝ	χ ɣ
Lateral approximant			l				

In Table 5, symbols to the right in a cell are voiced, to the left are voiceless. Shaded areas denote articulations that are not attested in Romeyka.

Some of the consonants shown in Table 5 are nonphonemic, i.e., some of them are in complementary distribution, with one variant appearing in a specific set of phonetic environments and the other being excluded from these contexts. In what follows, I list the distribution of these consonants as well as certain peculiarities of the remaining consonants. This inventory reveals that whether a word is borrowed from Turkish or inherited from previous stages of Greek has certain effects on the sounds it contains and the patterns of allophony.

First, [c] and [k] are allophones in borrowed words, in the sense that in the environments where [c] occurs [k], does not occur: [c] occurs only before [+front] vowels (see (5)a), [k] occurs elsewhere (see (5)b):

(5) Romeyka:

a. c^{h}itápin (< Turkish *kitap* ‘book’)
‘book’
(S07; 812_0004; 01:04)

b. ok^{h}úl:in (< Turkish *okul* ‘school’)
‘school’
(S07; 812_0048; 00:50)

In inherited words, [k] occurs before [−front] vowels (see (6)a) and in the consonant clusters where it is the first sound (see (6)b); however, in contrast to borrowed words, [k] in inherited words undergoes a sound change before [+front] vowels:

(6) Romeyka:

a. katsíte (< Greek *káθome* ‘I sit’)
‘sit’
(S07; 812_0048; 10:26)

b. mikrés:a (< Greek *mikrí* ‘young’)
‘young girl’
(S07; 812_0048; 00:01)

When /k/ occurs before [+front] vowels in inherited words, it undergoes a sound change that is explained in §2.1.2.2.

Second, [χ] occurs before [−front] vowels in inherited words (see (7)a) and in consonant clusters where it is the first sound (see (7)b):

(7) Romeyka:

a. χoríon (< Greek *χoríon* ‘village’)
‘village’
(S01; 812_0055; 01:19)

b. érχome (< Greek *érχome* ‘I come’)
‘I come’
(S01; 812_0061; 09:41)

When /χ/ occurs before [+front] vowels in inherited words, it undergoes a sound change that is explained in §2.1.2.2.

In borrowed words, irrespective of the frontness or backness of the following vowel, /χ/ is always /χ/ (see (8)a and (8)b):[6]

(8) Romeyka:

a. hastás (< Turkish *hasta* 'ill')
   'ill'
   (S07; 812_0004; 05:34)

b. her (< Turkish *her* 'every')
   'every'
   (S01; 150702_0013; 16:11)

Third, [k] in both inherited and borrowed words becomes [g] when it follows the nasal [n] (see (9)a and (9)b):

(9) Romeyka:

a. éton kalá > éton galá (< Greek *kalá* 'well')
   's/he was well'
   (S07; 812_0004; 00:47)

b. epíren kazáçin > epíren gazáçin (< Turkish *kazak* 'sweater')
   's/he got a sweater'
   (S01; 150702_0014; 04:52)

Fourth, in inherited words, [ʃ] occurs as a result of a palatalisation process that [χ] undergoes before [+front] vowels (see §2.1.2.2). It also occurs in words borrowed from Turkish in which it corresponds to the original [ʃ] (see (10)a and (10)b):

[6] Note that, throughout the dissertation, I use the voiceless fricative glottal consonant [h] in borrowed words from Turkish instead of the voiceless velar fricative consonant [χ], which I use in inherited words from Greek. Whether the two consonants are distinctive in Romeyka is not clear and is open for future research.

(10) Romeyka:

a. tšorbás (< Turkish *çorba* ‘soup’)
‘soup’
(S01; 150702_0014; 11:09)

b. konušéfko (< Turkish *konuşmak* ‘speak’)
‘I speak’
(S03; 150702_0015; 02:00)

Fifth, unlike the case in MG, [l] does not seem to have the palatal allophone [ʎ]. Independent of its position inside a word, it is always [l].

From the above inventory, we can already deduce that inherited and borrowed words may differ in their consonant inventories. There is one more phonological difference between inherited and borrowed words which should be mentioned here. Plosives are aspirated only in borrowed words (see (11)a). In inherited words, on the other hand, they are unaspirated (see (11)b):

(11) Romeyka:

a. $t^{h}op^{h}$lanéfkumen (< Turkish *toplamak* ‘gather’)
‘we are gathered’
(S01; 150703_0040; 12:33)

b. áuston (< Greek *ávɣustos* ‘August’)
‘August’
(S01; 150703_0041; 01:35)

Despite the phonemic differences between inherited and borrowed words, every word has a unique stress pattern in Romeyka, irrespective of whether it is inherited or borrowed, simplex or derived. Whether there is a difference between inherited and borrowed words in terms of their respective syllable structures remains to be seen.

2.1.2 Phonological processes

There are three notable phonological processes in Romeyka, which I briefly describe in the following sections, namely geminates (see §2.1.2.1), the palatalization of velars (see §2.1.2.2) and the deletion of unstressed [i] (see §2.1.2.3).

2.1.2.1 Geminates

In Ancient Greek (henceforth AG), consonant length was distinctive, e.g. *μέλω* [mélɔː] 'I am of interest' *versus μέλλω* [mélːɔː] 'I am going to'. The distinction has been lost in the standard and most other varieties, with the exception of Cypriot (where it might carry over from AG or arise from a number of synchronic and diachronic assimilatory processes, or even spontaneously). In Romeyka, geminates are attested in both inherited (see (12)a) and borrowed words (see (12)b), yet the consonant length does not seem to be distinctive (further investigation need to be carried out in the future):

(12) Romeyka:

a. χálasːa (< Greek *θάλασσα* [ˈθalasa] 'sea')
'sea'
(S05; 812_0117; 00:17)

b. kalːíon (< AG *κάλλιον* [ˈkalːion] 'better')
'better'
(S01; 812_0822; 01:02)

2.1.2.2 Palatalisation of velars

In Romeyka, it occurs palatalisation of the velar consonants [ɣ] (see (13)a and (13)b), [k] (see (14)a and (14)b), [g] (see (15)a and (15)b and [χ] (see (16)a and (16)b) when they are followed by a [+front] vowel /e/ or /i/. However, only velars in inherited words undergo such a process:

(13) Romeyka:

a. eɣó (< Greek *eɣó* 'I')
'I'
(S01; 812_0822; 01:02)

b. ejéndo (< AG *γίγνομαι* [ˈʝiɣnome] 'become')
'it became'
(S01; 140102_0009; 05:19

(14) Romeyka:

a. néka (< Greek *ʝinéka* 'woman')
'woman'
(S01; 140102_0009; 06:57)

b. stétši (< Greek *stékome* 'I live')
's/he lives'
(S07; 812_0074; 00:39)

(15) Romeyka:

a. egálisen (< Greek *agaʎázo* 'I hug')
's/he hugged'
(S01; 812_0065; 06:46)

b. spudžís (< Greek *sfuʝízo* 'I wipe')
's/he wipes'
(S01; 150702_0020; 00:24)

(16) Romeyka:

a. χálas:a (< Greek *θάλασσα* [ˈθalasa] ‘sea’)
‘sea’
(S05; 812_0117; 00:17)

b. vréši (< Greek *vréçi* ‘it rains’)
‘it rains’
(S01; 150702_0013; 16:30

2.1.2.3 Deletion of unstressed [i]

Unstressed final [i] when occurring in certain nominative nouns, adjectives and certain verb forms is usually deleted (see (17)a and (17)b):

(17) Romeyka:

a. spudžíz < spudžízi (< Greek *sfuɣízo* ‘I wipe’)
‘s/he wipes’
(S01; 150702_0020; 00:24)

b. eksér < ekséri (< Greek *kséro* ‘I know’)
‘s/he knows’
(S01; 150703_0041; 03:11)

2.2 Morphology

2.2.1 Nominal inflection

In Romeyka, nominals are inflected for (a) number (see §2.2.1.1), (b) gender (see §2.2.1.2) and (c) case (see §2.2.1.3). Within the nominal phrase, articles, adjectives and some numerals agree with the noun they modify in gender, case and number, although there is variation especially with regard to gender.

2.2.1.1 Number

In Romeyka, the grammatical number is either singular or plural. The plural is formed by adding a plural suffix, which is selected based on grammatical gender, i.e. neuter nouns inflect with *-æ* e. g. *ta ɣarðélæ* 'the children' (see (20)); and feminises and masculines by *-ðæ*, e.g. *ta patsí-ðæ* 'the girls' (see (19)) and *t andr-úðæ* 'the men' (see (18)) (see Table 6):

Table 6. Number paradigm of Romeyka.

Gender	SG	PL	English translation
M	*o ándr-as*	*t andr-úðæ*	'the man/men'
F	*i patsí-Ø*	*ta patsí-ðæ*	'the girl(s)'
N	*to ɣarðél-in*	*ta ɣarðél-æ*	'the child/children'

(18) Romeyka

a. ándras ates epíren ospítin.
man.NOM she.POSS buy.Past.3SG house.ACC
'Her husband bought a bought.'
(S01; 150702_0019; 09:29)

b. t andrúðæ érθane.
the.NOM men.NOM come.Past.3SG
'The men have arrived.'
(S07; 812_0071; 03:08)

(19) Romeyka:

a. jo, i patsí to mandrín epʰakʰláepsen.
no the.NOM girl.NOM the.NOM stable.NOM clean.Past.3SG
'No, it's the girl that cleaned the stable.'
(S01; 140102_0009; 08:09)

b. tsi θías m ta patsíðæ.
the.GENaunt.GEN I.POSS the.NOM girls.NOM.
'My aunt's daughters.'
(S01; 150702_0019; 09:15)

(20) Romeyka:

a. to ɣarðélin porí na mairévi.
the.NOM child.NOM can.3SG PRT.MOD cook.3SG
'The child can cook.'
(S01; 812_0131; 00:15)

b. úlːa ta ɣarðélæ páne s okʰúlːin.
all.NOM the.NOM children.NOMgo.3PL to.the.ACC school.ACC
'All the children go to school.'
(S01; 150702_0013; 23:19)

2.2.1.2 Gender

In Romeyka, nominals distinguish three grammatical genders: masculine, feminine and neuter, which are formed morphologically. For agreement purposes, the modifier of the noun agrees only if it is [masculine] and [+human] (see (21)). Otherwise, the nominal modifier appears in the neuter gender (see (22)):

(21) Romeyka:

o tranón o ándras érθen.
the.NOM.M strong.NOM.M the.NOM.M man.NOM.M come.Past.3SG
'The strong man came.'
(S07; 812_0071; 02:48)

(22) Romeyka:

etšíno	t	ómorfon	i	patsí
that.NOM.N	the.NOM.N	beautiful.NOM.N	the.NOM.F	girl.NOM.F

so	χoríon	stétši.
in.the.ACC	village.ACC	live.3SG

'That beautiful girl lives in the village.'

(S01; 812_0109; 00:34)

Originally, gender was assigned morphologically, but shifted toward a more semantically oriented assignment based on animacy (see Karatsareas 2014). This development, starting in AG and going further in Romeyka than in MG, includes the spread of neuter forms to masculine and feminine paradigms in both singular and plural declension of nouns, determiners and adjectives (Mackridge 1987, Özkan 2013).

The spread of neuter applies especially to the plural of [–human] feminine nouns and inanimate masculines (see (23)) and it may even extend to [+human] masculine/feminine nouns (see (24)). In [–human] feminine nouns, also mixed declensions occur consisting of a female determiner and neuter adjective (Mackridge 1987) (see (25)):

(23) Romeyka:

pol:á	šíndres	íne.
'many.NOM.N	lizards.NOM.F	be.3PL

'There are many lizards.'

(S07; 812_0044; 13:49)

(24) Romeyka:

to	paléon	o	χoríos.
the.NOM.N	old.NOM.N	the.NOM.M	village.NOM.M

'The old village.'

(S07; 812_0061; 01:44)

(25) Romeyka:

to	tranón	i	káta	éðaksen	me.
the.NOM.N	big.NOM.N	the.NOM.F	cat.NOM.F	bite.Past.3SG	I.ACC

'The big cat bit me.'

(S07; 812_0061; 07:42)

2.2.1.3 Case

Nominative, accusative and genitive cases of nominals are formed morphologically in Romeyka. The accusative is used for both the direct and the indirect object, unlike in other MG varieties in which the genitive is used to mark the indirect object (Mackridge 1987). The genitive expresses nominal determination and possession (Drettas 1999). Table 7 shows the nominal declension in Romeyka:

Table 7. Nominal declension of Romeyka.

Gender	Case	SG	PL	English translation
M	**NOM**	*o ándra-s*	*t andr-úðœ*	'the man/men'
	ACC	*ton ándra-n*	*t andr-úðœ*	
	GEN	*t ándra-Ø*	*t andr-uðion*	
F	**NOM**	*i patsí-Ø*	*ta patsí-ðœ*	'the girl(s)'
	ACC	*tin patsi-n*	*ta patsí-ðœ*	
	GEN	*tsi patsí-Ø*	*ta patsi-ðœ*	
N	**NOM**	*to ɣarðél-in*	*ta ɣarðél-œ*	'the child/children'
	ACC	*to ɣarðél-in*	*ta ɣarðél-œ*	
	GEN	*to ɣarðél-in*	*ta ɣarðel-íon*	

A word-final /n/ originally occurring in the accusative singular of all genders (in nouns of certain declension classes) is subject to sub-variation and may appear in some nouns, e.g. *ton tšírin* 'the father' (Mackridge 1987: 124).

The declension of some masculine nominative singular nouns (those of the second declension; in any case only animates) is sensitive to definiteness: the original *-os* ending becomes *-o(n)* when the noun is definite, e.g. o *škílos*, *o škílon* '(the) dog' (Mackridge 1987: 124; Dawkins 1931: 394, also for other Asia Minor Greek varieties). According to Dawkins

(1931), this phenomenon is caused by the merger of the second and third declension class whereby the nominative of the second declension in *-os* is used for indefinite and the nominative of the third declension in *-o(n)* for definite nouns.

2.2.2 Determiners, pronouns and quantifiers

2.2.2.1 Definite and indefinite articles

Romeyka has three definite articles: the masculine singular article *o*, the feminine singular article *i* and the neuter singular article *to* (see Table 8):

Table 8. Definite articles in Romeyka.

		SG	**PL**
Male	**NOM**	*o*	*i*
	ACC	*ton*	*ts(i)/tus*
	GEN	*ts(i)/tu*	*ts(i)*
Female	**NOM**	*ta*	*ta*
	ACC	*tin*	*ts(i)*
	GEN	*ts(i)*	*ts(i)*
Neuter	**NOM**	*to*	*ta*
	ACC	*to*	*ta*
	GEN	*ts(i)/tu*	*ta*

As regards the structure of a NP, an attributively used adjective preceding the noun requires its own definite article, so the NP contains two articles: before the adjective and the noun (Mackridge 1987). Masculine definite articles may be dropped before nouns with syllable-initial vowels (see (26)a and (26)b) or before Turkish loanwords (see (27)):

(26) Romeyka:

a. alís si mazirán stétši.
Alis.NOM in.the.ACC Mazira.ACC live.3SG
'Alis lives in Mazira.'
(S01; 150703_0040; 07:29)

b. árkon trói ta za.
bear.NOM eat.3SG the.ACCanimals.ACC
'The bear eats the animals.'
(S01; 150703_0040; 15:19)

(27) Romeyka:
fotodžís ésuren ts aišés to fotoráfin.
cameraman.NOM take.Past.3SG the.GENAyşe.GEN the.ACCpicture.ACC
'The cameraman took Ayşe's picture.'
(S01; 812_0062; 10:07)

In general, the definite determiners in Romeyka agree with the head in number, gender and case. However, the spread of the neuter plural declension to [–human] feminine and inanimate masculine nouns and the spread of neuter adjective inflection to adjectives qualifying a [–human] feminine noun, contribute towards a mixed system (Mackridge 1987: 128). Michelioudakis & Sitaridou (2012: 366, Fn. 3) note that in Romeyka, we find both an inherited grammatical gender agreement system and a semantic agreement system, the distribution of the two being conditioned by properties of the head such as animacy and gender. Following Karatsareas (2011), they note that position on the Animacy Hierarchy is relevant, with human behaving distinctly from non-human nouns (see also Karatsareas 2014). The initial article *to* indicating semantic agreement (if neuter is considered to reflect non-humanness), while the article immediately adjacent to the head exhibits syntactic (feminine) agreement. This would also be in line with the prediction that with stacked agreement targets, the target more distant from the controller will show semantic agreement, if any target does (see Karatsareas 2014). Singular definite articles immediately preceding their heads always exhibit syntactic agreement in Romeyka (Michelioudakis & Sitaridou 2013: 366, Fn. 3).

The indefinite article has the form of the numeral 'one' *éna* and inflects for case.

2.2.2.2 Pronominal system

2.2.2.2.1 Personal pronouns

Romeyka has both strong and weak personal pronouns, the latter being enclitically attached to the verb. Enclitic pronouns only occur in nonnominative functions. Strong and weak personal pronouns in Romeyka are presented in Table 9:

Table 9. Subject and object pronouns in Romeyka.

	Subject pronoun	Strong object pronoun	Weak object pronoun (clitic)
1SG	*eɣó*	*emenane*	*-m(e)*
2SG	*esí*	*esenane*	*-s(e)*
3SG	*atós*	*atonane*	*-anæ*
	até	*atenane*	
	ató		
1PL	*emís(t)*	*emasuna*	*-mas*
2PL	*esís(t)*	*esasuna*	*-sas*
3PL	*atiní*	*atinusa*	*-at*
	atiné		
	atinæ		

Object (accusative) pronouns have emerged from AG enclitic personal pronouns (Sitaridou 2014a: 30-31, Özkan 2013: 143) and are either strong or weak whereby, as stated by Michelioudakis & Sitaridou (2012: 219), unlike in other Greek varieties, the use of strong object pronouns seems to be an unmarked option. The third singular object clitic *-æ* seems to be the only third-person form with neutralised gender (Michelioudakis & Sitaridou 2012: 219).

Also *etšínos* 'that' which is actually a demonstrative pronoun varies according to whether they are used in transitive or ditransitive verbs, i.e. as to whether clitic clusters occur (but cf. Michelioudakis & Sitaridou 2012: 237, who argue that there are no clitic clusters in Romeyka, as they question the clitic nature of first and second person object pronouns, the only ones that occur in clusters).

As Michelioudakis & Sitaridou (2012: 218) argue, if two object suffixes occur together, (weak) Person-Case-Constraint (PCC)-like restrictions apply, such that the third person clitic -

æ cannot combine with any other clitic to form a cluster (see (28)a). However, the combination of a first- and a second-person pronoun (strong PCC) is, unlike in MG, acceptable (see (28)b):

(28) Romeyka:

a. *o mehmétis éðotše m æ.
the.NOM Mehmetis.NOM give.Past.3SG I.ACC. it.ACC.CL
'Mehmetis gave it to me.'
(Michelioudakis & Sitaridou 2012: 218)

b. éðiksane m ese. / *eðiksane s eme.
show.Past.3PL I.ACC you.ACC show.Past.3PL you.ACC I.ACC
'They showed me to you.'
(Michelioudakis & Sitaridou 2012: 238)

Furthermore, first- and second-person accusative pronouns cannot be interpreted as direct objects in combination with third-person pronouns irrespective of their order (see (29)):

(29) Romeyka:
éðiksan(e) æ / aton(a) emenan.
show.Past.3PL he.ACC.CL he.ACC I.ACC
'They showed him to me/ *They showed me to him.'
(Michelioudakis & Sitaridou 2012: 238)

Özkan (2013: 146) notes, though, that the third person neuter pronominal suffix can be combined with the neuter definite articles *to* and *ta* used as direct object pronouns after imperative forms like in *ip*é *ato* 'say it' when referring to something previously introduced to the context. For a more detailed discussion of constraints of clitic stacking in Romeyka see Michelioudakis & Sitaridou (2012).

2.2.2.2.2 Demonstrative pronouns

Demonstrative pronouns in Romeyka such as *atós*, *χatós* (see (30)), *χatšós* (see (31)), *avútos* (see (32)) and *etšínos* (see (34)) inflect in general for number, gender and case. Also, the Turkish Trabzon dialectal *havú* 'this' which has arisen due to Greek influence (Brendemoen 2002) can be used as a deictic pronoun as well (see (33)):

(30) Romeyka:

χató	i	néka	i	mánam	en.
this.NOM	the.NOM	woman.NOM	the.NOM	mother.NOM	be.3SG

(S07; 812_0074; 02:07)

(31) Romeyka:

χatšó	esí	les	me.
this.ACC	you.NOM	say.2SG	I.ACC

'You say that to me.'

(S01; 812_0058; 05:37)

(32) Romeyka:

avúton	to	faín	i	miné	epítšen.
this.NOM	the.NOM	food.NOM	the.NOM	Mine.NOM	make.Past.3SG

'Mine made this food.'

(S01; 150703_0042; 01:49)

(33) Romeyka:

havú	t	ospítin	eɣó	éχtisa	to.
this.ACC	the.ACC	house.ACC	I.NOM	build.Past.1SG	it.ACC

'I built this house.'

(S02; 150703_0039; 01:43)

(34) Romeyka:

etšíno	t	ómorfon	i	patsí
that.NOM	the.NOM	beautiful.NOM	the.NOM	girl.NOM
so	χoríon	stétši.		
in.the.ACC	village.ACC	live.3SG		

(S01; 812_0109; 00:34)

The third person neuter pronoun seems also to be used as a demonstrative (see (35)):

(35) Romeyka:

ató	i	néka	
this.NOM.N	the.NOM.F	woman.NOM.F	
i	mána	m	en.
the.NOM.F	mother.NOM.F	I.POSS	be.3SG

(S07; 812_0074; 02:24)

If a demonstrative determines a noun phrase, the definite article is obligatory before the noun.

2.2.2.2.3 Reflexive pronouns

Romeyka uses the Turkish prefix *kendi-* to show that the pronoun is reflexive (Sitaridou 2013: 166) (see (36)):

(36) Romeyka:

a. eɣó tiró to céndi m.
I.NOM look.after.1SG the.ACC self.ACC I.POSS
'I look after myself.'
(S07; 812_0061; 00:53)

b. atós tirí to cédin at.
he.NOM look.after.3SG the.ACC self.ACC he.POSS
'He looks after himself.'
(S07; 812_0061; 00:55)

c. emís tirum ta cédj æmuna.
we.NOM look.after.1PL the.ACC selves.ACC we.POSS
'We look after ourselves.'
(S07; 812_0061; 00:58)

d. esís tiríte ta cédj æsuna.
you.NOM look.after.2PL the.ACC selves.ACC you.POSS
'You look after yourselves.'
(S07; 812_0061; 01:01)

e. ati tirún ta cédin atuna.
they.NOM look.after.3PL the.ACC selves.ACC they.POSS
'They look after themselves.'
(S07; 812_0061; 01:04)

2.2.2.2.4 Possessive pronouns

Possessive pronouns in Romeyka are either enclitics, or independent possessive pronouns (see Table 10), which originate from AG possessives (Mackridge 1987; Bortone 2009; Sitaridou 2014a). The third person possessives derive from demonstratives and therefore inflect (in the singular) in accordance with the gender of the possessor (cf. Drettas 1997).

Table 10. Possessive pronouns and pronominal adjectives in Romeyka.

	Clitic pronoun	**Pronominal adjectives**
1SG	*-m(u)*	*t emón*
2SG	*-s(u)*	*t esón*
3SG	*-(n)at*	*t atinú*
1PL	*-(æ)mun(a)*	*t eméteron*
2PL	*-(æ)sun(a)*	*t eséteron*
3PL	*-(n)atun(a)*	*t atinús*

The initial vowel of the clitic pronouns may change according to the phonological shape of the noun they are attached to. Prepositions indicating direction, such as *s* 'to', fuse with the possessive pronoun, like in *s t eméteron* > *s eméteron to χoríon* 'to our village'.

Another third-person singular possessive suffix used for inanimate objects, animals and babies is *-(e)θe,* which can apparently be combined with the other possessive clitics (see (37)) (see also Sitaridou 2014a: 51-52, Özkan 2013: 146):

(37) Romeyka:

aftí	i	ðulían	to	maθín	emun	-eθe
this.NOM	the.NOM	job.NOM	the.NOM	learning.NOM	we.POSS	-it.POSS
ɣóla	en.					
easy.NOM	be.3SG					

'It is easy for us to learn how to do this job.'
(Sitaridou 2014a: 52)

2.2.2.2.5Relative pronouns

In Romeyka, a frequent relative pronoun is a form derived from the neuter article, *d(o)*/*d=*/*t=*. It seems to be used with inanimate and animate non-human. Other variable relativisers are those derived from the interrogative pronouns, e.g. *opíos* and *ótinan* only used for animates/humans (Schreiber 2018). Furthermore, Sitaridou (2014a: 30) notes the use of *ítina*, an ancient relativiser.

2.2.2.2.6Interrogative pronouns

wh-words in Romeyka are presented in Table 11:

Table 11. *wh*-words in Romeyka (see Michelioudakis & Sitaridou 2016: 6)

***wh*-words**	**Translation**
póte	'when?'
pu meréa / pu tšeka	'where? / which place?'
apó çen / ap éndžeka	'from where?'
kaš cisí	'how many?'
mo tinan	'with whom?'

According to Michelioudakis & Sitaridou (2016: 7), interrogatives in Romeyka have grammaticalised [±human] restrictions but lack number and gender distinctions (see Table 12):

Table 12. Number [+/-human] distinctions on interrogatives in Romeyka (see Michelioudakis & Sitaridou 2016: 7).

Number	**Case**	**[+human]**	**[-human]**
SG	**NOM**	*ts, píos, píon*	*píos, píon*
	ACC	*tínan*	*do, dóɣna*
	GEN	*tínos*	*tínos*
PL	**NOM**	*kaš cʰiší*	*píos*
	ACC	*tínan*	*do, dóɣna*
	GEN	*tínos*	*píon*

Romeyka has borrowed the Turkish interrogative particle *mI*, used in yes/no questions, albeit without vowel harmony (see (38)):

(38) Romeyka:

esí	to	mátšin	epíes	mi?
you.NOM	the.ACC	match.ACC	go.Past.2SG	PRT.Q

'Did you go to the football match?'

(S04; 150702_0018; 00:03)

Like in Turkish, the question particle is flexible to move for focalization purposes. Schreiber (2018: 910) also notes that the *wh*-word *láγa* 'how' is attested *in-situ* (see (39)). However, this is not *in-situ*; it is rather indirect speech.

(39) Romeyka of Çaykara

t	esón	i	patsí	láγa	en,
the.NOM	you.POSS	the.NOM	daughter.NOM	how	be.3SG

émorfo	mi	en?
beautiful.NOM	PRT.Q	be.3SG

'How is your daughter, is she beautiful?'

(Schreiber 2018: 910)

2.2.2.2.7 Quantifiers and numerals

In Romeyka, quantifiers are at least *úlos* 'all'. *kat* 'some', *kanis* 'nobody', *tipu* 'nothing'. However, further investigation is needed to be undertaken in the future.

Romeyka numerals exist only for the numbers 'one' to 'five', the rest is Turkish (Mackridge 1987). According to Schreiber (2018: 911), pronominal clitics can be attached to numerals, though the personal suffix resembles the Turkish first plural suffix *-Iz* (see (40):

(40) Romeyka of Çaykara:

na	tróγume	i	ði-jiz.
PRT.FUT	eat.1PL	the	two-1PL

'The two of us will eat.'

(Schreiber 2018: 911)

2.2.3 Verbal inflection

Romeyka finite verbs are inflected for person and number (see §2.2.3.1), tense (see §2.2.3.2), aspect (see §2.2.3.3), voice (see §2.2.3.4) and mood (see §2.2.3.5). Some grammatical functions are expressed periphrastically by the use of particles, such as the modal particle *na* and others by enclitic affixes and/or stem alternation. The verbal paradigm of Romeyka contains many archaisms absent in other Modern Greek varieties such as the infinitive (see Mackridge 1987; Sitaridou 2014a, 2014b).

2.2.3.1 Person and number inflection

Romeyka verbs are specified for person and number, realised through personal suffixes agreeing with the subject. The clitic pronouns for objects form a distinct paradigm; whether they should be considered "agreement" is an issue we do not take up here. In what follows, I refer exclusively to the subject-indexing suffixes as agreement. The paradigms vary according to verb stems. Some verbs are strong, i.e. they exhibit different stems with regard to tense.

The citation form used here is the first-person singular present. Verbs can be classified according to the endings in this form: verbs ending in *-o*, e. g. *léɣo* 'say' and verbs ending in *-me*, e. g. *tšimúme* 'sleep' (see Table 13). Within these classes, verbs can be subcategorised according to ultimate and penultimate stress, the latter verb class forming the major group (Drettas 1997: 205).

Table 13. Present tense inflectional paradigm in Romeyka.

1SG	*léɣ-o*	*tšim-úme*
2SG	*léj-is*	*tšim-áse*
3SG	*léj-i*	*tšim-áte*
1PL	*léɣ-umen*	*tšim-úmasten*
2PL	*léj-ete*	*tšim-úsaste*
3PL	*léɣ-une*	*tšim-únde*

2.2.3.2 Tense

Romeyka has two morphologically marked tenses: [-past] (or present) tense and [+past] (or past) tense. The latter, but not the former, distinguishes two aspects, i.e. [-perfective] (imperfect) and [+perfective] (aorist).

2.2.3.2.1 Present

The [-past] (or present) tense expresses a [-perfective] aspect. It is made up of the [-perfective] stem and the person and number suffixes. Table 14 shows the verbal paradigm of present tense:

Table 14. Present tense endings in Romeyka.

Person & Number		
1SG	*léɣ-o*	*tšim-úme*
2SG	*léj-is*	*tšim-áse*
3SG	*léj-i*	*tšim-áte*
1PL	*léɣ-umen*	*tšim-úmasten*
2PL	*léj-ete*	*tšim-úsaste*
3PL	*léɣ-une*	*tšim-únde*

2.2.3.2.2 Imperfect

The imperfect expresses an action which happened continuously or habitually in the past. Like the aorist, it is formed by the vocalic augment /e/ (Bortone 2009), like in *éleɣa* ‘I was saying’. As in the other tenses, the inflectional paradigm depends on the verb class (see Table 15):

Table 15. Imperfect endings in Romeyka.

Person & Number		
1SG	*é-leɣ-a*	*e-tšim-úmun*
2SG	*é-lej-es*	*e-tšim-ásun*
3SG	*é-lej-em*	*e-tšim-átun*
1PL	*e-léɣ-amen*	*e-tšim-úmasten*
2PL	*e-léj-ete*	*e-tšim-úsaste*
3PL	*é-leɣ-an*	*e-tšim-úndane*

2.2.3.2.3 Aorist

The aorist is used for perfective actions in the past. It is formed on the basis of perfective aspect, i.e. the aorist stem with the ancient temporal augment /e/ (see Table 16) (Sitaridou 2014a: 53):

Table 16. Aorist endings in Romeyka.

Person & Number		
1SG	*íp-a*	*e-tšim-éθa*
2SG	*íp-es*	*e-tšim-éθes*
3SG	*íp-en*	*e-tšim-éθe*
1PL	*íp-amen*	*e-tšim-éθame*
2PL	*íp-ete*	*e-tšim-éθete*
3PL	*íp-ane*	*e-tšim-éθane*

2.2.3.2.4 Future

There is no morphological future tense in Romeyka. Future is expressed periphrastically by means of the modal particle *na* (see (41)) and the present stem, like in other AMG varieties (see (42)) and partially Cypriot Greek (see (43)b) and unlike in MG (see (44)), which uses the particle *θa* (see Sitaridou 2014b: 122):

(41) Romeyka:

i	néka	na	pʰakʰlaévi	t	ospítin.
the.NOM	woman.NOM	PRT.FUT	clean.3SG	the.ACC	house.ACC

'The woman will clean the house.'
(S01; 812_0057; 00:16)

(42) Cappadocian Greek; Delmesó:

νά	το	πάρω.
na	to	páro
PRT.FUT	it.ACC	marry.1SG

'I will marry her.'
(Dawkins 1916: 304)

(43) Cypriot Greek:

a. θa páo.
PRT.FUT go.PNP.1SG
'I will go.'

b. en na páo.
be.3SG PRT.MOD go.PNP.1SG
'I will go.'

(44) Modern Greek:

θa páo.
PRT.FUT go.PNP.1SG
'I will go.'

2.2.3.3 Aspect

Perfective and imperfective aspect are in Romeyka only realised in the past tense indicative, i.e. by the aorist and the imperfect respectively (Mackridge 1987; Sitaridou 2014b). Present and future have no morphological aspectual distinctions.

2.2.3.4 Voice

Historically, the categorisation of verbs in two classes approximately corresponds to a voice distinction: verbs in *-o* are called actives whereas verbs in *-me* are called (medio)passives (Drettas 1997: 205). Whereas Romeyka of Çaykara seems to lack passives in general, Romeyka of Sürmene allows passives (Michelioudakis & Sitaridou 2012: 236). The syntax of datives under passivisation in Romeyka of Sürmene has been investigated by Michelioudakis & Sitaridou (2012). They note that a theme argument can be a regular subject of a passive verb (see (45) and (46)a), while a benefactive or recipient cannot advance to subject under passivisation (see (46)b):

(45) Romeyka of Sürmene:

to	χartí	eɣráfte	tin	aišé.
the.NOM	letter.NOM	be.written.Past.3SG	the.ACC	Ayşe.ACC

'The letter was written for Ayşe.'

(Michelioudakis & Sitaridou 2012: 236)

(46) Romeyka of Sürmene:

a.

i	pará	tin	aišé	eðóste.
the.NOM	money.NOM	the.ACC	Ayşe.ACC	be.given.Past.3SG

'The money was given to Ayşe.'

b.

*i	aišé	eðóste	tin	parán.
the.NOM	Ayşe.NOM	be.given.Past.3SG	the.ACC	money.ACC

'Ayşe was given the money.'

(Michelioudakis & Sitaridou 2012: 236)

2.2.3.5 Mood

In Romeyka, there are attested four moods; namely, indicative, imperative, subjunctive and optative. Indicative and subjunctive are formed morphologically, whereas subjunctive and optative are formed periphrastically. In particular, subjunctive is formed by a *na*-clause, while optative is formed by either a *na*-clause or an *as*-clause. Both *na*- and *as*-clauses are composed of the particle *na* or *as* and a finite verb.

The imperative covers the second persons and differs according to verb class: verbs ending in *-o* such as *tróɣo* 'eat', *fa* 'eat.IMP.2SG', *fáte* 'eat.IMP.2PL'; verbs in *-me* such as *tšimúme* 'sleep', *tšiméθ* 'sleep.IMP.2SG', *tšiméθisten* 'sleep.IMP.2PL' (see Drettas 1997: 227–232). Negation is formed by the negation particle *mi*. Some verbs in Romeyka retain the ancient imperative ending *(-s)on*, e. g. *ákuso(n)* 'listen' (Bortone 2009: 84), but also other archaic imperatives such as *ipe* 'say'(Mackridge 1987: 125).

Table 17. Imperative in Romeyka (see Sitaridou 2014b: 121, 129).

Person & Number		
1SG	-	-
2SG	*fa*	*tšiméθ*
3SG	-	-
1PL	-	-
2PL	*fáte*	*tšiméθisten*
3PL	-	-

2.2.3.6 The copula verb

The copula verb links the subject of a clause to a subject (*aka* predicative) complement. In the case of Romeyka, the verb *íme* 'I am' is the main copula verb. It always proceeds the predicative complement of the clause. Table 18 shows the inflection of the copula verb *íme* 'I am' in Romeyka (see Sitaridou & Kaltsa 2014: 19):

Table 18. Inflection of the copula verb *íme* 'I am' in Romeyka.

Person & Number	Present	Imperfect
1SG	*íme*	*ímun*
2SG	*íse*	*ísun*
3SG	*íne, ne, en*	*étun*
1PL	*ímaste*	*ímaste*
2PL	*ísaste*	*ísaste*
3PL	*íne, ne, en*	*étun*

(47) Romeyka:

a. eɣó hastás íme.
I.NOM ill.NOM be.1SG
'I am ill.'
(S07; 812_0004; 05:34)

b. esí íse tranós.
you.NOM be.2SG strong.NOM
'You are strong.'
(S01; 150702_0013; 14:16)

c. ksénos en.
foreigner.NOM be.3SG
'He is a foreigner.'
(S01; 812_0056; 01:17)

d. i mahal:ímena kalés:a en.
the.NOM female.teacher.NOM good.NOM be.3SG
'The (female) teacher is nice.'
(S01; 812_0062; 05:53)

e. pol:á arapáðes áspra pal íne kótšina pal íne.
many.NOM cars.NOM white.NOM PRT be.3PL red.NOM PRT be.3PL
'Many cars are both white and red.'
(S07; 812_0044; 19:13)

(48) Romeyka:

a. ánda ímun mikrés:a.
while be.IMPF.1SG young.NOM
'While I was young.'
(S07; 812_0048; 00:01)

b. to faín émnoston éton.
the.NOM food.NOM delicious.NOM be.IMPF.3SG
'The food was delicious.'
(S01; 150702_0013; 13:01)

Interestingly, the third person of the copula verb can be omitted. I suppose that this might be a contact-induced borrowing from Turkish (see (49)):

(49) Romeyka:

t árk^{h}u ta níšæ traná.
the.GENbear.GEN the.NOM nails.NOM sharp.NOM
'The nails of the bear are sharp.'
(S07; 812_0044; 05:40)

2.2.3.7 Auxiliary verbs and verbal constructions

There are four auxiliary verbs forming verbal constructions in Romeyka: (a) the verb *to be*, either in present or imperfect tense (see §2.2.3.7.1), (b) the verb *to have*, only in imperfect tense (see §2.2.3.7.2), (c) the verb *to have* in the third person, singular number, present tense and (d) the verbs *stand* or *sit* (see §2.2.3.7.4).

2.2.3.7.1 be + participle

The first verbal construction is made up of the auxiliary verb *íme* 'I am' in the present and imperfect tense and the passive perfect participle. This verbal construction is equivalent to the passive present and passive perfect tenses in MG. Note that the auxiliary verb always proceeds the participle. Table 19 shows the inflection of the auxiliary verb *íme* 'I am' in Romeyka:

Table 19. Inflection of the auxiliary verb *íme* 'I am' in Romeyka.

Person & Number	Present	Imperfect
1SG	*íme*	*ímun*
2SG	*íse*	*ísun*
3SG	*íne, ne, en*	*étun*
1PL	*ímaste*	*ímaste*
2PL	*ísaste*	*ísaste*
3PL	*íne, ne, en*	*étun*

(50) Romeyka:

a. aɣanaχtiménis:a íme.
tired.PART be.1SG
'I am tired.'
(S01; 812_0056; 02:58)

b. aɣanaχtiménis:a étone.
tired.PART be.IMPF.3SG
'She was tired.'
(S01; 150703_0041; 01:57)

2.2.3.7.2 had + infinitive

The second verbal construction is made up of the auxiliary verb *íχa* 'I had' only in the imperfect tense and the infinitive (see Sitaridou 2014a). This verbal construction is attested in counterfactuals such as wishes and exclamatives (see (51)) and conditionals as a complement of *íχa* 'I had' (see (52) and (53)). Note that the auxiliary verb always precedes the infinitive. Table 20 shows the inflection of the auxiliary verb *íχa* 'I had' in Romeyka:

Table 20. Inflection of the auxiliary verb *íχa* 'I had' in Romeyka.

Person & Number	Imperfect
1SG	*íχa*
2SG	*íšes*
3SG	*íšen*
1PL	*íχamen*
2PL	*íχate*
3PL	*íχan*

(51) Romeyka:

a. as íšen porpatesíni sa rašíæ!
PRT.OPT have.IMPF.3SG walk.INF to.the.ACC mountains.ACC
'S/He should have walked in the mountains.'

b. na íχame panini
PRT.MOD have.IMPF.3PL go.INF
χtisíni t ospít so parχár!
build.INF the.ACC house.ACC in.the.ACC pastures.ACC
'I wish we had gone to build the house in the highland pastures.'

(Sitaridou 2014b: 136)

(52) Romeyka:

as íχa elíɣon faníni,
PRT have.IMPF.1SG less eat.INF
χar ómorfa n ésteka.
now well PRT.MOD be.IMPF.1SG
'If only I had eaten less, I would have been well now.'
(S01; 812_0123; 01:34)

(53) Romeyka:

na íχa šíta piterupsíni tas ðulíæs,
PRT.MOD have.IMPF.1SG immediately finish.INF the.ACCchores.ACC
χar n epínamen parakáθin.
now PRT.MOD make.IMPF.1PL gathering.ACC
'If I had finished the chores immediately, we would have been gathered together now.'
(S01; 812_0123; 02:52)

2.2.3.7.3 have + finite verb

According to Özkan (2013) and Schreiber (2018), progressive aspect seems to be realised by two strategies: (a) by the invariable auxiliary *eš(i)* + a finite verb form of the present stem for present (see (54)a) and in the imperfect for past. According to Schreiber (2018: 915), Istanbulite Romeyka applies the aorist form instead (see (54)b). Invariable *eš(i)* resembles the 3SG form of the verb *eš* 'to have' and could therefore be a grammaticalised form of 'have' (see also Drettas 1997: 334).

However, it is not clear whether we are dealing with the verb *eš(i)* or the particle *as*, which seems more plausible, since the complement of the verb *eš(i)* is suggested that is a finite verb. Further investigation of this phenomenon is needed.

(54) Romeyka of Çaykara:

a. asó pazár eš érχume.
from.the.ACC market.ACC have.3SG come.1SG
'I am coming from the market.'

b. i aišé asó istanból eš érte.
the.NOM Ayşe.NOM from.the.ACC Istanbul.ACC be.3SG come.Past.3SG
'Ayşe has been coming from Istanbul.'

(Schreiber 2018: 915)

In Pontic Greek, *eš* + *k* 'and' (*tše* in Romeyka) + present/imperfect is used to express explicit processes which are near to completion (Drettas 1997).

2.2.3.7.4 finite verb + coordinate + stand or sit

Another periphrastic construction consists of a finite verb + coordinative *tše* 'and' + finite *stéko*/*stékume* 'stand' or *káχome* 'sit'. Both auxiliaries seem to be used interchangeably whereby *stéko* and *stékume* seem to constitute a mixed paradigm in themselves (see Table 21):

Table 21. Periphrastic progressive forms with *káχome* 'sit' and *stéko*/*stékume* 'stand' on the example of *tšimúme* 'sleep' (Schreiber 2018: 915).

	káχome 'sit'	*stéko*/*stékume* 'stand'
	'sleep' + 'and' + 'sit'	'sleep' + 'and' + 'stand'
1SG	*tšimúme tše káχome*	*tšimúme tše stéko*/*stékume*
2SG	*tšimáse tše káçese*	?
3SG	*tšimáte tše káçete*	*tšimáte tše stitš*
1PL	*tšimúmist tše káχomist*	?
2PL	*tšimáste tše káçesten*	?
3PL	*tšimún tše káχontane*	*tšimún tše stékontane*/*stékune*

(55) Romeyka of Çaykara:

a. to vacít érθe (tše stitš)
the.NOM time.NOM come.Past.3SG and stand.1SG
n eftá to fají.
PRT.FUT make.1SG the.ACC food.ACC
'The time has come, I will prepare the food.'
(Schreiber 2018: 916)

b. so mutfák tšališévo tše stéko.
in.the.ACC kitchen.ACC work.1SG and stand.1SG
'I am working in the kitchen.'
(Schreiber 2018: 916)

According to Sitaridou (2014a: 44), the construction with *stéko* bears inchoative aspect, though without a progressive function (cf. Drettas 1997: 336).

2.2.3.8 Nonfinite verbal forms

In Romeyka, three nonfinite verbal forms are attested, namely (a) gerund (see §2.2.3.8.1), (b) infinitive (see §2.2.3.8.2) and (c) participle (see §2.2.3.8.3).

2.2.3.8.1Gerund

As for the gerund, although Sitaridou (2014b: 121) suggests that there appears to be no gerund ending in *-ondas* in Romeyka, however adverbial past participles formed by the suffix *-ta* appended to the perfective stem are attested. This is an uninflected verb form used adverbially. Its subject is normally identical with the subject of the clause in which it occurs. This gerund-like participle expresses manner (see (56)):

(56) Romeyka:

porpateftá na pas so chichénin.
walking.GER PRT.MOD go.2SG to.the.ACC grocery.ACC
'You should get to the grocery on foot.'
(S01; 150702_0013; 11:17)

Schreiber (2018: 917) mentions that Özkan provides examples from the Romeyka spoken in Sürmene that are similar to the ones found in Romeyka of 'Anasta'; that is, they are adverbial past participles which are formed by the suffix *-ta* appended to the perfective stem (see (57)):

(57) Romeyka of Sürmene:
ɣelaχtá érθame.
lauging.GER come.Past.1PL
'We came laughing'.
(Scheiber 2018: 917)

2.2.3.8.2 Infinitive

In regard to the infinitive, although Mackridge (1987: 127, Fn. 17) considers that Romeyka lacks nonfinite forms of the verb, based on the assumption that finiteness equates to 'indexing person', however the infinitive found in certain varieties of Romeyka seems to run counter to this (see Sitaridou 2014a, 2014b).

In Romeyka, the infinitive consists of the aorist stem followed by the infinitival ending *-ini*, which can also bear the passive stem marker *-θ-*. The infinitive may have a distinct nominative subject, such as in inflected and personal infinitives (see Sitaridou 2014a). Furthermore, it can take either a predicate or a DP as a complement, can be coordinated and modified by adverbs (Sitaridou 2014b: 130).

According to Sitaridou (2014b), in Romeyka infinitives are attested (a) as complements to negated past tense modal verbs (Sitaridou 2014b: 126) (see (58)); (b) as complements to the negated past tense volitional verb *utš eθélesa* 'I did not want' (Sitaridou 2014b: 126) (see (59)); and (c) in *prin*-clauses ('*before*-clauses') (see (60)):

(58) Romeyka:
tš epórese mairepsíni.
NEG can.Past.3SG cook.INF
'S/he couldn't cook.'
(S07; 812_0004; 04:06)

(59) Romeyka:

utš eθélesa mairepsíni.

NEG want.Past.1SG cook.INF

'I didn't want to cook.'

(S01; 150702_0032; 01:28)

(60) Romeyka:

i aišé éplinen ta χapsíæ

the.NOM Ayşe.NOM wash.Past.3SG the.ACC anchovies.ACC

prin mairepsíni.

before cook.INF

'Ayşe had washed the anchovies before she cooked them.'

(S01; 812_0113; 17:24)

2.2.3.8.3 Participle

The passive perfect participle ends in *-menos* and is fully inflected for number, gender and case. It is formed from verbs, although not all verbs have this form. This participle is made up of the passive perfective stem and the suffix *-menos*. The passive perfect participle functions as an adjective, agreeing in gender, number and case with the noun it modifies (see (61)a and (61)b):

(61) Romeyka:

a. íða énan cʰytʰýphane

see.Past.1SG a.ACC bookshop.ACC

ɣomáton, fortoménon cʰitʰápæ.

full.ACC carrying.PART.ACC books.ACC

'I saw a bookshop being full of books.'

(S01; 812_0062; 09:11)

b. jedí tané χaména ɣarðél:æ éχo.

seven CLF lost.PART.ACC children.ACC have.1SG

'I have lost seven children (Seven of my children have died).'

(S07; 812_0048; 04:12)

2.3 Syntax

2.3.1 Syntax of the determiner phrase

Determine phrases (henceforth DPs) in Romeyka are head-initial (see Guardiano et al. 2016) (see (62)):

(62) Romeyka:

to	paléon	o	χoríos.
the.NOM	old.NOM	the.NOM	village.NOM

'The old village.'
(S01; 812_0061; 08:49)

The definite article co-occurs with all nouns, including proper nouns and co-occurs with both demonstratives and possessives (see (63)). Quantifiers are generally treated like adjectives and thus require their own article (see (64)):

(63) Romeyka:

to	tranón	i	káta	éðaksen	me.
the.NOM.N	big.NOM.N	the.NOM.F	cat.NOM.F	bite.Past.3SG	I.ACC

'The big cat bit me.'
(S01; 812_0061; 07:42)

(64) Romeyka:

úlːon	to	χoríos	epíen	so	džamín.
all.NOM	the.NOM	village.NOM	go.Past.3SG	to.the.ACC	mosque.ACC

'All the (people of the) village went to the mosque.'
(S01; 150702_0013; 19:56)

Adjectives in Romeyka precede the noun they modify and generally agree with their heads in person and gender, though there may also be semantic agreement (Michelioudakis & Sitaridou 2013: 366, Fn. 3). In contrast, Özkan (2013: 145) suggests that syntactic agreement is not consistent with adnominal adjectives (see (65)), while predicatively used adjectives exhibit semantic agreement if the noun denotes a human referent (at least this is how his example can

be interpreted). There is obviously dialectal variation here, but the general tendency for the neuter form to expand its range appears to be valid here (cf. Karatsareas 2011, 2014).

(65) Romeyka:

etšíno	t	ómorfon	i	patsí
that.NOM	the.NOM	beautiful.NOM	the.NOM	girl.NOM
so	χoríon	stétši.		
in.the.ACC	village.ACC	live.3SG		

'That beautiful girl lives in the village.;
(S01; 812_0109; 00:34)

Furthermore, a feminine animate noun may be modified by one neuter and one masculine adjective (Mackridge 1987: 128).

2.3.2 Negation and modality

Romeyka expresses negation by the use of negators with rich allomorphic and cross-dialectal variation, whereby the allomorphic distribution of the negators correspond to that of AG (Sitaridou 2014b). The following four negators can be distinguished: (a) *utš* for sentential negation in indicative sentences, (b) *mi(n)* and *me(n)* in imperatives and wishes, (c) *ç* and *tš* in subjunctives and future, (d) *mutš* in counterfactuals (Sitaridou 2014b: 121).

utš is derived from the ancient preverbal negative particle *oύκ*. Its phonologically conditioned allomorphic forms are presented in Table 22:

Table 22. Negators in Romeyka (Sitaridou 2014b: 121).

Negator		Conditions	Example
utš	*utš*	phonologically conditioned variation (Sitaridou 2014b: 121)	(see (66))
	tš		(see (67))
	tše		(see (68))
	tši		(see (69))
	u		(see (70))
	utše		(see (71))
	utši		(see (72))
mi(n)	*me(n)*	imperatives; wishes	(see (73) and (74))
ç	*tš*	subjunctives; future	(see (75), (76), (77), (78) and (79))
	utš		(see (80))
mutš		counterfactuals	(see (75))

(66) Romeyka:
χar utš enístaksa.
now NEG be.sleepy.Past.1SG
'I am not sleepy now.'
(S07; 812_0004; 01:47)

(67) Romeyka:
tš epórena n eporpátena.
NEG can.IMPF.1SG PRT.MOD work.IMPF.1SG
'I couldn't walk.'
(S07; 812_0004; 05:27

(68) Romeyka:
tše ɣrikó.
NEG understand.1SG
'I don't understand.'
(S01; 812_0113; 02:18)

(69) Romeyka:

tši pámen s okʰúlːin.

NEG go.1PL to.the.ACC school.ACC

‘We are not going to school (for the time being).’

(S03; 150702_0015; 09:14)

(70) Romeyka:

u fanerúndan

NEG appear.3PL

‘They are invisible.’

(S07; 812_0044; 23:13)

(71) Romeyka:

eɣó polːá utše konušéfko rúmdža.

I.NOM much NEG speak.1SG Romeyka.ACC

‘I don’t speak Romeyka much.’

(S03; 150702_0015; 02:00)

(72) Romeyka:

atós utši kratí pʰáokʰ.

he.NOM NEG support.3SG Paok.ACC

atós pʰanaθináikʰos kratí.

he.NOM Panathinaikos.ACC support.3SG

‘He doesn’t support Paok. He supports Panathinaikos.’

(S03; 150702_0018; 00:15)

(73) Romeyka:

mi pas.

NEG go.2SG

‘Do not go.’

(S07; 812_0048; 10:55)

(74) Romeyka:

fíete	tše	me	sindišén:ete.
go.IMP.2PL	and	NEG	talk.IMP.2PL

'Go away and don't talk aloud.'

(S01; 150702_0023; 15:50)

(75) Romeyka:

na	mutš	íχa	χasíni	ton	pʰarán,
PRT.MOD	NEG	have.IMPF.1SG	lose.INF	the.ACC	money.ACC
χar	n	éχtiza	ospítin.		
now	PRT.MOD	build.IMPF.1SG	house.ACC		

'If I hadn't lost the money, I would have built a house now.'

(S01; 812_0123; 00:18)

Interestingly, the verb in a *na*-clause is never attested negated. If the proposition of the *na*-clause is negated, then the verb *en* 'be.3SG' negated is used before the particle *na* (see (76), (77) and (78)):

(76) Romeyka:

θaró,	alís	ç	en	na	faízi	mas.
think.1SG	Alis.NOM	NEG	be.3SG	PRT.MOD	eat.PNP.3SG	we.ACC

'I think that Alis would not feed us.'

(S07; 812_0012; 00:04)

(77) Romeyka:

mehmétis	os	na	rti,
Mehmetis.NOM	until	PRT.MOD	come.PNP.3SG
ç	en	na	mairévo.
NEG	be.3SG	PRT.MOD	cook.1SG

'I won't cook anything until Mehmetis comes home.'

(S07; 812_0012; 01:10)

(78) Romeyka:

dohtóris	ípen	me,	ç	en	n	eftǽis	ðulíæs.
doctor.NOM	say.Past.3SG	I.ACC	NEG	be.3SG	PRT.MOD	do.2SG	chores.ACC

'The doctor told me not to do any chores.'

(S01; 812_0029; 00:03)

The same holds for the future; that is, the *na*-clause is negated by the verb *en* 'be.3SG' negated (see (79) and (80)):

(79) Romeyka:

tš	en	na	faízi	mas.
NEG	be.3SG	PRT.MOD	eat.PNP.3SG	we.ACC

'He won't feed us.'

(S07; 812_0012; 01:12)

(80) Romeyka:

utš	en	na	inanéfcese.
NEG	be.3SG	PRT.MOD	believe.2SG

'You won't believe it.'

(S01; 150702_0023; 19:51)

Interestingly, this phenomenon is also attested in the Ophitic spoken by Christians as is illustrated in the following examples (see (81) and (82)):

(81) Ophitic (Nea Trapezounta, Greece):

a. να γράφω.
na ɣráfo.
PRT.MOD write.1SG
'I will write.'
(Revythiadou & Spyropoulos 2009: 56)

b. (ουτσ̆εν) να γράφω
utš en na ɣráfo.
NEG be.3SG PRT.MOD write.3SG
'I will not write.'
(Revythiadou & Spyropoulos 2009: 56)

(82) Ophitic (Asia Minor):

καὶ ᾽ς σὴν ἐγκλεσία οὐκ᾽ ἔν᾽ νὰ πάνε.
ce s sin eŋglesía uc en na páne.
and to the.ACCchurch.ACC NEG be.3SG PRT.MOD go.PNP.3PL
'And they will not go to the church.'
(Dawkins 1931: 106)

It also occurs in Cappadocian Greek (see (83)):

(83) Cappadocian (Ulaghátsh):

a. Ἔγερ ὄν γενήῃς, ἕνα παιί [νε],
éjer on jeníis, éna peí ne,
if when give.birth.PNP.2SG a.NOM boy.NOM be.3SG
να ἔρτουμ dé νε.
na értum dé ne.
PRT.MOD come.PNP.1SG NEG be.3SG
'If, when you have a child, it is a boy, we would not come.'
(Dawkins 1916: 348)

b. Ἔγερ tüφέκ, να ἔρτουμ dé νε.
éjer tyfék, na értum dé ne.
if gun.NOM PRT.MOD come.PNP.1PL NEG be.3SG
'If [the sign be given with] a gun, we would not come.'
(Dawkins 1916: 348)

c. Ἀπαπέρα τράνσαν κι do tüφέκ,
apapéra tránsan ci do tyfék
from.over.there see.Past.3PL and the.NOM gun.NOM
να ἔρτουν dé νε.
na értun dé ne.
PRT.MOD come.PNP.3PL NEG be.3SG
'From over there they saw that [the sign be given with] a gun, for them not to come.'
(Dawkins 1916: 348)

Finally, it also appears in Cypriot Greek (see (84)):

(84) Cypriot Greek:

a. en na páo.
be.3SG PRT.MOD go.PNP.1SG
'I will go.'

b. endž en na páo.
NEG be.3SG PRT.MOD go.PNP.1SG
'I will not go.'

This is opposed to the MG future negation (see (85)):

(85) Modern Greek:

a. θa páo.
PRT.FUT go.PNP.1SG
'I will go.'

b. ðe θa páo.
NEG PRT.FUT go.PNP.1SG
'I will not go.'

Interestingly, Cypriot Greek exhibits both MG and AMG strategies to negate future tense (see (86)a and (86)b):

(86) Cypriot Greek:

a. en θa páo.
NEG PRT.FUT go.PNP.1SG
'I will not go.'

b. endž en na páo.
NEG be.3SG PRT.MOD go.PNP.1SG
'I will not go.'

For Romeyka of Sürmene as spoken in Beşköy, Özkan (2013: 147) also mentions a preverbal particle *ha* used to form future in negated sentences and interrogatives.

Based on the data presented above, the future particle na in AMG must derive from the verb 'to be', whereas the future particle *θa* in MG derives from the phrase *θélo ína* 'I want to', indicating a major difference between the development of future tense in the two language groups. Nevertheless, Cypriot Greek exhibits both strategies. However, the investigation of this phenomenon remains open for future investigation.

2.3.3 Subordinate clauses

Romeyka exhibits several different strategies for complementation, including finite and non-finite verbal forms (Sitaridou 2014b). The only complementiser in regular use is the all-purpose modal particle *na* (see §2.3.3.1.3). In what follows, a broad distinction is drawn between finite and non-finite subordination distinguishing several sub-types.

2.3.3.1 Finite subordinate clauses

2.3.3.1.1 Relative clauses

Relativisers in Romeyka are in general preverbal and different strategies of relativisation seem to apply according to the role of the head noun of the relative clause. Invariable relativisers are the following, although relative clauses without relativiser are possible, too (see Gandon 2016).

First, *op* (possibly > AG *ὅπου*, Gandon 2016: 221) occurs in prenominal relative clauses, immediately preceding the relative clause verb. It may be used for relativising the subject of the relative clause and can be used as a free relative (see (87)):

(87) Romeyka:

o	peðás	[op	érθen	asó	cic^{h}énin]
the.NOM	child.NOM	REL	come.Past.3SG	from.the.ACC	grocery's.ACC

t	emón	t	anépsin	en.
the	mine	the.NOM	nephew.NOM	be.3SG

'The child who came from the grocery's is my nephew'

(S01; 812_0822; 03:54)

Gandon (2016: 223, Fn. 130) notes the possibly nominalising function of the free relativisers.

Second, though likely, it is not fully clear whether *p(i)/p(u)* (Özkan), *pe* (> AG *ὅπερ*, Sitaridou 2014a: 30) derives from the same root as *op*; a clitic *p* is also found in relative clauses:

(88) Romeyka:

opsé	íða	alís	p	epíren	inéka.
yesterday	see.Past.1SG	Alis.NOM	REL	marry.Past.3SG	wife.ACC

'Yesterday I saw the woman that Alis married.'

(S01; 812_0058; 01:42)

Romeyka uses a typologically uncommon strategy for relativisation, involving prenominal, but finite relative clauses (Gandon 2016: 220) (see (89)). Right-branching structures (postnominal relative clauses) may be possible, too (see (90)), presumably a reflex of the type widely-attested in AG and still maintained in MG:

(89) Romeyka:

i	patsí	[así	mazirán	d	érθen]
the.NOM	girl.NOM	from.the.ACC	Mazira.ACC	REL	come.Past.3SG
t	emón	t	anépsin	éton.	
the.NOM	I.POSS.NOM	the.NOM	niece.NOM	be.IMPF.3SG	

'The girl who came from Mazira is my niece.'
(S01; 812_0112; 06:01)

(90) Romeyka:

[así	mazirán	d	érθen]	i	patsí
from.the.ACC	Mazira.ACC	REL	come.Past.3SG	the.NOM	girl.NOM
t	emón	t	anépsin	éton.	
the.NOM	I.POSS.NOM	the.NOM	niece.NOM	be.IMPF.3SG	

'The girl who came from Mazira is my niece.'
(S01; 812_0112; 06:06)

2.3.3.1.2 Finite complement clauses in verbs of saying

Subordinate clauses introduced by verbs of saying have a finite verb form in the dependent clause which is either centre-embedded as in relatives and embedded imperatives (see (91)), or post-posed (see (92)). Note that all the available examples could be interpreted as direct speech (cf. the second-person pronouns in the subordinate clause); to what extent direct and indirect speech are differentiated syntactically remains an open question. With verbs of saying, in general no complementisers are used (Sitaridou 2014b: 128).

(91) Romeyka:

a. mohal:ímis léi, i aišé edótšen.
teacher.NOM say.3SG the.NOM Ayşe.NOM give.Past.3SG
'The teacher said that Ayşe gave (something to someone).'
(S07; 812_0065; 04:22)

b. i mána s ípen,
the.NOM mother.NOM you.POSS say.Past.3SG
o tšíri s ípen, ékleps aten.
the.NOM father.NOM you.POSS say.Past.3SG steal.Past.3SG she.ACC
'Your mother said that your father said that (someone) stole (someone).'
(S07; 812_0065; 06:13)

c. éleen mas, tʰýrkʰtše ɣonušéfcʰete.
say.IMPF.3SG we.ACC Turkish speak.IMP.2PL
'S/he was telling us to speak Turkish.'
(S07; 812_0048; 05:09)

(92) Romeyka:
pápʰos mu aðátšaka stetš, léi.
grandfather.NOM I.POSS here live.3SG say.3SG
'S/he says that my grandfather lives here.'
(S07; 812_0074; 01:13)

Interestingly, Schreiber (2018: 924) mentions that in Romeyka of Çaykara the use of *ci* as complementiser in (see (93)). However, it seems that this is probably wrong, since *ci* is the coordinator and is very much used as means of subordination.

(93) Romeyka of Çaykara:
eɣó ton džíri m ípa ci.
I.NOM the.ACC father.ACC I.POSS say.Past.1SG that
'I told my father that I love him much.'
(Schreiber 2018: 925)

2.3.3.1.3*na*-clauses

In Romeyka, nonveridical predicates select *na*-clauses (Sitaridou 2014b). The head of *na*-clauses is the particle *na*, which derives from the AG complementiser *ἵνα* (*hína*). The complement of the head *na* is a clause with a finite verb. Said that, except of its function as a future marker, *na* is used as a complementiser (Mackridge 1987: 130). According to Sitaridou (2014b), in comparison to MG, its use in Romeyka is more restricted.

First, in Romeyka, negated present tense modals such as *tš eporó* 'I can't' (see (94)a) and *íle* 'must' —a loanword from Turkish which functions as an invariant modal— (see (94)b) select *na*-clauses as their complements (Sitaridou 2014b: 123):

(94) Romeyka:

a. tš eporó na porpató.
NEG can.1SG PRT.MOD walk.1SG
'I can't walk.'
(S07; 812_0004; 05:34)

b. íle na porpató / porpatís / porpatí.
must PRT.MOD walk.1SG / walk.2SG / walk.3SG
'I/you/s/he/it must walk.'
(Sitaridou 2014b: 123)

Second, in Romeyka, volitionals select *na*-clauses, such as *θélo* 'I want', on the non-controlled interpretation (see (95)a) and *tši θélo* 'I don't want' and *utš eθélna* 'I wasn't wanting' on the controlled interpretation (see (95)b and (95)c) only when negated. Importantly, *θélo* 'I want' and *eθélna* 'I was wanting' on the positive controlled interpretation (see (95)d and (95)e) do not (Sitaridou 2014b: 123):

(95) Romeyka:

a. esí θélis eɣó çe na tróɣo.
you.NOM want.2SG I.NOM NEG PRT.MOD eat.1SG
'You don't want me to eat.'
(Sitaridou 2014b:123)

b. tši θélo na porpató.
NEG want.1SG PRT.MOD walk.1SG
'I don't want to walk'
(Sitaridou 2014b: 123)

c. utš eθélna n emáireva.
NEG want.IMPF.1SG PRT.MOD cook.IMPF.1SG
'I didn't want to cook'
(Sitaridou 2014b: 123)

d. *polá eθélna (n) étroɣa.
very want.IMPF.1SG PRT.MOD eat.IMPF.1SG
'I wanted to eat a lot.'
(Sitaridou 2014b: 123)

e. *θélo na porpató.
want.1SG PRT.MOD walk.1SG
'I want to walk.'
(Sitaridou 2014b: 124)

In fact, in Romeyka, in contexts such as (95)d and (95)e, another volitional verb surfaces, namely *aɣapó* 'I love/like'. For the latter, the only available type of complement is a *na*-clause, which surfaces regardless of the presence of negation and the control properties. This is demonstrated in (96)a and (96)b where there is both negation and control and in (96)c and (96)d where there is non-obligatory control (NOC) (Sitaridou 2014b: 124):

(96) Romeyka

a. utš aɣapó na páɣo sa rašíæ.
NEG love.1SG PRT.MOD go.1SG to.the mountains.ACC
'I don't like to go in the mountains.'
(Sitaridou 2014b: 124)

b. utš aɣápena n emáireva.
NEG love.IMPF.1SG PRT.MOD cook.IMPF.1SG
'I wasn't fond of cooking.'
(Sitaridou 2014b: 124)

c. to peð m aso χorion aɣapo.
the.NOM child.NOM I.POSS from.the.ACC village.ACC want.1SG
'I want my child to leave the village.'
(Sitaridou 2014b: 124)

d. aɣapó na tšimáste.
love.1SG PRT.MOD sleep.2PL
'I want you to sleep.'
(Sitaridou 2014b: 124)

Third, in Romeyka, *na*-clauses are found as complements to causatives (Sitaridou 2014b: 125) (see (97)):

(97) Romeyka:

a. efíkane sas na skáftete ta χoráfæ suna.
let.Past.3PL you.ACC PRT.MOD dig.3PL the.ACC fields you.POSS
'They let you dig your fields.'
(Sitaridou 2014b: 125)

b. i džandarmáðes utš efíkane
the.NOM policemen.NOM NEG let.Past.3PL
na skáftete ta χoráfæ.
PRT.MOD dig.2PL the.ACC fields.ACC
'The policemen didn't let you dig the fields.'
(Sitaridou 2014b: 125)

Fourth, in Romeyka, *na*-clauses appear as complements to mental perception verbs such as *enéspala* 'I forgot' (see (98)a) —interestingly, the corresponding antonym is only rendered periphrastically, namely *érte so tšeféli m* 'it came to mind', which also selects a *na*-clause (see (98)b):

(98) Romeyka:

a. enéspala na léɣo
forget.Past.1SG PRT.MOD say.1SG
ti mamí ta habéræ.
the.ACC grandmother.ACC the.ACC news.ACC
'I forgot to tell the news to the grandmother.'
(Sitaridou 2014b: 125)

b. érte so tšefali m na léɣo se
come.Past.3SG to.the.ACC head I.POSS PRT.MOD say.1SG you.ACC
do epíce.
what.ACC do.Past.3SG
'It came to mind to tell you what he did.'
(Sitaridou 2014b: 125)

Fifth, in Romeyka, *na*-clauses appear as complements to emotive verbs (see (99)) (Sitaridou 2014b: 126):

(99) Romeyka
eχára na mairévo.
be-happy.Past.1SG PRT.MOD cook.1SG
'I was happy I had cooked.'
(Sitaridou 2014b: 126)

Note in (99) that volitional *θélo* only requires a *na*-clause as complement when negated. This does not hold true for the volitional *aɣapó* 'I love/like' which always requires a *na*-clause (see (100)) (Sitaridou 2014a: 124):

(100) Romeyka:

aɣapó n eftǽɣo ðulíæs.

like.1SGPRT.MOD make.1SG chores.ACC

'I like to do chores.'

(S07; 812_0004; 06:37)

2.3.3.2 Nonfinite complementation

2.3.3.2.1 Infinitives

According to Sitaridou (2014b: 122), the infinitive in Romeyka is defined as not bearing any agreement features and as not occurring independently, but only as the complement of a superordinate syntactic construction (Özkan 2013: 149).

According to Sitaridou (2014b), in Romeyka infinitives are attested (a) as complements to negated past tense modal verbs (Sitaridou 2014b: 126) (see (101)); and, (b) as complements to negated past tense volitional verb *utš eθélesa* 'I did not want' (Sitaridou 2014b: 126) (see (102)); To this end, according to Sitaridou the infinitive in Romeyka surfaces in a subset of nonveridical predicated, namely negated past tense modals and volitionals (Sitaridou 2014b: 127). and, (c) in counterfactuals such as wishes (see (103)a), exclamatives (see (103)b) and conditionals as a complement of *íχa* 'I had' (see (103)c).

(101) Romeyka:

tš epórese mairepsíni.

NEG can.Past.3SG cook.INF

'S/he couldn't cook.'

(S07; 812_0004; 04:06)

(102) Romeyka:

utš eθélesa mairepsíni.

NEG want.Past.1SGcook.INF

'I didn't want to cook.'

(S01; 150702_0032; 01:28)

(103) Romeyka:

a. as íšen porpatesíni sa rašíæ!
PRT.OPT have.IMPF.3SG walk.INF to.the.ACC mountains.ACC
'S/He should have walked in the mountains.'
(Sitaridou 2014b: 136)

b. na íχame panini
PRT.MOD have.IMPF.3PL go.INF
χtisíni t ospít so parχár!
build.INF the.ACChouse.ACC in.the.ACC pastures.ACC
'I wish we had gone to build the house in the highland pastures.'
(Sitaridou 2014b: 136)

c. na íχa šíta piterupsíni tas ðulíæs,
PRT.MOD have.IMPF.1SG immediately finish.INF the.ACCchores.ACC
χar n epínamen parakáθin.
now PRT.MOD make.IMPF.1PL gathering.ACC
'If I had finished the chores immediately, we would have been gathered together now.'
(S01; 812_0123; 02:52)

The syntactic constructions in which the infinitive occurs vary between dialects (Mackridge 1999: 102–103). For example, in Romeyka of Sürmene as spoken in Beşköy, the infinitive is not used in *prin* 'before'-clauses (Sitaridou 2014a: 49, Table 3, Özkan 2013: 149). Furthermore, in this variety, the infinitive is inflected by active past personal endings added to the infinitive when occurring after the negated past tense of *θélo* 'I want' and *poró* 'I can', like in *utš eθélesa porpátesna* 'I did not want to walk' (Özkan 2013: 148).

Finally, inflected infinitives occur as a strategy of nominalization whereby they are complements to (a) aspectuals such as *epitúrepsa* 'I finished' and (b) verbs of mental perception such as *enéspala* 'I forgot' (see (104)). The nominalised infinitive is used with a, possibly obligatory, complex possessive *(e)muneθe* which may be a calque from Turkish nominalizations of the type *oku-ma-sı-nı* 'read-INF-3SG-POSS-ACC' (Sitaridou 2014b: 130). Furthermore, nominalization by means of inflected infinitives before adjectives are formed as in (see (105)) (Sitaridou 2014b: 42):

(104) Romeyka:

to	tšimiθín	emun	-eθe	enéspala.
the.NOM	sleep.INF	we.POSS	-it.POSS	forget.Past.1SG

'I forgot to sleep.'

(Sitaridou 2014b: 131)

(105) Romeyka:

aftí	i	ðulían	to	maθíni	mu.
this.NOM	the.NOM	job.NOM	the.NOM	learning.INF	I.POSS

'It is easy for me to learn how to do this job.'

(Sitaridou 2014a: 42)

2.3.3.2.2 Deverbal nouns

Along with infinitives, Sitaridou (2014b) also recognises verbal nouns. They may be used (a) as complements to volitionals such as *θélo* 'I want' (see (106)a) and (b) as a complement to the phrase modal *epašláepsa* 'I started' ((106)b) which requires a deverbal noun introduced by the preposition *so* 'to the' (Sitaridou 2014b: 130):

(106) Romeyka:

a.

to	peðí	m		
the.NOM	child.NOM	I.POSS		
to	pánimon	asó	χoríon	θélo.
the.NOM	going.NOM	from.the.ACC	village.ACC	want.1SG

'I want my child to leave the village.'

b.

epašláepsa	pol:á	so	ðipsásimo.
start.Past.1SG	very	to.the.ACC	drinking.ACC

'I started to get very thirsty.'

(Sitaridou 2014b: 131, 130)

2.3.4 Clausal complements lacking an overt complementiser

According to Sitaridou (2014b: 127–129), several types of complement clause occur without an overt complementiser. These include the complements of (a) perception verbs (see (107)a); (b) some emotive verbs such as *efovéθa* 'I feared' (see (107)b) (but not in all emotive verbs; others are expressed by *na*-clauses); (c) epistemic predicates (see (107)c); and (d) verbs of saying (see (91) and (92) in §2.3.3.1.2):

(107) Romeyka:

a. eɣó ékusa o tšopánon
I.NOM hear.Past.1SG the.NOM shepherd.NOM
ton árko endóke.
the.ACCbear.ACC kill.Past.3SG
'I heard that the shepherd killed the wolf.'

b. efovéθa χánis ton pará s.
fear.Past.1SG lose.2SG the.ACCmoney.ACC you.POSS
'I feared you may lose your money.'

c. θaró hastás en.
think.1SG sick.1SG be.3SG
'I think s/he is sick.'

(Sitaridou 2014b: 127-128)

2.4 Conclusions

In this chapter, I presented a brief overview of the Romeyka grammar. The way I presented the grammar was not explicit at all and was developed in a traditional way. I specifically introduced the basic facts about phonology, morphology and syntax in Romeyka. The goal of the chapter was to introduce the reader to the basic grammatical rules of Romeyka so that they would follow the linguistic discussion in the remainder of this study.

3 Word order in Romeyka

3.0 Introduction

The focus of this chapter is word order variation in Romeyka. I specifically examine the respective position of the subject (S), verb (V) and object (O) in matrix and subordinate declarative and interrogative clauses in Romeyka. The goal of this chapter is to survey word order variation in Romeyka.

The chapter is structured as follows: in §3.1, I present the results of my survey of word order variation in matrix and subordinate declarative and interrogative clauses in Romeyka; in §3.2, I give five arguments for V^0-to-T^0 raising in Romeyka; in §3.3, I show that there are two subject positions in matrix and subordinate clauses in Romeyka. The chapter concludes in §3.4.

3.1 The breakdown of word orders in Romeyka

3.1.0 Introduction

In this section, I present the results of my survey of word order variation in matrix and subordinate declarative and interrogative clauses in Romeyka. The strategy I follow is to examine clauses that are syntactically similar. I therefore limit it to clauses with subject, verb and object.

This section is structured as follows: §3.1.1 examines word order in matrix declarative clauses in Romeyka; §3.1.2 examines word order in subordinate declarative clauses in Romeyka; §3.1.3 examines word order in direct questions in Romeyka; and §3.1.4 examines word order in indirect questions in Romeyka. The main findings of the section are summarised in §3.1.5.

3.1.1 Matrix declarative clauses in Romeyka

Only three permutations of S, V and O are found in matrix declarative clauses in Romeyka, namely SVO (see (1)), SOV (see (2)) and OSV (see (3)):

(1) SVO clause:

o	dohtóris	epíren	tin	aišé.
the.NOM	doctor.NOM	marry.Past.3SG	the.ACC	Ayşe.ACC

'The doctor married Ayşe.'

(S01; 140102_00080; 01:25)

(2) SOV clause:

o	dohtóris	tin	aišé	epíren.
the.NOM	doctor.NOM	the.ACC	Ayşe.ACC	marry.Past.3SG

'The doctor married Ayşe.'

(S01; 140102_0008; 01:41)

(3) OSV clause:

tin	aišé	o	dohtóris	epíren.
the.ACC	Ayşe.ACC	the.NOM	doctor.NOM	marry.Past.3SG

'The doctor married Ayşe.'

(S01; 140102_0008; 01:33)

Nevertheless, V-initial and S-final word orders are not attested in matrix declarative clauses in Romeyka, namely VSO (see (4)), VOS (see (5)), or OVS (see (6)):

(4) VSO clause:

?epíren	o	dohtóris	tin	aišé.
marry.Past.3SG	the.NOM	doctor.NOM	the.ACC	Ayşe.ACC

'The doctor married Ayşe.'

(5) VOS clause:

?epíren	tin	aišé	o	dohtóris.
marry.Past.3SG	the.ACC	Ayşe.ACC	the.NOM	doctor.NOM

'The doctor married Ayşe.'

(6) OVS clause:

?tin	aišé	epíren	o	dohtóris.
the.ACC	Ayşe.ACC	marry.Past.3SG	the.NOM	doctor.NOM

'The doctor married Ayşe.'

3.1.2 Subordinate declarative clauses in Romeyka

Romeyka does not employ an overt complementiser to introduce subordinate clauses. That is, subordinate clauses in Romeyka reflect a stage prior their grammaticalisation. However, word order in subordinate clauses is different than that in matrix clauses. In particular, only two permutations of S, V and O are found in subordinate declarative clauses in Romeyka, namely SOV (see (7)) and OSV (see (8)):

(7) SOV clause:

o	mohalːímis	ípen,		
the.NOM	teacher.NOM	say.Past.3SG		
i	ɟylsén	aténan	utš	aɣapá.
the.NOM	Gülsen.NOM	her.ACC	NEG	love.3SG

'The teacher said that Gülsen doesn't like her.'

(S07; 812_0065; 05:06)

(8) OSV clause:

o	mohalːímis	ípen,		
the.NOM	teacher.NOM	say.Past.3SG		
aténan	i	ɟylsén	utš	aɣapá.
her.ACC	the.NOM	Gülsen.NOM	NEG	love.3SG

'The teacher said that Gülsen doesn't like her.'

(S07; 812_0065; 05:01)

Nevertheless, O-final, V-initial and S-final word orders are not attested in matrix declarative clauses in Romeyka, namely SVO, (see (9)), VSO (see (10)), VOS (see (11)), or OVS (see (12)):

(9) SVO clause:

?o	mohalːímis	ípen,		
the.NOM	teacher.NOM	say.Past.3SG		
i	ɟylsén	utš	aɣapá	aténan.
the.NOM	Gülsen.NOM	NEG	love.3SG	her.ACC

'The teacher said that Gülsen doesn't like her.'

(10) VSO clause:

?o	mohalːímis	ípen,		
the.NOM	teacher.NOM	say.Past.3SG		
utš	aɣapá	i	ɟylsén	aténan.
NEG	love.3SG	the.NOM	Gülsen.NOM	her.ACC

'The teacher said that Gülsen doesn't like her.'

(11) VOS clause:

?o mohalːímis ípen,
the.NOM teacher.NOM say.Past.3SG
utš aɣapá aténan i ɟylsén.
NEG love.3SG her.ACC the.NOM Gülsen.NOM
'The teacher said that Gülsen doesn't like her.'

(12) OVS clause:

?o mohalːímis ípen
the.NOM teacher.NOM say.Past.3SG
aténan utš aɣapá i ɟylsén.
her.ACC NEG love.3SG the.NOM Gülsen.NOM
'The teacher said that Gülsen doesn't like her.'

3.1.3 Direct questions in Romeyka

Only three permutations of S, V and O are found in both yes/no questions and *wh*-questions in Romeyka, namely SVO (see (13)a for yes/no questions and (13)b for *wh*-questions), SOV (see (14)a for yes/no questions and (14)b and (14)c for *wh*-questions) and OSV (see (15)a for yes/no questions and (15)b for *wh*-questions):

(13) SVO clause:

a. i nífe efáisen ti mamíka?
the.NOM daughter-in-law.NOM feed.Past.3SG the.ACCmother-in-law.ACC
'Did the daughter-in-law feed the mother-in-law?'
(S01; 150702_0013; 13:53)

b. píos eðótšen tin kosːáran?
who.NOM give.Past.3SG the.ACChen.ACC
'Who gave the hen?
(S01; 812_0093; 00:03)

(14) SOV clause:

a. esís ta tsupáðæ θerízete?
you.NOM the.ACCcorn.ACC harvest.2PL
'Do you harvest the corn?'
(S07; 812_0067; 01:58)

b. alís dóɣna ðótšen?
Alis.NOM what.ACC give.Past.3SG
'What did Alis give?'
(S01; 812_0093; 00:16)

c. píos dóɣna aɣórasen?
who.NOM what.ACC buy.Past.3SG
'Who bought what?'
(S01; 150703_0042; 07:37)

(15) OSV clause:

a. ató o mehmétis éndžen æ?
this.ACC the.NOM Mehmetis.NOM bring.Past.3SG it.ACC
'Did Mehmetis bring that?'
(S01; 150703_0042; 06:36)

b. χavítsæ píos éfaen?
anchovies.ACC who.NOM eat.Past.3SG
'Who ate anchovies?'
(S01; 812_0057; 04:06)

Nevertheless, V-initial and S-final word orders are not attested either in yes/no questions or *wh*-questions in Romeyka, namely either VSO (see (16)a for yes/no questions and (16)b, (16)c and (16)d for *wh*-questions), or VOS (see (17)a for yes/no questions and (17)b, (17)c and (17)d for *wh*-questions), or OVS (see (18)a for yes/no questions and (18)b, (18)c and (18)d for *wh*-questions):

(16) VSO clause:

a. ?efáisen i nífe ti mamíka?
feed.Past.3SG the.NOM daughter-in-law.NOM the.ACCmother-in-law.ACC
'Did the daughter-in-law feed her mother-in-law?'

b. ?eðótšen píos tin kosːáran?
give.Past.3SG who.NOM the.ACChen.ACC
'Who gave the hen?

c. ?ðótšen alís dóɣna?
give.Past.3SG Alis.NOM what.ACC
'What did Alis give?'

d. ?aɣórasen píos dóɣna?
buy.Past.3SG who.NOM what.ACC
'Who bought what?'

(17) VOS clause:

a. ?efáisen ti mamíka i nífe?
feed.Past.3SG the.ACCmother-in-law.ACC the.NOM daughter-in-law.NOM
'Did the daughter-in-law feed her mother-in-law?'

b. ?eðótšen tin kosːáran píos?
give.Past.3SG the.ACChen.ACC who.NOM
'Who gave the hen?

c. ?ðótšen dóɣna alís?
give.Past.3SG what.ACC Alis.NOM
'What did Alis give?'

d. ?aɣórasen dóɣna píos?
buy.Past.3SG what.ACC who.NOM
'Who bought what?'

(18) OVS clause:

a. ?ti mamíka efáisen i nífe?
the.ACC mother-in-law.ACC feed.Past.3SG the.NOM daughter-in-law.NOM
'Did the daughter-in-law feed her mother-in-law?'

b. ?tin kosːáran eðótšen píos?
the.ACC hen.ACC give.Past.3SG who.NOM
'Who gave the hen?

c. ?dóɣna ðótšen alís?
what.ACC give.Past.3SG Alis.NOM
'What did Alis give?'

d. ?dóɣna aɣórasen píos?
what.ACC buy.Past.3SG who.NOM
'Who bought what?'

Although multiple *wh*-questions are attested in SOV orders (see (14)c), they are not attested in OSV orders (see (19)):

(19) OSV clause (multiple *wh*-question):
?dóɣna píos aɣórasen?
what.ACC who.NOM buy.Past.3SG
'Who bought what?'

3.1.4 Indirect questions in Romeyka

Romeyka does not employ an overt complementiser to introduce indirect questions. That is, indirect questions in Romeyka reflect a stage prior their grammaticalisation. However, word order in indirect questions is different than that in direct questions. In particular, Only two permutations of S, V and O are found in subordinate interrogative clauses in both yes/no questions and *wh*-questions in Romeyka, namely SOV (see (20)a for yes/no questions and (20)b and (20)c for *wh*-questions) and OSV (see (21)a for yes/no questions and (21)b for *wh*-questions):

(20) SOV clause:

a. rotás me, alís tin aišén efílisen?
ask.NOM.2SG I.ACC Alis.NOM the.ACCAyşe.ACC kiss.Past.3SG
'You ask me, did Alis kiss Ayşe?'
(S01; 150702_0022; 03:16)

b. do θarís, alís tínan efílisen?
what.ACC think.2SG Alis.NOM who.ACC kiss.Past.3SG
'What do you think, who did Alis kiss?'
(S01; 150703_0040; 19:24)

c. esí erotás me, píos dóɣna aɣórasen?
you.NOM ask.2SG I.ACC who.NOM what.ACC buy.Past.3SG
'You ask me, who bought what?'
(S01; 150703_0042; 08:16)

(21) OSV clause:

a. rotás me, to cʰitʰápin alís eχúɟepsen?
ask.2SG I.ACC the.ACCbook.ACC Alis.NOM read.Past.3SG
'You ask me, did Alis read the book?'
(S01; 150702_0022; 06:13)

b. rotás me, tin aišén ts epíren?
ask.2SG I.ACC the.NOM the.ACCAyşe.ACC who.NOM marry.Past.3SG
'You ask me, who married Ayşe?'
(S01; 140102_0008; 01:41)

Nevertheless, O-final, V-initial and S-final word orders are not attested in either yes/no questions or *wh*-questions in Romeyka, namely SVO (see (22)a for yes/no questions and (22)c and (22)d for *wh*-questions), VSO (see (23)a for yes/no questions and (23)c and (23)d for *wh*-questions), VOS (see (24)a for yes/no questions and (24)c and (24)d for *wh*-questions), or OVS (see (25)a for yes/no questions and (25)c and (25)d for *wh*-questions):

(22) SVO clause:

a. ?rotás me, alís efílisen tin aišén?
ask.NOM.2SG I.ACC Alis.NOM kiss.Past.3SG the.ACC Ayşe.ACC
'You ask me, did Alis kiss Ayşe?'

b. rotás me, tin aišén ts epíren?
ask.2SG I.ACC the.NOM the.ACC Ayşe.ACC who.NOM marry.Past.3SG
'You ask me, who married Ayşe?'

c. ?do θarís, alís efílisen tínan?
what.ACC think.2SG Alis.NOM kiss.Past.3SG who.ACC
'What do you think, who did Alis kiss?'

d. ?esí erotás me, píos aɣórasen dóɣna?
you.NOM ask.2SG me who.NOM buy.Past.3SG what.ACC
'You ask me, who bought what?'

(23) VSO clause:

a. ?rotás me, efílisen alís tin aišén?
ask.NOM.2SG I.ACC kiss.Past.3SG Alis.NOM the.ACC Ayşe.ACC
'You ask me, did Alis kiss Ayşe?'

b. rotás me, tin aišén ts epíren?
ask.2SG I.ACC the.NOM the.ACC Ayşe.ACC who.NOM marry.Past.3SG
'You ask me, who married Ayşe?'

c. ?do θarís, efílisen alís tínan?
what.ACC think.2SG kiss.Past.3SG Alis.NOM who.ACC
'What do you think, who did Alis kiss?'

d. ?esí erotás me, aɣórasen píos dóɣna?
you.NOM ask.2SG me buy.Past.3SG who.NOM what.ACC
'You ask me, who bought what?'

(24) VOS clause:

a. ?rotás me, efílisen tin aišén alís?
ask.NOM.2SG I.ACC kiss.Past.3SG the.ACC Ayşe.ACC Alis.NOM
'You ask me, did Alis kiss Ayşe?'

b. rotás me, tin aišén ts epíren?
ask.2SG I.ACC the.NOM the.ACC Ayşe.ACC who.NOM marry.Past.3SG
'You ask me, who married Ayşe?'

c. ?do θarís, efílisen tínan alís?
what.ACC think.2SG kiss.Past.3SG who.ACC Alis.NOM
'What do you think, who did Alis kiss?'

d. ?esí erotás me, aɣórasen dóɣna píos?
you.NOM ask.2SG I.ACC buy.Past.3SG what.ACC who.NOM
'You ask me, who bought what?'

(25) OVS clause:

a. ?rotás me, tin aišén efílisen alís?
ask.NOM.2SG I.ACC the.ACC Ayşe.ACC kiss.Past.3SG Alis.NOM
'You ask me, did Alis kiss Ayşe?'

b. rotás me, tin aišén ts epíren?
ask.2SG I.ACC the.NOM the.ACC Ayşe.ACC who.NOM marry.Past.3SG
'You ask me, who married Ayşe?'

c. ?do θarís, tínan efílisen alís?
what.ACC think.2SG who.ACC kiss.Past.3SG Alis.NOM
'What do you think, who did Alis kiss?'

d. ?esí erotás me, dóɣna aɣórasen píos?
you.NOM ask.2SG I.ACC what.ACC buy.Past.3SG who.NOM
'You ask me, who bought what?'

Although multiple *wh*-questions are attested in SOV orders (see (20)c), they are not attested in OSV orders (see (26)):

(26) OSV clause (multiple *wh*-question):

?esí	erotás	me,	dóɣna	píos	aɣórasen?
you.NOM	ask.2SG	I.ACC	what.ACC	who.NOM	buy.Past.3SG

'You ask me, who bought what?'

3.1.5 Summary

In this section, I presented the results of my survey of word order variation in matrix and subordinate declarative and interrogative clauses in Romeyka. The findings of this survey show that three word orders are attested in Romeyka, namely SVO, SOV and OSV. On the other hand, V-initial and S-final word orders, i.e. VSO, VOS and OVS, are not attested in Romeyka.

3.2 Verb positions in Romeyka

3.2.0 Introduction

In this section, I give four arguments in favour of V^0-to-T^0 raising in Romeyka. The first two arguments address the interaction between rich morphological inflection of the verb and syntactic raising. Romeyka shows two properties typically associated with V^0-to-T^0 movement: rich person and number agreement on the one hand and null subjects, or pro-drop on the other. I discuss these properties in §3.2.1 and §3.2.2 respectively. Another argument for V^0-to-T^0 raising comes from placement facts. In §3.2.3, I use the respective position of the focus-associate particle *dže* 'also' and verb as evidence for verb raising. In §3.2.4, I use the respective position of adverbs and auxiliary verbs as evidence for verb raising. The findings of the section are summarised in §3.2.5.

3.2.1 The Rich Agreement Hypothesis (RAH)

The Rich Agreement Hypothesis (RAH) generally refers to the generalisation that V^0-to-T^0 movement is conditioned by rich subject agreement on the finite verb. RAH has long been taken as an important argument in favour of a direct connection between syntax and morphology (see Pollock 1989, Roberts 1993, Vikner 1995, 1997). In recent years, however, the RAH has been disputed on both empirical and theoretical grounds. Empirically, data have been put forward that seem to suggest the existence of languages that are poorly inflected but still display V^0-to-T^0 movement.

Theoretically, under lexicalist approaches (see Chomsky 1995), the tight connection between rich agreement and V^0-to-T^0 movement has been taken as a strong argument for the idea that morphology drives syntax (see Vikner 1995). However, in more current generative models of grammar, morphological insertion is assumed to take place after the syntactic computation, suggesting that morphology can have no direct influence on the syntactic derivation.

Romeyka has distinct verbal forms for all persons and singular and plural numbers with no suppletion, at least in most tense-voice combinations. Table 23 shows the present active declension of *léɣo* 'I say':

Table 23. Subject agreement paradigm.

	Singular number	**Plural number**
1st person	*léɣ-o*	*léɣ-umen*
2nd person	*léʝ-is*	*léʝ-ete*
3rd person	*léʝ-i*	*léɣ-un(e)*

The rich person and number inflection of the subject agreement paradigm in Romeyka indicates that the verb raises to the head of a Tense projection.

3.2.2 Null subjects

The pro-drop, or null subject phenomenon refers to a clause in which no overt subject is expressed, as in the example in (27) from Romeyka:

(27) Romeyka:

opsé	χars	ípe	tes.
yesterday	now	say.Past.3SG	her

'Yesterday she told her.'
(S01; 0120713192027; 01:36)

Recent typologies of null subjects distinguish various types of null subjects. In some languages, not only subject pronouns, but also object pronouns can be dropped. One example is Chinese (see Huang 1984). This type of pro-drop has recently been referred to as radical pro-drop or discourse pro-drop (see Neeleman & Szendröi 2007).

There is long held typological correlation between rich person and number inflection and the type of pro-drop found in consistent null subject languages. Roberts & Holmberg (2010: 3) note that this observation was noted by AGs scholars, quoting a passage from Apollonius Dyscolus on AG.

The intuition is that verbs that are inflected for person and number do not require further specification as to what the subject is. This intuition has been formulated syntactically in various ways. One option is that the requirement that all clauses have a subject, where this subject occurs in a particular syntactic position, (the Extended Projection Principle of Chomsky 1982), is not universal (see a current variation in Alexiadou & Anagnostopoulou 1998). Under Alexiadou & Anagnostopoulou's (1998) analysis, verb movement to T^0 is sufficient to identify

the formal features on T and therefore subjects are not required in the [Spec, TP] subject position. Another intuition is that the empty category pro occupies the canonical [Spec, TP] subject position (Rizzi 1982, Chomsky 1982). When a verb moves to T^0, the person and number features of the verb are copied onto the empty pronominal, licensing it. These proposals imply a direct relationship between V^0-to-T^0 movement and pro-drop. Notice, however, that not all languages with V^0-to-T^0 movement have consistent null subjects. For example, French has V^0-to-T^0 but not pro-drop (see (28)):

(28) French:

J’	ai	mangé	une	soupe.
I	have.1SG	eat.INF	a.ACC	soup.ACC

‘I ate a soup.’

Romeyka shows all of the properties that other null subject languages show too. First, 3rd person subjects can be dropped (see (29)):

(29) Romeyka:

léǰi me.

say.3SG me

‘S/He says to me.’

(S01; 812_0113; 05:58)

Second, 1st and 2nd personal pronouns can be dropped (see (30)):

(30) Romeyka:

a. léɣo se.

say.1SG you.ACC

‘I say to you.’

(S08; 812_0067; 04:19)

b. léǰis me.

say.2SG me

‘You say to me.’

(S01; 150702_0023; 05:31)

In harmonic null subject languages, if the subordinate clause of a bi-clausal construction contains an overt subject pronominal, the subject of the main clause is not necessarily co-referential with the subject of the subordinate (see Fascarelli & Hinterhölzl 2007, i.a.) (see (31) from MG):

(31) Modern Greek:

a. i maria ɟelase afu iðe ton ɟani.
the.NOM Maria.NOM laugh.Past.3SG after saw the.ACCYannis.ACC
'Maria laughed after she saw Yannis.'

b. i maria ɟelase
the.NOM Maria.NOM laugh.Past.3SG
afu afti iðe ton ɟani.
after she see.Past.3SG the.ACCYiannis.ACC
'Maria laughed after she saw Yannis.'

(Roberts & Holmberg 2010: 7)

In Romeyka, it is the case that subordinate clauses whose subjects are co-referential with matrix clause subjects do not contain overt pronouns (see (32)):

(32) Romeyka:
θélo na páo.
want.1SG PRT.MOD go.1SG
'I want to go.'
(S06; 812_0117; 00:20)

Biberauer & Roberts (2010) put forth a proposal that covers more typological correlations concerning null subjects, verbal inflection and V^0-to-T^0 movement. In particular, they make a distinction between person and number inflection and tense inflection. In such a system, what triggers V^0-to-T^0 movement is tense inflection, rather than person/number inflection. Specifically, both T^0 and V^0 carry unvalued features, making them active in the derivation. While V^0 lacks a valued Tense feature, T^0 is valued for Tense. T^0, being a functional head, is not specified with respect to argument structure, whereas V^0 is specified as having argument structure. Within the Agree based system of Chomsky (2000, 2001), this means that T^0 and V^0 always establish an Agree relation. In languages like English (as well as V2 Germanic

languages), the tense on the verb is licensed in this way, with no movement to T^0. In null subject languages, on the other hand, T^0 bears an EPP feature, relating to rich tense synthesis, triggering V^0-to-T^0 movement.

Biberauer & Roberts (2010) discuss the contrast between Romance languages, which have V^0-to-T^0 and Germanic languages, which do not. A typical example of the latter is English, where verbs do not raise. The difference is that the Romance languages have more synthetic (non-periphrastic) tense distinctions than the Germanic languages. These tense distinctions also encompass aspect and mood. For example, Italian shows the distinctions in (33)a, French those in (33)b, while English shows only the distinctions in (33)c:

(33) Tense/aspect/mood forms:

a. Italian:
parlo (present), *parlerò* (future), *parlerei* (conditional), *parlavo* (imperfect), *parli* (present subjunctive), *parlassi* (past subjunctive), *parlay* (preterit)

b. French:
parle (present indicative/subjunctive), *parlerai* (future), *parlerais* (conditional), *parlais* (imperfect), *parlay* (preterite), *parlasse* (past subjunctive)

c. English:
speak (present), *spoke* (past)

Biberauer & Roberts' (2010) proposal accounts for more cross-linguistic variation concerning null subjecthood and V^0-to-T^0 movement. V^0-to-T^0 movement is available not due to rich person and number inflection, but to tense synthesis. Pro-drop, on the other hand, is available due to rich person and number inflection. This explains the contrast between English (also Mainland Scandinavian), French and Italian/Modern Greek (among other languages). The differences are summarised in Table 24:

Table 24. Cross-linguistic variation concerning V^0-to-T^0 movement.

	Rich person, number	**Pro-drop**	**Tense synthesis**	**V^0-to-T^0**
Modern Greek, Italian	Yes	Yes	Yes	Yes
French	No	No	Yes	Yes
English	No	No	No	No

The tense/aspect/voice system in Romeyka is similar to the very complex system in Modern Greek, as shown in Table 25:

Table 25. Attested tense/aspect/mood form of *leyo* 'I say'.

Tense/aspect, mood, voice	**Form**
Present	*léyo*
Past	*ípa*
Imperfect	*éleya*

To sum up, while the Rich Agreement Hypothesis claims that V^0-to-T^0 movement corresponds to rich person and number inflection, Biberauer & Roberts (2010) propose that V^0-to-T^0 movement is the consequence of a high degree of tense synthesis. Romeyka displays both of these properties and it is therefore expected that V^0-to-T^0 movement takes place.

3.2.3 Additive particle *dž(e)* 'also' placement

According to Chatzikyriakidis et al. (2015), *ce* 'also' in Modern Greek is a focal associate operator and it only surfaces in its base-generated position, as a sister to its associate. Therefore, whatever moves out of the associate of *ce* necessarily precedes it, a fact that makes *ce* a straightforward diagnostic for clause structure.

In Romeyka, *dž(e)* 'also' seems to be a focal associate operator as well, similar to the Modern Greek *ce*. If we assume that *dž(e)* occupies the [Spec, vP] position, then the verb appears to precede *dž(e)*, indicating that the verb raises to T^0 (see (34)):

(34) Romeyka:

θénumen dže álas.

put.1PL PRT salt.ACC

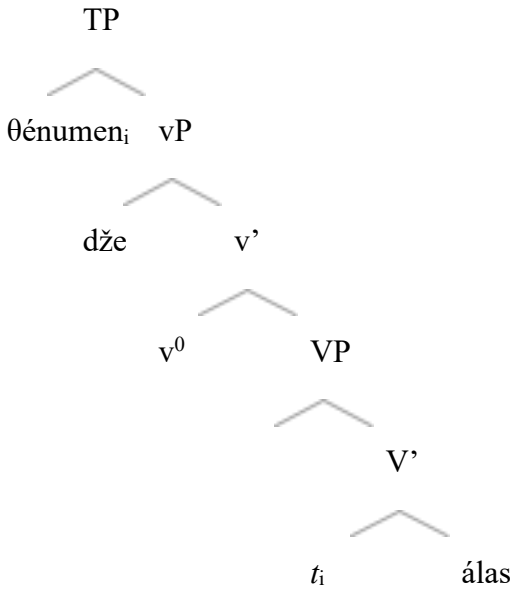
'We also put salt.'

(S01; 150703_0041; 05:09)

In the example (34), *dž(e)* is placed in the [Spec vP], indicating that the verb raises out of the vP.

The syntactic derivation of (34) should be like the one in (35):

(35)



3.2.4 Adverb placement

Interestingly, some adverbials are attested to interpolate between the auxiliary and the verb. Given that the following adverbs are placed in [Spec, vP], the auxiliary must be in T^0. Consider the example in (36):

(36) Romeyka:

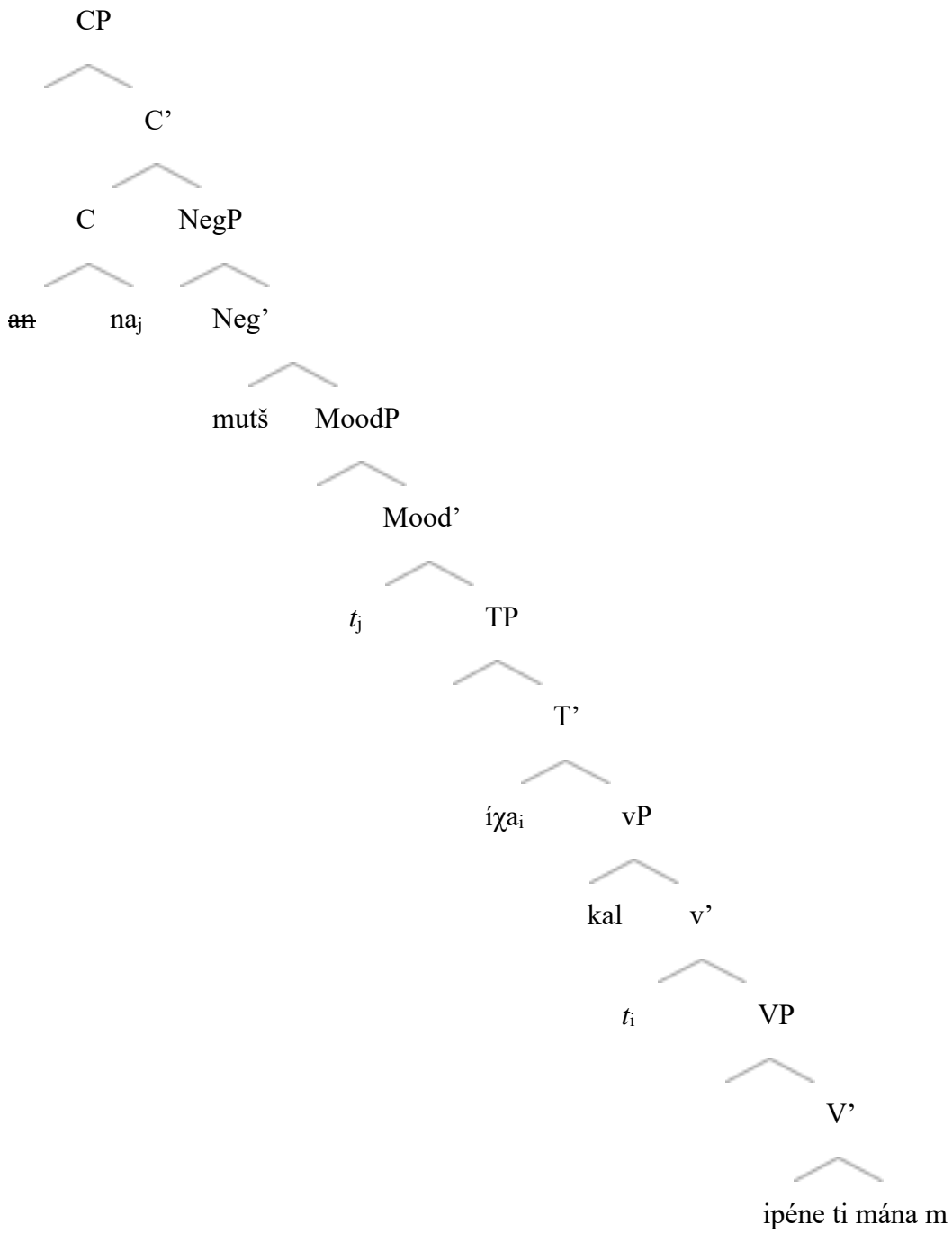
na	mutš	íχa	kal	ipéne	ti	mána	m,
PRT.MOD	NEG	have.IMPF.1SG	again	say.INF	the.ACC	mother.ACC	I.POSS
i	mána	m	n	éstetšen.			
the.NOM	mother.NOM	I.POSS	PRT.MOD	be.IMPF.3SG			

'If I hadn't said (that) to my mother again, my mother would have stayed (here).'
(S01; 812_0123; 01:04)

In the example (36) the adverb is placed in the [Spec vP], indicating that the verb raises out of the vP.

The syntactic derivation of (36), which according to Chatzopoulou & Sitaridou (2014) implies conditional inversion, should be like the one in (37):

(37)


CP
C’
C
NegP
~~an~~
na$_j$
Neg’
mutš
MoodP
Mood’
t_j
TP
T’
íχa$_i$
vP
kal
v’
t_i
VP
V’
ipéne ti mána m

3.2.5 Summary

In this section, I gave four arguments in favour of V^0-to-T^0 raising in Romeyka. The first two arguments addressed the interaction between rich morphological inflection of the verb and syntactic raising. Romeyka shows two properties typically associated with V^0-to-T^0 movement: rich person and number agreement on the one hand and null subjects, or pro-drop on the other. Another argument for V^0-to-T^0 raising came from placement facts; I used the respective position of the additive particle *dže* 'also' and verb as evidence for verb raising; and I used the respective position of adverbs and auxiliary verbs as evidence for verb raising.

3.3 Subject positions in Romeyka

3.3.0 Introduction

In §3.2, I showed that there is one position for finite verbs in main and subordinate clauses in Romeyka. In this section, I show that there is one subject position in main and subordinate clauses in Romeyka too. I specifically show that in Romeyka subjects in pragmatically unmarked orders are always topics, which are left-dislocated.

This section is organised as follows: in §3.3.1, I present Alexiadou & Anagnostopoulou's (1998) typology of subject placement; in §3.3.2, I present arguments against Alexiadou & Anagnostopoulou's (1998) typology; in §3.3.3, I use the respective position of adverbs to identify TopP in Romeyka; and, in §3.3.4, I provide a summary of the main findings of the section.

3.3.1 Alexiadou & Anagnostopoulou (1998)

Alexiadou & Anagnostopoulou (1998) take the view that in languages with rich person and number inflection, there is no requirement that an element be in the canonical [Spec, TP] position. They discuss facts from Celtic, Modern Greek, Icelandic and English, creating a typology of languages that allow VS orders, based on the parametrisation of the T^0 position (in their terminology Agr^0). Basically, they claim that the Null Subject Parameter is the source of the cross-linguistic variation. For the sake of simplicity, I discuss only the data from English, a non-null-subject language and MG, a null subject language.

There are a number of asymmetries between English and MG VS structures. First of all, in English VS orders an expletive *there* is required in [Spec, TP] (see (38)), unlike in Greek (see (39)):

(38) English:
  a. A man arrived.
  b. *(There) arrived a man.

(39) Modern Greek

éfije	o	pétros.
leave.Past.3SG	the.NOM	Petros.NOM

'Petros left.'
(Kirk 2012: 80)

Second, in English, only intransitive verbs can appear in VS orders, while in MG all types of predicates occur in VS(O) orders. The contrast is shown in (40) and (41):

(40) English:
*There built a man a house.

(41) Modern Greek:

éktise	i	maría	to	spíti.
build.Past.3SG	the.NOM	Maria.NOM	the.ACC	house.ACC

'Maria built the house.'
(Kirk 2012: 80)

A well-known property of expletive constructions in English, among other languages, is that they are ungrammatical if the associate of the expletive is definite. An example is given in (42):

(42) English:
There arrived three men / a man/ *the man / *all the men / *each man / *every man.

The phenomenon is known as the Definite Restriction (DR), or Definiteness Effect (Belletti 1988, Milsark 1977, Moro 1997).

Alexiadou & Anagnostopoulou (1998) take the DR to indicate that definite subjects are incompatible with an expletive in [Spec, TP]. They show (1998: 496) that this restriction is absent in MG (see (43)):

(43) Modern Greek:

írθe	to	káθe	peði.
arrive.Past.3SG	the.NOM	every.NOM	child.NOM

'Each child arrived.'

(Alexiadou & Anagnostopoulou 1998: 496)

They take the fact that there is no expletive in VS orders in MG as an indication that the [Spec, TP] position is not filled and thus, not projected. In their analysis, the verbal inflection in a null subject language is specified enough to satisfy the [EPP], which corresponds to an uninterpretable Definiteness [D] feature on T^0, when the verb moves to T^0.

The parametric difference then lies in what exactly the category that checks the [EPP] is. It can be checked either through Move/Merge XP or Move/Merge X^0 (Alexiadou & Anagnostopoulou 1998: 518). Languages with rich verbal inflection such as MG check the [EPP] through V^0 head (X^0) movement to T^0 and languages with poor agreement such as English check it through XP movement (Move XP), or expletive insertion (Merge XP). Therefore, the [EPP] as a feature is universal; however there is no [Spec, TP] position projected in null subject languages.

A consequence of Alexiadou & Anagnostopoulou's (1998) analysis is that preverbal subjects in null subject languages are left-dislocated to the left periphery, undergoing A' movement rather than A movement. Postverbal subjects stay *in-situ* in the VP. This corresponds to the fact that at least in MG, preverbal subjects have the interpretation of topics, while postverbal subjects are pragmatically neutral (i.e., the neutral order is VSO, not SVO). Consider the example (44) from Alexiadou & Anagnostopoulou (1998: 506):

(44) Modern Greek:

a.

éna	peði	ðjávase	to	'paramíθi	χorís	ónoma'.
a.NOM	child.NOM	read.Past.3SG	the.ACC	fairy.tale.ACC	without	title.ACC

'A certain child/one of the children read 'Fairy Tale without a Title'.'

b.

ðjávase	éna	peði	to	'paramíθi	χorís	ónoma'.
read.Past.3SG	a.NOM	child.NOM	the.ACC	fairy.tale.ACC	without	title.ACC

Alexiadou & Anagnostopoulou (1998: 506)

The preverbal subject in (44)a has a 'strong' partitive or specific interpretation, while the postverbal subject in (44)b favours a non-specific reading. Further evidence that preverbal

subjects in MG are A' moved comes from the contrast between MG and English with respect to scope ambiguities with indefinites and strong quantifiers. The examples in (45) and (46) illustrate this:

(45) English:
Some student filed every article.

(46) Modern Greek:

a.	kápços	fititís	stiçioθétise	káθe árθro.
	some.NOM	student.NOM	file.Past.3SG	every article.ACC
	'Some student filed every article.'			
b.	stiçioθétise	kápços	fititís	káθe árθro
	file.Past.3SG	some.NOM	student.NOM	every article.ACC

Alexiadou & Anagnostopoulou (1998: 506)

In English, an indefinite subject with a strong quantifier object has ambiguous scope; (45) can either mean that one single student filed every article, or that every article was filed by some student or another. In MG, on the other hand, when the indefinite subject is preverbal as in (46)a, the indefinite has to have wide scope; only the reading where one and the same student filed every single article is available. In the VSO order in (46)b, the scope is ambiguous as in English.

Alexiadou & Anagnostopoulou's (1998: 505) explanation is that if the preverbal subject in (46)a were in an A position, the interpretation would remain ambiguous. Alexiadou & Anagnostopoulou (1998) provide a number of arguments showing that preverbal subjects in MG are left-dislocated topics and I will not repeat them all here. In the following subsection I discuss some of the problems that have been brought up with Alexiadou & Anagnostopoulou's (1998) account.

3.3.2 Arguments against Alexiadou & Anagnostopoulou (1998)

Alexiadou & Anagnostopoulou's (1998) analysis makes a couple of very strong predictions. First, it predicts that all null subject languages have VSO orders, which is not true; take for example in Modern Hebrew (see Doron (2000, p. note 8)). Furthermore, even Italian, a consistent null subject language that Alexiadou & Anagnostopoulou treat as an exemplar of

their proposal does not easily allow VSO orders (see Belletti 2001, Cardinaletti 2004, Pinto 1997, Sheehan 2010).

Interestingly, although the pragmatically unmarked order in Romeyka is SVO, it does not allow VSO orders (see (47)):

(47) Romeyka:

a. SVO clause:

o	mustafás	epeLÁpsen	to	χoráfin.
the.NOM	Mustafas.NOM	put.fertiliser.Past.3SG	the.ACC	field.ACC

'Mustafas put fertiliser on the field.'

(S01; 150703_0040; 02:16)

b. ?VSO clause:

?epeLÁpsen	o	mustafás	to	χoráfin.
put.fertiliser.Past.3SG	the.NOM	Mustafas.NOM	the.ACC	field.ACC

'Mustafas put fertiliser on the field.'

The proposal also makes the very strong prediction that all preverbal subjects are left-dislocated in null subject languages, since [Spec, TP] is never projected. Many have shown that this prediction is not borne out for the Romance null subject languages (for example, see Goodall (2001) concerning Spanish; Costa (2004) concerning European Portuguese; Sheehan (2010) concerning Spanish, Italian and European Portuguese).

3.3.3 Identifying TopicP: intervening adverbs

Subjects are found separated from the verb by an adverb. In the default case, verbs move to T^0; this indicates that the subject does not occupy [Spec, TP], or at least that the subject and verb are not in a Spec-head configuration. Consider the example in (48):

(48) Romeyka:

a. alís mían eFÍlisen tin aišén.
Alis.NOM once kiss.Past.3SG the.ACCAyşe.ACC
'Alis kissed Ayşe once.'
(S01; 15702_0022; 00:44)

b. alís mían [tin aiŠÉN]$_{Foc}$ efilisen.
Alis.NOM once the.ACCAyşe.ACC kiss.Past.3SG
'Alis kissed Ayşe once.'
(S01; 150702; 06:24)

c. alís [MÍan]$_{Foc}$ efílisen tin aišén.
Alis.NOM once kiss.Past.3SG the.ACCAyşe.ACC
'Alis kissed Ayşe once.'
(S01; 15702_0022; 00:52)

In (48), it is obvious that the adverb *mían* 'once' is placed in the [Spec, TP] in a pragmatically unmarked order; thus the subject *alís* 'Alis' must be left-dislocated (see (48)a). The same holds in (48)0 and (48)c. The result of this diagnostic shows that the subject is left-dislocated in Romeyka.

3.3.4 Summary

In this section, I have shown that there is one subject position in main and subordinate clauses in Romeyka too. I have specifically shown that in Romeyka subjects in pragmatically unmarked orders are always topics, which are left-dislocated.

3.4 Conclusions

In this chapter, I focused on the word order variation in Romeyka. I specifically examined the respective position of the subject (S), verb (V) and object (O) in main and subordinate declarative and interrogative clauses in Romeyka.

First, I presented the results of my survey of word order variation in matrix and subordinate declarative and interrogative clauses in Romeyka. The findings of this survey show that three word orders are attested in Romeyka, namely SVO, SOV and OSV. On the other hand, V-initial and S-final word orders, i.e. VSO, VOS and OVS, are not attested in Romeyka. Second, I argued that V^0 raises to T^0 in Romeyka. Third, I showed that subjects in pragmatically unmarked orders are left-dislocated.

4 Information structure in Romeyka

4.0 Introduction

In this chapter, (a) I determine the pragmatically unmarked/neutral word order in Romeyka and (b) I examine the syntactic distribution and the semantic type of the constituents in pragmatically marked word orders in Romeyka.

As a matter of fact, I take into consideration cross-cutting generalisations regarding topics, foci and contrastive elements. These jointly motivate the following four-way typology (see Table 26):

Table 26. Neeleman et al. (2009: 15).

	Topic	**Focus**
Non-contrastive	aboutness topic [topic]	(new) information focus [focus]
Contrastive	contrastive topic [topic, contrast]	contrastive focus [focus, contrast]

What Table 26 expresses is that topic and focus are basic notions in information structure that can be enriched to yield a contrastive interpretation. In other words, a contrastive topic and a contrastive focus are an aboutness topic and a (new) information focus interpreted contrastively. A suggestion along these lines can be found in Neeleman et al. (2009), Giusti (2006), McCoy (2003), Molnár (2002) and Vallduví & Vilkuna (1998).

The chapter is structured as follows: in §4.1, I look for the pragmatically unmarked/neutral word order of monotransitive clauses with an overt subject in Romeyka; in §4.2, I investigate the distribution of topics in Romeyka; in §4.3, I examine the distribution of foci in Romeyka; in §4.4, I address the *wh*-questions, multiple *wh*-questions and multiple focus in Romeyka; and in §4.5, I deal with the syntactic distribution of topics and foci in Romeyka. The chapter concludes in §4.6.

4.1 The pragmatically unmarked word order in Romeyka

4.1.0 Introduction

In this section, I look for the pragmatically unmarked word order in Romeyka. The section is structured as follows: in §4.1.1, I determine the notion of the pragmatically unmarked word order; in §4.1.2, I investigate the pragmatically unmarked word order in matrix declarative clauses in Romeyka; and, in §4.1.3, I examine the pragmatically unmarked word order in subordinate declarative clauses in Romeyka. The main findings of the section are summarised in §4.1.4.

4.1.1 The notion of the pragmatically unmarked word order

All of the elements of a clause carry discourse information; discourse information can be either new or given. There are further sub-categories, which I will examine in the remainder of this chapter. The two perhaps most persistent intuitions that researchers have expressed about the new and given information distinction are as follows (see (1)):

(1) a. Question–Answer: The material in the answer that corresponds to the *wh*-constituent in the (constituent) question is focused.
b. Given/New: New material is focused, Given material is not.

I now provide formal implementations of both of these ideas and subsequently discuss the relation between them. I would like to begin with the first intuition, i.e. foci correspond to the *wh*-phrase in a question that precedes (see (2)):

(2) a. Question:
Who kicked the ball?
b. Answer:
[The boy]$_{Foc}$ kicked the ball.

To formalise this, I introduce the notion of a focus value (sometimes called an alternative value or P-set). The focus value for the answer in (2), written as *[the boy]*$_{Foc}$ *kicked the ball* is the set of propositions in (3)a, roughly those expressed by sentences of the form x kicked the ball,

where x is an individual (W is the set of all possible worlds, E the set of all individuals); we will informally write such sets as in (3)b:

(3) Propositions:

a. {{w ∈ W | x kicked the ball in w}|x ∈ E}

b. {x kicked the ball | x an individual}

A question–answer congruence condition makes use of the fact that question meanings, too, can be taken to be sets of propositions, roughly the set of all direct answers. Thus the question in (4) denotes the set of propositions indicated in (4)a, while a question like 'Which boy kicked the ball?, which likewise can be answered by the declarative in (3), denotes the set in (4)b (subscript 'O' indicates that this is the ordinary meaning —as opposed to the focus value— of the expression in brackets):

(4) Propositions:

a. $[\text{Who kicked the ball?}]_O$ = {x kicked the ball | x is a person}

b. $[\text{Which boy kicked the ball?}]_O$ = {x kicked the ball | x is a boy}

c. $[\text{What did the boy kick?}]_O$ = {the boy kicked x | x is an object}

The question 'What did the boy kick?', on the other hand, (3) cannot answer, is interpreted as in (4)c. To derive this pattern the question-answer condition needs to be stated as in (5):

(5) Question–Answer Congruence (QAC):

A is a felicitous answer to Q only if

a. $[Q]o \subseteq [A]f$ and

b. there is no alternative focusing A' of A, which has less F-markings and meets (5)a.

QAC predicts F-markings on complex constituents such as VPs (What did x do?) or clauses (What happened?).

Based on the aforementioned analysis, I propose that new information is per se divided into two major categories:

(6) a. The material in the answer that corresponds to the *wh*-constituent in the (constituent) question is pragmatically marked.

b. New material in any context apart from the one in (a) is pragmatically unmarked.

I therefore assume that a pragmatically marked clause is one in which at least one constituent is the answer that corresponds to the *wh*-constituent in the (constituent) question. On the other hand, I assume that a pragmatically unmarked clause is one in which no constituent is the answer that corresponds to any *wh*-constituent in the (constituent) question. Likewise, a pragmatically marked word order would be the word order of a pragmatically marked clause, while the pragmatically unmarked word order would be the word order of a pragmatically unmarked clause.

4.1.2 The pragmatically unmarked word order in matrix declarative clauses in Romeyka

A pragmatically unmarked order is said to be an 'all-focus sentence', *aka* 'a presentational focus sentence', containing neither old information nor any presuppositions. Several diagnostics can be applied to determine the canonical order in a language. The first one is a 'what happened?' question, which typically invokes a context in which all of the elements of the answer constitute new information and hence are equal in terms of their discourse-pragmatic properties (Büring 2009, van der Wal 2016).

The pragmatically unmarked word order in matrix declarative clauses in Romeyka is VO (see (7)) like in MG (see (8)):

(7) Romeyka:

a. Question:

DO ejéndo?
what.NOM happen.Past.3SG
'What happened?'

b. Answer:

o mustafás epeLÁpsen to χoráfin.
the.NOM Mustafas.NOM put.fertiliser.Past.3SG the.ACC field.ACC
'Mustafas put fertiliser on the field.'
(S01; 150703_0040; 02:16)

(8) Modern Greek:

a. Question:

‘What happened?’

b. Answer:

éspase	ti	lába	o	jánis.
break.Past.3SG	the.ACC	lamp.ACC	the.NOM	Yanis.NOM

‘Yanis broke the lamp.’

(Alexiadou & Anagnstopoulou 2000: 174)

Interestingly, in Romeyka existential constructions with the verb *en* ‘be.3SG/3PL’, the predicative complement always precede the verb *en* (see (9)). The order PC-V must be the result of phonological change, since *en* functions as an enclitic:

(9) Romeyka:

Predicative complements:

a.

atós	o	PÁp^{h}os	m	en.
he	the.NOM	grandfather.NOM	I.POSS	be.3SG

‘This is my grandfather.’

(S07; 812_0074; 00:13)

b.

até	i	inéka	i	MÁna	m	en.
this.NOM	the.NOM	woman.NOM	the.NOM	mother.NOM	I.POSS	be.3SG

‘This woman is my mother.’

(S07; 812_0074; 01:55)

c.

ató	to	zon	i	koS:Ára	m	en.
this.NOM	the.NOM	animal.NOM	the.NOM	hen.NOM	I.POSS	be.3SG

‘This animal is my hen.’

(S07; 812_0074; 04:33)

4.1.3 The pragmatically unmarked word order in subordinate declarative clauses in Romeyka

The pragmatically unmarked word order in subordinate declarative clauses in Romeyka is SOV when the verb is finite (see (10)a) in contrast to VSO in MG (see (11)) and SVO when the verb is an infinitive (see (10)b):

(10) Romeyka:

a. eɣó θaRÓ, alís pol:á ómorfa cʰitʰápæ eχÚjepsen.
I think.1SG Alis.NOM many.ACC nice.ACC books.ACCread.Past.3SG
'I think that Alis read many nice books.'
(S01; 812_0059; 00:20)

b. na mutš Íχa šíta spundžisíni t ospítin,
PRT.MOD NEG have.IMPF.1SG immediately clean.INF the.ACChouse.ACC
n Épezes me ta χómatæ.
PRT.MOD play.IMPF.2SG with the.ACCsolid.ACC
'If I hadn't cleaned the house immediately, you would have played with the soil.'
(S01; 812_0123; 03:32)

(11) Modern Greek:

i maría ípe
the.NOM Maria.NOM say.Past.3SG
óti éfage o yánis ta míla.
that eat.Past.3SG the.NOM Yanis.NOM the.ACC apples.ACC
'Maria said that Yanis ate the apples.'
(Tsimpli 1990: 228)

4.1.4 Summary

The goal of this section was to determine the pragmatically unmarked word order in Romeyka. In particular, I argued that in Romeyka the pragmatically unmarked word order is SVO in matrix declarative clauses and SOV in subordinate declarative clauses.

4.2 Topics in Romeyka

4.2.0 Introduction

Romeyka employs four syntactic strategies to convey given information in the discourse. First, a constituent may be left-dislocated and interpreted either as an aboutness topic (see (12)), or a contrastive topic (see (13)):

(12) Romeyka:

a. Question:

tin	aišén	TS	epíren?
the.ACC	Ayşe.ACC	who.NOM	marry.Past.3SG

'Who married Ayşe?'

b. Answer:

[tin	aišén]$_{\text{A-Top}}$	o	mohaL:Ímis	epíren.
the.ACC	Ayşe.ACC	the.NOM	teacher.NOM	marry.Past.3SG

'The teacher married Ayşe.'

(S01; 140102_0008; 01:10)

(13) Romeyka:

a. Context:

to pont^{h}ólin aLÍs epíren,
the.ACCtrousers.ACC Alis.NOM buy.Past.3SG
to kazáçin o mehMÉtis epíren.
the.ACCsweater.ACC the.NOM Mehmetis.NOM buy.Past.3SG
'Alis bought the trousers and Mehmetis bought the sweater.'

b. Question:

to pont^{h}ólin TS epíren
the.ACCtrousers.ACC who.NOM buy.Past.3SG
tše to kazáçin TS epíren?
and the.ACCsweater.ACC who.NOM buy.Past.3SG
'Who bought the trousers and who the sweater?'

c. Answer:

[to pont^{h}ólin]$_{\text{C-Top}}$ aLÍs epíren,
the.ACCtrousers.ACC Alis.NOM buy.Past.3SG
áma [to kazáçin]$_{\text{C-Top}}$ o mehMÉtis epíren.
but the.ACCsweater.ACC the.NOM Mehmetis.NOM buy.Past.3SG
'Alis bought the trousers, but Mehmetis bought the sweater.'
(S01; 150702_0014; 05:10)

Second, an aboutness topic —but not a contrastive topic— may be yielded through clitic left dislocation (ClLD) with the only clitic attested in Romeyka, i.e. *a* 'him/her/it/them' (see (14)):

(14) Romeyka:

ombrón [ta patsíðæ]$_{\text{A-Top}}$
in.the.past the.ACCgirls.ACC
s ok^{h}úl:in tš epóliɣan æ.
to school.ACC NEG send.IMPF.3PL them
'In the past, they did not send the girls to school.'
(S01; 150702_0019; 03:23)

Third, a topic particle, i.e. *pa(l)*, can assign contrastive (but not aboutness) topichood to the constituent with which it is associated, as illustrated in (15):

(15) Romeyka:

asón alín habérin tš éχo,
from.the.ACC Alis.ACC news.ACC NEG have.1SG
áma [i ailín pa]$_{\text{C-Top}}$ TIP tš éfaen.
but the.NOM Aylin.NOM PRT nothing NEG eat.Past.3SG
'I don't know about Alis, but Aylin didn't eat anything.'
(S01; 150702_0023; 06:41)

Fourth, (non-contrastive) given information may appear postverbally (see (16)):

(16) Romeyka:

a. Question:
PÍos epíren tin aišén?
who.NOM marry.Past.3SG the.ACCAyşe.ACC
'Who married Ayşe?'

b. Answer:
o dohTÓris epíren tin aišén.
the.NOM doctor.NOM marry.Past.3SG the.ACCAyşe.ACC
'The doctor married Ayşe.'
(S01; 140102_0008; 01:25)

This section is structured as follows: §4.2.1 examines aboutness topics in Romeyka; §4.2.2 examines contrastive topics in Romeyka; and §4.2.3 examines ClLD in Romeyka; The main findings of the section are summarised in §4.2.4.

4.2.1 Aboutness topics in Romeyka

The aboutness topic is identified in the literature as the constituent representing the theme of the predication, i.e. what the sentence is about (see Frascarelli & Hinterhölzl 2007).

An aboutness topic in Romeyka is yielded by left dislocation and can involve any type of phrase. For instance, any constituent can be an aboutness topic: a personal pronoun (see (17)b), a demonstrative pronoun (see (18)b), a subject NP (see (19)b), a subject DP (see (20)b and (21)b) and a direct object DP (see (22)b and (23)b):

(17) Romeyka:

a. Question:

esí	DO	epítšes?
you.NOM	what.ACC	do.Past.2SG

'What did you do?'

b. Answer:

[eɣó]$_{\text{A-Top}}$	ta	ŠCÉvæ	éplisa.
I	the.ACC	dishes.ACC	wash.Past.1SG

'I washed the dishes.'

(S01; 150703_0040; 00:19)

(18) Romeyka:

a. Question:

até	DO	epítšen?
she	what.ACC	do.Past.3SG

'What did she do?'

b. Answer:

[até]$_{\text{A-Top}}$	Éɣraften.
she	write.IMPF.3SG

'She was writing.'

(S01; 150703_0040; 00:24)

(19) Romeyka:

a. Question:

éna	líkon	DO	epítšen?
a.NOM	wolf.NOM	what.ACC	do.Past.3SG

'What did a wolf do?'

b. Answer:

[éna	líkon]$_{\text{A-Top}}$	tin	koS:Áran	éfaen.
a.NOM	wolf.NOM	the.ACC	hen.ACC	eat.Past.3SG

'A wolf ate the hen.'

(S01; 150703_0040; 01:41)

(20) Romeyka:

a. Question:

i	aišé	DO	epítšen?
the.NOM	Ayşe.NOM	what.ACC	do.Past.3SG

'What did Ayşe do?'

b. Answer:

[i	aišé]$_{A\text{-}Top}$	t	oSPÍtin	espúndžisen.
the.NOM	Ayşe.NOM	the.ACC	house.ACC	clean.Past.3SG

'Ayşe cleaned the house.'

(S01; 150703_0040; 07:54)

(21) Romeyka:

a. Question:

o	dohtóris	DO	epítšen?
the.NOM	doctor.NOM	what.ACC	do.Past.3SG

'What did the doctor do?'

b. Answer:

[o	dohtóris]$_{A\text{-}Top}$	ameLIÁtin	epítšen.
the.NOM	doctor.NOM	surgery.ACC	make.Past.3SG

'The doctor performed a surgery.'

(S01; 150703_0040; 01:36)

(22) Romeyka:

a. Question:

tin	aišén	TS	epíren?
the.ACC	Ayşe.ACC	who.NOM	marry.Past.3SG

'Who married Ayşe?'

b. Answer:

[tin	aišén]$_{A\text{-}Top}$	aLÍS	epíren.
the.ACC	Ayşe.ACC	Alis.NOM	marry.Past.3SG

'Alis married Ayşe.'

(S01;140102_0008; 01:03)

(23) Romeyka:

a. Question:

to	tšáin	TS	epítšen?
the.ACC	tea.ACC	who.NOM	make.Past.3SG

'Who made the tea?'

b. Answer:

[to	tšáin]$_{A\text{-}Top}$	i	MÁna	m	epítšen.
the.ACC	tea.ACC	the.NOM	mother.NOM	I.POSS	make.Past.3SG

'My mother made the tea.'

(S01; 150703_0040; 01:28)

Based on the previous data, if a phrase XP is an aboutness topic, then it is the leftmost element of the clause, as illustrated in (24):

(24) $XP_{A\text{-}Top} > TP$

4.2.2 Contrastive topics in Romeyka

A context that favours an interpretation of a constituent as a contrastive topic is the one in which the hearer answers a question that differs from the one being asked (see Büring 2003, 2009).

A contrastive topic in Romeyka is yielded by left dislocation and can involve any type of phrase. For instance, a subject (DP) (see (25)b and (26)c) and a direct object (DP) ((27)c) can be a contrastive topic:

(25) Romeyka:

a. Question:

o	mehmétis	d	epítšen
the.NOM	Mehmetis.NOM	what.ACC	do.Past.3SG

tš	esí	d	epítšes?
and	you.NOM	what.ACC	do.Past.2SG

'What did Mehmetis do and what did you do?'

b. Answer:

[o	mehmétis]$_{\text{C-Top}}$	ton	bahtsén	epéleipsen
the.NOM	Mehmetis.NOM	the.ACC	garden.ACC	put.fertiliser.Past.3SG

tš	[eɣó]$_{\text{C-Top}}$	éskapsa	ton.
and	I	dig.Past.1SG	it.ACC

'Mehmetis put fertiliser on the garden and I dug it.'

(S01; 150702_0022; 09:49)

(26) Romeyka:

a. Context:

alís	éfaen	éna	mílon,
Alis.NOM	eat.Past.3SG	a.ACC	apple.ACC

o	mehmétis	éfaen	énan	ap^{h}íðin.
the.NOM	Mehmetis.NOM	eat.Past.3SG	a.ACC	pear.ACC

'Alis ate an apple and Mehmetis ate a pear.'

b. Question:

alís	do	éfaen,
Alis.NOM	what.ACC	eat.Past.3SG

tš	o	mehmétis	do	éfaen?
and	the.NOM	Mehmetis.NOM	what.ACC	eat.Past.3SG

'What did Alis eat and what did Mehmetis eat?'

c. Answer:

[alís]$_{C\text{-}Top}$	éfaen	éna	mílon,
Alis.NOM	eat.Past.3SG	a.ACC	apple.ACC

[o	mehmétis]$_{C\text{-}Top}$	éfaen	énan	ap^{h}íðin.
the.NOM	Mehmetis.NOM	eat.Past.3SG	a.ACC	pear.ACC

'Alis ate an apple and Mehmetis ate a pear.'

(S01; 150703_0040; 03:21)

(27) Romeyka:

a. Context:

to	líkon	íðes	s	óros,		
the.ACC	wolf.ACC	see.Past.2SG	in.the.ACC	forest.ACC		
ton	árkon	íðes	s	alátæ	pu	ka.
the.ACC	bear.ACC	see.Past.2SG	in	trees.ACC	from	under

'You saw the wolf in the forest and the bear under the trees.'

b. Question:

pútšeka	íðes	to	líkon
where	see.Past.2SG	the.ACC	wolf.ACC

tše	pútšeka	íðes	ton	árkon?
and	where	see.Past.2SG	the.ACC	bear.ACC

'Where did you see the wolf and where did you see the bear?'

c. Answer:

[to	líkon]$_{C\text{-}Top}$	íða	s	óros,		
the.ACC	wolf.ACC	see.Past.1SG	in.the.ACC	forest.ACC		
[ton	árkon]$_{C\text{-}Top}$	íða	s	alátæ	pu	ka.
the.ACC	bear.ACC	see.Past.1SG	in	trees.ACC	from	under

'I saw the wolf in the forest and the bear under the trees.'

(S01; 150703_0040; 04:49)

Based on the previous data, if a phrase XP is a contrastive topic, then it is placed in the left periphery, as illustrated in (28):

(28) $XP_{C\text{-}TopP} > TP$

4.2.3 Clitic Left Dislocation (ClLD) in Romeyka

In Romeyka, ClLD does not have the same pragmatic import as in MG. While in MG a left dislocated constituent is interpreted as a topic if and only if it is ClLD'ed (see (29)a), otherwise it is interpreted as a focus (see (29)b), in Romeyka a left dislocated constituent can be interpreted as a topic even if it is not ClLD'ed (see a ClLD'ed topic in (30)a and a non ClLD'ed one in (30)b):

(29) Modern Greek:

a. [to jáni]$_{Top}$, ton sinádisa χθes.
the.ACC Yanis.ACC he.ACC meet.Past.1SG yesterday
'I met Yanis yesterday.'

b. [to jáni]$_{Foc}$, (*ton) sinádisa χθes.
the.ACC Yanis.ACC he.ACC meet.Past.1SG yesterday
'It is Yanis that I met yesterday.'
(Tsimpli 1995: 179)

(30) Romeyka:

a. [ta patátes]$_{A\text{-}Top}$ zimónum æ.
the.ACC potatoes.ACC knead.1PL them
'We knead the potatoes.'
(S01; 150702_0019; 05:52)

b. [ta patátes]$_{A\text{-}Top}$ zimónum.
the.ACC potatoes.ACC knead.1PL
'We knead the potatoes.'
(S01; 150702_0019; 06:25)

4.2.4 Summary

In this section, I have shown that Romeyka employs four syntactic strategies to convey given information in the discourse. First, a constituent may be left-dislocated and interpreted either as an aboutness or a contrastive topic. Second, aboutness topic —but not a contrastive topic— may be yielded through clitic left dislocation (ClLD) with the only clitic attested in Romeyka, i.e. *a* 'him/her/it/them'. Third, a topic particle, i.e. *pa(l)*, can assign contrastive (but not aboutness) topichood to the constituent with which it is associated. Fourth, (non-contrastive) given information may appear postverbally.

4.3 Foci in Romeyka

4.3.0 Introduction

In this section, I discuss the syntactic distribution of the focused constituents in Romeyka. I precisely examine variation in focus in terms of two parameters; first, I investigate the size of the focus; and second, I discuss the semantic-pragmatic type of focus. The size of the focus can comprise one or multiple arguments, adjuncts or verbs and can vary from just a nominal argument to a whole verb phrase. The type of focus establishes the semantic-pragmatic interpretation of a certain linguistic strategy that is suspected to express focus.

This section is structured as follows: in §4.3.1, I present the definition of focus that I follow in this study; §4.3.2 examines information foci in Romeyka; and, §4.3.3 examines contrastive foci in Romeyka. The main findings of the section are summarised in §4.3.4.

4.3.1 Focus

Throughout this study, I adopt the semantic definition of focus proposed by Rooth's (1985, 1992, 1996) Alternative Semantics, which states that focus "indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions" (Krifka 2007: 6).

According to Rooth (1985, 1992, 1996), the sentence in (31), in which the DP-subject *the boy* is focused, is associated with two semantic objects: on the one hand, there is the proposition expressed by the sentence —the set of possible worlds in (32)a. I will talk about this proposition informally as in (32)b:

(31) [The boy]$_{Foc}$ kicked the ball.

(32) Propositions:
  a. λw.the boy kicked the ball in w
  b. that the boy kicked the ball

Besides the ordinary semantic value of (31), the sentence makes salient a set of alternative propositions —for example the set in (33)a, which contains alternative propositions to the proposition *that the boy kicked the ball*. This is the focus-semantic value of the sentence, rendered more generally in (33)b and in the form of a semi-logical expression in (33)c:

(33) Alternative propositions:

a. {that x kicked the ball | x is an individual}

b. λp ∃x [p = λw.x kicked the ball in w]

c. {that the boy kicked the ball, that the girl kicked the ball, that John kicked the ball, …}

4.3.2 Information foci in Romeyka

The information focus asserts the membership of an individual in a set (see Gundel 1998). As I have already mentioned, the most widespread and accepted test for focus and a method of establishing the scope of focus is *wh*-questions and their answers (Beaver & Clark 2008, Krifka 2007, Lambrecht 1994, Roberts 1996, Rooth 1992, van der Wal 2016, i.a.). The basic idea is that a *wh*-question always yields new information. If focus is defined as the new information in a sentence, then it follows that the phrase that replaces the *wh*-constituent is focused.

Consider the sentences in (34) and (36) from Romeyka and the equivalent ones in (35) and (37) from MG:

(34) Romeyka:

a. Question:

alís DÓɣna éfaen?

Alis.NOM what.ACC eat.Past.3SG?

'What did Alis eat?'

Answers:

b. alís [χaVÍts]$_{\text{I-Foc}}$ éfaen.

Alis.NOM pudding.ACC eat.Past.3SG

'Alis ate a pudding.'

(S01; 150703_0040; 07:14)

c. #alís éfaen [χaVÍts]$_{\text{I-Foc}}$.

Alis.NOM eat.Past.3SG pudding.ACC

'Alis ate a pudding.'

(35) Modern Greek:

a. Question:

ti	éfaɟe	o	ɟóryos?
what.ACC	eat.Past.3SG	the.NOM	George.NOM

'What did George eat?'

b. Answer:

o	ɟóryos	éfaɟe	[tin	kobósta]$_{\text{I-Foc}}$.
the.NOM	Geroge.NOM	eat.Past.3SG	the.ACC	stewed-fruit.ACC

'George ate the stewed fruits.'

(Sitaridou & Kaltsa 2014: 12)

(36) Romeyka:

a. DÓ eχúɟepsen?

what.ACC read.Past.3SG

'What did s/he read?'

Answers:

b. [pol:á c^{h}iTÁpæ]$_{\text{I-Foc}}$ eχúɟepsen.

many.ACC books.ACC read.Past.3SG

'S/He read many books.'

(S01; 812_0059; 00:10)

c. #eχúɟepsen [pol:á c^{h}iTÁpæ]$_{\text{I-Foc}}$.

read.Past.3SG many.ACC books.ACC

'S/He read many books.'

(37) Modern Greek:

a. Question:

ti ðjávase?

what.ACC read.Past.3SG

'What did s/he read?'

Answers:

b. ðjávase [polá vivlía]$_{\text{I-Foc}}$.

read.Past.3SG many.ACC books.ACC

'S/He read many books.'

c. ?[polá vivlía]$_{\text{I-Foc}}$ ðjávase.

many.ACC books.ACC read.Past.3SG

'S/He read many books.'

(Sitaridou & Kaltsa 2014: 12)

From the examples in (34) and (36), it becomes obvious that Romeyka allows for information focus to the left of the verb. The degree of diversification of the Romeyka pattern from the MG attenuates if we consider Gryllia's (2008) findings, on the basis of experimental tests, which show that preverbal objects are not always either exhaustive or contrastive in MG. In both positions, VO and OV, the focused direct object is interpreted as a new information focus, as in (38)b and (38)c, respectively:

(38) Modern Greek:

a. Question:

ti	χárise	metaksí	álon	o	jánis
what.ACC	give.Past.3SG	among	others.GEN	the.NOM	Yanis.NOM

stin	iléktra?
to.the.ACC	Ilektra.ACC

'Among other things, what did Yanis give to Ilektra?'

Answers:

b. χárise [éna vivlío]$_{\text{I-Foc}}$ stin iléktra.
give.Past.3SG a.ACC book.ACC to.the.ACC Ilektra.ACC
'He gave a book to Ilektra.'

c. [éna vivlío]$_{\text{I-Foc}}$ χárise stin iléktra.
a.ACC book.ACC give.Past.3SG to.the.ACC Ilektra.ACC
'He gave a book to Ilektra.'

(Gryllia 2008: 21)

Nevertheless, the fact that MG may allow for either option does not alter the parametric difference with Romeyka, where the preverbal position is the only option. See the following judgements made by Romeyka speakers in (39) and (40):

(39) Romeyka:

a. Question:

ánda erotó se alís DÓɣna éfaen
if ask.1SG you.ACC Alis.NOM what.ACC eat.Past.3SG
esí léjis me, o alís éfaen MÍla
you.NOM say.2SG I.ACC the.NOM Alis.NOM eat.Past.3SG apples.ACC
jóksa, alís MÍla éfaen?
or Alis.NOM apples.ACC eat.Past.3SG
'If I ask you, what did Alis eat, what do you say to me? Alis ate apples, or Alis ate apples?'

b. Answer:

kalːíon, alís [MÍla]$_{\text{I-Foc}}$ éfaen.
better Alis.NOM apples.ACC eat.Past.3SG
'Alis ate apples, sounds better.'
(S01; 812_0055; 03:09)

(40) Romeyka:

a. Question:

ánda léɣo, alís DÓɣna éfaen?
if say.1SG Alis.NOM what.ACC eat.Past.3SG
esí léjis me, alís MÍla éfaen.
you.NOM say.2SG I.ACC Alis.NOM apples.ACC eat.Past.3SG
ánda léɣo, éfaen MÍla alís,
if say.1SG eat.Past.3SG apples.ACC Alis.NOM
émorfon en?
good.NOM be.3SG
alís MÍla éfaen,
Alis.NOM apples.ACC eat.Past.3SG
MÍla éfaen alís?
apples.ACC eat.Past.3SG Alis.NOM
'If I say, what did Alis eat? You say to me, Alis ate apples. If I say, Alis ate apples, does this sound good? Alis ate apples or Alis ate apples?'

b. Answer:

jok^{h}, alís [MÍla]$_{\text{I-Foc}}$ éfaen.
no Alis.NOM apples.ACC eat.Past.3SG
'No, Alis ate apples.'
(S01; 812_0055; 02:31)

The Romeyka pattern is reminiscent of what has recently been claimed about information focus, namely that it also commonly appears within the left periphery (see Sitaridou & Kaltsa 2014). No matter the type of phrase, in Romeyka any focused phrase appears before the verb: object in a focused predicate (see (41)), subject (NP) (see (42)), subject (DP) (see (43)), direct object (NP) (see (44)), direct object (DP) (see (45)), indirect object (Beneficiary) (DP) (see (46)), predicative complement (see (47)), prepositional phrase (see (48)), adverbial (see (49)) and quantifier (see (50)):

(41) Romeyka:

Object in a focused predicate:

a. Question:

o jusúfis DO eftǽi?

the.NOM Yusufis.NOM what.ACC do.3SG

'What is Yusufis doing?'

b. Answer:

o jusúfis [bilɟÍsajarin pez]$_{\text{I-Foc}}$.

the.NOM Yusufis.NOM computer.ACC play.3SG

'Yusufis is playing on the computer.'

(S01; 150702_001; 00:16)

(42) Romeyka:

Subject (NP) is focused:

a. Question:

tin aišén TS epíren?

the.ACCAyşe.ACC who.NOM marry.Past.3SG

'Who married Ayşe?'

b. Answer:

tin aišén [énan Áɣuros]$_{\text{I-Foc}}$ epíren.

the.ACCAyşe.ACC a.NOM young.man.NOMmarry.Past.3SG

'A young man married Ayşe.'

(S01; 140102_0008; 02:12)

(43) Romeyka:

Subject (DP) is focused:

a. Question:

tin	aišén	TS	epíren?
the.ACC	Ayşe.ACC	who.NOM	marry.Past.3SG

'Who married Ayşe?'

b. Answer:

tin	aišén	[aLÍS]$_{\text{I-Foc}}$	epíren.
the.ACC	Ayşe.ACC	Alis.NOM	marry.Past.3SG

'Alis married Ayşe.'

(S01; 140102_0008; 01:03)

(44) Romeyka:

Direct object (NP) is focused:

a. Question:

alís	DÓɣna	éfaen?
Alis.NOM	what.ACC	eat.Past.3SG

'What did Alis eat?'

b. Answer:

alís	[χaVÍtsin]$_{\text{I-Foc}}$	éfaen.
Alis.NOM	pudding.ACC	eat.Past.3SG

'Alis ate a pudding.'

(S01; 150703_0040; 07:14)

(45) Romeyka:

Direct object (DP) is focused:

a. Question:

i	aišé	TÍnan	epíren?
the.NOM	Ayşe.NOM	who.ACC	marry.Past.3SG

'Who did Ayşe marry?'

b. Answer:

i	aišé	[ton	dohTÓrin]$_{\text{I-Foc}}$	epíren.
the.NOM	Ayşe.NOM	the.ACC	doctor.ACC	marry.Past.3SG

'Ayşe married the doctor.'

(S01; 140102_0008; 01:37)

(46) Romeyka:

Indirect object (Beneficiary) (DP) is focused:

a. Question:

to	c^{h}itápin	TÍnan	éndžes?
the.ACC	book.ACC	who.ACC	bring.Past.2SG

'To whom did you give the book?'

b. Answer:

to	c^{h}itápin	[ton	ǰuSÚfin]$_{\text{I-Foc}}$	éŋga.
the.ACC	book.ACC	the.ACC	Yusufis.ACC	bring.Past.1SG

'I brought the book for Yusufis.'

(S01; 150703_0042; 00:54)

(47) Romeyka:

Predicative complement is focused:

a. Question:

o	šcʰíl:on	DO	en?
the.NOM	dog.NOM	what.NOM	be.3SG

'What is the dog?'

b. Answer:

[haiVÁnin]$_{\text{I-Foc}}$	en.
animal.NOM	be.3SG

'It's an animal.'

(S01; 140102_0009; 00:35)

(48) Romeyka:

Prepositional phrase is focused:

a. Question:

LÁɣana	páɣo	so	ɟunanistánin?	
how	PRT.MOD	go.1SG	to.the.ACC	Greece.ACC

'How can I go to Greece?'

b. Answer:

[me	to	uTŠAçin]$_{\text{I-Foc}}$	na	pas	so	ɟunanistánin.
by	the.ACC	airplane.ACC	PRT.MOD	go.2SG	to.the.ACC	Greece.ACC

'You can get to Greece by plane.'

(S01; 150702_0013; 10:38)

(49) Romeyka:

Adverbial phrase is focused:

a. Question:

i	mána	s	PÓte	efáisen	ton	musafírin?
the.NOM	mother.NOM	you.POSS	when	feed.Past.3SG	the.ACC	guest.ACC

'When did your mother feed the guest?'

b. Answer:

[oPSÉ]$_{\text{I-Foc}}$	efáisen	ton	musafírin.
yesterday	feed.Past.3SG	the.ACC	guest.ACC

'She fed the guest yesterday.'

(S01; 150703_0041; 07:10)

(50) Romeyka:

Quantifier is focused:

a. Question:

PÍos	epíen	so	džamín?
who.NOM	go.Past.3SG	to.the.ACC	mosque.ACC

'Who went to the mosque?'

b. Answer:

[Úl:ini]$_{\text{I-Foc}}$	epíɣane	so	džamín.
everyone.NOM	go.Past.3PL	to.the.ACC	mosque

'Everyone went to the mosque.'

(S01; 150702_0013; 20:03)

In the question 'what is x doing?', the verb phrase of the answer would be focused ("VP focus", "predicate focus"). Romeyka employs OV to focus VPs with monotransitive verbs, as illustrated in example (51):

(51) Romeyka:

a. Question:

o jusúfis DO eftæi?

the.NOM Yusufis.NOM what.ACC do.3SG

'What is Yusufis doing?'

b. Answer:

o jusúfis [biɟÍsajarin pez]$_{\text{I-Foc}}$.

the.NOM Yusufis.NOM computer.ACC play.3SG

'Yusufis is playing on the computer.'

(S01; 150702_0013; 00:16)

Furthermore, if the verb is intransitive, it can still be focused (see (52)):

(52) Romeyka:

a. Question:

até DO epítšen?

she.NOM what.ACC do.Past.3SG

'What did she do?'

b. Answer:

até [eɣRÁften]$_{\text{I-Foc}}$.

she.NOM write.IMPF.3SG

'She was writing.'

(S01; 150703_0040; 00:24)

Based on the previous data, if a VP is focused, then the verb stays *in-situ*, i.e. in T^0; if the verb is transitive, the object of the verb is placed to the left of the verb, as it is illustrated in (53):

(53) DP/NP-object > [TP verb$_{\text{I-Foc}}$]

In the question 'who VP?', the subject of the answer would be focused ("argument focus"). Romeyka displays SV to focus the subject. SV is always attested no matter what the syntactic type or semantic properties of the subject are.

First, the focused subject can be a personal pronoun (see (54)):

(54) Romeyka:

a. Question:

PÍos éfaen ta χapsíæ?

who.NOM eat.Past.3SG the.ACCanchovies.ACC

'Who ate the anchovies?'

b. Answer:

[eɣÓ]$_{\text{I-Foc}}$ éfaɣa ta χapsíæ.

I.NOM eat.Past.1SG the.ACCanchovies.ACC

'I ate the anchovies.'

(S01; 812_0058: 00:06)

Second, the subject may be a demonstrative pronoun (see (55)):

(55) Romeyka:

a. Question:

PÍos eðótšen tin kos:áran?

who.NOM give.Past.3SG the.ACChen.ACC

'Who gave the hen?'

b. Answer:

[aTÉ]$_{\text{I-Foc}}$ eðótšen me tin kos:áran.

she.NOM give.Past.3SG I.ACC the.ACChen.ACC

'She gave me the hen.'

(S01; 812_0093; 0:08)

Third, it could also be an indefinite pronoun (see (56)):

(56) Romeyka:

a. Question:

tin aišén TS epíren?

the.ACCAyşe.ACC who.NOM marry.Past.3SG

'Who married Ayşe?'

Answers:

b. [IS]$_{\text{I-Foc}}$ epíren.

someone.NOM marry.Past.3SG

'Someone married her.'

(S01; 140102_0008; 04:00)

c. [IS]$_{\text{I-Foc}}$ epíren æ.

someone.NOM marry.Past.3SG she.ACC

'Someone married her.'

(S01; 140102_0008; 04:01)

Fourth, the subject may be an NP (see (57)):

(57) Romeyka:

a. Question:

tin aišén TS epíren?
the.ACCAyşe.ACC who.NOM marry.Past.3SG
'Who married Ayşe?'

Answers:

b. tin aišén [énan Áɣuros]$_{\text{I-Foc}}$ epíren.
the.ACCAyşe.ACC a.NOM young.man.NOMmarry.Past.3SG
'A young man married Ayşe.'
(S01; 140102_0008; 02:12)

c. tin aišén [éna mohaL:Ímis]$_{\text{I-Foc}}$ epíren.
the.ACCAyşe.ACC a.NOM teacher.NOM marry.Past.3SG
'A teacher married Ayşe.'
(S01; 140102_0008; 02:19)

d. tin aišén [éna dohTÓris]$_{\text{I-Foc}}$ epíren.
the.ACCAyşe.ACC a.NOM doctor.NOM marry.Past.3SG
'A doctor married Ayşe.'
(S01; 140102_0008; 02:26)

e. tin aišén [is dohTÓris]$_{\text{I-Foc}}$ epíren.
the.ACCAyşe.ACC a.NOM doctor.NOM marry.Past.3SG
'A doctor married Ayşe.'
(S01; 140102_0008; 02:36)

f. tin aišén [énan Árθepos]$_{\text{I-Foc}}$ epíren.
the.ACCAyşe.ACC a.NOM man.NOM marry.Past.3SG
'A man married Ayşe.'
(S01; 140102_0008; 02:53)

Fifth, the subject may be a DP (see (58)):

(58) Romeyka:

a. Question:

džumártesi	PÍos	epíjen	so	parχárin?
Saturday	who.NOM	go.Past.3SG	to.the.ACC	pastures.ACC

'Who went to the pastures on Saturday?'

b. Answer:

džumártesi	[i	zeiNÉP]$_{\text{I-Foc}}$	epíjen	so	parχárin.
Saturday	the.NOM	Zeynep.NOM	go.Past.3SG	to.the.ACC	pastures.ACC

'Zeynep went to the pastures on Saturday.'

(S01; 150702_0013; 20:43)

Based on the previous data, if a subject is focused, no matter what the type of subject is, it is placed to the left of the verb, as illustrated in (59):

(59) DP/NP-subject$_{\text{I-Foc}}$ > TP

In the question 'whom/what VP?', the object of the answer would be focused ("argument focus"). Romeyka displays OV to focus the object. OV is always attested, no matter what the syntactic type or the semantic properties of the object are.

First, the focused object can be a demonstrative pronoun (see (60)):

(60) Romeyka:

a. Question:

i aišé TÍnan epíren?
the.NOM Ayşe.NOM who.ACC marry.Past.3SG
'Who did Ayşe marry?'

Answers:

b. i aišé [aTŠÓnan]$_{\text{I-Foc}}$epíren.
the.NOM Ayşe.NOM this.ACC marry.Past.3SG
'Ayşe married this one.'
(S01; 140102_0008; 07:24)

c. [aTÓnan]$_{\text{I-Foc}}$ epíren.
this.ACC marry.Past.3SG
'She married this one.'
(S01; 140102_0008; 07:30)

Second, the focused object can be an NP (see (61)):

(61) Romeyka:

a. Question:

i aišé TÍnan epíren?
the.NOM Ayşe.NOM who.ACC marry.Past.3SG
'Who did Ayşe marry?'

b. Answer:

i aišé [énan Áγuron]$_{\text{I-Foc}}$ epíren.
the.NOM Ayşe.NOM a.ACC man.ACC marry.Past.3SG
'Ayşe married a man.'
(S01; 140102_0008; 07:03)

Third, the focused object can be a DP (see (62)):

(62) Romeyka:

a. Question:

i	aišé	TÍnan	epíren?
the.NOM	Ayşe.NOM	who.ACC	marry.Past.3SG

'Who did Ayşe marry?'

Answers:

b.

i	aišé	[ton	dohTÓrin]$_{\text{I-Foc}}$	epíren.
the.NOM	Ayşe.NOM	the.ACC	doctor.ACC	marry.Past.3SG

'Ayşe married the doctor.'

(S01; 140102_0008; 01:37)

c.

i	aišé	[ton	aLÍN]$_{\text{I-Foc}}$	epíren.
the.NOM	Ayşe.NOM	the.ACC	Alis.ACC	marry.Past.3SG

'Ayşe married Alis.'

(S01; 140102_0008; 06:21)

Based on the previous data, if an object is focused, then the object is placed to the left of the verb, as illustrated in (63):

(63) DP/NP-object$_{\text{I-FocP}}$ > TP

In the question 'what is x?', the predicative complement that answers to 'what' in the question would be focused. Romeyka displays PC > VP to focus the predicative complement. PC > VP is always attested, no matter what the syntactic type or the semantic properties of the predicative complement are (see (64)):

(64) Romeyka:

a. Question:

alís DO en?

Alis.NOM what.NOM be.3SG

'What is Alis?'

Answers:

b. alís [Ándras m]$_{\text{I-Foc}}$ en.

Alis.NOM husband.NOM I.POSS be.3SG

'Alis is my husband.'

(S01; 140102_0009; 00:16)

c. alís [Áɣuros]$_{\text{I-Foc}}$ en.

Alis.NOM boy.NOM be.3SG

'Alis is a boy.'

(S01; 140102_0009; 00:20)

d. alís [ɣarðÉlin]$_{\text{I-Foc}}$ en.

Alis.NOM child.NOM be.3SG

'Alis is a child.'

(S01; 140102_0009; 00:09)

Based on the previous data, if a predicative complement is focused, then the focused predicative complement is placed to the left of the verb, as illustrated in (65):

(65) PC$_{\text{I-Foc}}$ > TP

In the question 'where VP?', the prepositional phrase that answers to 'where' would be focused. Romeyka displays PP > VP to focus the prepositional phrase. PP > VP is always attested, no matter what the syntactic type or the semantic properties of the prepositional phrase are (see (66)):

(66) Romeyka:

a. Question:

i aišé opsé PÚtšeka epíen?
the.NOM Ayşe.NOM yesterday where go.Past.3SG
'Where did Ayşe go yesterday?'

Answers:

b. i aišé opsé [s oK^{H}Úl:in]$_{I\text{-}Foc}$ epíen.
the.NOM Ayşe.NOM yesterday to school.ACC go.Past.3SG
'Ayşe went to school yesterday.'
(S01; 140102_0009; 01:23)

c. i aišé opsé [so istaMBÓlin]$_{I\text{-}Foc}$ epíen.
the.NOM Ayşe.NOM yesterday to.the.ACC Istanbul.ACC go.Past.3SG
'Ayşe went to Istanbul yesterday.'
(S01; 140102_0009; 01:28)

d. i aišé opsé [sa staLÍæ]$_{I\text{-}Foc}$ epíen.
the.NOM Ayşe.NOM yesterday to.the.ACC stalls.ACC go.Past.3SG
'Ayşe went to the stalls yesterday.'
(S01; 140102_0009; 01:32)

e. i aišé opsé [sa parχÁræ,
the.NOM Ayşe.NOM yesterday to.the.ACC pastures.ACC
sa staLÍæ]$_{I\text{-}Foc}$ epíen.
to.the.ACC stalls.ACC go.Past.3SG
'Ayşe went to the pastures and to the stalls yesterday.'
(S01; 140102_0009; 01:39)

Based on the previous data, if a prepositional phrase is focused, then the prepositional phrase is placed to the left of the verb, as illustrated in (67):

(67) $PP_{I\text{-}Foc}$ > TP

In the question 'how/when VP?', the adverbial phrase that answers to 'how/when' would be focused. Romeyka displays AdvP > VP to focus the adverbial phrase. AdvP > VP is always attested, no matter what the syntactic type or the semantic properties of the adverbial phrase are (see (68)):

(68) Romeyka:

a. Question:

alís LÁγa efílisen tin aišén?
Alis.NOM how kiss.Past.3SG the.ACC Ayše.ACC
'How did Alis kiss Ayşe?'

b. Answer:

alís [ŠÍta]$_{\text{I-Foc}}$ efílisen tin aišén.
Alis.NOM immediately kiss.Past.3SG the.ACC Ayşe.ACC
'Alis immediately kissed Ayše.'
(S01; 150702_0022; 04:07)

Based on the previous data, if an AdvP is focused, then the AdvP is placed to the left of the verb, as illustrated in (69):

(69) AdvP$_{\text{I-Foc}}$ > TP

The opposite of an exhaustive focus is when the answer necessarily or typically has more than one referent for which the proposition can be true. This can be tested when explicitly asking for a non-exhaustive answer. An exhaustive focus strategy is infelicitous (in questions and answers) if an exhaustive answer to the question is impossible or highly implausible for pragmatic reasons. This is not the case in Romeyka, as shown in (70) and (71):

(70) Romeyka:

a. Question:

ap Éndžeka na páo
from where PRT.MOD go.1SG
n aɣoráso gazéte s ató to χoríon?
PRT.MOD buy.1SG newspaper.ACC in this.ACC the.ACC village.ACC
'Where will I go to buy a newspaper from this village?'

b. Answer:

[so $c^{h}iC^{H}$Énin]$_{I\text{-}Foc}$ na pas
to.the.ACC grocery's.ACC PRT.MOD go.2SG
n aɣoráis gazet^{h}éðes.
PRT.MOD buy.2SG newspapers.ACC
'You will go to the grocery's to buy newspapers.'
(S01; 150702_0013; 15:25)

(71) Romeyka:

a. Question:

PÚtšeka en ta ómorfa ta méræ
where be.PL the.NOM nice.NOM the.NOM places.NOM
s ató to χoríon na elépo ata?
in this.ACC the.ACC village.ACC PRT.MOD see.1SG them
'Where are the most beautiful places to see in this village?'

Answers:

b. [si maziRÁN]$_{\text{I-Foc}}$ en ómorfon
to.the.ACC Mazira.ACC be.3SG nice
na pas elépis æ.
PRT.MOD go.2SG see.2SG it.ACC
'It's nice to go to see Mazira.'
(S01; 150702_0013; 15:38)

c. [ta staLÍæ]$_{\text{I-Foc}}$ en ómorfa
the.ACC stalls.ACC be.3PL nice
na pas elépis æ.
PRT.MOD go.2SG see.2SG it.ACC
'It's nice to go to see the pastures.'
(S01; 150702_0013; 15:41)

d. [ta kaTÚnes]$_{\text{I-Foc}}$ en ómorfa
the.ACC Katunes.ACC be.3PL nice
na pas elépis ata.
PRT.MOD go.2SG see.2SG they.ACC
'It's nice to go to see Katunes.'
(S01; 150702_0013; 15:44)

The use of quantifiers as focus tests is largely based on their entailment scales. Tests have been found that make use of quantifiers to establish the type of focus interpretation (exclusive/non-exclusive) that a certain strategy expresses. Naturally, whether these tests can be used depends on the existence of quantifiers and indefinites in the language (van der Wal 2016: 290).

First, numerals become exact. Numerals normally have an underspecified interpretation either as the given quantity, or as a lower boundary, at least this amount (Horn 1972, Levinson 2000). However, in (exhaustive) focus numerals refer only to the exact quantity. É. Kiss (2010)

shows this interpretation for the preverbal focus position in Hungarian. When a focus numeral follows the verb, or is topicalised in Hungarian, we get the lower-bound reading, but in the directly preverbal focus position, the meaning narrows down to only the value given in the focused constituent, that is, exactly the numeral. The latter appears in Romeyka as well (see (72)):

(72) Romeyka:

alís	s	éna	mína	[BIN	lirá]$_{\text{I-Foc}}$	ekazánepsen.
Alis.NOM	in	a.ACC	month.ACC	thousand.ACC	lira.ACC	earn.Past.3SG

'It is one thousand that Alis earned in a month.'

(S01; 150702_0013; 16:50)

In (72), the meaning narrows down to only/exactly one thousand.

Second, the universal quantifiers 'all' and 'every' are incompatible with exclusive focus (É. Kiss 1998); all referents are included and therefore there is no exclusion of alternatives in the same set. For example, if the preverbal focus position in Romeyka does not house a universal pronoun, this strategy/position can therefore be said to express exclusive focus. As shown by the following examples, this is not the case in Romeyka (see (73)):

(73) Romeyka:

a. Question:

PÍos	éši	ospítin	so	χoríon?
who.NOM	have.3SG	house.ACC	in.the.ACC	village.ACC

'Who has a house in the village?'

Answers:

b. [Úl:ini]$_{\text{I-Foc}}$ éχun ospítin so χoríon.

everyone.NOM have.3PL house.ACC in.the.ACC village.ACC

'Everyone has a house in the village.'

(S01; 150702_0013; 22:06)

c. [KÁθa is]$_{\text{I-Foc}}$ éši ospítin so χoríon.

everyone.NOM have.3SG house.ACC in.the.ACC village.ACC

'Everyone has a house in the village.'

(S01; 150702_0013; 22:09)

Third, Kenesei (1986, 2006) remarks that no alternatives exist for a unique referent and that hence no alternatives can be excluded. Therefore, he reasons, if a focus strategy is incompatible with a unique referent, it expresses exclusive focus. If the Romeyka preverbal focus position is associated with exclusivity, the referent 'the sun' is predicted to be ungrammatical in the immediate preverbal slot, as there is only one sun in our galaxy (see (74)):

(74) Romeyka:

a. Question:

DO eɣvéni asó doɣúnin?
what.ACC rise.3SG from.the.ACC east.ACC
'What rises from the east?'

b. Answer:

[o Ílon]$_{\text{I-Foc}}$ eɣvéni asó doɣúnin.
the.NOM sun.NOM rise.3SG from.the.ACC east.ACC
'The sun rises from the east.'
(S01; 150703_0040; 13:31)

The answer in (74)b shows that the preverbal focus position in Romeyka is associated with exclusivity. If the Romeyka preverbal focus position is associated with exclusivity, the referent 'the moon' is predicted to be ungrammatical in the immediate preverbal slot, as our planet has only one moon (see (75) and (76)):

(75) Romeyka:

a. Question:

t akšémin DÓɣna lámbi?
the.ACC night.ACC what.ACC shine.3SG
'What shines at night?'

b. Answer:

[o FÉŋgon]$_{\text{I-Foc}}$ lámbi.
the.NOM moon.NOM shine.3SG
'The moon shines.'
(S01; 150703_0040; 13;38)

(76) Romeyka:

a. Question:

DO ékripsen tin katsímalin?

what.NOM hide.Past.3SG the.ACCclouds.ACC

'What hid the clouds?'

b. Answer:

[o FÉŋgon]$_{\text{I-Foc}}$ ékripsen.

the.NOM moon.ACC hide.Past.3SG

'The moon hid the clouds.'

(S01; 150703_0040; 14:03)

Moreover, the focused constituent in questions of total ignorance which yield a yes/no reply also takes place to the left of the verb, as shown in (77). This is in contrast to MG which allows movement of the focused constituent in questions only with indefinite DPs (see (78)a), but not with definite ones (see (78)b):

(77) Romeyka:

a. [i NÍfe]$_{\text{I-Foc}}$ efáisen ti mamíka?

the.NOM daughter-in-law.NOM feed.Past.3SG the.ACCmother-in-law.ACC

'Did the daughter-in-law feed her mother-in-law?'

(S01; 150702_0013; 13:53)

b. esís [ta tsuPÁðæ]$_{\text{I-Foc}}$ θerízete?

you.NOM the.ACCcorn.ACC harvest.2PL

'Do you harvest the corn?'

(S07; 812_0067; 01:58)

c. ató [o mehMÉtis]$_{\text{I-Foc}}$ éndžen æ?

this.ACC the.NOM Mehmetis.NOM bring.Past.3SG it.ACC

'Did Mehmetis bring that?'

(S01; 150703_0042; 06:36)

(78) Modern Greek:

a. kseris ta aχlaðja?
know.2SG the.ACCpears.ACC
'Do you know the pears?'

b. *t aχlaðja kseris?
the.ACCpears.ACC know.2SG
'Do you know the pears?'

(Sitaridou & Kaltsa 2014: 14)

Furthermore, it has been claimed that movement of the focused constituent is unavailable within the left periphery of the subordinate clause (see Cruschina 2008). *Prima facie*, on the basis of (79), it appears as though the focused constituent in subordinate clauses in Romeyka can appear in the left periphery. Information focused constituent in subordinate clauses in Romeyka can involve any type of phrase: direct object (NP) (see (79)), direct object (DP) (see (80) and adverbial phrase (see (81)):

(79) Romeyka:

Direct object (NP) is focused:

a. Question:
alís DÓɣna ðótšen?
Alis.NOM what.ACC give.Past.3SG
'What did Alis give?'

b. Answer:
eɣó léɣo, alís [koS:Áran]$_{\text{I-Foc}}$ eðótšen.
I.NOM say.1SG Alis.NOM hen.ACC give.Past.3SG
'I say that Alis gave a hen.'
(S01; 812_0093; 00:19)

(80) Romeyka:

Direct object (DP) is focused:

a. Question:

DO θarís, alís TÍnan efílisen?

what.ACC think.2SG Alis.NOM who.ACC kiss.Past.3SG

'Who do you think that Alis kissed?'

b. Answer:

eɣó θaró, alís [tin aiŠÉN]$_{\text{I-Foc}}$ efílisen.

I.NOM think.1SG Alis.NOM the.ACCAyşe.ACC kiss.Past.3SG

'I think that Alis kissed Ayşe.'

(S01; 150703_0040; 19:07)

(81) Romeyka:

Adverbial phrase is focused:

a. Question:

DO θarís, alís PÓte eχújepsen to c^{h}it^{h}ápin?

what.ACC think.2SG Alis.NOM when read.Past.2SG the.ACCbook.ACC

'What do you think? When did Alis read the book?'

b. Answer:

eɣó θaró, [oPSÉ]$_{\text{I-Foc}}$ eχújepsen æ.

I.NOM think.1SG yesterday read.Past.3SG it.ACC

'I think that he read it yesterday.'

(S01; 150703_0041; 00:03)

Moreover, in indirect yes/no questions the focused constituent also takes place left to the verb (see (82)):

(82) Romeyka:

Focus movement in indirect yes/no question:

a. rotás me, alís [tin aiŠÉN]$_{\text{I-Foc}}$ efílisen?
ask.NOM.2SG I.ACC Alis.NOM the.ACC Ayşe kiss.Past.3SG
'You ask me, did Alis kiss Ayşe?'
(S01; 150702_0022; 03:16)

b. erotás me, to c^{h}it^{h}ápin [aLÍS]$_{\text{I-Foc}}$ eχújepsen?
ask.2SG I.ACC the.ACC book.ACC Alis.NOM read.Past.3SG
'You ask me, did Alis read the book?'
(S01; 150702_0022; 06:13)

4.3.3 Contrastive foci in Romeyka

Contrastive focus involves the selection of a subset from a set of alternatives (see Molnar 2006). Contrastive focus in Romeyka is attested preverbally as in MG. Consider the sentence in (83) from Romeyka and the equivalent one in (84) from MG:

(83) Romeyka:

a. Question:
kahVÉN ɟóksa TŠÁin θélis?
coffee.ACC or tea.ACC want.2SG
'Do you want coffee or tea?'

Answers:

b. eɣó [kahVÉN]$_{\text{C-Foc}}$ θélo.
I.NOM coffee.ACC want.1SG
'I want coffee.'
(S01; 150702_0013; 12:15)

c. manaχón [kahVÉN]$_{\text{C-Foc}}$ θélo.
only coffee.ACC want.1SG
'I only want coffee.'
(S01; 150702_0013; 12:22)

(84) Modern Greek:

a. Question:

θélis kaFÉ i TSÁI?

want.2SG coffee.ACC or tea.ACC

'Do you want coffee or tea?'

Answers:

b. [kaFÉ]$_{C\text{-}Foc}$ θélo.

coffee.ACC want.1SG

'I want coffee.'

c. móno [kaFÉ]$_{C\text{-}Foc}$ θélo.

only coffee.ACC want.1SG

'I only want coffee.'

(Sitaridou & Kaltsa 2014: 12)

From the examples in (83) and (84), it becomes obvious that both Romeyka and MG allow for contrastive focus to the left of the verb. In particular, the preverbal position is the only option for marking contrastive focus in Romeyka. See the following grammatical judgement made by a Romeyka speaker in (85):

(85) Romeyka:

a. Question:

eɣó léɣo se alís apʰíðæ aɣórasen,
I.NOM say.1SG you.ACC Alis.NOM pears.ACC buy.Past.3SG
áma esí eksérts alís míla aɣórasen.
but you.NOM know.2SG Alis.NOM apples.ACC buy.Past.3SG
eɣó érχome léɣo se alís apʰíðæ aɣórasen.
I.NOM come.1SG say.1SG you.ACC Alis.NOM pears.ACC buy.Past.3SG
esí dóɣna léjis me?
you.NOM what.ACC say.2SG I.ACC?
'Alis bought pears, but you know that he bought apples. I came and told you that Alis bought pears. What do you reply to me?'

b. Answer:

alís [MÍla]$_{\text{C-Foc}}$ aɣórasen.
Alis.NOM apples.ACC buy.Past.3SG
'Alis bought apples.'
(S01; 812_0055; 01:54)

Any type of phrase can be contrastively focused in Romeyka: predicate (see (86)), gerund (see (87)), subject (NP) (see (88)), subject (DP) (see (89)), object (NP) (see (90)), object (DP) (see (91)), predicative complement (see (92)), prepositional phrase (see (93)), adverbial phrase (see (94)) and quantifier (see (95)):

(86) Romeyka:

Predicate is focused:

a. Question:

na eχujÉvis cʰitʰápæ

PRT.MOD read.2SG books.ACC

jóksa na teRÍS tʰelevizjónin aɣapás?

or PRT.MOD watch.2SG television.ACC love.2SG

'Do you like to read books or watch television?

b. Answer:

eɣó [tʰelevizjónin na teRÓ]$_{\text{C-Foc}}$.

I.NOM television.ACC PRT.MOD watch.1SG

'I like to watch television.'

(S01; 150702_0013; 12:25)

(87) Romeyka:

Gerund is focused:

a. Question:

LÁɣa na páɣo son bakʰálin?

how PRT.MOD go.1SG to.the.ACC grocery's.ACC

me t arapán jóksa me ta poðáræ m?

by the.ACC car.ACC or on the.ACC feet.ACC I.POSS

'How can I go to the grocery's? By car or on foot?'

b. Answer:

[porpateFTÁ]$_{\text{I-Foc}}$ na pas so cʰicʰénin.

walking.GER PRT.MOD go.3SG to.the.ACC grocery's.ACC

'You can get to the grocery's on foot.'

(S01; 150702_0013; 11:17)

(88) Romeyka:

Subject (NP) is focused:

a. Question:

éna	Ándras	pʰakʰlaévi	t	ospítin?
a.NOM	man.NOM	clean.3SG	the.ACC	house.ACC

'Does a man clean the house?'

b. Answer:

jo, [éna iNÉka]$_{C\text{-}Foc}$		pʰakʰlaévi	t	ospítin.
no a.NOM	woman.NOM	clean.3SG	the.ACC	house.ACC

'No, it's a woman that cleans the house.'

(S01; 140102_0009; 06:57)

(89) Romeyka:

Subject (DP) is focused:

a. Question:

o	ramaZÁnis	epíren	tin	aišén?
the.NOM	Ramazanis.NOM	marry.Past.3SG	the.ACC	Ayşe.ACC

'Did Ramazanis marry Ayşe?'

b. Answer:

jo, [o	œMÉris]$_{C\text{-}Foc}$	epíren	tin	aišén.
no the.NOM	Ömeris.NOM	marry.Past.3SG	to.the.ACC	Ayşe.ACC

'No, it's Ömeris that married Ayşe.'

(S01; 140102_0009; 06:13)

(90) Romeyka:

Direct object (NP) is focused:

a. Question:

o	mehmétis	MÍla	jóksa	aP^{H}Íðæ	aɣórasen?
the.NOM	Mehmetis.NOM	apples.ACC	or	pears.ACC	buy.Past.3SG

'Did Mehmetis buy apples or pears?'

b. Answer:

o	mehmétis	[MÍla]$_{C\text{-}Foc}$	aɣórasen.
the.NOM	Mehmetis.NOM	apples.ACC	buy.Past.3SG

'It's apples that Mehmetis bought.'

(S01 ;150702_0013; 12:05)

(91) Romeyka:

Direct object (DP) is focused:

a. Question:

o	ramazánis	ti	zeiNÉP	epíren?
the.NOM	Ramazanis.NOM	the.ACC	Zeynep.ACC	marry.Past.3SG

'Did Ramazanis marry Zeynep?'

b. Answer:

o	ramazánis	[tin	aiŠÉN]$_{C\text{-}Foc}$	epíren.
the.NOM	Ramazanis.NOM	the.ACC	Ayşe.ACC	marry.Past.3SG

'It's Ayşe that Ramazanis married.'

(S01; 140102_0009; 07:50)

(92) Romeyka:

Predicative complement is focused:

a. Question:

DÓɣna en avúto? vútiron?

what.NOM be.3SG this.NOM butter.NOM

'What is this? Butter?'

b. Answer(s):

[anΘÓɣalan]$_{\text{C-Foc}}$ en.

buttermilk.ACC be.3SG

'This is buttermilk.'

(S01; 812_0055; 00:54)

(93) Romeyka:

Prepositional phrase is focused:

a. Question:

o dohtóris so istamBÓlin epíen?

the.NOM doctor.NOM to.the.ACC Istanbul.ACC go.Past.3SG

'Did the doctor go to Istanbul?'

b. Answer:

jo, [sin iŋɟilT^{H}Éran]$_{\text{C-Foc}}$ epíen.

no to.the.ACC England.ACC go.Past.3SG

'No, he went to England.'

(S01; 140102_0009; 08:21)

(94) Romeyka:

Adverbial phrase is focused:

a. Question:

alís oSÍm:eron érθen asin tšáikaran?

Alis.NOM today come.Past.3SG from.the.ACC Çaykara.ACC

'Did Alis come from Çaykara today?'

b. Answer:

jok^{h}, [oPSÉ]$_{C\text{-}Foc}$ érθen.

no yesterday come.Past.3SG

'No, it's yesterday that he came.'

(S01; 150703_0040; 08:46)

(95) Romeyka:

Quantifier is focused:

a. Question:

manaχón eSÍST ešete ospítin so χoríon?

only you.NOM have.2PL house.ACC in.the.ACC village.ACC

'Only you have a house in the village?'

b. Answer:

jo, [Úl:ini]$_{C\text{-}Foc}$ éχun ospítin so χoríon.

no everyone.NOM have.3PL house.ACC in.the.ACC village.ACC

'No, it's everyone that has a house in the village.'

(S01; 150702_0013; 22:03)

Romeyka displays SV to contrastively focus the subject. SV is always attested, no matter what the syntactic type or the semantic properties of the subject are.

First, the contrastively focused subject can be a personal pronoun (see (96)):

(96) Romeyka:

a. Question:

i aiŠÉ epíen sa stalíæ?
the.NOM Ayşe go.Past.3SG to.the.ACC stalls.ACC
'Did Ayşe go to the stalls?'

b. Answer:

jo, [eɣÓ]$_{\text{C-Foc}}$ epíɣa sa stalíæ.
no I.NOM go.Past.3SG to.the.ACC stalls.ACC
'No, I went to the stalls.'
(S01; 140102_0009; 06:43)

Second, the contrastively focused subject can be an NP (see (97)):

(97) Romeyka:

a. Question:

éna Ándras p^{h}ak^{h}laévi t ospítin?
a.NOM man.NOM clean.3SG the.ACChouse.ACC
'Does a man clean the house?'

b. Answer:

jo, [éna iNÉka]$_{\text{C-Foc}}$ p^{h}ak^{h}laévi t ospítin.
no a.NOM woman.NOM clean.3SG the.ACChouse.ACC
'No, it's a woman that cleans the house.'
(S01; 140102_0009; 06:57)

Third, the contrastively focused subject can be a DP (see (98)):

(98) Romeyka:

a. Question:

o	ramaZÁnis	epíren	tin	aišén?
the.NOM	Ramazanis.NOM	marry.Past.3SG	the.ACC	Ayşe.ACC

'Did Ramazanis marry Ayşe?'

b. Answer:

jo, [o	œMÉris]$_{C\text{-}Foc}$	epíren	tin	aišén.
no the.NOM	Ömeris.NOM	marry.Past.3SG	the.ACC	Ayşe.ACC

'No, it's Ömeris that married to Ayşe.'

(S01; 140102_0009; 06:13)

Based on the previous data, if a subject is contrastively focused, no matter what the type of subject is, it is placed to the left of the verb, as illustrated in (99):

(99) DP/NP-subject$_{C\text{-}Foc}$ > TP

Romeyka displays OV to contrastively focus the object. OV is always attested, no matter what the syntactic type or the semantic properties of the object are.

First, the focused object can be a demonstrative pronoun (see (100)):

(100) Romeyka:

a. Question:

alís	tin	aiŠÉN	epíren?
Alis.NOM	the.ACC	Ayşe.ACC	marry.Past.3SG

'Did Alis marry Ayşe?'

b. Answer:

jo, alís	[eMÉN]$_{C\text{-}Foc}$	epíren.
no Alis.NOM	I.ACC	marry.Past.3SG

'No, it's me that Alis married.'

(S01; 140102_0009; 09:01)

Second, the focused object can be a DP (see (101)):

(101) Romeyka:

a. Question:

o	ramazánis	ti	zeiNÉP	epíren?
the.NOM	Ramazanis.NOM	the.ACC	Zeynep.ACC	marry.Past.3SG

'Did Ramazanis marry Zeynep?'

b. Answer:

o	ramazánis	[tin	aiŠÉn]$_{C\text{-}Foc}$	epíren.
the.NOM	Ramazanis.NOM	the.ACC	Ayşe.ACC	marry.Past.3SG

'It's Ayşe that Ramazanis married.'

(S01; 140102_0009; 07:50)

Based on the previous data, if a subject is contrastively focused, no matter what the type of object is, it is placed to the left of the verb, as illustrated in (102):

(102) DP/NP-object$_{C\text{-}Foc}$ > TP

Romeyka displays PP > VP to contrastively focus the PP. PP > VP is always attested (see (103)):

(103) Romeyka:

a. Question:

i	aišé	sa	parχáræ	epíen?
the.NOM	Ayşe.NOM	to.the.ACC	pastures.ACC	go.Past.3SG

'Did Ayşe go to the pastures?'

b. Answer:

jo, i	aišé	[s	éna	χoRÍon]$_{C\text{-}Foc}$	epíen.
no the.NOM	Ayşe.NOM	to	a.ACC	village.ACC	go.Past.3SG

'No, it's to a village that Ayşe went.'

(S01; 140102_0009; 08:49)

Based on the previous data, if a PP is contrastively focused, no matter what the type of PP is, it is placed to the left of the verb, as illustrated in (104):

(104) PP$_{C\text{-}Foc}$ > TP

Romeyka displays AdvP > VP to contrastively focus the AdvP. AdvP > VP is always attested (see (105)):

(105) Romeyka:

a. Question:

alís tšaBÚχa eχújepsen to c^hitápin?

Alis.NOM quickly read.Past.3SG the.ACCbook.ACC

'Did Alis read the book quickly?'

b. Answer:

ɟo, [ŠÍta]$_{C\text{-}Foc}$ eχújepsen æ.

no immediately read.Past.3SG it.ACC

'No, it's immediately that he read it'

(S01; 150702_0022; 05:43)

Based on the previous data, if an adverbial phrase is contrastively focused, no matter what the type of adverbial phrase is, it is placed to the left of the verb, as illustrated in (106):

(106) AdvP$_{C\text{-}Foc}$ > TP

Apart from *wh*-questions, another type of question is also used as a focus test. These are the *so-called* 'alternative questions' of the form 'do you want coffee or tea?'. This special type of yes/no question requires a selection from among a set of given alternatives. The answer can be said to display selective focus. A focus constituent X is used selectively if "it introduces an element of [the alternative set] into the common ground and is chosen from a restricted subset of [the alternative set] the members of which have been explicitly mentioned in the preceding context" (Zimmermann & Onea 2011: 1663).

Consider (107) from the Romeyka data:

(107) Romeyka:

a. Question:

kahVÉN jóksa TŠÁin θélis?

coffee.ACC or tea.ACC want.2SG

'Do you want coffee or tea?'

Answers:

b. [TŠÁin]$_{C\text{-}Foc}$ aɣapó.

tea.ACC want.1SG

'I want tea.'

(S01; 150703_0042; 05:38)

c. [kahVÉN]$_{C\text{-}Foc}$ na píno.

coffee.ACC PRT.MOD drink.1SG

'I want to drink coffee.'

(S01; 150702_0013; 12:15)

The fact that the alternatives are present and one of the alternatives is selected, e.g. 'I want tea', excludes the other alternative, 'not coffee'. A question is whether this exclusion is necessarily present, having truth-conditional effects (semantics), or just implicature (pragmatics) (see van der Wal 2016). After all, choosing tea does not necessarily mean that one does not want coffee.

In the answers to the alternative questions in Romeyka, the focused constituent can be any type of phrase: gerund (see (108)), VP (see (109)), subject (personal pronoun) (see (110)), subject (DP) (see (111)), object (NP) (see (112)), object (DP) (see (113)), predicative complement (see (114)), prepositional phrase (see (115)) and adverbial phrase (see (116)):

(108) Romeyka:

a. Question:

LÁɣana páɣo son bakʰálin?

how PRT.MOD go.1SG to.the.ACC grocery's.ACC

me t araPÁN ɟóksa me ta poðÁræ m?

by the.ACCcar.ACCor on the.ACCfeet.ACC I.POSS

'How should I get to the grocery's? By car or on foot?'

b. Answer:

[porpateFTÁ]$_{C\text{-}Foc}$ na pas so cʰicʰénin.

walking.GER PRT.MOD go.2SG to.the.ACC grocery's.ACC

'You should get to the grocery's on foot.'

(S01; 150702_0013; 11:17)

(109) Romeyka:

a. Question:

na eχuɟÉvis cʰitʰápæ

PRT.MOD read.2SG books.ACC

ɟóksa na teRÍS tʰelevizɟónin aɣapás?

or PRT.MOD watch.2SG television.ACC love.2SG

'Do you like to read books or watch television?

b. Answer:

eɣó [tʰelevizɟónin na teRÓ]$_{C\text{-}Foc}$.

I.NOM television.ACC PRT.MOD watch.1SG

'I like to watch television.'

(S01; 150702_0013; 12:25)

(110) Romeyka:

a. Question:

PÍos éfaen ton tšorbán?
who.NOM eat.Past.3SG the.ACC soup.ACC?
eSÍ ɟóksa i aiŠÉ?
you.NOM or the.NOM Ayşe.NOM
'Who ate the soup? You or Ayşe?'

b. Answer:

ton tšorbán [eɣÓ]$_{\text{C-Foc}}$ éfaɣa.
the.ACCsoup.ACC I.NOM eat.Past.1SG
'It's me that ate the soup.'
(S01; 150702_0013; 14:08)

(111) Romeyka:

a. Question:

o mehMÉtis ɟóksa aLÍS éfaen to χavítsin?
the.NOM Mehmetis.NOM or Alis.NOM eat.Past.3SG the.ACCpudding.ACC
'Did Mehmetis or Alis eat the pudding?'

b. Answer:

[o mehMÉtis]$_{\text{C-Foc}}$ éfaen to χavítsin.
the.NOM Mehmetis.NOM eat.Past.3SG the.ACCpudding.ACC
'It's Mehmetis that ate the pudding.'
(S01; 150702_0013; 10:23)

(112) Romeyka:

a. Question:

o mehmétis MÍla ɟóksa aP^{H}Íðæ aɣórasen?
the.NOM Mehmetis.NOM apples.ACC or pears.ACC buy.Past.3SG
'Did Mehmetis buy apples or pears?'

b. Answer:

o mehmétis [MÍla]$_{\text{C-Foc}}$ aɣórasen.
the.NOM Mehmetis.NOM apples.ACC buy.Past.3SG
'It's apples that Mehmetis bought.'
(S01; 150702_0013; 12:05)

(113) Romeyka:

a. Question:

i aišé ton mehMÉtin
the.NOM Ayşe.NOM the.ACC Mehmetis.ACC
jóksa ton aLÍN epíren?
or the.ACC Alis.ACC marry.Past.3SG
'Did Ayşe marry Mehmetis or Alis?'

b. Answer:

i aišé [ton aLÍN]$_{C\text{-}Foc}$ epíren,
the.NOM Ayşe.NOM the.ACC Alis.ACC marry.Past.3SG
jok^{h} ton mehmétin.
no the.ACC Mehmetis.ACC
'It's Alis that Ayşe married, not Mehmetis.'
(S01; 150702_0013; 11:46)

(114) Romeyka:

a. Question:

o šc^{h}íl:on éna haiVÁnin jóksa énan inSÁnin en?
the.NOM dog.NOM a.NOM animal.NOM or a.NOM human.NOM be.3SG
'Is a dog an animal or a human?'

b. Answer:

o šc^{h}íl:on [haiVÁnin]$_{C\text{-}Foc}$ en.
the.NOM dog.NOM animal.NOM be.3SG
'The dog is an animal.'
(S01; 150702_0013; 13:11)

(115) Romeyka:

a. Question:

t aðelfó s sin TŠAIkaran
the.NOM brother.NOM you.POSS in.the.ACC Çaykara.ACC
jóksa sin trapeZÚndan stétši?
or in.the.ACC Trabzon.ACC stay.3SG
'Does your brother stay in Çaykara or in Trabzon?'

b. Answer:

aðelfó m [sin trapeZÚndan]$_{\text{C-Foc}}$ stétši.
brother.NOM I.POSS in.the.ACC Trabzon.ACC stay.3SG
'My brother stays in Trabzon.'
(S01; 150702_0013; 12:36)

(116) Romeyka:

a. Question:

i nífe oSÍm:eron jóksa oPSÉ
the.NOM daughter-in-law.NOM today or yesterday
efáisen ti mamíka?
feed.Past.3SG the.ACC mother-in-law.ACC
'Did the daughter-in-law feed her mother-in-law today or yesterday?'

b. Answer:

i nífe [oPSÉ]$_{\text{C-Foc}}$
the.NOM daughter-in-law.NOM yesterday
efáisen ti mamíka.
feed.Past.3SG the.ACC mother-in-law.ACC
'It's yesterday that the daughter-in-law fed her mother-in-law.'
(S01; 150702_0013; 14:02)

Second, the universal quantifiers 'all' and 'every' are incompatible with exclusive focus (É. Kiss 1998); all referents are included and therefore there is no exclusion of alternatives in the same set. For example, if the preverbal focus position in Romeyka does not house a universal pronoun, this strategy/position can be therefore said to express exclusive focus. As shown by the following examples, this is not the case in Romeyka (see (117) and (118)):

(117) Romeyka:

a. Question:

manaχón	i	aišé	epíen	sa	parχáræ?
only	the.NOM	Ayşe	go.Past.3SG	to.the.ACC	pastures.ACC

'Did only Ayşe go to the pastures?'

b. Answer:

ɟokʰ,	[Úl:ini]$_{C\text{-}Foc}$	épiɣane	sa	parχáræ.
no	everyone.NOM	go.Past.3PL	to.the.ACC	pastures.

'It's everyone that went to the pastures.'

(S01; 150702_0013; 20:55)

(118) Romeyka:

a. Question:

manaχón	alís	epíen	so	kurbétin?
only	Alis.NOM	go.Past.3SG	to.the.ACC	abroad.ACC

'Did only Alis go abroad?'

b. Answer:

[Úl:ini	i	aɣúri]$_{C\text{-}Foc}$	píɣane	so	kurbétin.
all.NOM	the.NOM	men.NOM	go.Past.3PL	to.the.ACC	abroad.ACC

'It's all the men that went abroad.'

(S01; 150702_0013; 23:06)

Similarly, the indefinite quantifiers 'some' and 'few' "are upward entailing, i.e. they imply that the denoted quantity reaches at least a minimum from a scale of potential quantities" (Skopeteas & Franselow 2010: 1387). As with the numerals, when these indefinite quantifiers are in exclusive or exhaustive focus, the alternative quantities are excluded (see (119)):

(119) Romeyka:

a. Question:

PÍos	epíen	so	istambólin?
who.NOM	go.Past.3SG	to.the.ACC	Istanbul.ACC

'Who went to Istanbul?'

b. Answer:

[kaNÍS]$_{\text{C-Foc}}$	tš	epíen	so	istambólin.
no-one.NOM	NEG	go.Past.3SG	to.the.ACC	Istanbul.ACC

'No-one went to Istanbul.'

(S01; 150702_0013; 21:22)

The term 'focus particles' has been used to refer to two categories of particles. On the one hand, there are languages, which have a dedicated particle marking the focus of the sentence. On the other hand, there are focus particles or focus-sensitive operators, which trigger a focused reading on the element they modify, or associate with the focus of the constituent in their environment and do not have an influence on the propositional content of the sentence, but may influence the truth-conditional values (König 1991, Rooth 1985, 1992, Krifka 2006, Beaver & Clark 2008, i.a.). While in some languages all particles behave the same in terms of the linguistic expression (e.g. the interaction with stress), in others there are important differences between them, in terms of their effect on the sentence (see (120) and (121)):

(120) Romeyka:

a. Question:

PÍos	epíen	so	kurbétin?
who.NOM	go.Past.3SG	to.the.ACC	abroad.ACC

'Who went abroad?'

b. Answer:

[manaχós	aLÍS]$_{\text{C-Foc}}$	epíen	so	kurbétin.
only	Alis.NOM	go.Past.3SG	to.the.ACC	abroad.ACC

'Only Alis went abroad.'

(S01; 150702_0013; 23:01)

(121) Romeyka:

a. Question:

PÍos epíen so džamín?

who.NOM go.Past.3SG to.the.ACC mosque.ACC

'Who went to the mosque?'

b. Answer:

[manaçés:a i aiŠÉ]$_{\text{C-Foc}}$ epíen so džamín.

only the.NOM Ayşe.NOM go.Past.3SG to.the.ACC mosque.ACC

'Only Ayşe went to the mosque.'

(S01; 150702_0013; 20:31)

Finally, it appears that preverbal contrastive focus in subordinate clauses in Romeyka is possible (see (122)):

(122) Romeyka:

a. Question:

alís DÓɣna éfaen?

Alis.NOM what.ACC eat.Past.3SG

mÍlon éfaen i aP^{H}Íðin éfaen?

apple.ACC eat.Past.3SG or pear.ACC eat.Past.3SG

'What did Alis eat? An apple or a pear?'

b. Answer:

eɣó léɣo, alís [aP^{H}Íðin]$_{\text{C-Foc}}$ éfaen

I.NOM say.1SG Alis.NOM pear.ACC eat.Past.3SG

'I say that it's a pear that Alis ate.'

(S01; 812_0055; 00:43)

4.3.4 Summary

In this section, I have shown that focused constituents in Romeyka always occupy the immediate preverbal position, no matter what the type of focus or the syntactic category of the constituent are. In §4.4, I will investigate the syntactic distribution of *wh*-phrases in *wh*-questions in Romeyka and whether Romeyka allows for multiple *wh*-questions and multiple focus.

4.4 *wh*-questions, multiple *wh*-questions and multiple focus in Romeyka

4.4.0 Introduction

In this section, I address *wh*-questions in §4.4.1, multiple *wh*-questions in §4.4.2 and multiple focus in §4.4.3. In §4.4.4, I provide a theory for multiple *wh*-questions and multiple focus in Romeyka. The findings of the section are summarised in §4.4.5.

4.4.1 *wh*-questions in Romeyka

MG displays *wh*-questions, (see (123)). Likewise, Romeyka also employs *wh*-questions (see (124)) (see Michelioudakis & Sitaridou 2016):

(123) Modern Greek:

a. pços fílise ti maría?
who.NOM kiss.Past.3SG the.ACCMaria.ACC
'Who kissed Maria?'
(Alexopoulou & Baltazani 2012)

b. pçon fílise i maría?
whoACC kiss.Past.3SG the.NOM Maria.NOM
'Who did Maria kiss?'
(Alexopoulou & Baltazani 2012)

(124) Romeyka:

a. PÍos eðótšen tin kos:áran?
who.NOM give.Past.3SG the.ACChen.ACC
'Who gave the hen?
(S01; 812_0093; 00:03)

b. χavítsæ PÍos éfaen?
anchovies.ACC who.NOM eat.Past.3SG
'Who ate anchovies?'
(S01; 812_0057; 04:06)

c. alís DÓγna ðótšen?
Alis.NOM what.ACC give.Past.3SG
'What did Alis give?'
(S01; 812_0093; 00:16)

Crucially, the order of *wh*-questions is strictly order-preserving in Romeyka (see (125) and (126)) (see Michelioudakis & Sitaridou 2016):

(125) Romeyka:

a. PÍos eðótšen tin kos:áran?
who.NOM give.Past.3SG the.ACChen.ACC
'Who gave the hen?
(S01; 812_0093; 00:03)

b. ?PÍos tin kos:áran eðótšen?
who.NOM the.ACChen.ACC give.Past.3SG
'Who gave the hen?

(126) Romeyka:

a. χavítsæ PÍos éfaen?
anchovies.ACC who.NOM eat.Past.3SG
'Who ate anchovies?'
(S01; 812_0057; 04:06)

b. ?PÍos χavítsæ éfaen?
who.NOM anchovies.ACC eat.Past.3SG
'Who ate anchovies?'

wh-phrases are obligatorily left-dislocated (see (127)), with no option to leave any *wh*-phrase *in-situ* (Michelioudakis & Sitaridou 2016):

(127) Romeyka:

a. alís DÓγna ðótšen?
Alis.NOM what.ACC give.Past.3SG
'What did Alis give?'
(S01; 812_0093; 00:16)

b. ?alís ðótšen DÓγna?
Alis.NOM give.Past.3SG what.ACC
'What did Alis give?'

4.4.2 Multiple *wh*-questions in Romeyka

While MG does not allow left-dislocated multiple *wh*-questions (henceforth MWQ), see (128), Romeyka does (see (129)) (Michelioudakis & Sitaridou 2016):

(128) Modern Greek:

pços χtípise pçon
who.NOM hit.Past.3SG who.ACC
'Who hit who?'
(Alexopoulou & Baltazani 2012)

(129) Romeyka:

PÍos DÓγna aγórasen?
who.NOM what.ACC buy.Past.3SG
'Who bought what?'
(S01; 150703_0042; 07:37)

Crucially, Romeyka seems to exhibit Superiority effects, which show that MWQ is strictly order-preserving, as in Bulgarian, although it is not otherwise identical (see (130)) (see Bošković 1997, Michelioudakis & Sitaridou 2016):

(130) Romeyka:

a. PÍos DÓγna aγórasen?
who.NOM what.ACC buy.Past.3SG
'Who bought what?'
(S01; 150703_0042; 07:37)

b. ?DÓγna PÍos aγórasen?
what.ACC who.NOM buy.Past.3SG
'Who bought what?'

MWQ are obligatorily left-dislocated, with no option to leave any *wh*-phrase *in-situ* (see (131)) (Michelioudakis & Sitaridou 2016: 9):

(131) Romeyka:

a. PÍos DÓγna aγórasen?
who.NOM what.ACC buy.Past.3SG
'Who bought what?'
(S01; 150703_0042; 07:37)

b. ?PÍos aγórasen DÓγna?
who.NOM buy.Past.3SG what.ACC
'Who bought what?'

4.4.3 Multiple focus in Romeyka

While MG does not allow left-dislocated multiple focus (henceforth MF) (see (132)), Romeyka does (see (133)) (Michelioudakis & Sitaridou 2016):

(132) Modern Greek:

a. Question:

pite	mu	pços	parakolúθise	pçon?
tell.IMP.2PL	I.GEN	who.NOM	follow.Past.3SG	who.ACC

'Tell me who followed who?'

b. Answer:

[o	petros]$_{Foc}$	parakoluθise	[ti	maria]$_{Foc}$,
the.NOM	Petros.NOM	follow.Past.3SG	the.ACC	Maria.ACC,
o	stavros	tin	eleni…	
the.NOM	Stavros.NOM	the.ACC	Eleni.ACC	

'Petros followed Maria, Stavros Eleni…'

(Alexopoulou & Baltazani 2012)

(133) Romeyka:

a. Question:

PÍos	DÓɣna	aɣórasen?
who.NOM	what.ACC	buy.Past.3SG

'Who bought what?'

(S01; 150703_0042; 07:37)

b. Answer:

[o	mehMÉtis]$_{Foc}$	[araPÁn]$_{Foc}$	aɣórasen.
the.NOM	Mehmetis.NOM	car.ACC	buy.Past.3SG

'Mehmetis bought a car.'

(S01; 150703_0042; 07:39)

MF is strictly order-preserving in Romeyka (see (134)) (see Michelioudakis & Sitaridou 2016):

(134) Romeyka:

a. [o mehMÉtis]$_{Foc}$ [araPÁn]$_{Foc}$ aɣórasen.
the.NOM Mehmetis.NOM car.ACC buy.Past.3SG
'Mehmetis bought a car.'
(S01; 150703_0042; 07:39)

b. ?[araPÁn]$_{Foc}$ [o mehMÉtis]$_{Foc}$ aɣórasen.
car.ACC the.NOM Mehmetis.NOM buy.Past.3SG
'Mehmetis bought a car.'

MF is obligatorily left-dislocated with no option to leave any focused phrase *in-situ* (see (135)) (Michelioudakis & Sitaridou 2016: 9):

(135) Romeyka:

a. '[o mehMÉtis]$_{Foc}$ [araPÁn]$_{Foc}$ aɣórasen.
the.NOM Mehmetis.NOM car.ACC buy.Past.3SG
'Mehmetis bought a car.'
(S01; 150703_0042; 07:39)

b. ?[o mehMÉtis]$_{Foc}$ aɣórasen [araPÁn]$_{Foc}$.
the.NOM Mehmetis.NOM buy.Past.3SG car.ACC
'Mehmetis bought a car.'

4.4.4 A theory for the multiple *wh*-questions and the multiple focus in Romeyka

Following Bošković's (2002, 2007) proposal, Michelioudakis & Sitaridou (2016) propose (a) the availability of an Attract-1 Focus head below C^0, (b) *wh*-fronting as focus-movement and (c) the availability of *wh*-movement to [Spec, CP] in matrix questions. Thus, while C^0 is an Attract-1 head, in Bošković's terms, which attracts just the highest *wh*-phrase, Focus is an Attract-all head that attracts all *wh*-phrases available to the same position.

Schematically, then, a partial representation of the left periphery of Romeyka is proposed by Michelioudakis & Sitaridou (2016), as in (136):

(136) [$_{CP}$ C[+Q=mɪ]/[+*wh*]=Attract-1 [$_{FocusP}$ Foc^0 Attract-1 [$_{FocusP}$ Foc^0[new info] ...

Based on Beck's (2006) account of intervention effects following the focus interpretation, I provide a principled explanation for MWQ and MF in Romeyka, which is supplementary to Michelioudakis & Sitaridou's (2016) analysis.

Beck's account begins with Kim's (2002) study, which identifies a core intervention effect (see (137)), which is attested cross-linguistically:

(137) *[Q_i [... [FocP [... *wh*-phrase$_i$ ...]]]]
(Kim 2002)

According to (137), a focused phrase cannot intervene between a *wh*-phrase and its licensing complementiser. As I have already shown in §4.3.1, according to Rooth's (1985, 1992, 1996) Alternative Semantics, focus is associated with two semantic objects: On the one hand, there is the proposition expressed by the sentence —the set of possible worlds. On the other hand, the sentence makes salient a set of alternative propositions, which contains alternative propositions to the proposition. This is the focus-semantic value of the sentence.

Turning now to the interrogative in (138), according to the standard semantic theory of questions (Hamblin 1973, Karttunen 1977), the denotation of a question is the set of answers to the question —for example (139)a. More generally, this is the set of propositions in (139)b (rendered in more formal terms in (139)c):

(138) Who left?

(139) Alternative propositions:
a. {that John left, that Bill left, that Amelie left,…}
b. {that x left | x is an individual}
c. $\lambda p \exists x[p= \lambda w.x \text{ left in } w]$

wh-phrases, like focus, introduce a set of alternatives. Unlike a focused phrase, introducing alternatives seems to be the only semantic role of a *wh*-phrase. It is not surprising that this parallel has inspired semanticists to derive the interpretations of questions and focus in the same way. I will develop a particular way of doing that.

I follow Rooth in attributing a twofold semantic contribution to focused phrases: their ordinary semantic value on the one hand and a set of alternatives of the same type on the other. A *wh*-phrase shares with focus the second role. Unlike focus, the *wh*-phrase makes no ordinary

semantic contribution. I propose that the ordinary semantics of the *wh*-phrase is in fact undefined. Since *wh*-phrases occur in expressions that have a perfectly well-defined ordinary semantic value, something must rescue the structure as a whole from indefiniteness. This is the role of the question operator. Thus, I propose that the LF of (140) and that the semantics of Q let it ignore the ordinary semantic value of its sister and elevate its focus semantic value to ordinary semantics:

(140) [Q [who left]]

Things go wrong when there is also a focus in the question whose contribution is evaluated within the question, i.e. within the scope of the Q operator. This situation is schematised in (141):

(141) [Q ... [Op [φ ... XPF ... wh ...]]]

For the focus on XP to be evaluated within the scope of the Q operator means that there is a focus sensitive operator, here Op, which uses the semantic contribution of the focus. Op could be 'only' or 'even' or the like, or, in Rooth's (1992) Alternative Semantics, which states that focus "indicates the presence more indirect framework for association with focus, it could be the ~ operator". We know that when focus is evaluated at the level of a phrase φ, focus semantic values enter into ordinary semantics.

This analysis explains why Romeyka does not give rise to intervention effects.

4.4.5 Summary

In this section, I have shown that *wh*-phrases in *wh*-questions in Romeyka occupy the same position that focused constituents occupy. Furthermore, Romeyka allows for MWQ and MF, while intervention effects are not allowed. As for the latter, I have provided an explanation based on the semantic interpretation of focus and *wh*-phrases.

So far, I have focused on the identification of the positions of topics, foci and *wh*-phrases in Romeyka. In §4.5, I will investigate the syntactic distribution of foci and topics in Romeyka.

4.5 The syntactic distribution of topics and foci in Romeyka

4.5.0 Introduction

The aim of this section is to offer a detailed presentation of the possible discourse-related positions in the Romeyka clausal structure. So far, I have shown that foci and topics in Romeyka are always attested in the left periphery of the verb. I have also shown that the verb raises to T^0. I have therefore concluded that marked constituents move to the left of T^0. The question I address in this section is, where is the landing site of those moved constituents in the Romeyka clausal structure?

The section is structured as follows: in §4.5.1, I examine how many topic positions there are in Romeyka; in §4.5.2, I examine how many focus positions there are in Romeyka; in §4.5.3, I investigate the distribution of topics and foci in Romeyka; and I summarise the findings of the section in §4.5.4.

4.5.1 How many topic positions are there in Romeyka?

A crucial question, which needs to be answered, is the following: if both the preverbal and the postverbal domain can accommodate given information, then what differentiates the preverbal domain from the postverbal domain in terms of given information? Interestingly, in Romeyka, while given information can be contrastive in the preverbal domain, it cannot be contrastive in the postverbal domain. Thus, what differentiates the preverbal domain from the postverbal domain in terms of given information is the [contrast] feature. For instance, consider the object *tin aišén* 'Ayşe' in the preverbal domain in (142)a and in the postverbal domain in (142)b. In both sentences, the object *tin aišén* 'Ayşe' carries [noncontrastive] given information. However, in (143)a the object *dolmán* 'dolma' carries [contrastive] given information and can appear only in the preverbal domain, but not in the postverbal one (see (143)b):

(142) Romeyka:

a. [tin aišén]$_{A\text{-}Top}$ o dohTÓris epíren.
the.ACC Ayşe.ACC the.NOM doctor.NOM marry.Past.3SG
'The doctor married Ayşe.'
(S01; 140102_0008; 01:15)

b. o dohTÓris epíren [tin aišén]$_{A\text{-}Top}$.
the.NOM doctor.NOM marry.Past.3SG the.ACC Ayşe.ACC
'The doctor married Ayşe.'
(S01; 140102_0008; 01:25)

(143) Romeyka:

a. [dolmán]$_{C\text{-}Top}$ o mehMÉtis éfaen.
dolma.ACC the.NOM Mehmetis.NOM eat.Past.3SG
'Mehmetis ate dolma.'
(S01; 150702_0014; 11:46)

b. #o mehMÉtis éfaen [dolmán]$_{C\text{-}Top}$.
the.NOM Mehmetis.NOM eat.Past.3SG dolma.
'Mehmetis ate dolma.'

4.5.2 How many focus positions are there in Romeyka?

In order to examine whether the information focus and the contrastive focus share the same position, I use an adverb placement test. I specifically test to see whether a preverbal adverb can interpolate between information focus and the verb on the one hand and between contrastive focus and the verb on the other. If so, this would indicate that there are two positions for focus in Romeyka, one for information focus and one for contrastive focus. If not, this would indicate that there is a single position that can host both information and contrastive focus.

As for the placement of the adverbs in the clause structure, I follow Alexiadou's (1997) proposal, as stated below (see (144)):

(144) Adverbial Licensing Principle (ALP):

Adverbs are licensed either as Specifiers of Functional Projections or via incorporation into the verbal head by the relevant (semantic) feature associated with the head.

(Alexiadou 1997: 41)

According to the ALP, adverbs are split into two types: (a) specifier-type adverbs, such as quantifier or degree adverbs, e.g. *purely*, *unique*, *nearly*, *always*; and, (b) complement-like adverbs: *completely*, *easily*, *badly*, *well*, *lovingly*.

Specifier-type adverbs are base generated adjuncts. Complement-type ones are mainly manner adverbs, e.g. qualifying adverbs. Moreover, the following descriptive generalisation holds (see (145)):

(145) Generalisation:

Specifier-type adverbs have their base position to the left of the verb (non-thematic, specifiers of NegPs, AspectPs), hence they are VP-external. Complement-like ones have their base position to the right of the verb; hence they are VP internal.

(Alexiadou 1997: 129)

Having shown that (a) the verb raises to T^0 (see §3.3) and (b) focus is always preverbal, no matter what the type of focus is (see §4.2), I test specifier-type adverbs to the left of T^0, in particular aspectual and temporal adverbs.

Alexiadou (1997) shows that aspectual and temporal adverbs are licensed as specifiers of AspP and TP respectively. A result of such an analysis is that it distinguishes between operator type adverbs like aspectual ones and nominal type adverbs like nominal ones. The latter behave similarly to other argumental DPs. Moreover, she proposes that only XPs marked [$+T^0$], [$+Asp^0$] can occupy the specifier positions of TP and AspP respectively in languages such as Greek and Spanish. In particular, she argues that [Spec, TP] is parameterised cross-linguistically, hosting subjects in some languages, e.g. Celtic and Icelandic and TAs in some others, e.g. Greek and Chinese.

In Romeyka, the aspectual distinction perfective vs. imperfective holds in all tenses, moods and both voices.

Aspectual adverbs can be classified into: (a) durative, indefinite frequency ones; this class includes adverbs like *kondá kondá* ‘every now and then’ and *her jyn* ‘daily’ and (b)

cardinal count/definite frequency, point ones; this class subsumes adverbs like *šíta* 'immediately', *mían* 'once', *ðío tané* 'twice' etc.

The adverb I use for the test is *šíta* 'immediately', which is an aspectual adverb. As such, it is placed low in the phrase structure, specifically in the specifier of TP. In the case of information focus, strict adjacency seems to hold between the information focused object and the verb in (146)b. In (146)c, we see that the adverb *šíta* cannot interpolate between information focus to the left and the verb, whereas in (146)b, the adverb is positioned before the focused object, which is adjacent to the verb:

(146) Romeyka:

a. Question:

alís	šíta	TÍnan	efílisen?
Alis.NOM	immediately	who.ACC	kiss.Past.3SG

'Who did Alis kiss immediately?'

Answers:

b.

alís	šíta	[ti	MÁnan	at]$_{\text{I-Foc}}$	efílisen.
Alis.NOM	immediately	the.ACC	mother.ACC	his	kiss.Past.3SG

'Alis kissed his mother immediately.'

(S01; 150702_0013; 05:27)

c.

#alís	[ti	MÁnan	at]$_{\text{I-Foc}}$	šíta	efílisen.
Alis.NOM	the.ACC	mother.ACC	his	immediately	kiss.Past.3SG

Furthermore, when the adverb is information focused, it occupies the immediate preverbal slot as shown in (147)b, which we have seen being occupied by the information focused object in (146)b, indicating that informational focus is immediately adjacent to the left of the verb:

(147) Romeyka:

a. Question:

alís LÁɣa efílisen tin aišén?
Alis.NOM how kiss.Past.3SG the.ACC Ayşe?
'How did Alis kiss Ayşe?'

Answers:

b. alís [ŠÍta]$_{\text{I-Foc}}$ efílisen tin aišén.
Alis.NOM immediately kiss.Past.3SG the.ACCAyşe.ACC
'Alis kissed Ayşe once.'
(S01; 150702_0022; 04:07)

c. #alís efílisen [ŠÍta]$_{\text{I-Foc}}$ tin aišén.
Alis.NOM kiss.Past.3SG immediately the.ACCAyşe.ACC

Turning now to contrastive focus and adverb placement, in (148)c we see that the adverb cannot interpolate between the verb and the contrastively focused object and thus (148)b is the only grammatical option:

(148) Romeyka:

a. Question:

alís šíta tin aišÉN efílisen?
Alis.NOM immediately the.ACCAyşe.ACC kiss.Past.3SG
'Did Alis kiss Ayşe immediately?'

Answers:

b. alís šíta [tin birɟÝL]$_{\text{C-Foc}}$ efílisen.
Alis.NOM immediately the.ACCBirgül.ACC kiss.Past.3SG
'It's Birgül that Alis kissed immediately.'
(S01; 150702_0013; 09:23)

c. #alís [tin birɟÝL]$_{\text{C-Foc}}$ šíta efílisen.
Alis.NOM the.ACCBirgül.ACC immediately kiss.Past.3SG

Interestingly, when the adverb is contrastively focused, it occupies the immediate preverbal slot as shown in (149)b, indicating that contrastive focus is immediately adjacent to the left of the verb and thus yielding identical results to information focus:

(149) Romeyka:

a. Question:

alís tšaBÚuχa eχújepsen to c^{h}itápin?

Alis.NOM quickly read.Past.3SG the.ACCbook.ACC

'Did Alis read the book quickly?'

Answers:

b. ɟo, [ŠÍta]$_{C\text{-}Foc}$ eχújepsen a.

no immediately read.Past.3SG it.ACC

'No, he read it immediately.'

(S01; 150702_0022; 05:43)

c. #ɟo, alís [ŠÍta]$_{C\text{-}Foc}$ to c^{h}itápin eχújepsen

no Alis.NOM immediately the.ACCbook.ACC read.Past.3SG

Therefore, on the basis of the data above I conclude that (a) constituents carrying information or contrastive focus occupy the same position and (b) strict adjacency holds between the focused constituent and the verb.

4.5.3 The syntactic distribution of topics and foci in Romeyka

I now move to the investigation of the syntactic distribution of topics and foci in Romeyka. In order to examine their distribution, I apply the tests offered in Neeleman & van de Koot (2008) in their investigation of Dutch, as well as in Șener (2010) in his investigation of Turkish.

Consider now example (150):

(150) Romeyka:

a. Question:

alís DO epítšen?
Alis.NOM what.ACC do.Past.3SG
DO éfaen so bairámin?
what.ACC eat.Past.3SG in.the.ACC Bayram.ACC
'What did Alis do? What did he eat at Bayram?'

Answers:

vál:ahi, utš ekséro alís do epítšen, áma …
frankly NEG know.1SG Alis.NOM what.ACC do.Past.3SG but
'Frankly, I don't know about Alis, but …'

b. [o mehmétis]$_{\text{C-Top}}$ [dolMÁN]$_{\text{C-Foc}}$ éfaen.
the.NOM Mehmetis.NOM dolma.ACC eat.Past.3SG
'Mehmetis ate dolma.'
(S01; 150702_0014; 09:06)

c. #[dolMÁN]$_{\text{C-Foc}}$ [o mehmétis]$_{\text{C-Top}}$ éfaen.
dolma.ACC the.NOM Mehmetis.NOM eat.Past.3SG
'Mehmetis ate dolma.'

The contrast of felicity in the responses in (150) demonstrates that a C-Foc may follow a C-Top, whereas a C-Foc cannot move across a C-Top.

Below, I consider a pair, where the context is set up so as to favour an interpretation of the object as C-Top and the subject as C-Foc (see (151)):

(151) Romeyka:

a. Question:

o	tšorbás	DO	ejéndo?
the.NOM	soup.NOM	what.ACC	happen.Past.3SG
atón	kaNÍS	éfaen	a?
this.ACC	anyone.NOM	eat.Past.3SG	it.ACC

'What about the soup? Has anyone eaten that?'

Answers:

vál:ahi,	utš	ekséro	o	tšorbás	do	ejéndo,	áma …
frankly	NEG	know.1SG	the.NOM	soup.NOM	what.ACC	happen.Past.3SG	but

'Frankly, I don't know about the soup, but …'

b. [dolmán]$_{\text{C-Top}}$ [o mehMÉtis]$_{\text{C-Foc}}$ éfaen.

dolma.ACC the.NOM Mehmetis.NOM eat.Past.3SG

'Mehmetis ate dolma.'

(S01; 150702_0014; 11:46)

c. #[o mehMÉtis]$_{\text{C-Foc}}$ [dolmán]$_{\text{C-Top}}$ éfaen.

the.NOM Mehmetis.NOM dolma.ACC eat.Past.3SG

'Mehmetis ate dolma.'

The infelicity between (151)b and (151)c supports the assumption that a C-Top cannot follow a C-Foc in Romeyka. Therefore, the only licit order would be C-Top > C-Foc.

Last but not least, I introduce some observations on the interaction between VP-internal objects and Topic and Focus. What follows is an investigation of the order of DO and IO, showing that the linear order of DO and IO depends on the discourse function of the objects. The sentences in example (152) have a ditransitive verb, where the context is set up so as to favour the interpretation of the IO as C-Foc and the DO as C-Top:

(152) Romeyka:

a. Question:

i	antíka	tše	i	sandalía
the.NOM	antique.NOM	and	the.NOM	chair.NOM

DO ejéndo?
what.ACC happen.Past.3SG

o	pápʰos	TÍnan	éðocen	a?
the.NOM	grandfather.NOM	who.ACC	give.Past.3SG	it.ACC

'What about the antique table and the chair? Who has your granddad bequeathed that to?'

Answers:

vál:ahi	i	antíka	do	ejéndo
frankly	the.NOM	antique.NOM	what.ACC	happen.Past.3SG

utš ekséro, áma …
NEG know.1SG but
'Frankly, I don't know about the antique table, but …'

b. [ti sandalían]$_{\text{C-Top}}$[ton TŠÍri m]$_{\text{C-Foc}}$ eðótšen.
the chair.ACC the father.ACC I.POSS give.Past.3SG
'my granddad bequeathed the chair to my dad.'
(S01; 150702_0014; 14:01)

c. #[ton TŠÍri m]$_{\text{I-Foc}}$ [ti sandalían]$_{\text{C-Top}}$eðótšen.
the father.ACC I.POSS the chair.ACC give.Past.3SG
'my granddad bequeathed the chair to my dad.'

The contrast in the felicity in (152) provides more evidence on the order C-Top > C-Foc. Below, I consider a pair, where the context is set up so as to favour an interpretation of the IO as C-Top and the DO as C-Foc (see (153)):

(153) Romeyka:

a. Question:

o	tšíri	s	DO	ejéndo?
the.NOM	father.NOM	you.POSS	what.ACC	happen.Past.3SG

o	pápʰos	DO	éðocen	aton?
the.NOM	grandfather.NOM	what.ACC	give.Past.3SG	he.ACC

'What about your dad? What has granddad bequeathed to him?'

Answers:

vál:ahi	o	tšíri	m	DO	ejéndo
frankly	the.NOM	father.NOM	I.POSS	what.ACC	happen.Past.3SG

utš	ekséro,	áma …
NEG	know.1SG	but

'Frankly, I don't know about my dad, but …'

b. [ti mána m]$_{\text{C-Top}}$ [to saÁtʰin]$_{\text{C-Foc}}$ efítšen.
the mother.ACC I.POSS the watch.ACC bequeath.Past.3SG
'my granddad bequeathed the watch to my mother.'
(S01; 150702_0023; 08:18)

c. #[to saÁtʰin]$_{\text{C-Foc}}$ [ti mána m]$_{\text{C-Top}}$ efítšen.
the watch.ACC the mother.ACC I.POSS bequeath.Past.3SG
'my granddad bequeathed the watch to my mother.'

The felicity of the sentences in (153) supports the generalisation that C-Top precedes C-Foc in Romeyka, no matter what grammatical function they bear.

Information focus must be left adjacent to the verb. Thus, information focus is not different from contrastive focus in terms of its distribution; thus, it follows contrastive topics (see (154)):

(154) Romeyka:

a. Question:

tsi birjýlis t aðélfæ
the.GEN Birgül.GEN the.NOM brothers.NOM
D epíkane so pártin?
what.ACC do.Past.3PL at.the.ACC party.ACC
'What did Birgül's brothers get to drink at the party?'

Answers:

vál:ahi as aðélfæ tes utš ekséro, áma …
frankly from.the.ACC brothers.ACC her NEG know.1SG but
'Frankly, I do not know about all the brothers but …'

b. [úl:unon o mikrón]$_{C\text{-}Top}$ [raCÍN]$_{I\text{-}Foc}$ epíen.
all.GEN the.NOM young.NOM rakı drink.Past.3SG
'Birgül's youngest brother drank rakı.'
(S01; 150702_0023; 23:26)

c. #[raCÍN]$_{I\text{-}Foc}$ [úl:unon o mikrón]$_{C\text{-}Top}$ epíen.
rakı all.GEN the.NOM young.NOM drink.Past.3SG
'Birgül's youngest brother drank rakı.'

Information focus follows aboutness topics too (see (155):

(155) Romeyka:

a. Question:

avúto to faín PÍon patsín epítšen?
this.ACC the.ACC food.ACC which.NOM girl.NOM make.Past.3SG
'Which girl made this food?'
(S01; 150703_0042; 03:32)

b. Answer:

[avúton to faín]$_{A\text{-}Top}$ [i miNÉ] $_{I\text{-}Foc}$ epítšen.
this.ACC the.ACC food.ACC the.NOM Mine.NOM make.Past.3SG
[t ál:on]$_{C\text{-}Top}$ [i aiŠÉ]$_{C\text{-}Foc}$ epítšen.
the.ACC other.ACC the.NOM Ayşe.NOM make.Past.3SG
'Mine made this food; Ayşe made the other one.'
(S01; 150703_0042; 03:45)

To conclude, it is obvious from the reasoning provided above that in Romeyka topics always precede foci. This hierarchy is shown in (156):

(156) Hierarchy of discourse features in Romeyka:

a. Topic > Focus

b. #Focus > Topic

On the basis of the hierarchy of discourse features in Romeyka in (156), the articulation of foci and topics in Romeyka should be the one shown in (157):

(157) Articulation of discourse-related features in Romeyka:

A-/C-Topic > I-/C-Focus > T^0 > Given (noncontrastive) information

4.5.4 Summary

In this section, I have shown that in Romeyka (a) topics are always left-dislocated, whereas constituents carrying given (and noncontrastive) information may appear in the postverbal domain too; (b) focus always appears to the left of the verb and occupies the immediate preverbal slot, no matter what the type of focus is; and (c) topics are hierarchically higher than foci.

4.6 Conclusions

In this chapter, (a) I determined the pragmatically unmarked/neutral word order in Romeyka and (b) examined the syntactic distribution and the semantic type of the constituents in pragmatically marked word orders in Romeyka. Overall, I argued that (a) I the pragmatically unmarked order is SVO in matrix declarative clauses and SOV in subordinate declarative clauses and (b) focused constituents and *wh*-phrases always occupy the immediate preverbal position, while topicalised elements are always left dislocated. Romeyka also allows multiple *wh*-questions and multiple focus. The postverbal domain can only be occupied by elements carrying given (noncontrastive) information.

The word order pattern in matrix clauses is summarised in Table 27 and the word order pattern in subordinate clauses is summarised in Table 28:

Table 27. Word order pattern in matrix clauses in Romeyka.

SVO	pragmatically unmarked word order or pragmatically marked word order (subject is focused / a *wh*-phrase)
SOV	pragmatically marked word order (subject is topicalised and object is focused / a *wh*-phrase or both subject and object are focused / *wh*-phrases)
OSV	pragmatically marked word order (object is topicalised and subject is focused / a *wh*-phrase)

Table 28. Word order pattern in subordinate clauses in Romeyka.

SOV	pragmatically unmarked word order or pragmatically marked word order (subject is topicalised and object is focused / a *wh*-phrase or both subject and object are focused / *wh*-phrases)
OSV	pragmatically marked word order (object is topicalised and subject is focused / a *wh*-phrase)

Interestingly, the word order pattern in matrix clauses in Table 27 shows that word order variation in matrix clauses is structure-dependent. In chapter 5, I put forth my proposal to account for word order variation in Romeyka.

5 Order in a minimalist system

5.0 Introduction

In chapters 3 and 4, I examined word order variation in Romeyka. I would now like to move on to account for such word order variation in a minimalist system. In order to account for word order variation in Romeyka, I aim to pursue a third factor (principled) explanation for word order variation, within the minimalist program. After that, I will test my predictions for every potential syntactic derivation of the subject (S), verb (V) and object (O) in Romeyka and map those derivations into PF and LF rules. In particular, I will examine (a) matrix declarative clauses, (b) subordinate declarative clauses, (c) direct questions and (d) indirect questions.

The chapter is organised as follows: in §5.1, I put forth my proposal; §5.2 examines matrix declarative clauses. §5.3 provides a summary of the chapter.

5.1 Order in a minimalist system

5.1.1 Order in Romeyka is not structure-dependent

It has long been believed that Merge is assumed not to impose order, i.e. {x, y} = {y, x}. As such, order is structure-dependent, i.e. no syntactic operation can make reference to it. It has also been claimed that hypothetical languages, in which syntactic operations are defined in linear terms, such that Merge creates an ordered pair <x, y>, are outside of the spectrum of variation defined by UG (see Musso et al. 2003, Smith & Tsimpli 1995).

For instance, Smith & Tsimpli (1995) attempted to teach a polyglot savant, Christofer, Epun, a language they invented to enable them to test Christofer's reaction to structures, which, they hypothesised, could not occur in the world real languages. Epun is an SVO language (Smith & Tsimpli 1995: 139). Interestingly, Epun displays operations defined in linear terms, which, to our knowledge, are not attested in natural languages. They specifically illustrate two examples of such constructions (Smith & Tsimpli 1995: 146). In particular, Epun displays different word orders to express negation on the one hand and tense on the other (see Table 29):

Table 29. Word order pattern in Epun (Smith & Tsimpli 1995: 146).

SV(O)	Positive (present and future)
VS(O)	Negative (present and future)
(O)SV	Positive (past)
(O)VS	Negative (past)

According to Table 29, in Epun there is a contrast between SV(O) word order in positive sentences and V(S)O word order in negative sentences. Moreover, the past tense is characterised by the object being moved to the initial position (Smith & Tsimpli 1995: 146).

The findings of Smith & Tsimpli's (1995) study show that Christopher had considerable difficulty with those operations, indicating that these operations are outside of the spectrum of UG (Smith & Tsimpli 1995: 154-155).

Similarly, Musso et al. (2003) investigated the neural correlate of acquiring new linguistic competence with two fMRI studies. First, German native speakers learned a sample of 'real' grammatical rules from different languages (Italian or Japanese), which, although parametrically different, follow the UG principles. The activity during this task was compared with that during a task that involved learning 'unreal' rules of language. The 'unreal' language

that they used was constructed by manipulating Italian. In the 'unreal' grammar, the subjects could not relate the nominal and verbal elements by means of any hierarchical order, as the new rules defined a more linear order of the single words (Musso et al. 2003: 775) (see Table 30):

Table 30. Unreal Italian (artificial rules violating UG) (Musso et al. 2003: 775).

Negative construction	Paolo mangia la no pera
	"Paolo eats the no pear"
Interrogative construction	Pera la mangia Paolo
	"Pear the eats paolo"
Use of indefinite article	Una bambino mangia una pera
	"A (fem.) child (masc.) eats a (fem.) pear (fem.)"

According Table 30, in the first rule, negative sentences were built by always putting the negation word *no* after the third word of the clause. The second rule required that the interrogative construction be built by inverting the linear sequence of the words in a sentence. The third rule, arbitrarily, emphasised a specific word position for choosing the correct indefinite article; i.e. indefinite articles within a sentence always agree with the last noun of the clause (Musso et al. 2003: 775).

The results of Musso et al.'s (2003) study show that the increase of activation over time in Borca's area was specific to 'real' language acquisition only, independent of the kind of language. Thus, in Borca's area, biological constraints and language experience interact to enable linguistic competence in a new language.

It might be concluded from Smith & Tsimpli (1995) and Musso et al.'s (2003) findings that syntactic operations that are defined in linear terms are outside of the spectrum of variation defined by UG.

However, against all odds, the findings of chapters 3 and 4 show that word order variation (at least in matrix clauses, but see §7.3 for a similar account of subordinate clauses) in Romeyka is defined in linear terms (see the word order pattern in matrix clauses in Romeyka in Table 31 and the word order pattern in subordinate clauses in Romeyka in Table 32):

Table 31. Word order pattern in matrix clauses in Romeyka.

SVO	pragmatically unmarked word order or pragmatically marked word order (subject is focused / a *wh*-phrase)
SOV	pragmatically marked word order (subject is topicalised and object is focused / a *wh*-phrase or both subject and object are focused / *wh*-phrases)
OSV	pragmatically marked word order (object is topicalised and subject is focused / a *wh*-phrase)

Table 32. Word order pattern in subordinate clauses in Romeyka.

SOV	pragmatically unmarked word order or pragmatically marked word order (subject is topicalised and object is focused / a *wh*-phrase or both subject and object are focused / *wh*-phrases)
OSV	pragmatically marked word order (object is topicalised and subject is focused / a *wh*-phrase)

That is to say, in this section I put forth a syntax-semantics interface analysis to account for the word order pattern attested in matrix (and subordinate) clauses in Romeyka. My proposed analysis is precisely based on two premises: in particular, I postulate that the order of the constituents of the clause in Romeyka (a) contributes to the mapping of syntactic units (phases) from narrow syntax to both SEM and PHON interfaces and (b) plays a role in the semantic interpretation of focus, yes/no questions and *wh*-questions at SEM. Contrary to Chomsky (1995), who assumes that "there is no clear evidence that order plays a role at LF or in the computation from N to LF"[7] (Chomsky 1995: 334), in the rest of this section, I provide evidence that order plays a role at SEM and in the computation from narrow syntax to SEM in Romeyka. In the remainder of this chapter, I develop my arguments in favour of such an analysis.

5.1.2 Against Cartography

In the minimalist literature, there have been various attempts to account for a principled explanation of discourse-related elements, with cartography (see Cinque 1999, Rizzi 1997) being the most prominent. Cartography attempts to satisfy two interface conditions, namely the

[7] In this extraction, LF is for Chomsky the SEM.

Condition of Inclusiveness (CI) (see (1)) on the one hand and the Condition of Full Interpretation (CFI) (see (2)) on the other:

(1) Condition of Inclusiveness (CI):
"Any structure formed by the computation (in particular, π and λ) is constituted of elements already present in the lexical items selected for N [i.e. the numeration – SC]; no new objects are added in the course of computation apart from rearrangements of lexical properties" (Chomsky 1995: 228).

According to the CI, a formal feature cannot be assigned in the computation. By contrast, formal features must be present in the lexicon.

(2) Condition of Full Interpretation (CFI):
a. A syntactic expression is PF-interpretable iff it can be assigned a phonological representation (i.e., iff it can "read" by the phonology)
b. A syntactic expression is LF-interpretable iff it can be assigned a semantic representation (iff it can be read by the semantics)
(Chomsky 1995: 194)

The CFI implies that a direct interaction between the phonological component (PHON) and the semantic component (SEM) is not possible. Both components interface and communicate only with the narrow syntax. Given that discourse-related phenomena have both a phonological and an interpretive impact, it follows that the relevant features must already be present in the syntactic component.

Based on these interface conditions, cartography suggests that discourse-related features, such as [focus], [topic] and [contrast], are present in the lexicon. These features are considered to force their projection to narrow syntax as formal heads.

In contrast to the cartographic approaches, Chomsky (2000) introduces the concept of "core functional categories", by which he means C^0, T^0 and v^0 in the clause. Chomsky (2005: 18) speculates that "the more elaborate structures revealed by the cartographic inquiries are based on linearisation of features in these [i.e. CP and vP, TB/IR] labels and possibly labels closely linked to them (as in the C-T connection)". Biberauer & Roberts (2015) develop this idea by exploiting the distinction between formal and semantic features. The formal features are, as proposed in Chomsky (1995), interpretable or uninterpretable and, as such, are visible

for syntactic operations such as Agree and Merge. The semantic features, on the other hand, are invisible to the core computational system, but presumably visible at the semantic interface.

As a matter of principle, Chomsky (2005) proposes a reinterpretation of the relation between the functional heads C^0 and T^0: the Agree (ϕ) and Tense features associated with the inflectional system are not an inherent property of T^0; instead, they belong to the phase head C^0. Richards (2007) elaborates Chomsky's ideas, arguing that feature inheritance follows by conceptual necessity from two basic assumptions about the phase-based derivational system: the Value-Transfer Simultaneity (VTS) (see (3)) and the Phase Impenetrability Condition (PIC) (see (4)):

(3) Value-Transfer Simultaneity (VTS):
Value and Transfer of *u*Fs must happen together.
(Richards 2007: 566)

(4) Phase Impenetrability Condition (PIC):
The edge and nonedge (complement) of a phase are transferred separately.
(Richards 2007: 568)

As a result, *u*Fs must descend from the phase head onto that part of the phase that actually gets transferred (and deleted), namely the complement domain (see (5)):

(5) *u*F must descend from edge to nonedge (i.e. from C to T, v* to V, etc.).
(Richards 2007: 569)

That is, Richards (2007) derives Chomsky's (2005) mechanism of feature inheritance, which now follows from 'good design': it ensures that full interpretation (FI) is met by enabling Agree-features to be valued and deleted immediately, as part of Transfer, at the phase level. This line of third-factor reasoning, proceeding from the premise that FI must be optimally met (i.e. without delay or lookback), thus yields the conclusion that the efficient satisfaction of interface conditions requires three components: VTS, PIC and FI.

Given Richards' (2007) assumption that a system that is in conformance with a maximally empty UG is one in which phases are pairs of phase heads and nonphase heads (hence the core sequence C–T–v*–V), the possible expansions of the core functional sequence

into more richly articulated hierarchies (i.e. the cartographic) have to be constrained. Therefore, in a minimalist system, discourse-related features do not project onto the narrow syntax.

5.1.3 Linear order

Before I begin with my proposed account of order in Romeyka and since I assume that the order of the constituents of the clause in Romeyka is defined by a linear relation, I provide a formal definition of the notions of linear order. Note that a linear order is necessarily a binary relation (Partee et al. 1987: 28-30).

First, a binary relation on a set A is a collection of ordered pairs of elements of A. In other words, it is a subset of the Cartesian product $A^2 = A \times A$. More generally, a binary relation between two sets A and B is a subset of $A \times B$ (Partee et al. 1987: 28-30).

Second, a linear order is a binary relation on some set A, which is antisymmetric, transitive and total (this relation is denoted here by infix $\leq$). In particular, a set A is linearly ordered under $\leq$ if the following statements hold for all α, β and γ in A (Partee et al. 1987: 28-30):

(6) A set A is linearly ordered under $\leq$ if the following statements hold for all α, β and γ in A:
   a. If $\alpha \leq \beta$ and $\beta \leq \alpha$ then $\alpha = \beta$ (antisymmetry);
   b. If $\alpha \leq \beta$ and $\beta \leq \gamma$ then $\alpha \leq \gamma$ (transitivity);
   c. $\alpha \leq \beta$ or $\beta \leq \alpha$ (totality).
   (Partee et al. 1987: 28-30)

Based on (6), I assume that an antisymmetric view of order implies that the syntax does not make a categorical distinction between specifiers and complements, (see Lohndal 2012, 2014). The main syntactic relation, *modulo* adjuncts, should be that of a merged head and a non-head. In line with Lohndal (2012, 2014), I define specifiers and complements as follows:

(7) a. A complement is the sister of a head
   b. A specifier is the sister of a head
   (Lohndal 2012: 120)

The statement in (7) satisfies the total view of order as well. It apparently reduces the relational difference between complement and specifier and as such they are syntactically equal. Moreover, the relation between the specifier and the complement of a specific head is transitive, for the specifier is hierarchically higher than the complement. That said, for the specifier α and the complement γ of a head β, the order in (8) must hold:

(8) $<\alpha, \beta, \gamma> = \{\alpha, \beta, \gamma \mid \alpha$ is the complement of the head β, β is the head, γ is the complement of the head $\beta\}$

To sum up, I assume that in a minimalist system the order is linear. A consequence of this assumption is that the specifier-head relation is the same as the head-complement relation in syntactic terms, such that the specifier does not bear any relational difference from the complement; rather, the former is higher than the latter, such that the order specifier > head > complement is rigid. Having defined the properties of linear order, I now need to explore how heads are hierarchically articulated in a minimalist system.

So far, I have shown that in a minimalist system, in which the order of the constituents of the clause is linear, narrow syntax consists of phase-heads and nonphase-heads and specifiers and complements. Furthermore, in a minimalist system the order of the heads in a clause is rigidly C > T > v* > V, whereas each head projects a complement and a specifier in the rigid order specifier > head > complement. I have also argued that the possible expansions of the core functional sequence into more richly articulated hierarchies (i.e. the cartographic) are not a property of a minimalist system.

If my argumentation is on the right track, I now need to answer a crucial question: if order is a purely phonological feature and focus a purely semantic feature, such that both the latter and the former do not have information that project onto narrow syntax, how could we then account for Romeyka data, in which linear order plausibly plays a role in focus interpretation?

5.1.4 Focus interpretation constraint in Romeyka

As I have shown in chapters 3 and 4, while the pragmatically unmarked word order in Romeyka is VO, OV is, at least in matrix clauses, pragmatically marked. As such, OV order in Romeyka is pragmatically marked in relation to VO. This asymmetry shows that there is good reason to

assume that OV phrases are pragmatically marked than VO phrases, such that OV might be derived from VO. What I am therefore looking for is a theory to capture this assumption.

Indeed, such a claim is made by Biberauer et al. (2014), who argue that head-final orders are formally marked in relation to head-initial ones. According to them, head-final order refers to the presence of a diacritic ^, which, when associated with the categorial feature of a head, triggers the movement of the complement of that head to its specifier, i.e. to what Biberauer et al. (2014) refer to as "L(inearisation)-movement". The main evidence for this stems from the Final-over-Final Condition (FOFC) (see Biberauer et al. 2014).

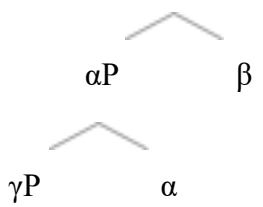
The notion of formal markedness that Biberauer et al. (2014) adopt should not be taken to imply relative frequency; hence they do not predict that head-initial languages are more frequent than head-final ones. Their notion of formal marking relates simply to formal properties visible to the computational system.

It has been observed that there is an asymmetry in possible word orders cross-linguistically, both synchronically and diachronically. This observation is captured by the FOFC, which is a universal constraint on phrase-structure configurations that was proposed by Biberauer et al. (2014) (see (9)):

(9) The Final-over-Final Condition (FOFC):
A head-final phrase αP cannot dominate a head-initial phrase βP where α and β are heads in the same Extended Projection.
(Biberauer et al. 2014: 171)

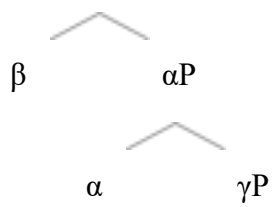
The import of the formulation of FOFC in (9) is illustrated in (10):

(10) a.

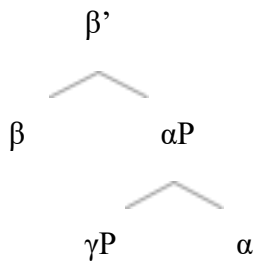


Consistent head-final
(harmonic)

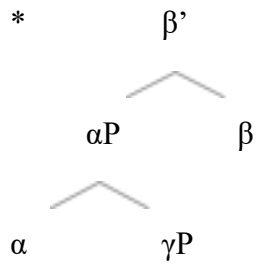
b.



Consistent head-initial
(harmonic)

c. 

Initial-over-final
(disharmonic)

d. * 

Final-over-initial
(disharmonic)

(Biberauer et al. 2014: 171)

The import of the formulation of FOFC in (9) rules out structures where αP is the complement β and γP is the complement of α (Biberauer et al. 2014: 171). Empirical evidence in favour of FOFC comes from the fact that certain disharmonic word orders are not attested cross-linguistically, for instance in the clausal domain and in the nominal domain, like the ones in (11):

(11) Disharmonic word orders that are not attested cross-linguistically:

a. *V-O-Aux *$[_{AuxP}[_{VP}$ V DP] Aux]

b. *V-O-C *$[_{CP}$ $[_{TP}$ T VP] C] or *$[_{CP}$ $[_{TP}$ $[_{VP}$ V O] T] C]

c. *C-TP-V *$[_{VP}$ $[_{CP}$ C TP] V]

d. *N-O-P *$[_{PP}$ $[_{DP/NP}$ D/N PP] P]

e. *Num-NP-D(em) *$[_{D(em)P}$ $[_{NumP}$ Num NP] D(em)]

f. *Pol-TP-C *$[_{CP}$ $[_{PolP}$ Pol TP] C]

(Biberauer et al. 2014: 196)

The asymmetry of the formulation of FOFC shows that there is good reason to assume that head-final phrases are more marked than head-initial phrases, such that OV might be derived from VO.

Indeed, within the generative framework, Roberts & Roussou (2003) propose that the presence of an extra EPP feature makes a representation more marked. Therefore, a derivation in which a constituent is moved is more marked than one in which there is no movement. The hierarchy in (12) is given as a markedness scale, where > means 'more marked than':

(12) Markedness Hierarchy:
F*Move/Merge > F*Move > F*Merge > F

F is the least marked option, as it has no feature that takes part in Merge, Agree (and hence not Move). F*Merge is more marked, since there are two elements being merged, which both have phonological matrices. F*Merge is less marked than F*Move because the former lacks the EPP feature driving movement. Finally, F*Move/Merge is the most marked since it involves the merging of two phonological feature matrices, as well as the EPP feature.

In more recent minimalist approaches to parametric syntax, markedness has been defined in terms of simplicity of derivations (see Roberts 2007, Roberts & Roussou 2003: 201) (see (13)):

(13) Simplicity of derivations:
Given two structural representations R and R' for a substring of input text S, R is simpler than R' if R contains fewer formal feature syncretisms than R'
(Roberts & Roussou 2003: 201)

Feature syncretism refers to more than one formal feature occurring in a particular structural position. Here formal features include ϕ-features like person and number, Case features, as well as features that trigger movement (the EPP feature as defined in Chomsky 2001).

Biberauer et al. (2014) point out that FOFC is based on the way in which movement is triggered (see (14)):

(14) Movement theory derived from FOFC:
Movement is triggered by a general movement-triggering feature. We use ^ (caret) as a symbol for this feature.
(Biberauer et al. 2014: 209)

They take ^ to be a purely formal, arbitrary diacritic. In itself, it has no semantic content and no connection to phonological or morphological properties beyond simply causing movement. Moreover, although it can be seen as a kind of formal feature, ^ differs in several important respects from formal features like ϕ-features. Unlike ϕ-features, which are arguably best seen as attribute-value pairs, it is privative, has no internal structure, cannot be valued or in any

obvious way "checked off" and, as already mentioned, has no semantic or morphophonological effects.

The idea that movement is triggered by a purely formal diacritic is widespread in the current literature. In different versions and with different notations, it appears in, among others, Roberts and Roussou (2003).

Biberauer et el. (2014) argue that the properties of different types of movement depend on the features that ^ is associated with. Where the movement-trigger ^ is associated with the uninterpretable ϕ-features of an active Probe, for instance finite T^0, it gives rise to A-movement; in this respect it replaces the EPP features of Chomsky. Where ^ is associated with a phase head, for instance C^0, it triggers A'-movement. Finally, where ^ is associated with the categorial, Extended Projection-defining feature [±V], 'linearisation movement' takes place, i.e. movement of the sister of a head to its specifier. These types of movement-triggers are illustrated in (15):

(15) a. $T_{[u\phi, ^\wedge]}$ triggers movement of the goal of the probe [uϕ] to Spec, TP.
b. $C_{[EF, ^\wedge]}$ triggers Ā-movement to Spec, CP.
c. $V_{[+V, ^\wedge]}$ triggers movement of the sister of V to Spec, VP.
(Biberauer et al. 2014: 210)

If head-final phrases are derived through movement, then this observation suggests that there is a restriction on where the ^ feature can occur. In particular, the disallowed order involves the movement of a phrase that does not itself contain any movement. Biberauer et al (2014) capture this intuition by proposing that whether or not the ^ feature is associated with a given head's c-selection feature depends on whether the lexical head (e.g. V^0) also carries the ^ feature. The central idea is that the ^ feature can 'spread' upwards from head to head, starting from the lexical head (thus providing the feature configurations that give rise to roll-up movement), but this spreading must take place monotonically, in line with Rizzi's (2001) Relativised Minimality. In other words, lower heads' c-selectional features must carry the ^ feature in order for higher heads to do so. Furthermore, no intermediate head's c-selection feature can be without a ^ feature if both the lower and higher heads in the structure carry the ^ feature —i.e. the ^ feature cannot 'skip' a head.

Very much in line with Biberauer et al. (2014), I propose that the ^ feature could account for word order variation in Romeyka too. However, in Romeyka the ^ feature also has internal structure, since it carries both phonological and semantic information, visible at the interfaces.

That said, I interpret the notion of markedness used by Biberauer et al. (2014) differently; while they interpret markedness as a formal notion, I interpret it as both phonological and semantic. In my case, the ^ feature enters derivation carrying both phonological and semantic information as follows:

(16) a. $T_{[u\phi, ^]}$ triggers movement of the goal of the probe [uϕ] to Spec, TP | ^ carries *Foc* or *wh*.

b. $C_{[EF, ^]}$ triggers Ā-movement to Spec, CP | ^ carries either *Top*, or *Foc*, or *wh*.

c. $V_{[+V, ^]}$ triggers movement of the sister of V to Spec, VP | ^ carries either *Foc* or *wh*.

While ^ is associated with the formal features of a phase head, it triggers movement of its complement to its specifier position. Therefore, when ^ is associated with the c-selection feature inherited to the lexical head V^0 from the phase head v^0, it triggers movement of the complement of the V^0 to its specifier. Like in Biberauer et al. (2014), I call this movement 'linearisation movement' and I consider it as the movement that is responsible for the VO and OV alternation in Romeyka.

I now move on to the argumentation in favour of my proposal. In order to explain how the ^ feature contributes to the efficient computation of interface conditions in Romeyka, I take into consideration the two aforementioned interface conditions, namely the Condition of Inclusiveness (CI) (see (1)) and the Condition of Full Interpretation (CFI) (see (2)).

Based on CI and CFI, I propose that discourse-related features are encoded in the ^ feature and they are present in the lexicon. That is, the ^ feature (a) drives the computation in the narrow syntax and (b) contributes to the mapping of syntactic units (phases) to the interfaces. A transfer (phase) unit is transparently mapped onto the SEM as an ordered logical form LF. A spell-out (phase) unit is transparently mapped onto the PHON as an ordered phonological form PF. Thus, a transfer and a spell-out unit are equal in terms of the carried information mapped onto the SEM and PHON respectively. While it is obvious why a PF must be an ordered set, it is not clear why an LF must be an ordered set too. The Romeyka data shed light on the latter, as I will show below.

On the syntactic side, the basic operation is Merge (Chomsky 1995). In particular, Set Merge gives an unordered set {α, β}, whereas Pair Merge gives an ordered set <α, β>. In my theory, I assume that Merge can only create an ordered set <α, β>. This ordered set A can be merged with another ordered set B and so on, resulting in a Cartesian product. Since syntactic

units (phases) are transparently mapped onto the interfaces, I therefore assume that Cartesian products are mapped onto SEM as LFs.

Continuing the focus on the semantic side of this architecture, I follow Lohndal (2012) and argue that the LF is just a Transfer rule. These rules determine the mapping from narrow syntax to the SEM. Lohndal (2012) assumes that the SEM is what you get after the narrow syntactic representation has been handed over to the semantic interface. Furthermore, Lohndal (2012) shows that the syntactic representations can be mapped onto conjuncts. These conjuncts are then conjoined and existential closure is added. That is, while on the syntactic side the basic operation is Merge, on the semantic side I follow Lohndal (2012) by assuming that the main operation is Conjunction.

As such, at the SEM, an LF, which is a Cartesian product, is an open statement. In order to be interpreted at the SEM, it is made into a statement (a) by prefixing an existential quantifier, which binds the event variable and (b) by chaining all of the predicates together, as illustrated in (17):

(17) {∃e (Ae & Te) | e is a verb}

The statement in (17) would be interpreted as in (18):

(18) There exists an event e such that A is the agent of e and T is the Theme of e.[8]

For instance, in the sentence in (19), all of the predicates are chained together and then existential closure binds the event variable (see (20)):

(19) The boy kicked the ball.

(20) {∃e (Agent (e, the boy) & Theme (e, the ball)) | e is *kicked*}

[8] Throughout this study, the existence of event variables is taken for granted. Lexical items or roots can either be born with an event variable in the encyclopedia, or they can get an event variable from their categoriser. The latter view entails that the categoriser contributes the event variable, as in the following illustration for a verb.

i. V

V(e) √KILL

It follows that a syntactic derivation of a clause Σ, like the one in (21), is mapped onto the PHON as a PF, as in (22) and onto the SEM as an LF, as in (23):

(21) Syntactic derivation of a clause Σ:
Σ = {the boy} × <kicked, the ball>

(22) Phonological Form (PF) of the Σ:
PF_{Σ} = {the boy} × <kicked, the ball>

(23) Logical Form (LF) of the Σ:
LF_{Σ} = {the boy} × <kicked, the ball>

Afterwards, the semantic interpretation of the LF_{Σ} in (23) would be like the one in (24):

(24) Semantic interpretation of the LF_{Σ}:
{∃e (Agent (e, the boy) & Theme (e, the ball)) | e is *kicked*}

Finally, the semantic interpretation of an LF of the Σ, such as the one in (25), which I call 'ordinary interpretation' after Rooth (1985, 1992, 1996) and Beck (2006), interfaces with a pragmatic component, taking either a focus (see (26)) or a *wh*-question interpretation (see (27)):

(25) Ordinary interpretation:
{∃e (Agent (e, the boy) & Theme (e, the ball) | e is *kicked*}

(26) Focus interpretation:
a. {∃e ([Agent (e, someone)]$_F$ & Theme (e, the ball)) | e is *kicked*}
b. {∃e ([Agent (e, the boy)]$_O$ & Theme (e, the ball)) | e is *kicked*}

(27) *wh*-question interpretation:
{∃e ([Agent (e, someone)]$_F$ & Theme (e, the ball)) | e is *kicked*}

Based on the Romeyka data, I propose a semantic constraint on focus interpretation in Romeyka, which provides a principled explanation of the contribution of the ^ feature to a computationally efficient satisfaction of interface conditions.

First, consider the sentences in (28) and their LF representation in (29):

(28) Romeyka:

a. o mustafás epeLÁ́Épsen to χoráfin.
the.NOM Mustafas.NOM put.fertiliser.Past.3SG the.ACCfield.ACC
'Mustafas put fertiliser on the field.'
(S01; 150703_0040; 02:16)

b. #o mustafás to χoráfin epeLÁ́Épsen.
the.NOM Mustafas.NOM the.ACCfield.ACC put.fertiliser.Past.3SG
'Mustafas put fertiliser on the field.'

c. alís [ti MÁnan at]$_{Foc}$ efílisen.
Alis.NOM the.ACCmother.ACC his kiss.Past.3SG
'Alis kissed his mother.'
(S01; 150702_0013; 06:08)

d. #alís efílisen [ti MÁnan at]$_{Foc}$.
Alis.NOM kiss.Past.3SG the.ACCmother.ACC his
'Alis kissed his mother.'

(29) Logical Form (LF):

a. LF_{Σ} = {o mustafás} × {epelǽpsen} × {to χoráfin}
b. LF_{Σ} = {o mustafás} × <to χoráfin, epelǽpsen>
c. LF_{Σ} = {alís} × <ti mánan at, efílisen>
d. LF_{Σ} = {alís} × <efílisen, ti mánan at>

The felicity of the sentences in (28) shows that OV order in Romeyka is pragmatically marked in relation to VO. Moreover XP-V order in Romeyka is pragmatically marked in relation to V-XP —consider the sentences in (30), their LF representation in (31):

(30) Romeyka:

a. mían [alís]$_{Foc}$ efílisen tin aišén.
once Alis.NOM kiss.Past.3SG the.ACC Ayşe.ACC
'Alis kissed Ayşe once.'
(S01; 150702_0013; 06:24)

b. #[alís]$_{Foc}$ mían efílisen tin aišén.
Alis.NOM once kiss.Past.3SG the.ACC Ayşe.ACC
'Alis kissed Ayşe once.'

c. alís [mían]$_{Foc}$ efílisen tin aišén.
Alis.NOM once kiss.Past.3SG the.ACC Ayşe.ACC
'Alis kissed Ayşe once.'
(S01; 150702_0022; 02:46)

d. #[mían]$_{Foc}$ alís efílisen tin aišén.
once Alis.NOM kiss.Past.3SG the.ACC Ayşe.ACC
'Alis kissed Ayşe once.'

(31) Logical Form (LF):

a. LF_{Σ} = {mían} × <alís, efílisen> × {tin aišén}
b. LF_{Σ} = {alís} × <mían, efílisen> × {tin aišén}
c. LF_{Σ} = {alís} × <mían, efílisen> × {tin aišén}
d. LF_{Σ} = {mían} × <alís, efílisen> × {tin aišén}

This felicity of the sentences in (30) shows that there is good reason to assume that head-final phrases in Romeyka are more pragmatically marked than head-initial phrases, such that OV might be derived from VO and XP-V from V-XP. This asymmetry between a focused and a nonfocused clause can be captured in the semantic constraint in (32):

(32) Focus interpretation constraint in Romeyka:

A phrase α is focused iff

a. $\alpha \in LF$
b. $LF = \{<\alpha, \beta> \mid \beta \text{ is a verb}\}$

According to (32), a phrase α is focused if and only if (a) it is an element of a set LF and (b) LF is a set of the phrases α and β, in which α precedes β and β is a verb.

The constraint in (32) makes the following predictions for a monotransitive matrix clause with an overt subject, which can be seen in the Romeyka data:

First, consider a pragmatically unmarked clause, like the one in (33):

(33) Romeyka:

o	mustafás	epeLÆpsen	to	χoráfin.
the.NOM	Mustafas.NOM	put.fertiliser.Past.3SG	the.ACC	field.ACC

'Mustafas put fertiliser on the field.'

(S01; 150703_0040; 02:16)

The LF representation of (33) would be like the one in (34):

(34) LF_{Σ} = {o mustafás} × {epelæpsen} × {to χoráfin}

The semantic interpretation of (34) would be like the one in (35):

(35) {∃e (Agent (e, o mustafás) & Theme (e, to χoráfin)) | e is *epelæpsen*}

The constraint in (32) predicts that no phrase of the LF_{Σ} in (34) is focused, since although (a) is an element of a set LF, (b) LF is not a set of the phrases α and β, in which α precedes β and β is a verb.

Second, consider a pragmatically marked clause, like the one in (36), in which the object is focused:

(36) Romeyka:

alís	[ti	MÁnan	at]$_{\text{I-Foc}}$	efílisen.
Alis.NOM	the.ACC	mother.ACC	his	kiss.Past.3SG

'Alis kissed his mother.'

(S01; 150702_0013; 06:08)

The LF representation of (36) would be like the one in (37):

(37) LF_{Σ} = {alís} × <ti mánan at, efílisen>

The semantic interpretation of (37) would be like the one in (38):

(38) a. {∃e (Agent (e, alís) & [Theme (e, someone)]$_F$)) | e is *efílisen*}
b. {∃e (Agent (e, alís) & [Theme (e, ti mánan at)]$_O$)) | e is *efílisen*}

The constraint in (32) predicts that the DP-object *ti mánan at* 'his mother' in the LF$_\Sigma$ in (37) is focused, since (a) it is an element of a set LF and (b) LF is a set of the phrases α and β, in which α precedes β and β is a verb, i.e. *efílisen* 'he kissed'.

Third, consider a yes/no question, like the one in (39), in which the object is focused:

(39) Romeyka:
ta tsuPÁðæ θerízete?
the.ACC[pears.ACC]$_{I\text{-}Foc}$ harvest.3SG
'Do you harvest the pears?'
(S07; 812_0067; 01:58)

The LF representation of (39) would be like the one in (40):

(40) LF$_\Sigma$ = <tsupáðæ, θerízete>

The semantic interpretation of (40) would be like the one in (41):

(41) a. {∃e (Agent (e, esíst) & [Theme (e, something)]$_F$)) | e is *θerízete*}
b. {∃e (Agent (e, esíst) & [Theme (e, tsupáðæ)]$_O$)) | e is *θerízete*}

The constraint in (32) predicts that the DP-object *tsupáðæ* 'pears' in the LF$_\Sigma$ in (40) is focused, since (a) it is an element of a set LF and (b) LF is a set of the phrases α and β, in which α precedes β and β is a verb, i.e. *θerízete* 'you harvest'.

Fourth, consider a pragmatically marked clause, like the one in (42), in which the object is a *wh*-phrase:

(42) Romeyka:

alís	DÓɣna	aɣórasen?
Alis.NOM	what.ACC	buy.Past.3SG

'What did Alis buy?'

(S01; 812_0056; 04:13)

The LF representation of (42) would be like the one in (43):

(43) LF_{Σ} = {alís} × <dóɣna, aɣórasen>

The semantic representation of (43) would be like the one in (44):

(44) {∃e (Agent (e, alís) & [Theme (e, dóɣna)$_F$)) | e is *aɣórasen*}

The constraint in (32) predicts that the *wh*-phrase *doɣna* 'what' in the LF_{Σ} in (43) is focused, since (a) it is an element of a set LF and (b) LF is a set of the phrases α and β, in which α precedes β and β is a verb, i.e. *aɣórasen* 'he bought'.

Fifth, consider a pragmatically marked clause, like the one in (45), in which the subject is focused:

(45) Romeyka:

[o	mehMÉtis]$_{Foc}$	aɣórasen	to	biljísajarin.
the.NOM	Mehmetis.NOM	buy.Past.3SG	the.ACC	computer.ACC

'Alis bought the computer.'

(S01; 150703_0040; 05:32)

The LF representation of (45) would be like the one in (46):

(46) a. LF_{Σ} = <o mehmétis, aɣórasen> × {to biljísajarin}

The semantic representation of (46) would be like the one in (47):

(47) a. {∃e (Agent (e, someone)]$_F$ & Theme (e, to biljísajarin)) | e is *aɣórasen*}

b. {∃e (Agent (e, o mehmétis)]$_O$ & Theme (e, to biljísajarin)) | e is *aɣórasen*}

The constraint in (32) predicts that the DP-subject *o mehmétis* 'Mehmetis' in the LF_{Σ} in (46) is focused, since (a) it is an element of a set LF and (b) LF is a set of the phrases α and β, in which α precedes β and β is a verb, i.e. *ayórasen* 'he bought'.

Sixth, consider a yes/no question, like the one in (48), in which the subject is focused:

(48) Romeyka:
[eSÍ]$_{Foc}$ pézis futbólin?
you.NOM play.2SG football.ACC
'Do you play football?'
(S03; 150702_0015; 08:35)

The LF representation of (48) would be like the one in (49):

(49) a. LF_{Σ} = <esí, pézis> × {futbólin}

The semantic interpretation of (49) would be like the one in (50):

(50) a. {∃e (Agent (e, someone)]$_F$ & Theme (e, futbólin)) | e is *pézis*}
b. {∃e (Agent (e, esí)]$_O$ & Theme (e, futbólin)) | e is *pézis*}

The constraint in (32) predicts that the DP-subject *esí* 'you' in the LF_{Σ} in (49) is focused, since (a) it is an element of a set LF and (b) LF is a set of the phrases α and β, in which α precedes β and β is a verb, i.e. *pézis* 'you play'.

Seventh, consider a *wh*-question, like the one in (51), in which the subject is a *wh*-phrase:

(51) Romeyka:
TS Épiren to bilɟísajarin?
who.NOM buy.Past.3SG the.ACCcomputer.ACC
'Who bought the computer?'
(S01; 140102_0007; 05:36)

The LF representation of (51) would be like the one in (52):

(52) LF_{Σ} = <ts, epíren> × {to bilɟísajarin}

The semantic interpretation of (52) would be like the one in (53):

(53) {∃e (Agent (e, ts)]$_F$ & Theme (e, to bilɟísajarin)) | e is *epíren*}

The constraint in (32) predicts that the *wh*-phrase *ts* 'who' in the LF$_\Sigma$ in (52) is focused, since (a) it is an element of a set LF and (b) LF is a set of the phrases α and β, in which α precedes β and β is a verb, i.e. *epíren* 'he bought'.

Eighth, consider a pragmatically marked clause, like the one in (54), in which the object is left-dislocated and the DP-subject is focused:

(54) Romeyka:

tin	aišén	[o	dohTÓris]$_{I\text{-}Foc}$	epíren.
the.ACC	Ayşe.ACC	the.NOM	doctor.NOM	marry.Past.3SG

'The doctor married Ayşe.'

(S01; 140102_0008; 01:15)

The LF representation of (54) would be like the one in (55):

(55) LF$_\Sigma$ = {tin aišén} × <o dohtóris, epíren>

The semantic interpretation of (55) would be like the one in (56):

(56) a. {∃e (Theme (e, tin aišén) & Agent (e, someone)]$_F$) | e is *epíren*}
b. {∃e (Theme (e, tin aišén) & Agent (e, o dohtóris)]$_O$) | e is *epíren*}

The constraint in (32) predicts that the DP-subject *o dohtóris* 'the doctor' in the LF$_\Sigma$ in (55) is focused, since (a) it is an element of a set LF and (b) LF is a set of the phrases α and β, in which α precedes β and β is a verb, i.e. *epíren* 'he married'.

Ninth, consider a yes/no question, like the one in (57), in which the object is left-dislocated and the DP-subject is focused:

(57) Romeyka:

ató	o	mehmétis	éndžen	æ?
this.ACC	the.NOM	Mehmetis.NOM	bring.Past.3SG	it.ACC

'Did Mehmetis bring that?'

(S01; 150703_0042; 06:36)

The LF representation of (57) would be like the one in (58):

(58) LF_{Σ} = {ató} × <o mehmétis, éndžen>

The semantic interpretation of (58) would be like the one in (59):

(59) a. {∃e (Theme (e, ató) & Agent (e, o mehmétis)]$_F$) | e is *éndžen*}

b. {∃e (Theme (e, ató) & Agent (e, o mehmétis)]$_O$) | e is *éndžen*}

The constraint in (32) predicts that the DP-subject *o mehmétis* 'Mehmetis' in the LF_{Σ} in (58) is focused, since (a) is an element of a set LF and (b) LF is a set of the phrases α and β, in which α precedes β and β is a verb, i.e. *éndžen* 'he brought'.

Tenth, consider a *wh*-question, like the one in (60), in which the object is left dislocated and the subject is a *wh*-phrase:

(60) Romeyka:

tin	aišén	TS	epíren?
the.ACC	Ayşe.ACC	who.NOM	marry.Past.3SG

'Who married Ayşe?'

(S01; 140102_0008; 00:35)

The LF representation of (60) would be like the one in (61):

(61) LF_{Σ} = {tin aišén} × <ts, epíren>

The semantic interpretation of (61) would be like the one in (62):

(62) ∃e (Theme (e, DP-object) & [Agent (e, DP-subject)]$_F$, verb (e)) | e is *epíren*}

The constraint in (32) predicts that the *wh*-phrase *ts* 'who' in the LF_{Σ} in (61) is focused, since (a) it is an element of a set LF and (b) LF is a set of the phrases α and β, in which α precedes β and β is a verb, i.e. *epíren* 'he married'.

However, the constraint in (32) does not make any predictions for the focus interpretation in clauses with MF, like the one in (63), or in clauses with MWQ, like the one in (66).

Let us begin with a clause with MF, like the one in (63):

(63) Romeyka:

[o	mehMÉtis]$_{Foc}$	[araPÁN]$_{Foc}$	aɣórasen.
the.NOM	Mehmetis.NOM	car.ACC	buy.Past.3SG

'Mehmetis bought a car.'

(S01; 150703_0042; 07:39)

The LF representation of (63) would be like the one in (64):

(64) a. LF_{Σ} = {o mehmétis} × <arapán, aɣórasen>

The semantic interpretation of (64) would be like the one in (65):

(65) a. {∃e ([Agent (e, someone)]$_F$ & [Theme (e, something)]$_F$) | e is *aɣórasen*}

b. {∃e ([Agent (e, o mehmétis)]$_O$ & [Theme (e, arapán)]$_O$) | e is *aɣórasen*}

Let us now turn to a clause with MWQ, like the one in (66):

(66) Romeyka:

PÍos	DÓɣna	aɣórasen?
who.NOM	what.ACC	buy.Past.3SG

'Who bought what?'

(S01; 150703_0042; 07:37)

The LF representation of (66) would be like the one in (67):

(67) LF_{Σ} = {píos} × <dóɣna, aɣórasen>

The semantic interpretation of (67) would be like the one in (68):

(68) {∃e ([Agent (e, píos)]$_F$ & [Theme (e, dóɣna)]$_F$) | e is *ayórasen*}

Based on the semantic interpretation of the LF_Σ in (64) and (67), I propose the following prediction regarding pragmatically unmarked and marked word orders, which constrains the semantic interpretation of focused phrases in Romeyka (see (69)):

(69) Focus interpretation constraint in Romeyka:

a. A phrase α is focused iff

i. $\alpha \in LF$

ii. $LF = \{<\alpha, \beta> \mid \beta \text{ is a verb}\}$

b. A phrase α is focused iff

i. $\alpha \in LF_1$

ii. $LF_1 = \{\alpha\}$

iii. $LF_1 \times LF_2 = \{<\alpha, \beta, \gamma> \mid \alpha \in LF_1 \text{ and } \beta \in LF_2 \text{ and } \gamma \in LF_2\}$

iv. $LF_2 = \{<\beta, \gamma> \mid \gamma \text{ is a verb}\}$

v. $\beta \in LF_2 \mid \beta$ is [+focus] or [+*wh*]

According to (69)a (a), a phrase α is focused if and only if (i) is an element of a set LF and (ii) LF is a set of the phrases α and β, in which α precedes β and β is a verb. According to (69)b, a phrase α is focused if and only if (i) is an element of the set LF_1, (ii) LF_1 is a set of the phrase α, (iii) LF_1 precedes LF_2, such that α precedes β and β precedes γ.the phrase α is an element of the set $LF_{1\ \text{and}}$ (iv) LF_2 is a set of the phrases β and γ, in which β precedes γ and γ is a verb and (v) the phrase β is either [+focus] or [+*wh*] and is an element of the set LF_2.

However, the last revision of the focus interpretation constraint in Romeyka in (69) cannot capture the infelicity of the sentences in (70):

(70) Romeyka:

a. ?[o mehMÉtis]$_{Foc}$ DÓɣna aɣórasen?
the.NOM Mehmetis.NOM what.ACC buy.Past.3SG
'What did Mehmetis buy?'

b. ?[tšorBÁN]$_{Foc}$ PÍos éfaen?
soup.ACC who.NOM eat.Past.3SG
'Who ate only soup?'

The LF representation of (70) would be like the one in (71):

(71) Logical Form (LF) representation:

a. LF_{Σ} = {o mehmétis} × <dóɣna, aɣórasen>

b. LF_{Σ} = {tšorbán} × <píos, éfaen>

The semantic interpretation of (71) would be like the one in (72):

(72) Semantic interpretation:

a. i. {∃e ([Agent (e, someone)]$_F$ & [Theme (e, dóɣna)]$_F$) | e is *éfaen*}

ii. {∃e ([Agent (e, alís)]$_O$ & [Theme (e, dóɣna)]$_F$) | e is *éfaen*}

b. i. {∃e ([Theme (e, something)]$_F$ & [Agent (e, píos)]$_F$) | e is *éfaen*}

ii. {∃e ([Theme (e, tšorbán)]$_O$ & [Agent (e, píos)]$_F$) | e is *éfaen*}

Based on the infelicity of the sentences in (70), I propose the following prediction regarding pragmatically unmarked and marked word orders, which constrains the semantic interpretation of focused phrases in Romeyka (see (73)):

(73) Focus interpretation constraint in Romeyka:

a. A phrase α is focused iff

i. $\alpha \in LF$

ii. $LF = \{<\alpha, \beta> \mid \beta \text{ is a verb}\}$

b. A phrase α is focused iff

i. $\alpha \in LF_1$

ii. $LF_1 = \{\alpha\}$

iii. $LF_1 \times LF_2 = \{<\alpha, \beta, \gamma> \mid \alpha \in LF_1 \text{ and } \beta \in LF_2 \text{ and } \gamma \in LF_2\}$

iv. $LF_2 = \{<\beta, \gamma> \mid \gamma \text{ is a verb}\}$

v. $\beta \in LF_2 \mid \beta$ is [+focus] or [+*wh*]

vi. α carries the same semantic value as β; i.e. α is [+focus] as is β, or α is [+*wh*] as is β

According to (73)a, a phrase α is focused if and only if (i) is an element of a set LF and (ii) LF is a set of the phrases α and β, in which α precedes β and β is a verb. According to (73)b, a phrase α is focused if and only if (i) is an element of the set LF_1, (ii) LF_1 is a set of the phrase α, (iii) LF_1 precedes LF_2, such that α precedes β and β precedes γ.the phrase α is an element of the set LF_1 and (iv) LF_2 is a set of the phrases β and γ, in which β precedes γ and γ is a verb, (v) the phrase β is either [+focus] or [+*wh*] and is an element of the set LF_2 and (vi) α carries the same semantic value as β; i.e. α is [+focus] as is β, or α is [+*wh*] as is β.

So far, I have shown that the constraint in (73) can make predictions regarding the semantic interpretation of focused phrases in declarative and interrogative clauses in Romeyka.

In conclusion, there is good reason to assume that the linearisation feature in Romeyka contributes to the efficient computation of interface conditions, since it efficiently maps phases as ordered LFs and PFs onto the SEM and PHON respectively. In the remainder of this chapter, I test the predictions made by the constraint in (73) for every potential syntactic derivation of the subject (S), verb (V) and object (O) in Romeyka and I map those derivations into PF and LF rules. In particular, I examine (a) matrix declarative clauses, (b) subordinate declarative clauses, (c) direct questions and (d) indirect questions.

5.2 Application of the focus interpretation constraint in syntactic derivations in Romeyka

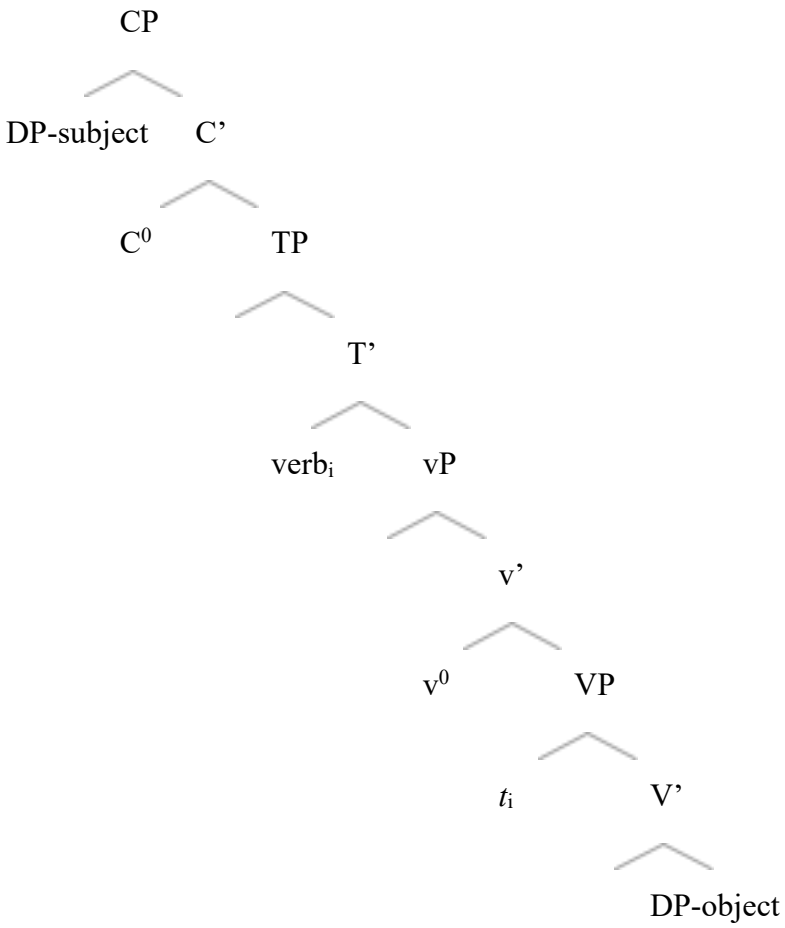
Based on the focus interpretation constraint proposed in (70), in this section I test the predictions made by this constraint in a pragmatically unmarked and a marked word order. I specifically map those syntactic derivations into PF and LF rules (for a detailed application of the focus interpretation constraint in syntactic derivations in Romeyka see Appendix B).

5.2.1 SVO in matrix declarative clauses in Romeyka

In a pragmatically unmarked SVO word order, the verb undergoes V^0-to-T^0 movement, while a ^ feature is associated with the EF of the phase head C^0, resulting in the left dislocation of the DP-subject (see (74) and (75)):

(74) $[_{CP}$ DP-subject C^0 $[_{TP}$ $verb_i$ $[_{vP}$ $[_{VP}$ t_i DP-object]]]]

(75)



For (74) and (75), the Spell-Out rule in (76) is applied:

(76) a. $PF_{\Sigma} = PF_1 \times PF_2 \times PF_3$
b. PF_{Σ} = CP × TP × VP
c. PF_{Σ} = {DP-subject} × {verb} × {DP-object}

For (74) and (75), the Transfer rule in (77) is applied:

(77) a. $LF_{\Sigma} = LF_1 \times LF_2 \times LF_3$
b. LF_{Σ} = CP × TP × VP
c. LF_{Σ} = {DP-subject} × {verb} × {DP-object}

The semantic interpretation of (77) would be like the one in (78):

(78) {∃e (Agent (e, DP-subject) & Theme (e, DP-object)) | e is a verb}

Consider the clause in (79):

(79) Romeyka:

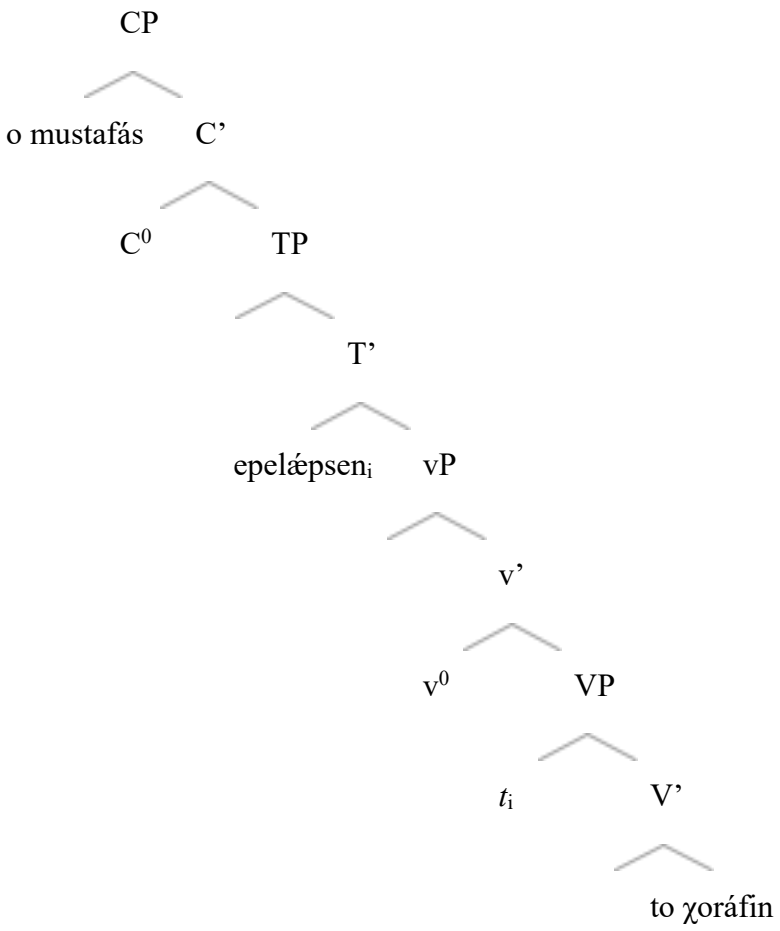
o	mustafás	epeLÆpsen	to	χoráfin.
the.NOM	Mustafas.NOM	put.fertiliser.Past.3SG	the.ACC	field.ACC

'Mustafas put fertiliser on the field.'
(S01; 150703_0040; 02:16)

In the clause in (79), the verb *eplæpsen* 'he put fertiliser' undergoes V^0-to-T^0 movement, while a ^ feature is associated with the EF of the phase head C^0, resulting in the left dislocation of the DP-subject *o mustafás* 'Mustafas' (see (80) and (81)):

(80) $[_{CP}$ o mustafás C^0 $[_{TP}$ epelæpsen$_i$ $[_{vP}$ $[_{VP}$ t_i to χoráfin]]]]

(81)



For (80) and (81), the Spell-Out rule in (82) is applied:

(82) a. $PF_{\Sigma} = PF_1 \times PF_2 \times PF_3$

b. $PF_{\Sigma} = CP \times TP \times VP$

c. PF_{Σ} = {o mustafás} × {epeláepsen} × {to χoráfin}

For (80) and (81), the Transfer rule in (83) is applied:

(83) a. $LF_{\Sigma} = LF_1 \times LF_2 \times LF_3$

b. $LF_{\Sigma} = CP \times TP \times VP$

c. LF_{Σ} = {o mustafás} × {epeláepsen} × {to χoráfin}

The semantic interpretation of (83) would be like the one in (84):

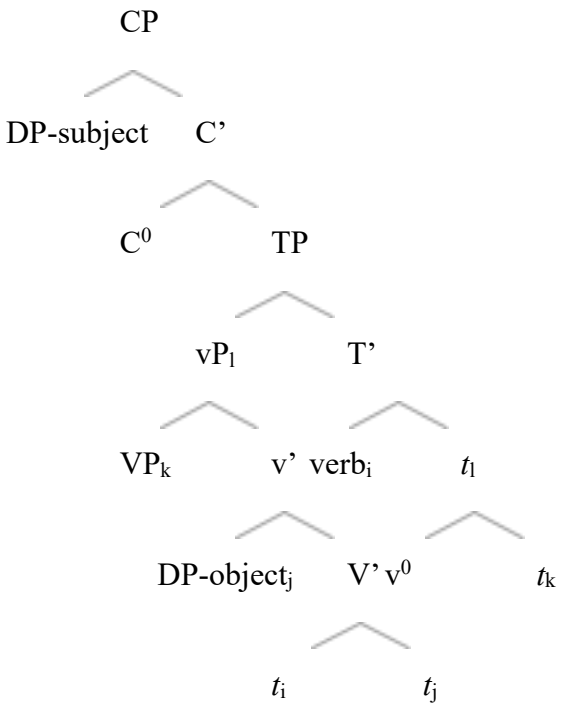
(84) {∃e (Agent (e, o mustafás) & Theme (e, to χoráfin)) | e is *epeláepsen*}

5.2.2 SOV in matrix declarative clauses in Romeyka

When the DP-object is focused in an SOV word order, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-object to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (85) and (86)):

(85) $[_{CP}$ DP-subject C^0 $[_{TP}$ $[_{vP}$ $[_{VP}$ DP-object$_j$ t_i $t_j]]$ verb$_i]]$

(86)



For (85) and (86), the Spell-out rule in (87) is applied:

(87) a. $PF_\Sigma = PF_1 \times PF_2$

b. $PF_\Sigma = CP \times TP$

c. PF_Σ = {DP-subject} × <DP-object, verb>

For (85) and (86), the Transfer rule in (88) is applied:

(88) a. $LF_{\Sigma} = LF_1 \times LF_2$
b. LF_{Σ} = CP × TP
c. LF_{Σ} = {DP-subject} × <DP-object, verb>

The semantic interpretation of (88) would be like the one in (89):

(89) a. {∃e (Agent (e, DP-subject) & [Theme (e, DP-object)]$_F$) | e is a verb}
b. {∃e (Agent (e, DP-subject) & [Theme (e, DP-object)]$_O$) | e is a verb}

Consider the clause in (90):

(90) Romeyka:

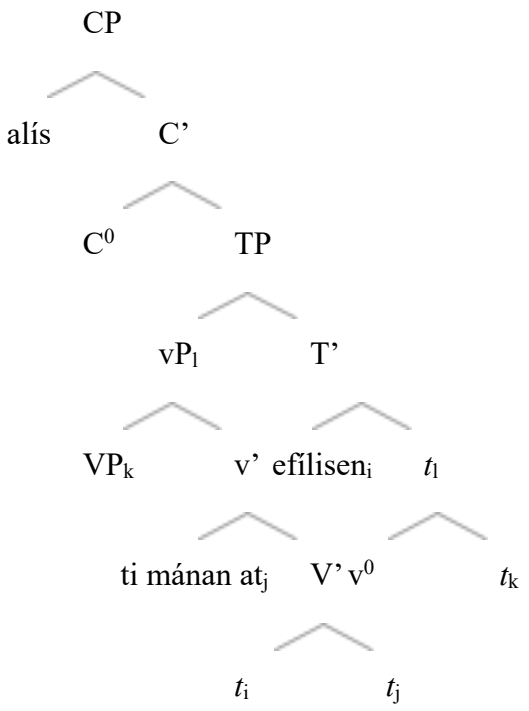
alís	[ti	MÁnan	at]$_{I\text{-}Foc}$	efílisen.
Alis.NOM	the.ACC	mother.ACC	his	kiss.Past.3SG

'Alis kissed his mother.'
(S01; 150702_0013; 06:08)

In the clause in (90), the verb *efílisen* 'he kissed' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-object *ti mánan at* 'his mother' to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (91) and (92)):

(91) [$_{CP}$ alís C^0 [$_{TP}$ [$_{vP}$ [$_{VP}$ ti mánan at$_j$ t_i t_j]] efílisen$_i$]]

(92)



For (91) and (92), the Spell-out rule in (93) is applied:

(93) a. $PF_{\Sigma} = PF_1 \times PF_2$

b. $PF_{\Sigma} = CP \times TP$

c. PF_{Σ} = {alís} × <ti mánan at, efílisen>

For (91) and (92), the Transfer rule in (94) is applied:

(94) a. $LF_{\Sigma} = LF_1 \times LF_2$

b. $LF_{\Sigma} = CP \times TP$

c. LF_{Σ} = {alís} × <ti mánan at, efílisen>

The semantic interpretation of (94) would be like the one in (95):

(95) a. {∃e (Agent (e, alís) & [Theme (e, ti mánan at)]$_F$) | e is *efílisen*}

b. {∃e (Agent (e, alís) & [Theme (e, ti mánan at)]$_O$) | e is *efílisen*}

5.3 Conclusions

In this chapter, I proposed a semantic constraint on focus interpretation, which makes predictions regarding the semantic interpretation of focused phrases in declarative and interrogative clauses in Romeyka. The predictions made by that constraint were tested for every potential syntactic derivation of the subject (S), verb (V) and object (O) in Romeyka and those derivations were mapped into PF and LF rules. The findings of this chapter show that there is good reason to assume that the linearisation feature in Romeyka contributes to the efficient computation of interface conditions, since it efficiently maps phases as ordered LFs and PFs onto the SEM and PHON respectively.

6 Typological classification of Romeyka word order

6.0 Introduction

The goal of this chapter is twofold: (a) it aims at typologically classifying Romeyka word order; and (b) it compares word order in Romeyka with word order (i) in Turkish, (ii) in Georgian and (iii) in Pontic Greek. The chapter is organised as follows: §6.1 typologically classifies Romeyka word order; §6.2 compares the word order in Romeyka with the one in Turkish; §6.3 compares the word order in Romeyka with the one in Georgian; and §6.4 compares the word order in Romeyka with the one in Pontic Greek. §6.5 provides a summary of the chapter.

6.1 Typological classification of Romeyka word order

6.1.0 Introduction

In this section, I put word order variation in Romeyka in a typological context. In particular, I show that, like German, Romeyka can be classified to "a third subtype of language lacking a dominant order, which consists of languages in which different word orders occur but the choice is syntactically determined" (Dryer 2005: 330-331). In Romeyka, the dominant order is SVO in matrix clauses and SOV in subordinate clauses. Matrix and subordinate clauses may contain an auxiliary, in which case the order is rigidly AuxVO.

The section is organised as follows: §6.1.1 classifies the order of subject, object and verb in Romeyka; §6.1.2 classifies the order of subject and verb in Romeyka; and §6.1.3 classifies the order of object and verb in Romeyka. §6.1.4 provides a summary of the section.

6.1.1 Order of Subject, Object and Verb in Romeyka

According to Dryer (2005: 338), there are three types of languages in terms of the order of object (O) and verb (V): (a) languages that are OV (in which the object precedes the verb), like Turkish (see (1)); (b) languages that are VO (in which the verb precedes the object) like Gulf Arabic (see (2)); (c) languages with both orders where neither order is dominant. Languages in which neither OV nor VO is dominant fall into two sorts: on the one hand, there are languages with a flexible word order where both orders are common and the choice is determined by extragrammatical factors. A second class of languages in which both OV and VO are common are languages in which word order is primarily determined syntactically, but in which there are competing OV and VO constructions. German is an instance of this, in that the VO order is used in matrix clauses in which there is no auxiliary verb, while the OV order is used in clauses with an auxiliary verb and in subordinate clauses introduced by a subordinator (see (3)):

(1) Turkish:
Mehmed-i gör-dü-m.
Mehmet-ACC see-Past-1SG
'I saw Mehmet.'
(Dryer 2005: 338 *apud* Underhill 1976: 51)

(2) Gulf Arabic:

ʔakalaw sandwiich-aat.

eat.3PL sandwich-PL

'They ate sandwiches.'

(Dryer 2005: 338 *apud* Holes 1990: 119)

(3) German:

a. Anna trink-t Wasser.

Anna drink-3SG water

'Anna is drinking water.'

(Dryer 2005: 338)

b. Anna ha-t Wasser getrunken.

Anna have-3SG water drink.Past.PART

'Anna has drunk water.'

(Dryer 2005: 338)

c. Hans sag-t, dass Anna Wasser trink-t.

Hans say-3SG that Anna water drink-3SG

Hans says that Anna is drinking water.'

(Dryer 2005: 339)

Another language, whose word order depends both on whether there is an auxiliary and whether the clause is a matrix clause, is Dinka (Nilotic; Sudan); like German, the order is SVO in matrix clauses without an auxiliary, SAuxOV in main clauses with an auxiliary, but it is VSO in subordinate clauses without an auxiliary and AuxSOV in subordinate clauses with an auxiliary (Dryer 2005: 331 *apud* Nebel 1948: 9, 25, 42, 75, 82) (see (4)):

(4) Kisi:

a. kɛ̀ùwó lɔ̀wá sàá

snake bite Saa

'The snake bit Saa.'

b. Fàlà có Lɛ́ɛ́ŋndó yìkpàá

Fallah PROG machete sharpen

'Fallah is sharpening the machete.'

(Dryer 2005: 339 *apud* Childs 1995: 249-250)

Like German, Romeyka can be classified as "a third subtype of language lacking a dominant order, which consists of languages in which different word orders occur but the choice is syntactically determined" (Dryer 2005: 330-331).[9]

In Romeyka, the dominant order is SVO in matrix clauses and SOV in subordinate clauses. Matrix and subordinate clauses in Romeyka may contain an auxiliary, in which case the order is rigidly AuxVO. In matrix clauses, the auxiliary *éχo* 'I have' is found in wishes and exclamatives. They are counterfactual optatives and they take an infinitive as their complement. Consistent with what holds for wishes and exclamatives is the use of *éχo* 'I have' in subordinate clauses, in particular, in the protasis of counterfactual conditionals.

6.1.2 Order of Subject and Verb in Romeyka

Regarding the dominant order of lexical (or nonpronominal) subject and verb, the primary types are languages that are SV (in which the subject precedes the verb), a type represented by English and Turkish, as illustrated in (5) and languages that are VS (in which the subject follows the verb), exemplified by Welsh, as illustrated in (6) (Dryer 2005):

(5) Turkish:
Su kayna-dı.
water boil-Past
'The water boiled.'
(Dryer 2005 *apud* Kornfilt 1997: 90)

(6) Welsh:
Daeth y dyn.
come.Past.3SG the man
'The man came.'
(Dryer 2005 *apud* Williams 1980: 165)

The VSO word order is attested cross-linguistically in languages where the SVO word order is the basic order. Greenberg (1966) was the first to capture this generalisation as a language

[9] The other two are (a) languages with a flexible order, in which there is one order that is most common that can be described as the dominant order and (b) languages with a flexible order, in which the flexibility is greater and there is no-one order that is dominant in terms of frequency of usage or pragmatic neutrality (see Dryer 2005: 330).

universal: "All languages with dominant VSO order have SVO as an alternative or as the only alternative basic order" (Greenberg 1966: 79). Greenberg (1966: Appendix II) lists Welsh, Hebrew and Berber among the VSO languages.

Although Greenberg classifies MG as an SVO language, more recent research seems to agree that it is a VSO language, with SVO as an alternative (see Alexiadou & Anagnostopoulou 1998, Roussou & Tsimpli 2006, i.a.).

In our case, it is interesting that Romeyka, which is an SVO language as I have shown in §3.1, lacks VSO orders. I suppose that such a mismatch does not necessarily raise typological queries, but rather witnesses an on-going macroparametric change in Romeyka (see chapter 7).

6.1.3 Order of Object and Verb in Romeyka

The order of object and verb has received considerable attention because of the fact that a large number of other features are predictable from it, at least in a statistical sense (Dryer 1992, Greenberg 1963, Hawkins 1983). For example, OV languages tend to be postpositional, genitive before noun, adverb before verb, complementiser at end of a clause and standard-marker-adjective order in comparative clauses, while VO languages tend to exhibit the opposite orders. The patterns are sometimes more complex than this. For example, while VO languages almost exclusively place relative clauses after nouns, both orders of relative clause and noun are common among OV languages. In addition, there are some word order features that do not correlate with the order of object and verb. For example, contrary to some claims, the order of adjective and noun does not correlate with the order of object and verb (Dryer 1988, 1992).

While the order of genitive and noun correlates with the order of object and verb (OV languages tending to be GenN and VO languages NGen), it differs from other pairs of elements whose order correlates in that SVO languages are intermediate between OV and verb-initial languages: SVO&GenN languages are as common as SVO&NGen languages.

In Romeyka, the GenN order is rigid, as illustrated in the following examples (see (7))[10]:

[10] For an account of the history of genitive in Greek, see Mertyris (2014).

(7) Romeyka:

íða	t	alí	ton	arapán.
see.Past.1SG	the.GEN	Alis.GEN	the.ACC	car.ACC

'I saw Alis' car.'

(S07; 812_0062; 00:09)

Furthermore, both orders of relative clause and noun are common in Romeyka; it places relative clauses both after the noun (see (8)a) and before the noun (see (8)b)[11]:

(8) Romeyka:

a.

i	patsí	[así	mazirán	d	érθen]
the.NOM	girl.NOM	from.the.ACC	Mazira.ACC	REL	come.Past.3SG

t	emón	t	anépsin	éton.
the.NOM	mine.NOM	the.NOM	niece.NOM	be.IMPF.3SG

'The girl who came from Mazira is my niece.'

(S01; 812_0112; 06:01)

b.

[así	mazirán	d	érθen]
from.the.ACC	Mazira.ACC	REL	come.Past.3SG

i	patsí	t	emón	t	anépsin	éton.
the.NOM	girl.NOM	the.NOM	mine.NOM	the.NOM	niece.NOM	be.IMPF.3SG

'The girl who came from Mazira is my niece.'

(S01; 812_0112; 06:06)

6.1.4 Summary

In this section, I have shown that, like German, Romeyka can be classified as "a third subtype of language lacking a dominant order, which consists of languages in which different word orders occur but the choice is syntactically determined" (Dryer 2005: 330-331). In Romeyka, the dominant order is SVO in matrix clauses and SOV in subordinate clauses. Matrix and subordinate clauses may contain an auxiliary, in which case the order is rigidly AuxVO.

[11] For a tentative presentation of relative clauses in Romeyka, see Gandon (2016).

6.2 Is Romeyka like Turkish?

6.2.0 Introduction

Turkish is a predominantly head-final language. Thus, it manifests an OV (object-verb) word order, a head-postposition order, a possessor-possessed order and a dependent verbal form–main verb order. It is also a pro-drop language and a complement-drop language, since agreement and possessive relationships are obligatorily marked on the head. Case marking in Turkish is extremely important for verb valency, as verbs take direct object complements in the nominative (indefinite/non-referential objects) and the accusative (definite objects), as well as the dative and the ablative, according to their properties. Modifiers (adjectives, pronouns, numerals, indefinite articles) generally precede nouns. Control structures are also evident in dependent verbal forms with no morphological marking with respect to their subject, as well as in infinitival constructions. Finally, one can observe a strictly fixed constituent order in Turkish and any derivations from it are explained on a pragmatic basis (see Göçmen et al. 1995).

The goal of this section is to compare word order in Turkish with that in Romeyka. The findings of this section show that Romeyka seems to be like Turkish, in that it has the focus to the left of the verb, but unlike Turkish in that VO is the pragmatically unmarked order in matrix clauses.

The section is organised as follows: §6.2.1 looks for the pragmatically unmarked word order in Turkish; §6.2.2 examines the syntactic distribution of pragmatically marked word orders in Turkish. §6.2.3 provides a summary of the findings of the section.

6.2.1 Pragmatically unmarked word orders in Turkish

The pragmatically unmarked word order in Turkish is argued to be OV (see Erguvanlı 1984, Erkü 1983, Kornfilt 1997, Şener 2010, i.a.) (see (9)):

(9) Turkish:

a. Question:

Ne ol-du?

what happen-Past.3SG

'What happened?'

b. Answer:

Cadı hırsız-ı lanetle-di.

witch-NOM thief-ACC curse-Past.3SG

'The witch cursed the thief.'

(Şener 2010: 10)

Unlike in Turkish, in Romeyka the pragmatically unmarked word order is VO.

6.2.2 Pragmatically marked word orders in Turkish

In Turkish, the focused constituent is argued to be placed immediately preverbally no matter what sub-type of focus it conveys (see Göksel & Kerslake 2005, Kornfilt 1997, Şener 2010, i.a.). It should also be noted that nothing that bears information or contrastive focus can be placed in the postverbal field (see Erguvanlı 1984, Göksel & Kerslake 2005, Kornfilt 1997, Şener 2010, i.a.). Likewise, in Romeyka, the focus is always to the left of the verb. That is, contrastive focus in both Turkish and Romeyka cannot move across contrastive topic (see (10)):

(10) Turkish:

a. Question:

Can'dan n'aber? O ne yedi partide?

'What about John? What did he eat at the party?'

Answers:

Valla Can-'ı bil-mi-yor-um, ama …

frankly Can-ACC know-NEG-PROG-1SG but

'Frankly, I don't know about John, but …'

b. [Aylın]$_{C\text{-}Top}$ [dolma-lar-dan]$_{C\text{-}Foc}$ ye-di.

Aylin-NOM dolma-PL-ABL eat-Past-3SG

'Aylin ate from the dolmas.'

c. #[dolma-lar-dan]$_{C\text{-}Foc}$ [Aylın]$_{C\text{-}Top}$ ye-di.

dolma-PL-ABL Aylin-NOM eat-Past-3SG

'Aylin ate from the dolmas.'

(Şener 2010: 19)

In the same way, information focus must be left adjacent to the verb in both Turkish and Romeyka. Thus, information focus is not different from contrastive focus in terms of its distribution, hence follows contrastive topics (see (11)):

(11) Turkish:

a. Question:

Filiz-in	kardeş-ler-i	ne	iç-ti	parti–de?
Filiz-GEN	sister-PL-POSS	what	drink-Past-3SG	party-LOC

'What did Filiz's sisters get to drink at the party?'

Answers:

Valla	tüm	kardeş-ler-den	haberim	yok,	ama …
frankly	all	sister-PL-ABL	news-POSS-1SG	NEG	but

'Frankly, I do not know about all the sisters but …'

b. [Filiz-in en küçük kardeş-i]$_{\text{C-Top}}$ [rakı-dan]$_{\text{I-Foc}}$ iç-ti.

Filiz-GEN	most	young	sister-3SG-POSS	rakı-ABL	drink-Past-3SG

'Filiz's youngest sister drank (from the) rakı.'

c. #[rakı-dan]$_{\text{I-Foc}}$ [Filiz-in en küçük kardeş-i]$_{\text{C-Top}}$ iç-ti.

rakı-ABL	Filiz-GEN	most	young	sister-3SG-POSS	drink-Past-3SG

'Filiz's youngest sister drank (from the) rakı.'

(Şener 2010: 35)

6.2.3 Summary

The findings of this section show that Romeyka seems to be like Turkish, in that it has the focus to the left of the verb, but unlike Turkish in that VO is the pragmatically unmarked order in matrix clauses.

6.3 Is Romeyka like Georgian?

6.3.0 Introduction

The goal of this section is to compare word order in Georgian with that in Romeyka. The findings of this section show that Romeyka seems to be like Georgian, in that it has the focus to the left of the verb, but unlike Georgian, in that it has both VO and OV as the pragmatically unmarked orders.

This section is organised as follows: §6.3.1 examines the unmarked word orders in Georgian; §6.3.2 examines the syntactic distribution of pragmatically marked word orders in Georgian. §6.3.3 provides a summary of the findings of the section.

6.3.1 Pragmatically unmarked word orders in Georgian

Georgian is characterised as a 'free word order' language: all permutations between major clausal constituents are grammatical (see Aronson 1982: 47, Boeder 1989: 160, Harris 1981: 22, Hewitt 1995: 528). The following examples illustrate the most frequent orders of clauses, SOV in (12)a and SVO in (12)b. Both orders may occur in all-new contexts:

(12) Georgian:

a. ʒarisk'ac-i monadire-s da-cˇ'r-i-s.
soldier-NOM hunter-DAT PR(FUT)-cut-THM-S.3.SG

b. ʒarisk'ac-i da-cˇ'r-i-s monadire-s.
soldier-NOM PR(FUT)-cut-THM-S.3.SG hunter-DAT
'The soldier will wound the hunter.'

(Skopeteas et al. 2009: 103)

Unlike in Georgian, in Romeyka the pragmatically unmarked word order is VO.

Despite this "free word order" it has been claimed that the basic order is V-final (see Aronson 1982: 47, Harris 2000: 141-146, Boeder 2005: 64). The syntactic evidence for the assumption of V-finality is admittedly weak, but all of the available criteria (e.g. the order in sequences of finite and non-finite verbs in Harris (2000: 145), object placement with coordinated verbs in Skopeteas & Fanselow (2010), suggest a V-final word order.

Regarding the organisation of information structure in Georgian, according to Asatiani (2007) and Skopeteas et al. (2009), VO order seems to result from (a) the predicate being

focused and (b) for an element to receive information focus postverbally. Skopeteas' (2012: 132) crucial generalisation for the study of information structure in Georgian is that the alternation between OV and VO orders does not depend on a semantic or a pragmatic trigger.

Skopeteas et al. (2009) suggest that Georgian has an optional syntactic operation that fronts the finite verb to a high position within the thematic layer of the clause. This operation is an instance of head movement, which however, is not triggered by a discourse feature and has the effect of giving Georgian mixed OV/VO properties. They assume that these sentences involve an optional head-fronting operation. The assumption of 'optional' V-fronting implies that VO orders are not necessarily the result of a movement operation that targets a position that is associated with a discrete information structural function. V-fronting is a semantically vacuous operation that may be optionally selected in discourse in order to meet preferences on the linearisation of the involved constituents.

According to Skopeteas et al. (2009), the difference between a head-final language such as Turkish that consistently prohibits postverbal focus and a head-final language such as Georgian, is that the latter (but not the former) has an optional operation of verb fronting with the following effects: (a) VO order may appear in all focus contexts and (b) *in-situ* focus surfaces as postverbal whenever the verb is fronted. The intuition behind this syntactic operation is that to the extent that the verb is fronted without a discrete trigger, Georgian behaves like a head-initial language. The consequence of optional verb fronting is the possibility of postverbal focus, which is available in head-initial languages. This is the crucial difference between a consistently head-final and a head-initial language: In an OV language without verb fronting, postverbal focus is not possible (which is the case for OV languages like Turkish).

6.3.2 Pragmatically marked word orders in Georgian

The following word orders have been encountered as valid permutations: SOV (see (13)), SVO (see (14)), OSV (see (15)), OVS (see (16)) and two orders containing argument ellipsis, SV (see (17)) and OV (see (18)):

(13) Georgian:

a. Question:

In the scene, in front of the well: is a/the boy pushing a/the bus?

b. Answer:

ara, bič'-i mankana-s a-c'v-eb-a.

no boy-NOM car-DAT (IO.3)PV-push-THM-PRS.S.3.SG

'No, a/the boy is pushing a/the car.'

(Condition: contrastive/object)

(Skopeteas & Franselow 2010: 7)

(14) Georgian:

a. Question:

In the scene, in front of the fence: what is a/the girl hitting?

b. Answer:

gogo u-rt'q'-am-s mankana-s.

girl(NOM) PV(IO.3)-hit-THM-PRS.S.3.SG car-DAT

'A/the girl is hitting a/the car.'

(Condition: non contrastive/object)

(Skopeteas & Franselow 2010: 7)

(15) Georgian:

a. Question:

In the scene with the blue sky: who is looking at a/the lamp?

b. Answer:

lamp'a-s k'ac-i u-q'ur-eb-s.

lamp-DAT man-NOM PV(IO.3)-ear-THM-PRS.S.3.SG

'No, a/the man and not a/the woman is cutting a/the melon.'

(Condition: non contrastive/subject)

(Skopeteas & Franselow 2010: 7)

(16) Georgian:

a. Question:

In the scene in the room: what is a/the man kicking?

b. Answer:

sk'am-s u-rt'q'-am-s igi.

chair-DAT PV(IO.3)-hit-THM-PRS.S.3.SG 3.SG.DIST:NOM

'He is hitting a/the chair.'

(Condition: non contrastive/object)

(Skopeteas & Franselow 2010: 7-8)

(17) Georgian:

a. Question:

In the scene, inside the house: is a/the woman cutting the melon?

b. Answer:

ara, k'ac-i č'r-i-s.

no man-NOM cut-THM-PRS.S.3SG

'No, a/the man is cutting it.'

(Condition: contrastive/subject)

(Skopeteas & Franselow 2010: 8)

(18) Georgian:

a. Question:

In the scene, in front of the blue wall: whom is the man pulling?

b. Answer:

kal-s e-kač-eb-a.

woman-DAT PV(IO.3)-pull-THM-PRS.S.3.SG

'(She) is pulling a/the woman.'

(Condition: non contrastive/object)

(Skopeteas & Franselow 2010: 8)

Likewise, in Romeyka, the focus is always to the left of the verb.

6.3.3 Summary

The findings of this section show that Romeyka seems to be like Georgian, in that the focus is to the left of the verb, but unlike Georgian, in that it has both VO and OV as the pragmatically unmarked orders.

6.4 Is Romeyka like Pontic Greek?

6.4.0 Introduction

In this section, I compare word order in Romeyka with word order in Pontic Greek. First consider the following extract from Sitaridou & Kaltsa (2014):

> "Pontic Greek is a variety of Asia Minor Greek spoken both within and outside of Greece by Greek nationals. Within Greece, it is mainly spoken in Macedonia (especially in Thessaloniki, Kozani, Imathia, Kilkis, Pieria and Drama), Thrace and to a lesser extent in Attica. Outside of Greece, it is spoken within Greek immigrant communities in diaspora, especially in Germany and in the USA. Despite the robustness of Pontic Greek speakers in Greece numerically (roughly estimated at 300.000 speakers), in real terms, large numbers of speakers have suffered attrition or are heritage speakers and, consequently, only a fraction of the estimated Pontic Greek-speaking population can be claimed to be native speakers of the variety. Due to the geographical dispersion of Pontic Greek, it is important to note that the term 'Pontic Greek', synchronically, can only be used as an "umbrella" term for the various sub-dialects, which, crucially, can diverge significantly from one other in terms of syntax" (Sitaridou & Kaltsa 2014: 3-4).

For the purposes of this study, I focus exclusively on the Pontic Greek varieties of Northern Greece and, in particular, on the variety used in the area of Thessaloniki (Sitaridou & Kaltsa 2014). Romeyka and Rumeic, i.e. the varieties spoken in the Azov region (Ukraine, Russia) (see Pappou-Zhouravliova 1995), are considered to be distinct from Pontic Greek, though Pontic Greek, Romeyka and Rumeic fall under AMG (along with Cappadocian).

Sitaridou & Kaltsa (2014) argue that in Pontic Greek, which is a VO language, discourse operations result in a great deal of OV word order. Specifically, they argue that (a) information focus is obligatorily in the left periphery and (b) a Contrast projection is in the CP domain, which can host both topics and foci.

The findings of this section show that Romeyka seems to be like Pontic Greek, in that it has VO as the pragmatically unmarked order in matrix declarative clauses and the focus is to the left of the verb, but unlike Pontic Greek, in that it has a single focus position.

This section is organised as follows: §6.4.1 examines the unmarked word orders in Pontic Greek; §6.4.2 examines the syntactic distribution of topics in Pontic Greek; and §6.4.3

examines the syntactic distribution of foci in Pontic Greek. §6.4.4 provides a summary of the findings of the section.

6.4.1 Pragmatically unmarked word order in Pontic Greek

The pragmatically unmarked word order in Pontic Greek is VO (see (19)) like in Romeyka:

(19) Pontic Greek:

énas	ɟinéka	pánda	epérine	kréas.
a.NOM	woman.NOM	always	buy.Past.3SG	meat.ACC

'A woman always bought meat.'
(Mackridge 1990: 119)

6.4.2 Topics in Pontic Greek

There are two main strategies for conveying old information in Pontic Greek: (a) ClLD, as in Romeyka and (b) usage of a particle, like in Romeyka (Sitaridou & Kaltsa 2014: 6).

First, let us consider ClLD in Pontic Greek (see (20)a and (20)b). Like in Romeyka (see (21)a), clitic doubling (CD) with right dislocation is ungrammatical both in Pontic Greek (see (20)c and (20)d) and in Romeyka (see (21)b) (Sitaridou & Kaltsa 2014: 6):

(20) Pontic Greek:

a.

tin	elean	eðek	aten	to	ɟitonan.
the.ACC	olive.ACC	give.Past.1SG	her	the.ACC	neighbour.ACC

'I gave the olive to the neighbour.'

b.

?ton	ɟitonan	eðek	aton	din	elean.
the.ACC	neighbour.ACC	give.Past.1SG	he.ACC	the.ACC	olive.ACC

c.

*eðek	aten	to	ɟitonan	din	elean.
give.Past.1SG	her	the.ACC	neighbour.ACC	the.ACC	olive.ACC

d.

*eðek	aton	to	ɟitonan	din	elean.
give.Past.1SG	he.ACC	the.ACC	neighbour.ACC	the.ACC	olive.ACC

(Sitaridou & Kaltsa 2014: 6 *apud* Drettas 1997: 278)

(21) Romeyka:

a. ombrón ta patsíðæ s okʰúlːin tš epóliɣan æ.
then the.ACC girls.ACC to school.ACC NEG send.Past.3PL them
'Then, they were not sending the girls to school.'
(S01; 150702_0019; 03:26)

b. ?ombrón tš epóliɣan æ s okʰúlːin ta patsíðæ.
then NEG send.Past.3PL them to school.ACC the.ACC girls.ACC
'Then, they were not sending the girls to school.'

Sitaridou & Kaltsa (2014) discuss the frequent occurrence of the *pa*-particle in Pontic Greek and argue that it functions as a contrastive topic particle. Despite the relatively frequent usage of *pa(l)*-phrases in Romeyka, the *pa(l)*-particle does not seem to express contrast, as in Pontic Greek (see Sitaridou & Kaltsa 2014); rather it seems to reflect a stage prior to the one we find in Pontic Greek where grammaticalisation of contrast into the particle seems not to have taken place, as in Phárasiot and Rumeic (see Agouraki 2010, Dawkins 1916, Kisilier 2007).

6.4.3 Foci in Pontic Greek

Focus in Pontic Greek is expressed (a) by a fronting operation in the case of information focus, or (b) with particles attached to the verbal constituent undergoing contrastive focalisation (Sitaridou & Kaltsa 2014: 12).

Consider the sentences in (22) and (24) from Pontic Greek and the equivalent ones in (23) and (25) from Romeyka:

(22) Pontic Greek:

a. Question:

o jorikas do efane?
the.NOM Yorikas.NOM what.ACC eat.Past.3SG?
'What did Yorikas eat?'

Answers:

b. (o jorikas) [to χošaf]$_{\text{I-Foc}}$ efaen.
the.NOM Yorikas.NOM the.ACC stewed.fruit.ACC eat.Past.3SG
'Yorikas ate a (traditional) soup.'

c. #o jorikas efaen [do χošaf]$_{\text{I-Foc}}$.
the.NOM Yorikas.NOM eat.Past.3SG the.ACC stewed.fruit.ACC
'Yorikas ate a (traditional) soup.'

(Sitaridou & Kaltsa 2014: 12)

(23) Romeyka:

a. Question:

alís DÓγna éfaen?
alis.NOM what.ACC eat.Past.3SG?
'What did Alis eat?'

Answers:

b. alís [χaVÍts]$_{\text{I-Foc}}$ éfaen.
Alis.NOM pudding.ACC eat.Past.3SG
'Alis ate a pudding.'
(S01; 150703_0040; 07:14)

c. #alís éfaen [χaVÍts]$_{\text{I-Foc}}$.
Alis.NOM eat.Past.3SG pudding.ACC
'Alis ate a pudding.'

(24) Pontic Greek:

a. Question:

do eðevasen?

what read.Past.3SG?

'What did he read?'

Answers:

b. pola vivlia eðevasen.

many.ACC books.ACCread.Past.3SG

'He read many books.'

c. #eðevasen pola vivlia.

read.Past.3SG many.ACC books.ACC

'He read many books.'

(Sitaridou & Kaltsa 2014: 12)

(25) Romeyka

a. DÓ eχújepsen?

what.ACC read.Past.3SG

'What did he read?'

Answers:

b. [pol:ǽ c^{h}iTÁpæ]$_{\text{I-Foc}}$ eχújepsen.

many.ACC books.ACC read.Past.3SG

'He read many books.'

(S01; 812_0059; 00:10)

c. #eχújepsen [pol:ǽ c^{h}iTÁpæ]$_{\text{I-Foc}}$.

read.Past.3SG many.ACC books.ACC

'He read many books.'

From the examples in (22) and (24) from Pontic Greek and the equivalent ones in (23) and (25) from Romeyka, it becomes obvious that Romeyka allows for information focus to the left of the verb.

The Pontic Greek pattern is reminiscent of what has recently been claimed about information focus, namely that it also commonly appears within the left periphery (Sitaridou & Kaltsa 2014: 13). This operation is dubbed focus-fronting and is different from contrastive fronting since a contrastive interpretation of the focus constituent is not necessary. Analogous

to Romeyka, information focus-fronting in Pontic Greek can involve any type of phrase: direct object (NP) (see (26) from Pontic Greek and (27) from Romeyka), direct object (DP) (see (28) from Pontic Greek and (29) from Romeyka), indirect object (beneficiary) (DP) (see (30) from Pontic Greek and (31) from Romeyka), predicative (adjective) (see (32) from Pontic Greek and (33) from Romeyka), adverbial (NP) (see (34) from Pontic Greek and (35) from Romeyka) and existential constructions (see (36) from Pontic Greek and (37) from Romeyka):

(26) Pontic Greek:

a. Question:

do efaes?

what.ACC eat.Past.2SG

'What did you eat?'

b. Answer:

[χavits]$_{\text{I-Foc}}$ efaa.

pudding.ACC eat.Past.1SG

'I ate pudding.'

(Sitaridou & Kaltsa 2014: 13 *apud* Drettas 1997: 280)

(27) Romeyka:

a. Question:

alís DÓɣna éfaen?

alis.NOM what.ACC eat.Past.3SG?

'What did Alis eat?'

b. Answer:

alís [χaVÍts]$_{\text{I-Foc}}$ éfaen.

Alis.NOM pudding.ACC eat.Past.3SG

'Alis ate a pudding.'

(S01; 150703_0040; 07:14)

(28) Pontic Greek:

a. Question:

do eplises?

what.ACC wash.Past.2SG

'What did you wash?'

b. Answer:

[ta poðaræ m]$_{\text{I-Foc}}$ eplisa.

the.ACCfeet.ACC I.POSS wash.Past.1SG

'I washed my feet.'

(Sitaridou & Kaltsa 2014: 13 *apud* Drettas 1997: 280)

(29) Romeyka:

a. Question:

i aišé TÍnan epíren?

the.NOM Ayşe.NOMwho.ACC marry.Past.3SG

'Who did Ayşe marry?'

b. Answer:

i aišé [ton doHTÓrin]$_{\text{I-Foc}}$ epíren.

the.NOM Ayşe.NOMthe.ACC doctor.ACC marry.Past.3SG

'Ayşe married the doctor.'

(S01; 140102_0008; 01:37)

(30) Pontic Greek:

epita ti nifæn θa eniγane lutron.

then the.ACCbride.ACC PRT.MOD open.IMPF.3PL bath.ACC

'Then they would prepare the bath for the married girl.'

(Sitaridou & Kaltsa 2014: 13 *apud* Drettas 1997: 280)

(31) Romeyka:

a. Question:

to c^{h}itápin TÍnan éndžes?
the.ACCbook.ACC who.ACC bring.Past.2SG
'To whom did you give the book?'

b. Answer:

to c^{h}itápin [ton ɟuSÚfin]$_{\text{I-Foc}}$ éŋga.
the.ACCbook.ACC the.ACC Yusufis.ACC bring.Past.1SG
'I brought the book for Yusufis.'
(S01; 150703_0042; 00:54)

(32) Pontic Greek:

a. Question:

do en atos?
what.ACC be.3SG he
'What is he like?'

b. Answer:

palalos en.
crazy.NOM be.3SG
'He is crazy.'

(Sitaridou & Kaltsa 2014: 14 *apud* Drettas 1997: 555)

(33) Romeyka:

a. Question:

alís DO en?
Alis.NOM what.NOM be.3SG
'What is Alis?'

b. Answer:

alís [Áɣuros]$_{\text{I-Foc}}$ en.
Alis.NOM boy.NOM be.3SG
'Alis is a boy.'
(S01; 140102_0009; 00:20)

(34) Pontic Greek:

mesaniχts eton.

midnight be.Past.3SG

'It was midnight.'

(Sitaridou & Kaltsa 2014: 14 *apud* Drettas 1997: 555)

(35) Romeyka:

a. Question:

i	mána	s	PÓte	efáisen	ton	musafírin?
the.NOM	mother.NOM	you.POSS	when	feed.Past.3SG	the.ACC	guest.ACC

'When did your mother feed the guest?'

b. Answer:

[oPSÉ]$_{\text{I-Foc}}$	efáisen	ton	musafírin.
yesterday	feed.Past.3SG	the.ACC	guest.ACC

'She fed the guest yesterday.'

(S01; 150703_0041; 07:10)

(36) Pontic Greek:

χorafæ c^{h} ine.

fields NEG exist.3PL

'There are no fields'

(Sitaridou & Kaltsa 2014: 14)

(37) Romeyka:

a. Question:

o	šc^{h}íl:on	DO	en?
the.NOM	dog.NOM	what.NOM	be.3SG

'What is the dog?'

b. Answer:

[haiVÁnin]$_{\text{I-Foc}}$	en.
animal.NOM	be.3SG

'It's an animal.'

(S01; 140102_0009; 00:35)

Analogous to Romeyka, in Pontic Greek focus-fronting also applies to questions of "total ignorance" that yield a yes/no reply (Sitaridou & Kaltsa 2014: 14). See (38) from Pontic Greek and (39) from Romeyka:

(38) Pontic Greek:

a. Question:

t apiðæ ekserts?

the.ACCpears.ACC know.2SG

'Do you know the pears?'

b. ???ekserts ta apiðæ?

know.2SG the.ACCpears.ACC

'Do you know the pears?'

(Sitaridou & Kaltsa 2014: 14)

(39) Romeyka:

a. [i NÍfe]$_{\text{I-Foc}}$ efáisen ti mamíka?

the.NOM daughter-in-law.NOM feed.Past.3SG the.ACCmother-in-law.ACC

'Did the daughter-in-law feed her mother-in-law?'

(S01; 150702_0013; 13:53)

b. esís [ta tsuPÁðæ]$_{\text{I-Foc}}$ θerízete?

you.NOM the.ACCcorn.ACC harvest.2PL

'Do you harvest the corn?'

(S07; 812_0067; 01:58)

c. ató [o mehMÉtis]$_{\text{I-Foc}}$ éndžen æ?

this.ACC the.NOM Mehmetis.NOM bring.Past.3SG it.ACC

'Did Mehmetis bring that?'

(S01; 150703_0042; 06:36)

Like in Romeyka, in Pontic Greek, strict adjacency seems to hold between the moved constituent and the predicate, especially in cases where the predicate is the verb *be* or *have* (Sitaridou & Kaltsa 2014: 14). See (40) from Pontic Greek and (41) from Romeyka:

(40) Pontic Greek:

a. Question:

aiksa esne panda.

like.this be.IMPF.2SG always

'You were always like this.'

b. Answer:

*aiksa panda esne.

like.this always be.IMPF.2SG

'You were always like this.'

(Sitaridou & Kaltsa 2014: 14 *apud* Drettas 1997: 182)

(41) Romeyka:

a. Question:

alís LÁɣa efílisen tin aišén?

Alis.NOM how kiss.Past.3SG the.ACC Ayşe?

'How did Alis kiss Ayşe?'

Answers:

b. alís [ŠÍta]$_{\text{I-Foc}}$ efílisen tin aišén.

Alis.NOM immediately kiss.Past.3SG the.ACCAyşe.ACC

'Alis kissed Ayşe once.'

(S01; 150702_0022; 04:07)

c. #alís efílisen [ŠÍta]$_{\text{I-Foc}}$ tin aišén.

Alis.NOM kiss.Past.3SG immediately the.ACCAyşe.ACC

Furthermore, like in Romeyka, movement of the focused constituent in subordinate clauses in Pontic Greek is possible (Sitaridou & Kaltsa 2014: 14). See (42) from Pontic Greek and (43) from Romeyka:

(42) Pontic Greek:

a. eθaresen oti tšantarmas eton.
think.Past.3SG that policeman be.IMPF.3SG
'He thought (that) he was a policeman.'

b. eθaresen džantarmas eton.
think.Past.3SG policeman be.IMPF.3SG
'He thought he was a policeman.'

(Sitaridou & Kaltsa 2014: 15 *apud* Drettas 1997: 370)

(43) Romeyka:

a. Question:
DO θarís, alís TÍnan efílisen?
what.ACC think.2SG Alis.NOM who.ACC kiss.Past.3SG
'Who do you think that Alis kissed?'

b. Answer:
eɣó θaró, alís [tin aiŠÉN]$_{\text{I-Foc}}$ efílisen.
I think.1SG Alis.NOM the.ACCAyşe.ACC kiss.Past.3SG
'I think that Alis kissed Ayşe.'
(S01; 150703_0040; 19:07)

Contrastive focus in Pontic Greek is realised through (a) the use of discourse particles unlike in Romeyka, or (b) focus movement like in Romeyka. Let us begin with (a). One of the focus particles in the particle *cela* (Sitaridou & Kaltsa 2014: 11). It is always in postposition, but not enclitic to the verb (see (44) and (45)):

(44) Pontic Greek:

a. kit ecekace c^{h} eleps ato cela.
lie.3SG there and NEG see.2SG it.ACC PRT
'It is there and you don't even see it.'

b. efaen do fain atun c edoken atsen cela.
eat.Past.3SG the.ACCfood.ACC their and strike.Past.3SG them PRT
'He ate their food and beat them as well.'

(Sitaridou & Kaltsa 2014: 11 *apud* Drettas 1997: 410)

(45) Pontic Greek

atos … eperane ci ti marian eksenkan aten
he take.Past.3PL PRT the.ACC Maria.ACC take.out.Past.3PL her.ACC
aso plan din bortan.
from.the.ACC sides.ACC the.ACC door.ACC
'He … they took Maria and forced her to exit through the side door.'
(Sitaridou & Kaltsa 2014: 11 *apud* Drettas 1997: 481)

Moving now to focus strategy (b), consider (46) from Pontic Greek and (47) from Romeyka:

(46) Pontic Greek:

a. Question:
θelts na pseno sen gaiven
want.2SG PRT.MOD make.1SG you.ACC coffee.ACC
ci ena ðio otia na vukuse?
and one.ACC two.ACC sweets.ACC PRT.MOD eat.PNP.2SG
'Do you want me to make you some coffee and a couple of sweets to eat?'

Answers:

b. kaiven pseson.
coffee.ACC bake.IMP.2SG
'Make coffee (and not something else).'

b'. manaχon kaiven pseson.
only coffee.ACC bake.IMP.2SG
'Only make coffee.'

c. *manaχon kaiven pa pseson.
only coffee.ACC PRT bake.IMP.2SG
'Only make coffee.'

d. kaiven pa θelo.
coffee.ACC PRT want.1SG
'I want coffee.'

e. kaiven pa θelo, otia pa θelo.
coffee.ACC PRT want.1SG sweets.ACC PRT want.1SG
'I want both coffee and cookies.'

(Sitaridou & Kaltsa 2014: 11-12)

(47) Romeyka:

a. Question:

kahVÉN jóksa TŠáin θélis?

coffee.ACC or tea.ACC want.2SG

'Do you want coffee or tea?'

b. Answer:

eɣó [kahVÉN]$_{\text{C-Foc}}$ θélo.

I coffee.ACC want.1SG

'I want coffee.'

(S01; 150702_0013; 12:15)

In (46) and (47), we conclude that focus movement in Pontic Greek is on a par with Romeyka.

However, in the case of contrastive focus, strict adjacency between the focused constituents and the verb holds in Romeyka, but not in Pontic Greek (see (48) from Pontic Greek and (49) from Romeyka) (Sitaridou & Kaltsa 2014: 15):

(48) Pontic Greek:

aika emorfa peðja esis kamian iðeten?

such beautiful.ACC children.ACC you.NOM ever see.Past.2PL

'Have you ever seen such beautiful children?'

(Sitaridou & Kaltsa 2014: 15 *apud* Drettas 1997: 183)

(49) Romeyka:

a. Question:

alís tšaBÚuχa eχújepsen to c^{h}itápin?

Alis.NOM quickly read.Past.3SG the.ACCbook.ACC

'Did Alis read the book quickly?'

Answers:

b. ɟo, [ŠÍta]$_{\text{C-Foc}}$ eχújepsen a.

no immediately read.Past.3SG it.ACC

'No, he read it immediately.'

(S01; 150702_0022; 05:43)

c. #ɟo, alís [ŠÍta]$_{\text{C-Foc}}$ to c^{h}itápin eχújepsen

no Alis.NOM immediately the.ACCbook.ACC read.Past.3SG

Sitaridou & Kaltsa (2014) argue that *pa* in Pontic Greek is the exponent of Contrast^{0}. They also discuss how the *pa* and *ci* markers are, in fact, in complementary distribution. *ci* is one of the contrastive focus particles. Consider (50):

(50) Pontic Greek:

ar aets pontiaka pe aton ci na ekser.

so this.way Pontic.ACC tell.IMP.2SG he.ACC PRT PRT.MOD know.3SG

'Hence, tell him in Pontic Greek so that he understands.'

(Sitaridou & Kaltsa 2014: 22 *apud* Drettas 1997: 523)

In (50) the *ci* particle appears attached to a verb+clitic complex and focalises the entire predicate. *ci* does not attach enclitically to any other element except for predicates. For this reason, (51) is ungrammatical:

(51) Pontic Greek:

*tin anasta ci iða.

the.ACCAnasta.ACC PRT see.Past.1SG

'I saw ANASTA.'

(Sitaridou & Kaltsa 2014: 22)

According to Sitaridou & Kaltsa (2014), the presence of *ci* under Contrast implies high verb movement to FocusP. Evidence for such an analysis comes from (a) the incompatibility of a *pa*-phrase and *ci*-phrase, as shown in (52)a, because a *pa*-constituent is not compatible with a focused verb; and (b) the complementarity of distribution between V-to-C and negation, as shown in (52)b, since it is well known that a Neg head blocks V-to-C, irrespective of the trigger:

(52) Pontic Greek:

a. *tin anasta pa pe aten ci.
the.ACCAnasta.ACC PRT tell.IMP.2SG she.ACC PRT
'Tell Anasta!'
(Sitaridou & Kaltsa 2014: 22)

b. *cʰ iða ci tin anasta
NEG see.Past.1SG PRT the.ACCAnasta.ACC
(ekusa tin anasta).
hear.Past.1SG the.ACCAnasta.ACC
'I didn't see Anasta (I heard Anasta).'
(Sitaridou & Kaltsa 2014: 22)

It follows that the discourse particles *pa* and *ci* have specialised selectional requirements. This clearly demarcates *ci* from *cela*. Therefore, the claim that Sitaridou & Kaltsa (2014) put forward is that *pa* attaches to XPs and *ci* to X^0s.

Consider the orderings in (53), (54), (55), (56), which give us an insight into the overall articulation of the information structure in Pontic Greek:

(53) CLlD-Object - Subject pa – V

ton memet eγo pa aɣapo aton.
the.ACCMemet.ACC I PRT love.1SG he.ACC
'It is Mehmet that I love.'
(Sitaridou & Kaltsa 2014: 23 *apud* Melanofrydis 2001: 13)

(54) Subject – Object pa – I-Foc – V

i	nazlu	xanum	ecinon	pa
the.NOM	Nazlu.NOM	lady.NOM	this.ACC	PRT

efcero	ci	θ	afin.
empty.ACC	NEG	PRT.FUT	leave.3SG

'Nazlu-hanum wouldn't leave this empty.'

(Sitaridou & Kaltsa 2014: 23 *apud* Melanofrydis 2001: 43)

(55) Object pa – I-Foc – V

eplirosam	efta	çilæðes;	c	ecina	pa
pay.Past.1PL	seven	thousand;	and	these.ACC	PRT

o	popas	eton	c	epicen	ato.
the.NOM	priest.NOM	be.Past.3SG	and	do.Past.3SG	it.ACC

'We paid 7000 (drachmas); and as for these, it was thanks to the priest that we managed (to pay so little).'

(Sitaridou & Kaltsa 2014: 23 *apud* Drettas 1997: 442)

(56) Subject pa - Top – I-Foc - V

eɣo	pa	osimeron	pola	stenaχorementza	ime.
I	PRT	today	very	sad	be.1SG

'Today I am very sad.'

(Sitaridou & Kaltsa 2014: 23 *apud* Andreadis 1990: 27)

Sitaridou & Kaltsa (2014: 23) argue for a low focus position in the vP-periphery of the clause, in line with Belleti (2004). The diagnostics for a low focus position in the vP periphery are presented below:

First, consider the position of postverbal subjects (see (57) and (59) from Pontic Greek and (58) and (60) from Romeyka):

(57) Pontic Greek:

a. Question:

pios erθen?

who.NOM come.Past.3SG

'Who came?'

b. Answer:

o ʝorikas erθen.

the.NOM Yorikas.NOM come.Past.3SG

'Yorikas came.'

(Sitaridou & Kaltsa 2014: 23)

(58) Romeyka:

a. Question:

PÍos érθen?

who.NOM come.Past.3SG

'Who came?'

b. Answer:

aLÍS érθen.

Alis.NOM come.Past.3SG

'Alis came.'

(S01; 812_0056; 08:52

(59) Pontic Greek:

a. Question:

do eɣomosen?

what fill.Past.3SG

'What got filled?'

b. Answer:

to potir eɣomosen.

the.NOM glass.NOM fill.Past.3SG

'The glass got filled.'

(Sitaridou & Kaltsa 2014: 23)

(60) Romeyka:

a. Question:

DÓɣna ɣómosen?

what fill.Past.3SG

'What got filled?'

b. Answer:

[to parTÁçin] eɣómosen.

the.NOM glass.NOM fill.Past.3SG

'The glass got filled.'

(S01; 812_0056; 03:47)

Second, consider the position of the focused adverbial (see (61) from Pontic Greek and (62) from Romeyka):

(61) Pontic Greek:

opse o juras efien.

yesterday the.NOM Yuras.NOM leave.Past.3SG

'Yesterday Yuras left.'

(62) Romeyka

alís [oPSÉ]$_{Foc}$ érθen asín tšáikaran.

Alis.NOM yesterday come.Past.3SG from.the.ACC Çaykara.ACC

'Alis came from Cayraka yesterday.'

(S01; 140102_0007; 03:26)

Third, consider the position of *wh*-phrases in multiple *wh*-questions (see (63) from Pontic Greek and (64) from Romeyka):

(63) Pontic Greek:

o	juras	tinan	pote	efilise?
the.NOM	Yuras.NOM	who.ACC	when	kiss.Past.3SG
(*o	juras)?			
the.NOM	Yuras.NOM			

'When Yuras kissed whom?'

(Sitaridou & Kaltsa 2014: 24)

(64) Romeyka:

alis	PÓte	TÍnan	efilisen?
Alis.NOM	when	who.ACC	kiss.Past.3SG

'When Alis kissed whom?'

(S01; 812_0056; 04:33)

The diagnostics for a low focus position in the vP periphery, which were presented above, are all consistently not obtained either in Pontic Greek or in Romeyka. I therefore conclude that all focus positions in Romeyka are in the left periphery, analogous to Pontic Greek.

6.4.4 Summary

The findings of this section show that Romeyka seems to be like Pontic Greek, in that it has VO as the pragmatically unmarked order in matrix declarative clauses and it has the focus to the left of the verb, but unlike Pontic Greek, in that it has a single focus position.

6.5 Conclusions

In this chapter, (a) I typologically classified Romeyka word order; (b) I compared word order in Romeyka with word order (i) in Turkish, (ii) in Georgian and (iii) in Pontic Greek; and (c) I examined the evolution of VO and OV alternation in matrix and subordinate clauses in Romeyka.

Beginning with (a), I have shown that, like German, Romeyka can be classified as "a third subtype of language lacking a dominant order, which consists of languages in which different word orders occur but the choice is syntactically determined" (Dryer 2005: 330-331). In Romeyka, the dominant order is SVO in matrix clauses and SOV in subordinate clauses. Matrix and subordinate clauses may contain an auxiliary, in which case the order is rigidly AuxVO.

As for (b), I have shown that Romeyka seems to be (i) like Turkish, in that it has the focus to the left of the verb, but unlike Turkish in that it has VO as the pragmatically unmarked order in matrix clauses; (ii) like Georgian, in that it has the focus to the left of the verb, but unlike Georgian in that it has both VO and OV as the pragmatically unmarked orders; and (iii) like Pontic Greek, in that it has VO as the pragmatically unmarked order in matrix clauses and has the focus to the left of the verb, but unlike Pontic Greek in that it has a single focus position.

7 The evolution of VO and OV alternation in Romeyka

7.0 Introduction

In this chapter, I investigate the evolution of VO and OV alternation in matrix and subordinate clauses in Romeyka. Essentially, the diachronic puzzle can be summarised as follows: is the development of the VO and OV alternation in Romeyka (a) the result of internal (endogenous) change or continuity, or rather (b) the result of external change due to contact with local Turkish varieties?

The main findings of the chapter indicate that (a) the pragmatically unmarked VO order in matrix clauses in Romeyka is the result of continuity from previous stages of Greek and (b) the pragmatically unmarked OV order in subordinate clauses in Romeyka is the result of external change due to contact with local Turkish varieties.

The chapter is structured as follows: in §7.1, I discuss the reconstruction method I use; §7.2 explores the evolution of VO and OV alternation in matrix clauses in Romeyka; and, §7.3 depicts the evolution of VO and OV alternation in subordinate clauses in Romeyka. The main findings of the chapter are summarised in §7.4.

7.1 The reconstruction method

The reconstruction of syntax did not emerge until recently, as syntactic parameters are not as rich as lexicon variety, that is similarities are less probative (Longobardi & Guardiano 2009: 1684). However, Campbell & Harris (2002), Willis (2011) and Longobardi & Guardiano (2009), i.a., have claimed that syntactic patterns can be compared cross-linguistically.

In this study, I essentially follow Sitaridou's (2016: 13) reconstruction method, as illustrated in the following extract:

> "I have been approaching the syntactic classification of Pontic Greek by means of, first, comparing the syntax of specific phenomena in Romeyka to the ones in Hellenistic, Medieval and Pontic Greek (Sitaridou 2014a) to see which one Romeyka matches best; second, assessing whether changes/innovations could have sprung out of a Hellenistic or Medieval Greek pool of grammatical cues (in the sense of Lightfoot 2010). Crucial to this modus operandi is the idea that reanalysis takes place during child language acquisition and the distinction between the abstract grammatical system and the surface output of that system. On this view, it follows that reanalysis is constrained both by pre- and post-reanalysis grammars and that it must be acquirable on the basis of the same primary linguistic data. This imposes limits on the possible hypotheses that can be entertained (see also Willis 2011) —the same would hold even if we simply adopt Sapir's (1921) drift, which refers to the predisposition to undergo certain changes given certain precursor traits" (Sitaridou 2016: 13).

Let us take a look at how this actually works: I compare the VO and OV alternation in matrix and subordinate clauses in Romeyka with the VO and OV alternation in (a) HelGr, (b) MedGr and (c) Anasta Turkish. Depending on the findings of this comparison, I make two predictions: First, if VO and OV alternation in Romeyka is similar to that found in previous stages of Greek and not to that found in Anasta Turkish, then we can safely conclude that this is the result of continuity from previous stages of Greek. Second, if the VO and OV alternation in Romeyka is similar to that found in Anasta Turkish and not to that found in previous stages of Greek, then we can conclude that this is the result of external change due to contact with Anasta Turkish.

7.2 The evolution of VO and OV alternation in matrix clauses in Romeyka

I begin my investigation with VO and OV alternation in matrix clauses in Romeyka. To start with, the pragmatically unmarked word order in matrix clauses in Romeyka is VO (see chapter 4), while OV results from movement of the focused object, regardless of whether it is information focus or contrastive focus (see chapter 4).

Likewise, in Hellenistic Greek (henceforth HelGr), the pragmatically unmarked word order in matrix clauses is VO (see (1)), while OV results from movement of the focused object, regardless of whether it is information focus (see (2)) or contrastive focus (see (3)) (Kirk 2012):

(1) Hellenistic Greek:

ἄνθρωπός	τις	ἐποίησε	δεῖπνον	μέγα.
ánthropós	tis	epoíe:se	deîpnon	méga.
man.NOM	INDEF.NOM	make.Past.3SG	dinner.ACC	large.ACC

'A certain man made a large dinner.'
(Kirk 2012: 41 *apud* Lk 14:16)

(2) Hellenistic Greek:

εἰ ἐμὲ	ᾔδειτε,	καὶ	[τὸν	πατέρα	μου]$_{\text{I-Foc}}$	ἄν	ᾔδειτε.
ei emè	é:ideite	kaì	tòn	patéra	mou	án	é:ideite.
if me	know.PNP.2PL	and	the.ACC	father.ACC	I.POSS	PCL	know.PNP.2PL

'If you had known me, you would have known my father.'
(Kirk 2012: 107 *apud* Jn 8:19)

(3) Hellenistic Greek:

[ἔλεος]$_{\text{C-Foc}}$	θέλω	καὶ	οὐ	θυσίαν.
éleos	thélo:	kaì	ou	thusían.
mercy.ACC	want.1SG	and	NEG	sacrifice.ACC

'I want mercy and not sacrifice.'
(Kirk 2012: 108 *apud* Mt 9:13, 12:7)

Unlike in Romeyka and HelGr, in Medieval Greek (henceforth MedGr) and MG, VO is both the pragmatically unmarked and marked word order. For instance, for MedGr see the

pragmatically unmarked VO order in (4), an *in-situ* information focused object in (5), as well as an *in-situ* contrastive focused object in (6):

(4) Medieval Greek:

Πατέρα	μου,	οἱ	θειοῦδες	τοῦ	ῥηγὸς	ηὗράν	τον
patéra	mu,	i	θiúðes	tu	riɣós	ívran	ton
father	I.POSS	the	uncles.NOM	the	king.GEN	find.Past.3SG	he.ACC

ὀρφανὸν	καὶ	πτωχὸν,	καὶ	ἐμοιράστησαν	τὸ	νησσίν	του
orfanón	ce	ptoχón,	ce	emirástisan	to	nis:ín	tu.
orphan.ACC	and	poor.ACC	and	divide.Past.3PL	the.ACC	island.ACC	his

'The uncles of the king found him orphaned and poor and they divided his island.'
(Chronicle of Machairas: 468)

(5) Medieval Greek:

Ἤξευρε	[καὶ	τοῦτον]$_{\text{I-Foc}}$.
íksevre	ce	túton.
know.IMPF.3SG	PRT	this.ACC

'He also knew that.'
(Chronicle of Machairas: 6)

(6) Medieval Greek:

Καὶ	πέψε	φρενίμους	μαντατοφόρους·
ce	pépse	frénimus	mandatofórus;
and	send.IMP.2SG	restrained.ACC	messengers.ACC

μηδὲν	πέψῃς	[ἄτυχους]$_{\text{C-Foc}}$.
miðén	pépsis	atíχus.
NEG	send.PNP.2SG	impetuous.ACC

'Also, send restrained messengers; do not send impetuous (messengers).'
(Chronicle of Machairas: 25)

In contrast to Romeyka, in Anasta Turkish, the pragmatically unmarked word order is OV (see (7) and (8)):

(7) Anasta Turkish:

a. Question:

Dün ne yap-tı-nız?

yesterday what do-Past-2PL

'What did you do yesterday?'

b. Answer:

Rumca konuş-tu-k.

Romeyka talk-Past-1PL

Yaz-dı-k.

write-Past-1PL

Ortu-du-k.

sit-Past-1PL

Televizyon-e izle-di-k.

television-DAT watch-Past-1PL

yemek ye-di-k.

food eat-Past-1PL

Çay iş-ti-k.

tea drink-Past-1PL

'We talked in Romeyka. We wrote. We sat. We had a meal. We had tea.'

(S01; 150703_0043_0008; 01:06)

(8) Anasta Turkish:

a. Question:

Ayşe dün ne yap-tı?

Ayşe yesterday what do-Past-3SG

'What did Ayşe do yesterday?'

b. Answer:

Ayşe dün yemek yap-ti.

Ayşe yesterday food make-Past-3SG

Bakala git-ti.

grocery-DAT go-Past-3SG

Ekmek al-dı.

bread buy-Past-3SG

'Yesterday, Ayşe made food. She went to the grocery. She bought bread.'

(S01; 150703_0043; 01:22)

On the contrary, like in Romeyka, OV in Anasta Turkish is the result of movement of the focused object (see (9)):

(9) Anasta Turkish:

a. Question:

Ali bir kere kim-i öp-tü?

Ali one time who-ACC kiss-Past-3SG

'Who did Ali kiss once?'

b. Answer:

Ali bir kere [Ayşe-'yi]$_{\text{I-Foc}}$ öptü.

Ali one time Ayşe-ACC kiss-Past-3SG

'Ali kissed Ayşe once.'

(S01; 150703_0044; 00:54)

I now move to the investigation of the syntactic distribution of topics and foci in Anasta Turkish and compare it to Romeyka.

In the first place, consider the example (10) from Anasta Turkish:

(10) Anasta Turkish:

a. Question:

Can-'dan n' aber? O ne ye-di parti-de?

Can-ABL what news he what eat-Past-3SG party-LOC

'What about Can? What did he eat at the party?'

Answers:

Can-'ın vallahi ne yap-tı-nı bil-mi-yor-um ama …

Can-GEN frankly what do-PART-POSS know-NEG-PROG-1SG but

'Frankly, I don't know about Alis, but …'

b. [Aylin]$_{\text{C-Top}}$[dolma]$_{\text{C-Foc}}$ ye-di.

Aylin dolma eat-Past-3SG

'Aylin ate dolma.'

(S01; 150702_0023; 01:59)

c. #[dolma]$_{\text{C-Foc}}$ [Aylin]$_{\text{C-Top}}$ ye-di.

dolma Aylin eat-Past.3SG

'Aylin ate dolma.'

The contrast of felicity in the responses in (10) demonstrates that, like in Romeyka, in Anasta Turkish a C-Foc may follow a C-Top, whereas a C-Foc cannot move across a C-Top.

Below, I consider a pair, where the context is set up so as to favour an interpretation of the object as C-Top and the subject as C-Foc (see (11)):

(11) Anasta Turkish:

a. Question:

Çorba-dan n' aber? On-dan içen ol-du mu peki?

'What about the soup? Has anyone eaten that?'

Answers:

Vallahi çorba-dan haber-im yok, ama …

frankly soup-ABL news-POSS-1SG NEG but

'Frankly, I don't know about the soup, but …'

b. [dolma-lar-ı]$_{\text{C-Top}}$ [Aylin]$_{\text{C-Foc}}$ ye-di.

but dolma-PL-ACC Aylin eat-Past-3SG.

'Aylin ate dolma.'

(S01; 150702_0023; 02:21)

c. #[Aylin]$_{\text{C-Foc}}$ [dolma-lar-ı]$_{\text{C-Top}}$ ye-di.

Aylin dolma-PL-ACC eat-Past-3SG

'Aylin ate dolma.'

The infelicity in the answers in (11) supports the assumption that a C-Top cannot follow a C-Foc in Anasta Turkish, like in Romeyka. Therefore, the only licit order would be C-Top > C-Foc in both Anasta Turkish and Romeyka.

In the same way, information focus must be left adjacent to the verb. Thus, information focus is not different from contrastive focus in terms of its distribution, hence follows contrastive topics (see (12)):

(12) Anasta Turkish:

a. Question:

Birgül-'ün kardeş-ler-i dün parti-de ne iç-ti?
Birgül-GEN sibling-PL-POSS yesterday party-LOC what drink-Past-3SG
'What did Birgül's siblings get to drink at the party?'

Answers:

Vallahi kardeş-ler-i haber-im yok, ama …
frankly sibling-PL-POSS news-POSS-1SG NEG but
'Frankly, I do not know about all her siblings but …'

b. [küçük kardeşi]$_{\text{C-Top}}$ [rakı]$_{\text{I-Foc}}$ içti.
young brother-POSS rakı drink-Past-3SG
'Birgül's youngest brother drank (from the) rakı.'
(S01; 150702_0023; 23:07)

c. #[rakı]$_{\text{I-Foc}}$ [küçük kardeşi]$_{\text{C-Top}}$ içti.
rakı young brother-POSS drink-Past-3SG
'Birgül's youngest brother drank (from the) rakı.'

Additional data can be drawn from the Anatolian Turkish variety of Cappadocia (Karamanlidika), which is genetically related to the one spoken in 'Anasta' and bears the same structure to Standard Modern Turkish, rather than that employed by HelGr. First, the pragmatically unmarked order in Karamanlidika Turkish is OV (see (13)b), like in Standard Modern Turkish (see (13)c) and unlike in HelGr, in which it is VO (see (13)a):

(13) Pragmatically unmarked order:

a. Hellenistic Greek:

ἄνθρωπός	τις	ἐποίησε	δεῖπνον	μέγα
ánthropós	tis	epoíe:se	deîpnon	méga
man.NOM	INDEF.NOM	make.Past.3SG	dinner.ACC	large.ACC

'A certain man was preparing a great banquet.'

(Kirk 2012: 41 *apud* Lk 14:16)

b. Karamanlidika Turkish:

Πὶρ	ἀτὰμ	ἀζὶμ	μετζλὶς	ἐ̈ἵλετί.
Bir	adam	azim	meclis	eyle-di.
a	man	large	gathering	give-Past-3SG

'A certain man was preparing a great banquet.'

(Lk 14:16)

c. Standard Modern Turkish:

Adam-ın	bir-i	büyük	bir	şölen	hazır-la-yıp
man-GEN	one-ACC	large	one	gathering	prepare-Past-3SG

'A certain man was preparing a great banquet.'

(Lk 14:16)

As for the information focus, Karamanlidika Turkish manifests OV (see (14)c), like in Standard Modern Turkish (see (14)c) and HelGr (see (14)a):

(14) Information focus:

a. Hellenistic Greek:

εἰ ἐμὲ ᾔδειτε,
ei emè é:ideite,
if me know.PNP.2PL
καὶ [τὸν πατέρα μου]$_{\text{I-Foc}}$ ἄν ᾔδειτε.
kaì tòn patéra mou án é:ideite.
and the.ACC father.ACC I.POSS PCL know.PNP.2PL
'If you knew me, you would know my Father also.'
(Kirk 2012: 107 *apud* Jn 8:19)

b. Karamanlidika Turkish:

ἔγερ πενὶ πιλέϊτινιζ
eğer ben-i bileidiniz
if I-ACC know-2PL
[πετεριμὶ]$_{\text{I-Foc}}$ ταχὴ πιλίριτινιζ.
peder-im-i dahı biliridiniz.
father-I.POSS-ACC also know-2PL
'If you knew me, you would know my Father also.'
(Jn 8:19)

c. Standard Modern Turkish:

Ben-i tanısaydınız,
I-ACC know-2PL
[Baba-m-ı]$_{\text{I-Foc}}$ da tanırdınız.
father-I.POSS-ACC also know-2PL
'If you knew me, you would know my Father also.'
(Jn 8:19)

Regarding the contrastive focus, Karamanlidika Turkish bears OV (see (15)b), like in Standard Modern Turkish (see (15)c) and HelGr (see (15)a):

(15) Contrastive focus:

a. Hellenistic Greek:

[ἔλεος]$_{\text{C-Foc}}$	θέλω	καὶ	οὐ	θυσίαν
éleos	thélo:	kaì	ou	thusían.
mercy.ACC	want.1SG	and	NEG	sacrifice.ACC

'I desire mercy, not sacrifice.'

(Kirk 2012: 108 *apud* Mt 9:13, 12:7)

b. Karamanlidika Turkish:

[Μερχαμὲτ]$_{\text{C-Foc}}$	ἰστέριμ,	βὲ	κουρπὰν	ἰστεμέμ.
Merhamet	isteri-m,	ve	kurban	iste-me-m.
mercy	want-1SG	and	sacrifice	want-NEG-1SG

'I desire mercy, not sacrifice.'

(Mt 9:13, 12:7)

c. Standard Modern Turkish:

Ben	kurban	değil,	[merhamet]$_{\text{C-Foc}}$	isterim.
I	sacrifice	NEG	mercy	want-1SG

'I desire mercy, not sacrifice.'

(Mt 9:13, 12:7)

Overall, our data show that the pragmatically unmarked word order in Romeyka is VO and it is similar to HelGr and MedGr and different from Anasta Turkish. This suggests that it is safe to assume that the pragmatically unmarked word order in Romeyka derives from previous stages of Greek and not from Turkish. However, due to the homoplasy of information focus and contrastive focus in Romeyka, HelGr and Anasta Turkish, it is difficult to come up with a safe conclusion regarding their diachronic development.

If the pragmatically unmarked word order in matrix clauses in Romeyka had changed due to its contact with Anasta Turkish, we would have expected OV to be the pragmatically unmarked word order in Romeyka. Additional evidence derives from the fact that the word order in a variety of Laz, a Kartvelian language spoken in an area close to 'Anasta', i.e. in Pazar, is not the same as that in Georgian, a language to which it is genetically related (see

Öztürk & Pöchtrager 2011). Nevertheless, it is similar to Turkish. Note that Pazar Laz is in massive contact with Turkish.[12]

In Laz, the pragmatically unmarked word order is OV, like in Turkish and unlike in Georgian, in which it is both VO and OV (see (16)):

(16) Pazar Laz:

a. Question:

'What happened?'

b. Answer:

Alik	çitabi	dót'k'u.
Ali.NOM	book.ACC	read.Past.3SG

'Ali read the book.'

(Göksel 2011: 146)

The information focus in Laz results in OV orders, like in both Anasta Turkish and Georgian (see (17)):

(17) Pazar Laz:

Alik	çitabi	[Ayşes]$_{\text{I-Foc}}$	kómeçu.
Ali.NOM	book.ACC	Ayşe.DAT	give.Past.3SG

'Ali gave the book to Ayşe.'

(Göksel 2011: 149)

Finally, the contrastive focus in Laz results in OV orders, like in both Anasta Turkish and Georgian (see (18)):

[12] Laz is a Caucasian language mainly spoken in Turkey. It belongs to the subgroup called the South-Caucasian branch along with Megrelian, Georgian and Svan. The data used in this study come from one of the varieties of Laz, viz. Pazar Laz, also known as Atinan. Laz is an endangered language. It is hard to estimate the number of speakers of Laz, but it is assumed to be between 50.000 and 500.000 (Öztürk & Pöchtrager 2011: 3).

(18) Pazar Laz:

Alik	çitabi	[Ayşes]$_{C\text{-}Foc}$	kómeçu.
Ali.NOM	book.ACC	Ayşe.DAT	give.Past.3SG

'Ali gave the book to Ayşe.'

(Göksel 2011: 149)

The pragmatically unmarked word order in Pazar Laz is a clue in favour of a non-contact explanation for the Romeyka counterpart, since Laz must have reshaped its word order schema in accordance with the Turkish one. This is a strong indication that the pragmatically unmarked VO order in Romeyka derives from previous stages of Greek.

Under those circumstances, it becomes obvious that the pragmatically unmarked order in matrix clauses in Romeyka developed from previous stages of Greek (see Table 33):

Table 33. Inherit development (Sitaridou 2016).

a. In Hellenistic Greek, the pragmatically unmarked word order is VO with OV resulting from movement of the focused object (see Kirk 2012)
b. In Medieval Greek, the pragmatically unmarked word order is VO with focus *in-situ*
c. In Anasta Turkish, the pragmatically unmarked word order is OV with the focused object *in-situ*
d. In Romeyka, the pragmatically unmarked word order is VO with OV resulting from movement of the focused object
e. The pragmatically unmarked VO order in Romeyka descends either from Hellenistic Greek or Medieval Greek
f. *Therefore, the pragmatically unmarked VO order is inherited (from previous stages of Greek)*

Crucially, data from other AMG varieties, namely Cappadocian, Phárasiot, Pontic (as they were presented in §6.4) and Sílliot, support this analysis. In all AMG varieties, the pragmatically unmarked word order is VO, like in Romeyka. However, except for Pontic Greek, the organisation of information structure in Cappadocian, Phárasiot and Sílliot more closely resembles the one found in MedGr and MG, rather than the one found in Romeyka and Pontic Greek (see Table 34):

Table 34. Distribution of information structure in AMG varieties.

	Romeyka	**Pontic Greek**	**Cappadocian**	**Phárasiot**	**Sílliot**
Pragmatically unmarked order	SVO	SVO	SVO (see (19))	SVO (see (23))	SVO (see (27))
Aboutness topic	$[O]_{A\text{-}Top}SV$	$[O]_{A\text{-}Top}SV$	$[O]_{A\text{-}Top}SV$ (see (20))	n/a	n/a
Contrastive topic	$[O]_{C\text{-}Top}SV$	$[O]_{C\text{-}Top}SV$	$[O]_{C\text{-}Top}SV$ (see (21))	$[O]_{C\text{-}Top}SV$ (see (24))	n/a
Information focus	$S[O]_{I\text{-}Foc}V$	$S[O]_{I\text{-}Foc}V$	$SV[O]_{I\text{-}Foc}$ (see (22))	$SV[O]_{I\text{-}Foc}$ (see (25))	$SV[O]_{I\text{-}Foc}$ (see (28)) $S[O]_{I\text{-}Foc}V$ (see (29))
Contrastive focus	$S[O]_{C\text{-}Foc}V$	$S[O]_{C\text{-}Foc}V$	n/a	$SV[O]_{C\text{-}Foc}$ (see (26))	n/a

As for Cappadocian, like in Romeyka the pragmatically unmarked word order is attested to be VO (see (19)):

(19) Cappadocian; Ghúrzono:

Ένα	πατισᾶχος	εἴχε	τρία	παιριά.
éna	patišáχos	íçe	tría	perjá.
a.NOM	king.NOM	have.IMPF.3SG	three.ACC	sons.ACC

'A king had three sons.'
(Dawkins 1916: 340)

Analogous to Romeyka, aboutness and contrastive topics in Cappadocian are placed in the preverbal domain. Consider the aboutness topics in (20) and the contrastive topics in (21):

(20) Cappadocian; Delmesó:

a. Να πήρα γώ του πατισσάχου το παιδί.
na píra γó tu patišáχu to peðí.
PRT.MOD marry.Past.1SG I the.GENking.GEN the.ACCson.ACC
'I would marry the king's son.'

b. [Του πατισσάχου το παδί]$_{A\text{-}Top}$ ἐγώ να το πήρα.
tu patišáχu to peðí]$_{A\text{-}Top}$ eγó na to píra.
the.GENking.GEN the.ACCson.ACC I PRT.MOD he.ACC marry.Past.1SG
'The king's son, I would marry him.'

(Dawkins 1916: 316)

(21) Cappadocian; Ulaghátsh:

a. Question:
κρέψε ἕνα καλὸ σέι,
krépse éna kaló šéi,
ask.IMP.2SG a.ACC good.ACC thing.ACC
ἰτὸ dò ἔκρεψες ἕνα σέι dé 'ναι.
itó dó ékrepses éna šéi dé ne.
this.NOM REL ask.Past.2SG a.NOM thing.NOMNEG be.3SG
'Ask for something else; what you have asked for is nothing.'

b. Answer:
ὀγώνα [ἰτὸ]$_{C\text{-}Top}$ κρέω.
oγóna itó kréo.
I this.ACC ask.1SG
'I ask for that.'

(Kesisoglou 1951: 138)

In Cappadocian, constituents that are information focused stay *in-situ*, i.e. in the postverbal domain (see (22)):

(22) Cappadocian; Ulaghátsh:

a. Question:

διλέda	τί	κρέεις	νὰ	σὲ	déκω.
diléda	ti	kréis	na	se	déko.
ask.IMP.2SG	what.ACC	want.2SG	PRT.MOD	you.ACC	give.PNP.1SG

'Ask what you want and I will give it to you.'

b. Answer:

κρέω	[ἕνα	ἄλοχο]$_{\text{I-Foc}}$.
kréo	éna	áloχo.
ask.1SG	a.ACC	horse.ACC

'I want a horse.'

(Kesisoglou 1951: 138)

In Phárasiot, the pragmatically unmarked word order is VO (see (23)):

(23) Phárasiot:

Κανείς	ǰó	πήρεν	dα	μισαφούρ.
kanís	džo	píren	da	misafúr.
no-one.NOM	NEG	receive.Past.3SG	the.ACC	guests.ACC

'No-one received the guests.'

(Dawkins 1916: 492)

In Phárasiot, contrastive topics are preverbal (see (24)):

(24) Phárasiot:

a. Question:

Σύ	τατάς	čαι	μάνα	ἔς;
sí	tatás	dže	mána	és?
you.NOM	father.ACC	and	mother.ACC	have.2SG

'Have you a father and mother?'

b. Answer(s):

[Τατάς	πάλ]$_{\text{C-Top}}$	ἔχω,	ǰαι	[μάνα	πάλ]$_{\text{C-Top}}$	ἔχω.
tatás	pal	éχo	dže	mána	pal	éχo.
father.ACC	PRT	have.1SG	and	mother.ACC	PRT	have.1SG

'A father I have and a mother I have.'

(Dawkins 1916: 536)

Phárasiot information foci are always postverbal (see (25)):

(25) Phárasiot:

a. Question:

Ἔ,	υἱό	μου,	τίνα	ἀ	πάρ;
e,	ió	mu	tína	a	par?
hey	son.VOC	I.POSS	who.ACC	PRT.MOD	take.PNP.2SG

'My son, whom will you marry?'

b. Answer:

Ἀ	πάρω	[το	γαϊρίδι	μας]$_{\text{I-Foc}}$.
a	páro	to	γairíði	mas.
PRT.MOD	take.PNP.1SG	the.ACC	donkey.ACC	our

'I will marry our donkey.'

(Dawkins 1916: 482)

In Phárasiot, contrastive focused constituents are placed in a postverbal position (see (26)):

(26) Phárasiot:

a. Question:

Ἀdé to φσ̆ὸκκο δώσετέ dα
adé to fš̆ók:o ðóseté da
this.ACC the.ACCboy.ACC give.IMP.2PL he.ACC
το μουτσούκο το χαζνά.
to mutsúko to χazná.
the.ACClittle.fellow.ACC the.ACCtreasure.ACC
'Give this boy, the little fellow, the treasure.'

b. Answer:

Γώ χαζνάς ǰό ὑρέβω,
γó χaznás džó irévo,
I.NOM treasure.ACC NEG seek.1SG
μά ὑρέβω [τη ἀροσύνη σου]$_{C\text{-}Foc}$.
ma irévo ti arosíni su].
but seek.1SG the.ACChealth.ACC you.POSS
'I do not ask for treasure, but I ask for your health.'

(Dawkins 1916: 540)

In Sílliot, the pragmatically unmarked word order is VO (see (27)):

(27) Sílliot:

Πήραμι ἕνα μικρὸ γαΐχ.
pírami éna mikró γaíχ.
get.Past.1PL a.ACC small.ACC boat.ACC
'We got onto a small boat.'
(Costakis 1968: 116)

Information focus in Sílliot is either postverbal (see (28)) or preverbal (see (29)):

(28) Sílliot:

a. Question:

κόρη	του	čína	σε	πάρη;
kóri	tu	tšína	se	pári?
daughter.NOM	his	who.ACC	PRT.FUT	marry.PNP.3SG

'Whom will his daughter marry?'

b. Answer:

Kó	σου	κόρη	σε	πάρη
kó	su	kóri	se	pári
yours.NOM	you.POSS	daughter.NOM	PRT.FUT	marry.PNP.3SG

[του σταχτηǰή]$_{\text{I-Foc}}$.

tu staχtidží.

the.ACCman.who.sells.ashes.ACC

'Your daughter will marry the man who sells ashes.'

(Dawkins 1916: 284)

(29) Sílliot:

a. Question:

Ποῦτ' εἴστιγις;

pút ístijis?

where be.IMPF.2PL

'Where are you from?'

b. Answer:

[Ὀπ' Κάστουρου]$_{\text{I-Foc}}$ ἤρταμι.

op kásturu írtami.

from Konya come.Past.1PL

'We came from Konya'.

(Costakis 1968: 116)

In a nutshell, the distribution of VO and OV alternation in other AMG varieties indicate that: (a) the pragmatically unmarked word order in AMG varieties derives from previous stages of Greek, (b) the information focus and contrastive focus in Cappadocian, Phárasiot and Sílliot could derive from MedGr and (c) it is still a mystery whether the information focus and

contrastive focus in Romeyka and Pontic Greek derive (i) from HelGr, or (ii) from MedGr and whether had reshaped after the Turkish equivalent due to its contact with it.

7.3 The evolution of VO and OV alternation in subordinate clauses in Romeyka

I now move on to account for the evolution of VO and OV alternation in subordinate clauses in Romeyka. In Romeyka, the pragmatically unmarked word order in subordinate clauses is OV (see chapter 4).

Unlike in Romeyka, in HelGr the pragmatically unmarked word order in subordinate clauses is VO (see (30)):

(30) Hellenistic Greek:

	λέγων	ὅτι	ἐντραπήσονται	τὸν	ὑιόν	μου.
…	légo:n	hóti	entrapé:sontai	tòn	hyión.ACC	mou.
	say.PART	that	respect.3PL	the.ACC	son.ACC	I.POSS

'… saying, 'They will respect my son.'
(Mk 12:6)

Moreover, unlike in Romeyka, but like in HelGr, in MedGr (see (31)) and MG (see chapter 4) the pragmatically unmarked word order in subordinate clauses is VO:

(31) Medieval Greek:

	λαλῶντα	τους	ὅτι	ἔχομεν	καλὴν	ἀγάπην	μεσόν	μας·
…	lalónta	tus	óti	éχomen	kalín	aɣápin	mesón	mas;
	say.GER	them	that	have.1PL	good.ACC	love.ACC	between	us.

'… saying to them that we have peace between us.'
(Chronicle of Machairas: 230)

In Anasta Turkish, the pragmatically unmarked word order in subordinate clauses is OV, like in Romeyka (see (32)):

(32) Anasta Turkish:

Hiç	iste-mi-yor-um	yemek	yap-ma-yacağ-ım.
no	want-NEG-PROG-1SG	food	make-NEG-FUT-1SG

'I don't want to make food.'
(S01; 150702_0032; 02:05)

Since the pragmatically unmarked word order in subordinate clauses in Romeyka is OV, like that in Anasta Turkish and unlike that in HelGr, MedGr and MG, I would argue that OV in subordinate clauses in Romeyka is the result of its contact with Anasta Turkish.

In order to account for the grammatical mechanism that have triggered such a change in the pragmatically unmarked word order in subordinate clauses in Romeyka, I follow a feature-based analysis. In particular, I build my proposal on Tsimpli (2003) and Tsimpli & Mastropavlou's (2007) distinction between formal features that are visible at the syntax-semantics interface because of their semantic import, i.e. SEM-interpretable features and those whose role is restricted to syntactic derivations and possibly have PHON-realisation but no role at SEM, i.e. the SEM-uninterpretable features. Based on this distinction, there are the following (un)interpretability possibilities between SEM and PHON (see also Neocleous & Sitaridou submitted) (see (33)):

(33) (Un)interpretability possibilities between SEM and PHON:

a. SEM-interpretable/PHON-uninterpretable features (e.g. animacy distinctions on Greek nouns and pronouns are not grammaticalised due to grammatical gender differences)
b. SEM-interpretable/PHON-interpretable (e.g. animacy distinctions on English *wh*- and personal pronouns)
c. SEM-uninterpretable/PHON-interpretable (e.g. resumptive uses of subject–verb agreement and object clitics in Greek)
d. SEM-uninterpretable/PHON-uninterpretable (e.g. case and subject–verb agreement in English)

(Tsimpli & Dimitrakopoulou 2007: 223)

In the previous chapters, I have shown that Romeyka word order variation is discourse-driven. Discourse-related features are encoded in a formal feature, i.e. the linearisation feature. The linearisation feature drives the computation in the narrow syntax and maps syntactic units (phases) as LFs onto the SEM and as PFs onto the PHON. The pragmatic interpretation of the clause takes place at the SEM, in which the linear order of the constituents plays a vital role in the interpretation of their pragmatic value.

That is to say, for the linearisation feature there must be the following (un)interpretability possibilities between the SEM and the PHON (see (34)):

(34) (Un)interpretability possibilities between the SEM and the PHON of the linearisation feature:

a. SEM-interpretable/PHON-uninterpretable (VO, the object is focused)
b. SEM-interpretable/PHON-interpretable (OV, the object is focused)
c. SEM-uninterpretable/PHON-interpretable (OV, the object is not focused)
d. SEM-uninterpretable/PHON-uninterpretable (VO, the object is not focused)

In other words, what differentiates a pragmatically unmarked VO order (see (34)a) from a pragmatically unmarked OV order (see (34)c) is (a) the presence of the linearisation feature in the latter but not in the former and (b) the lack of internal semantic interpretation in the linearisation feature in the latter (i.e. SEM-uninterpretable). Essentially, a formal linearisation feature without internal semantic structure (SEM-uninterpretable and PHON-interpretable) is reminiscent of Biberauer et al.'s (2014) analysis.

As for the distribution of the linearisation feature in subordinate clauses in Romeyka, I argue that the linearisation feature in focused OV orders is SEM-interpretable and PHON-interpretable, whereas in nonfocused OV orders it is SEM-uninterpretable and PHON-interpretable (see Table 35):

Table 35. Interface (un)interpretability of the linearisation feature in focused and nonfocused order subordinate clauses in Romeyka.

Focused	SEM-interpretable	PHON-interpretable
Nonfocused	SEM-uninterpretable	PHON-interpretable

In Anasta Turkish, I argue that, like in Romeyka, the linearisation feature in focused OV orders is SEM-interpretable and PHON-interpretable, whereas in nonfocused OV orders it is SEM-uninterpretable and PHON-interpretable (see Table 36):

Table 36. Interface (un)interpretability of the linearisation feature in focused and nonfocused order subordinate clauses in Anasta Turkish.

Focused	SEM-interpretable	PHON-interpretable
Nonfocused	SEM-uninterpretable	PHON-interpretable

Essentially the problem is, what is the initial state grammar that yielded the current state in Romeyka? In particular, in an early stage, I argue that the linearisation feature in focused OV

orders was SEM-interpretable and PHON-interpretable, whereas in nonfocused VO orders it was SEM-uninterpretable and PHON-uninterpretable (see Table 37):

Table 37. Early stage of the interface (un)interpretability of the linearisation feature in focused and nonfocused order subordinate clauses in Romeyka.

Focused	SEM-interpretable	SEM-interpretable
Nonfocused	SEM-uninterpretable	PHON-uninterpretable

I therefore assume that the pragmatically unmarked VO order in subordinate clauses in Romeyka shifted to OV as a result of language contact with Anasta Turkish. That is to say, the trajectory word order change in subordinate clauses in Romeyka is shown in (35):

(35) SEM-uninterpretable/PHON-uninterpretable (VO, the object is not focused) > SEM-uninterpretable/PHON-interpretable (OV, the object is not focused)

According to (35), the original absence of the linearisation feature in Romeyka nonfocused orders (SEM- and PHON-uninterpretable) changed into the appearance of a linearisation feature without semantic structure (SEM-uninterpretable) because of its contact with the equivalent Anasta Turkish structure.

In order to resolve the contact trajectory, I depart from the following premises on contact-induced syntactic change:

First, contact was possible because Romeyka allowed both orders to start with. That is, interference due to language contact was triggered in Romeyka, because both parameters values of the linearisation feature existed in both Romeyka and Anasta Turkish. This supports the Resistance Principle (Guardiano et al. 2016: 54) (see (36)):

(36) Resistance Principle:
Resetting of parameter α from value X to Y in language A as triggered by interference of language B only takes place if a subset of the strings that contribute to constituting a trigger for value Y of parameter α in language B already exists in language A.
(Guardiano et al. 2016: 54)

In other words, the resetting of a parameter under the influence of interference data is possible only if the new triggers are similar enough to triggers already unmistakably present in the

interfered language, though of course not sufficient on their own to trigger the new value (Guardiano et al. 2016: 54). The informal idea is that interference data in parametric syntax must appear at least in part as "familiar" in the interfered language, in order to be used as triggers; thus "contact may exacerbate/reinforce existing tendencies" (Sitaridou 2014a).

Second, I pursue a theoretical approach in that in multilingual environments it is the SEM-uninterpretable features of a language *x* that are not instantiated in the language *y* or vice versa that cause learnability problems. I therefore assume transfer from Turkish into Romeyka whereby the SEM-uninterpretable features of Romeyka cause learnability problems. Based on these assumptions, I propose a generalisation of contact-induced word order change in Romeyka (see (37)):

(37) Generalisation of contact-induced change in word order in Romeyka:
The PHON-realisation of the linearisation feature which is SEM-uninterpretable is sensitive to contact-induced change.

Third and equally important, as the Feature Economy states, acquirers tend to generalise the input from the above alternations; children conclude that the linearisation feature must always be present, even if there is no semantic interpretation (see (38)):

(38) Input Generalisation:
If a functional head sets parameter p_j to value v_i then there is a preference for similar functional heads to set p_j to value v_i.
(Biberauer & Roberts 2015)

This account explains the directionality of cross-linguistic effects: it is always the language that instantiates the more restrictive option that affects the other, not vice versa (see Feature Economy in Biberauer & Roberts 2015). Hence, it is Anasta Turkish that affects Romeyka regardless of whether the latter is the L1 (attrited or heritage) or L2.

In the final analysis, as shown above, the pragmatically unmarked OV order in subordinate clauses in Romeyka is assumed to have been reshaped after Anasta Turkish (see Table 38):

Table 38. Contact-induced change (Sitaridou 2016).

a.	In Hellenistic Greek, the pragmatically unmarked word order is VO (see Kirk 2012)
b.	In Medieval Greek, the pragmatically unmarked word order is VO
c.	In Anasta Turkish, the pragmatically unmarked word order is OV
d.	In Romeyka, the pragmatically unmarked word order is OV
e.	*Therefore, Romeyka OV is contact induced by Anasta Turkish*

Interestingly, the fact that the pragmatically unmarked word order in subordinate clauses in Romeyka changed from VO to OV, whereas the pragmatically unmarked order in matrix clauses in Romeyka did not change, can be explained by an acquisitional perspective of reanalysis (see Lightfoot 1991).

The asymmetries between matrix and subordinate clauses were first captured by Ross' (1973) Penthouse Principle, in which any syntactic phenomena treat matrix clauses differently from subordinate clauses (see (39)):

(39) The Penthouse Principle:
The rules are different if you live in the penthouse.
(Ross 1973)

Perhaps the best known example of the penthouse principle effect is the distribution of subject-auxiliary inversion in constituent questions in English, which in many (but not all) varieties of English is restricted to matrix clauses (see (40)):

(40) English:
a. What can Sam do about it?
b. I'll find out what Sam can do about it.

Compare it with (41):

(41) English:
b. *What Sam can do about it?
c. *I'll find out what can Sam do about it.

According to the Penthouse Principle, more goes on upstairs than downstairs. C^0 is said to be obligatory in subordinate clauses, because sentences must be turned into CPs in order to function as arguments and modifiers. Sentences with a fronted constituent of this type are therefore CPs, headed by a C^0 position obligatorily filled by a complementiser or finite verb. Otherwise, where C^0 is not required for the satisfaction of some principle of grammar, its presence is language-specifically determined. Thus, in subordinate clauses in German, the obligatory C^0 position is usually filled by a lexical complementiser, which blocks V^0-to-C^0 movement.

Reanalysis based on main clause V^0-to-C^0 movement in German is also advocated by Lightfoot (1991), in the context of a parameter-setting model of change which crucially assumes that acquisition must be based on main clauses only, so that the verb-final syntax overtly displayed by subordinate clauses is in principle not accessible to the learner. Here the triggering factor of the reanalysis is seen as random variation in the input, resulting in increasingly frequent exercise of the V2 option: "This no more reflects a difference in grammars than if some speaker were shown to use a greater number of passive or imperative sentences. Rather, it reflects the kind of accidental variation that is familiar from studies in population genetics. Nonetheless, changes in the primary linguistic data, if they show a slight cumulative effect, might have the consequence of setting a grammatical parameter differently" (Lightfoot 1991: 67-68).

If we followed the parameter-setting model of change, word order in subordinate clauses in Romeyka would not be accessible to the learner; hence word order in subordinate clauses in Romeyka is more sensitive to language contact change than that one in matrix clauses.

7.4 Conclusions

In this chapter, I investigated the evolution of VO and OV alternation in matrix and subordinate clauses in Romeyka. The main findings of the section indicate that (a) the pragmatically unmarked VO order in matrix clauses in Romeyka is the result of continuity from previous stages of Greek and (b) the pragmatically unmarked OV order in subordinate clauses in Romeyka is the result of external change due to contact with Turkish.

8 Conclusions

8.1 Summary of the study and conclusions

The goal of this dissertation was to investigate word order and information structure in the light of recent developments within the minimalist program. The language examined was Romeyka, the only Asia Minor Greek variety still spoken in the area historically known as Asia Minor (present-day Anatolia, Turkey). The objective of this study was twofold: (a) descriptively, to examine word order variation in Romeyka and (b) theoretically, to investigate whether such word order variation could be a language specific property or, rather, could be accommodated in a minimalist system. Descriptively, I aimed to (a) determine the pragmatically unmarked and marked word orders in Romeyka, (b) examine their typological classification and (c) identify their diachronic evolution. Theoretically, this study was fundamentally about the role that word order plays in the efficient computation of interface conditions, mainly of the syntax and semantics interface. The question I asked was (a) whether the order of the constituents of a clause plays a role in the computation from narrow syntax to the semantics interface and (b) whether order plays a role in the semantic interpretation at the SEM. I pursued an approach where the order of the constituents of a clause plays such a role and asked what the implications are for the theory of syntax and semantics interface.

In chapter 1, I presented (a) the scope of the dissertation, (b) the theoretical framework I followed in this study, (c) an introduction to Romeyka (d) the objectives of the study, (e) the methodology of the study and (f) a roadmap of the dissertation.

In chapter 2, I presented a brief overview of the Romeyka grammar. The way I presented the grammar was not explicit at all and was developed in a traditional way. I specifically introduced the basic facts about phonology, morphology and syntax in Romeyka. The goal of the chapter was to introduce the reader to the basic grammatical rules of Romeyka so that they would follow the linguistic discussion in the remainder of this study.

In chapter 3, I focused on the word order variation in Romeyka. I specifically examined the respective position of the subject (S), verb (V) and object (O) in main and subordinate declarative and interrogative clauses in Romeyka. First, I presented the results of my survey of

word order variation in matrix and subordinate declarative and interrogative clauses in Romeyka. The findings of this survey showed that three-word orders are attested in Romeyka, namely SVO, SOV and OSV. On the other hand, V-initial and S-final word orders, i.e. VSO, VOS and OVS, are not attested in Romeyka. Second, I argued that V^0 raises to T^0 in Romeyka. Third, I showed that subjects in pragmatically unmarked word orders are left-dislocated.

In chapter 4, (a) I determined the pragmatically unmarked/neutral word order in Romeyka and (b) examined the syntactic distribution and the semantic type of the constituents in pragmatically marked word orders in Romeyka. Overall, I argued that (a) the pragmatically unmarked order in Romeyka is SVO in matrix declarative clauses and SOV in subordinate declarative clauses and (b) focused constituents and *wh*-phrases always occupy the immediate preverbal position, while topicalised elements are always left dislocated. Romeyka also allows multiple *wh*-questions and multiple focus. The postverbal domain can only be occupied by elements carrying given (noncontrastive) information.

In chapter 5, I proposed a semantic constraint on focus interpretation, which makes predictions regarding the semantic interpretation of focused phrases in declarative and interrogative clauses in Romeyka. The predictions made by that constraint were tested for every potential syntactic derivation of the subject (S), verb (V) and object (O) in Romeyka and those derivations were mapped into PF and LF rules. The findings of this chapter show that there is good reason to assume that the linearisation feature in Romeyka contributes to the efficient computation of interface conditions, since it efficiently maps phases as ordered LFs and PFs onto the SEM and PHON respectively.

In chapter 6, (a) I typologically classified Romeyka word order and (b) I compared word order in Romeyka with word order (i) in Turkish, (ii) in Georgian and (iii) in Pontic Greek. Beginning with (a), I have shown that, like German, Romeyka can be classified as "a third subtype of language lacking a dominant order, which consists of languages in which different word orders occur but the choice is syntactically determined" (Dryer 2005: 330-331). In Romeyka, the dominant order is SVO in matrix clauses and SOV in subordinate clauses. Matrix and subordinate clauses may contain an auxiliary, in which case the order is rigidly AuxVO. As for (b), I have shown that Romeyka seems to be (i) like Turkish, in that it has the focus to the left of the verb, but unlike Turkish in that it has VO as the pragmatically unmarked order in matrix clauses; (ii) like Georgian, in that it has the focus to the left of the verb, but unlike Georgian in that it has not both VO and OV as the pragmatically unmarked orders; and (iii) like Pontic Greek, in that it has VO as the pragmatically unmarked order in matrix clauses and has the focus to the left of the verb, but unlike Pontic Greek in that it has a single focus position.

In chapter 7, I investigated the evolution of VO and OV alternation in matrix and subordinate clauses in Romeyka. The main findings of the section indicate that (a) the pragmatically unmarked VO order in matrix clauses in Romeyka is the result of continuity from previous stages of Greek and (b) the pragmatically unmarked OV order in subordinate clauses in Romeyka is the result of external change due to contact with Turkish. Based on the contact-induced change in (b), I proposed a generalisation of contact-induced change that which predicts that the semantic uninterpretable features are sensitive to change phonologically due to language contact.

8.2 Implications on broader issues

Overall, the goal of this dissertation was to account for the role of order in a minimalist system. My theory was based on two tenets: (a) order contributes to the mapping of syntactic units (phases) from narrow syntax to both the SEM and PHON and (b) order plays a role in the semantic interpretation of focus at the SEM.

The findings of my study indicate that there is good reason to assume that order in Romeyka (a) contributes to the efficient computation of interface conditions, since it efficiently maps phases as ordered LFs to the SEM and (b) contributes to the semantic interpretation of focus in declarative and interrogative clauses.

The proposal of the study (a) provides an alternative principled explanation of information structure within the minimalist program rather than the cartographic one (see Cinque 1999, Rizzi 1997) and (b) offers a principled explanation of previous theories on linearisation, such as the Kayne's Linear Correspondence Axiom (LCA) (1994) and the Biberauer et al.'s (2014) Final-Over-Final Condition (FOFC), as well as (c) a principled explanation of word order change due to language contact and (d) provides evidence against the claim that word order does not play any role in the efficient computation of interface conditions and mainly in relation to the syntax and semantics interface (see Chomsky et al. 2017, i.a.).

8.3 Limitations and future work

To begin with, in chapter 1 I noted that this study is in essence theoretical and not experimental, in the sense that I do not provide any statistics to analyse my data. The reason is crucial and fundamental for the nature of my study, which focuses on hierarchy and not on frequency. As such, hypotheses are developed based on empirical observations on primary language data, which lead to the development and proposal of a theoretical model within the minimalist framework, which makes predictions about (a) the semantic interpretation of different word orders in Romeyka and (b) the word orders in Romeyka that changed due to their contact with local Turkish varieties. However, the merits of future experimental studies on these phenomena would validate or not the theoretical claims of this study.

To add to this, I narrowed my study mainly to clauses consisting of a(n overt) subject (S), verb (V) and object (O) for the sake of the better comparison of clauses with similar structure throughout the investigation of the phenomena in examination. Nevertheless, future work on more complicated clauses would provide us with further insights on the credentials of the theoretical model developed in this study.

In conclusion, if the arguments in the present dissertation are on the right track, the predictions that have been made, will need to be tested cross-linguistically, both synchronically and diachronically, in order to evaluate their validity. Obviously, only future work can tell us whether the views in the present dissertation can be sustained or not.

References

Agouraki, Y. (2010). It-clefts and stressed operators in the preverbal field of Cypriot Greek. *Lingua , 120* (3), 527-554.

Alexiadou, A. (1997). *A Case Study in Antisymmetric Syntax.* Amstmerdam & Philadelphia: John Benjamins B.V.

Alexiadou, A., & Anagnostopoulou, E. (1998). Parametrizing AGR: Word order, V movement and EPP-checking. *Natural Language and Linguistic Theory , 16*, 491- 539.

Alexiadou, A, & Anagnostopoulou, E. (2000). Greek Syntax: a Principles and Parameters Perspective. *Journal of Greek Linguistics , 1*, 169-221.

Alexopoulou, T., & Baltazani, M. (2012). Focus in Greek *wh*-questions. In I. Kučerová & A. Neeleman, *Contrasts and Positions in Information Structure* (pp. 206-246). Cambridge: Cambridge University Press.

Andreadis, R. (1990). *Ποντιακό Θέατρο. Η Ξενιτεία.* Θεσσαλονίκη.

Andrews, P. A. (1989). *Ethnic Groups in the Republic of Turkey.* Wiesbaden: Reichert.

Aronson, H. I. (1982). *Georgian: A Reading Grammar.* Chicago: Slavica.

Asatiani, R. (2007). Georgian. In S. Skopeteas, S. Hellmuth, G. Franselow, & C. Féry, *The Expression of Information Structure: The Interaction of Syntax and Phonology in Cross-Linguistic Perspective.* Berlin: Mouton De Gruyter.

Beaver, D., & Clark, B. (2008). *Sense and sensitivity.* Oxford: Blackwell.

Beck, S. (2006). Intervention Effects Follow from Focus Interpretation. *Natural Language Semantic , 14* (1), 1-56.

Belletti, A. (2001). Inversion as Focalization. In A. Hulk, & J.-Y. Pollock, *Subject Inversion in Romance and the Theory of Universal Grammar* (pp. 60-90). Oxford: Oxford University Press.

Belletti, A. (1988). The Case of Unaccusatives. *Linguistic Inquiry , 19*, 1-34.

Biberauer, T. (2017). Factors 2 and 3: A principled approach. *Cambridge Occasional Papers in Linguistics , 10*, 38-65.

Biberauer, T., & Roberts, I. (2015). Rethinking formal hierarchies: A proposed unification. *Cambridge Occasional Papers in Linguistics , 7*, 1-31.

Biberauer, T., & Roberts, I. (2010). Subjects, Tense and verb-movement. In T. Biberauer, A. Holmberg, M. Sheehan, & I. Roberts, *Parametric Variation: Null Subjects in Minimalist Theory* (pp. 263-302). Cambridge: Cambridge University Press.

Biberauer, T., Holmberg, A., & Roberts, I. (2014). A syntactic universal and its consequences. *Linguistic Inquiry* (45.2), 169-225.

Boeder, W. (2005). The South Caucasian languages. *Lingua , 115*, 5–89.

Boeder, W. (1989). Verbal person marking, noun phrase and word order in Georgian. In L. Marácz, & P. Muysken, *Configurationality: The Typology of Asymmetries* (pp. 159–184). Dordrecht: Foris.

Bošković, Ž. (2007). A note on *wh*-typology. In P. Kosta, & L. Schürcks, *Linguistic investigations into formal description of Slavic languages: Contributions of the sixth European conference held at Potsdam university* (pp. 159-170). 2007, Frankfurt am Main: Peter Lang.

Bošković, Ž. (2002). On multiple *wh*-fronting. *Linguistic Inquiry , 33* (3), 351–383.

Bošković, Ž. (1997). Superiority effects with multiple *wh*-fronting in Serbo-Croatian. *Lingua, 102*, 1-20.

Bortone, P. (2009). Greek with No History, No Standard, No Models: Muslim Pontic Greek. In A. Georgakopoulou, & M. Silk, *Standard Languages and Language Standards: Greek, Past and Present* (pp. 67-89). London: Ashgate.

Brendemoen, B. (2006). Aspects of Greek-Turkish Language Contact in Trabzon. In H. Boeschoten, & L. Johanson, *Turkic Languages in Contact* (pp. 63-73). Wiesbaden: Harrassowitz.

Brendemoen, B. (2002). *The Turkish Dialects of Trabzon. Their Phonology and Historical Development.* Wiesbaden: Harrassowitz Verlag.

Büring, D. (2009). Towards a Typology of Focus Realization. In M. Zimmermann, & C. Fery, *Information Structure* (pp. 177-205). Oxford: Ofxord University Press.

Büring, D. (2003). On D-trees, beans and B-accents. *Linguistics and Philosophy , 26*, 511–545.

Campbell, L., & Harris, A. C. (2002). Syntactic Reconstruction and Demythologizing Myths and the Prehistory of Grammars. *Journal of Linguistics , 38* (3), 599–618.

Cardinaletti, A. (2004). Towards a Cartography of Subject Positions. In L. Rizzi, *The Structure of CP and IP: The Cartography of Syntactic Structures* (Vol. 2, pp. 115-165). Oxford: Oxford University Press.

Chatzikyriakidis, S., Michelioudakis, D., & Spathas, G. (2015). Greek Focus Operators and Their Associates. *Studies for the Greek Language. 35*, pp. 167-179. Thessaloniki: Aristotle University of Thessaloniki.

Chatzopoulou, K., & Sitaridou, I. (2014). Jespersen's cycle for NEG2 and conditional inversion in the history of Greek: Evidence from Romeyka conditionals. 16th Diachronic Generative Syntax Conference, Research Institute for Linguistics. *Hungarian Academy of Sciences, Budapest, 3-5 July 2014.*

Childs, T. G. (1995). *A Grammar of Kisi, a Southern Atlantic Language. (Mouton Grammar Library, 16.).* Berlin: Mouton de Gruyter.

Chomsky, N. (2005). Three Factors in Language Design. *Linguistic Inquiry , 36* (1), 1-22.

Chomsky, N. (2004). Beyond explanatory adequacy. In A. Belletti, *Structures and beyond. The cartography of syntactic structures* (Vol. 3, pp. 104-131). Oxford: Oxford University Press.

Chomsky, N. (2001). Derivation by phase. In M. Kenstowicz, *Ken Hale: a life in language* (pp. 1-53). Cambridge, MA: MIT Press.

Chomsky, N. (2000). Minimalist inquiries: The framework. In R. Martin, D. Michaels, & J. Uriagereka, *Step by Step: Essays on Minimalist Syntax in honor of Howard Lasnik* (pp. 89-156). Cambridge, MA: MIT Press.

Chomsky, N. (1995). *The Minimalist Program.* Cambridge, MA: MIT Press.

Chomsky, N. (1982). *Some concepts and consequences of the theory of government and binding.* Cambridge, MA: MIT Press.

Chomsky, N. (1981). *Lectures on Government and Binding.* Dordrecht: Foris.

Chomsky, N. (1976). Conditions on rules of grammar. *Linguistic Analysis , 2*, 302-350.

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press.

Chomsky, N., Gallego, Á. J., & Ott, D. (2017). Generative Grammar and the Faculty of Language: Insights, Questions and Challenges. In Á. J. Gallego, & D. Ott, *Generative Syntax: Questions, Crossroads and Challenges.* Special issue of Catalan Journal of Linguistics.

Cinque, G. (2006). *Restructuring and Functional Heads*. New York: Oxford University Press.

Cinque, G. (1999). *Adverbs and functional heads: A cross-linguistic perspective.* Oxford: Oxford University Press.

Costa, J. (2004). *Subject positions and interfaces. The case of European Portuguese.* Berlin: Mouton de Gruyter.

Costakis, A. P. (1968). *Τὸ γλωσσικὸ ἰδίωμα τῆς Σίλλης.* Athens: Centre for Asia Minor Studies.

Cruschina, S. (2008). *Discourse-Related Features and the Syntax of Peripheral Positions. A Comparative Study of Sicilian and Other Romance Languages.* Cambridge: PhD dissertation, University of Cambridge.

Dawkins, R. M. (1931). Notes on the study of the Modern Greek of Pontos. *Byzantion , 6*, 389-400.

Dawkins, R. M. (1916). *Modern Greek in Asia Minor: a Study of the Dialects of Silli, Cappadocia and Phárasa with Grammar, Texts, Translations and Glossary.* Cambridge: Cambridge University Press.

Deffner, M. (1878). *Die Infinitive in den pontischen Dialekten und die zusammengesetzten Zeiten im Neugriechischen.* Berlin: Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin.

Doron, E. (2000). VSO and left-conjunct agreement: Bibilical Hebrew vs. Modern Hebrew. In A. Carnie, & E. Guilfoyle, *The syntax of verb-initial languages* (pp. 75-95). Oxford: Oxford Univeristy Press.

Drettas, G. (1997). *Aspects pontiques.* Paris: Association de recherches pluridisciplinaires.

Drettas, G. (1999). The Greek-Pontic Dialect Group. In A. F. Christidis, M. Arapopoulou, & G. Giannoulopoulou, *Dialect Enclaves of the Greek Language* (pp. 91-100). Athens: Hellenic Republic, Ministry of National Education and Religious Affairs.

Dryer, M. S. (1988). Object-Verb Order and Adjective-Noun Order: Dispelling a Myth. *Lingua , 74*, 77-109.

Dryer, M. S. (1992). The Greenbergian Word Order Correlations. *Language , 68*, 81-138.

Dryer, M. S. (2005). V. In M. S. Dryer, & M. Haspelmath, *The World Atlas of Language Structures.* Leipzig: Max Planck Institute for Evolutionary Anthropology.

É. Kiss, K. (2010). An adjunction analysis of quantifiers and adverbials in the Hungarian sentence. *Lingua , 120*, 506-526.

É. Kiss, K. (1998). Identificational focus versus information focus. *Language , 74* (2), 245–273.

Erguvanlı, E. (1984). *The function of word order in Turkish grammar.* Berkeley: University of California Press.

Erkü, F. (1983). *Discourse pragmatics and word order in Turkish.* Unpublished PhD dissertation, University of Minnesota.

Frascarelli, M., & Hinterhölzl, R. (2007). Types of Topics in German and Italian. In K. Schwabe, & S. Winkler, *On Information Structure, Meaning and Form* (pp. 87-116). Amsterdam and Philadelphia: Benjamins.

Gandon, O. (2016). *La relativisation dans une perspective areale: l'aire caucase – Anatolie de l'est – Iran de l'ouest.* Paris: PhD dissertation, Universite Sorbonne Nouvelle – Paris 3.

Giusti, G. (2006). Parallels in clausal and nominal periphery. In M. Fascarelli, *Phases of Interpretation.* Berlin: Mouton de Gruyter.

Goodall, G. (2001). The EPP in Spanish. In W. D. Davies, & S. Dubinsky, *Objects and other subjects* (pp. 93-223). Dordrecht: Kluwer.

Göçmen, E., Şehitoğlu, O., & Bozşahin, C. (1995). *An Outline of Turkish Syntax.*

Göksel, A. (2011). Word order, sentential stress and intonation. In B. Öztürk, & M. A. Pöchtrager, *Pazar Laz* (pp. 146-159). Munich: Lincom.

Göksel, A., & Kerslake, C. (2005). *Turkish: A comprehensive grammar.* London: Routledge.

Greenberg, J. H. (1966). *Language Universals: With Special Reference to Feature Hierarchies.* The Hague: Mouton & Co.

Greenberg, J. H. (1963). *Universals of Human Language.* Cambridge, MA: MIT Press.

Gryllia, S. (2008). *On the Nature of Preverbal Focus in Greek: A Theoretical and Experimental Approach* (Vol. LOT Dissertation Series 200). Netherlands Graduate School of Linguistics: Landelijke.

Guardiano, C., Michelioudakis, D., Ceolin, A., Irimia, M.-A., Longobardi, G., Radkevich, N., et al. (2016). South by Southeast. A syntactic approach to Greek and Romance microvariation. *L'Italia dialettale , LXXVII*, 95-166.

Gundel, J. K. (1998). On three kinds of focus. In P. Bosch, & R. van der Sandt, *Focus: Linguistic, Cognitive and Computational Perspectives* (pp. 293-305). Cambridge: Cambridge University Press.

Hamblin, C. L. (1973). Questions in Montague English. *Foundations of Language , 10*, 41-53.

Harris, A. C. (1981). *Georgian Syntax: A Study in Relational Grammar.* Cambridge: Cambridge University Press.

Harris, A. C. (2000). Word order harmonies and word order change in Georgian. In R. Sornicola, E. Poppe, & A. Shisha-Haley, *Stability, Variation and Change of Word-Order Patterns over Time* (pp. 133–163). Amsterdam/Philadelphia: Benjamins.

Hawkins, J. A. (1983). *Word Order Universals.* Academic Press.

Hewitt, B. G. (1995). *Georgian: A Structural Reference Grammar.* Amsterdam/Philadelphia: Benjamins.

Holes, C. (1990). *Gulf Arabic. (Croom Helm Descriptive Grammars).* London/New York: Routledge.

Horn, L. R. (1972). *On the semantic properties of logical operators in English.* Los Angeles: University of California.

Horvath, J. (1986). *Focus in the Theory of Grammar and the Structure of Hungarian.* Dordrecht: Foris.

Huang, J. (1982). *Logical relations in Chinese and the theory of grammar.* Cambridge, MA: PhD dissertation, MIT.

Huang, J. (1984). Move *wh* in a language without *wh* movement. *The Linguistic Review , 1*, 369-416.

Hunter, T. (2015). Deconstructing Merge and Move to Make Room for Adjunction. *Syntax , 18*, 266-319.

Karatsareas, P. (2011). *A Study of Cappadocian Greek Nominal Morphology from a Diachronic and Dialectological Perspective.* Cambridge: Unpublished PhD dissertation, University of Cambridge.

Karatsareas, P. (2014). On the diachrony of gender in Asia Minor Greek: The development of semantic agreement in Pontic. *Language Sciences , 43*, 77–101.

Karttunen, L. (1977). Syntax and Semantics of Questions. *Linguistics and Philosophy , 1*, 3-44.

Kayne, R. (1994). *The Antisymmetry of Syntax.* Cambridge, MA: MIT Press.

Kenesei, I. (2006). Focus as identification. In V. Molnár, & S. Winkler, *The architecture of focus* (pp. 137–168). Berlin: Mouton de Gruyter.

Kenesei, I. (1986). On the logic of Hungarian word order. In A. Werner, & S. de Meij, *Topic, focus, configurationality* (pp. 143–159). Amsterdam: John Benjamins.

Kesisoglou, I. I. (1951). *Τὸ γλωσσικὸ ἰδίωμα τοῦ Οὐλαγὰτς (Le dialecte d'Oulagatch).* Athens: Institut Français d'Athènes.

Kim, S.-S. (2002). Intervention Effects are Focus Effects. In S. Strauss, & N. Akatsuka, *Japanese/Korean Linguistics* (Vol. 10, pp. 615–628). Stanford: CSLI.

Kirk, A. (2012). *Word order and information structure in New Testament Greek.* Leiden: Unpublished dissertation, University of Leiden.

Kisilier, M. (2007). Word-Order Patterns in the Greek Dialect of Mariupolis. *Proceedings of the 3rd International Conference on Modern Greek Dialects and Linguistic Theory, 14--16 June 2007.* Nicosia.

Kornfilt, J. (1997). On rightward movement in Turkish. In L. Johanson (Ed.), *he Mainz Meeting, Proceedings of the Seventh International Conference on Turkish Linguistics* (pp. 107-123). Wiesbaden: Harrassowitz.

König, E. (1991). *The meaning of focus particles: A comparative perspective.* London: Routledge.

Krifka, M. (2007). The semantics of questions and the focusation of answers. In C. Lee, M. Gordon, & D. Büring, *Topic and focus, crosslinguistic perspectives on meaning and intonation* (pp. 139-150). Dordrecht: Springer.

Krifka, M. (2006). Association with focus phrases. In V. Molnár, & S. Winkler, *The architecture of focus* (pp. 105-136). Berlin: Mouton de Gruyter.

Lambrecht, K. (1994). *Information structure and sentence form.* Cambridge: Cambridge University Press.

Ledgeway, A. (2010). The clausal domain: CP structure and the left periphery. In R. D'Alessandro, A. Ledgeway, & I. Roberts, *Syntactic variation. The dialects of Italy* (pp. 38-51). Cambridge: Cambridge University Press.

Levinson, S. C. (2000). *Presumptive meanings.* Cambridge, MA: MIT Press.

Lightfoot, D. W. (2010). Language acquisition and language change. *WIREs Cogn Sci , 1*, 677–684.

Lightfoot, D. W. (1991). *How to Set Parameters: Arguments from Language Change.* Cambridge, MA: MIT Press.

Lohndal, T. (2014). *Phrase structure and argument structure: A case study of the syntax-semantics interface.* Oxford: Oxford University Press.

Lohndal, T. (2012). *Without Specifiers: Phrase Structure and Events.* University of Maryland: Unpublished PhD dissertation.

Longobardi, G., & Guardiano, C. (2009). Evidence for syntax as a signal of historical relatedness. *Lingua , 119* (11), 1679–1706.

Mackridge, P. (1999). The Greek Spoken in the Region of Pontus. In A. F. Christidis, M. Arapopoulou, & G. Giannoulopoulou, *Dialect Enclaves of the Greek Language* (pp. 101-105). Athens: Hellenic Republic, Ministry of National Education and Religious Affairs.

Mackridge, P. (1990). Unpublished Pontic stories: Collected be R.M.Dawkins. *Δελτίο Κέντρου Μικρασιατικών Σπουδών , 8*, 107-122.

Mackridge, P. (1987). Greek-speaking Moslems of north-east Turkey: prolegomena to a study of the Ophitic sub-dialect of Pontic. *Byzantine and Modern Greek Studies* (11), 115-137.

McCoy, S. (2003). Connecting information structure and discourse structure through "Kontrast": The case of colloquial Russian particles -to, že and ved'. *Journal of Logic, Language and Information , 12*, 319-335.

Melanofrydis, P. (2001). *Οι Κλωστοί.* Θεσσαλονίκη: Αδελφοί Κυριακίδη.

Mertyris, D. (2014). *The loss of the genitive in Greek. A diachronic and dialectological analysis.* Unpublished PhD dissertation, La Trobe University, Melbourne, Australia.

Michelioudakis, D., & Sitaridou, I. (2016). Recasting the typology of multiple *wh*-fronting: Evidence from Pontic Greek. *Glossa: a journal of general linguistics , 1* (1), 40-72.

Michelioudakis, D., & Sitaridou, I. (2013). Multiple *wh*-fronting in Romeyka. In A. Ralli, B. Joseph, & M. Janse (Ed.), *Proceedings of the Fifth International Conference of Modern Greek Dialects and Linguistic Theory (MGDLT5), Ghent, 20-22 September 2012*, (pp. 353-378).

Michelioudakis, D., & Sitaridou, I. (2012). Syntactic microvariation: Dative Constructions in Greek. In R. Etxepare, & B. Fernández, *Datives in variation: a micro-comparative perspective* (pp. 212-255). Oxford: Oxford University Press.

Milsark, G. L. (1977). Toward an explanation of certain peculiarities of the existential construction in English. *Linguistic Analysis , 3*, 1-29.

Molnár, V. (2006). On different kinds of contrast. In V. Molnár, & S. Winkler, *Architecture of Focus. Studies in Generative Grammar* (Vol. 82, pp. 197-233). Berlin/New York: Mouton de Gruyter.

Molnár, V. (2002). Contrast -from a contrastive perspective. In H. Hallelgard, S. Johansson, B. Behrens, & F.-H. Cathrine, *Information Structure in a Cross-Linguistic Perspective* (pp. 147-161). Amsterdam/New York: Rodopi.

Moro, A. (1997). *The raising of predicates. Predicative noun phrases and the theory of clause structure.* Cambridge: Cambridge University Press.

Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Büchel, C., et al. (2003). Broca's area and the language instinct. *Nature Neuroscience , 6*, 774-781.

Nebel, A. (1948). *Dinka Grammar (Rek-Malual Dialect) with Texts and Vocabulary.* Verona: Missioni Africane.

Neeleman, A., & van de Koot, H. (2008). Dutch scrambling and the nature of discourse templates. *The Journal of Comparative Germanic Linguistics , 11*, 137-189.

Neeleman, A., & Szendrői, K. (2007). Radical pro drop and the morphology of pronouns. *Linguistic Inquiry , 38*, 671-714.

Neeleman, A., Titov, E., van de Koot, H., & Vermeulen, R. (2009). A syntactic typology of topic, foucs and contrast. In J. van Craenenbroeck, *Alternatives to Cartography* (pp. 15-51). Berlin: Mouton de Gruyter.

Neocleous, N. & I. Sitaridou (submitted). *The rise of final auxiliaries in Asia Minor Greek*.

Özkan, H. (2013). The Pontic Greek Spoken by Muslims in the Villages of Beşköy in the Province of Present-Day Trabzon. *Byzantine and Modern Greek Studies , 37* (1), 130–150.

Öztürk, B., & Pöchtrager, M. A. (2011). *Pazar Laz.* Muenchen: Lincom Europa.

Pappou-Zhouravliova, E. (1995). Η ταυροροουμαίικη διάλεκτος των Ελλήνων της Αζοφικής στη Νοτιοανατολική Ουκρανία (περιοχή Μαριούπολης). *Archeion Pontou , 46*, 165-274.

Parcharidis, I. (1880). *Γραμματικὴ τῆς διαλέκτου Τραπεζοῦντος: Περὶ ἐπιθέτων, ἀντωνυμιῶν και βαρυτόνων ρημάτων.* (Vol. Χειρόγραφο 335). Athens: Κέντρο Ἐρεύνης τῶν Νεοελληνικῶν Διαλέκτων και Ἰδιωμάτων, Ἀκαδημία Αθηνών.

Partee, B. H., ter Meulen, A., & Wall, R. E. (1987). *Mathematical Methods in Linguistics.* Dordrecht: Kluwer Academic Publishers.

Pinto, M. (1997). *Licensing and Interpretation of Inverted Subjects in Italian.* PhD dissertation, Utrecht University.

Pollock, J.-Y. (1989). Verb-movement, Universal Grammar and the Structure of IP. *Linguistic Inquiry , 20*, 365-424.

Revythiadou, A. & Spyropoulos, V. (2009). *Οφίτικη Ποντιακή: Έρευνα Γλωσσικής Καταγραφής με Έμφαση στη Διαχρονία και Συγχρονία της Διαλέκτου*. Report submitted to the John S. Latsis Public Benefit Foundation.

Richards, M. (2007). On feature-inheritance: an argument from the Phase Impenetrability Condition. *Linguistic Inquiry , 38* (3), 563-572.

Rizzi, L. (2001). Relativized Minimality effects. In by M. Baltin & C. Collins, *The handbook of contemporary syntactic theory* (pp. 89–110). Oxford: Blackwell.

Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman, *Elements of Grammar.* Dordrecht: Kluwer.

Rizzi, L. (1982). *Issues in Italian syntax.* Dordrecht: Foris.

Roberts, I. (2007). *Diachronic syntax. Oxford.* Oxford: Oxford University Press.

Roberts, I. (1997). *Comparative Syntax*. London: Arnold.

Roberts, I. (1993). *Verbs and Diachronic Syntax: A Comparative History of English and French.* Dordrecht: Kluwer.

Roberts, I., & Holmberg, A. (2010). Introduction: Parameters in Minimalist theory. In T. Biberauer, A. Holmberg, I. Roberts, & M. Sheehan, *Parametric variation: Null subjects in minimalist theory* (pp. 1-57). Cambridge: Cambridge University Press.

Roberts, I., & Roussou, A. (2003). *Syntactic change. A minimalist approach to grammaticalization.* Cambridge: Cambridge University Press.

Rochemont, M. S. (1986). *Focus in Generative Grammar.* Amsterdam: John Benjamins.

Rooth, M. (1996). Focus. In S. Lappin, *The Handbook of Contemporary Semantic Theory* (pp. 271-298). Oxford: Blackwell.

Rooth, M. E. (1992). A theory of focus interpretation. *Natural Language Semantics , 1* (1), 75-116.

Rooth, M. E. (1985). *Association with focus.* Unpublished dissertation, University of Massachussets, Amherst, MA, USA.

Ross, J. R. (1973). The penthouse principle and the order of constituents. In C. T. Corum, T. C. Smith-Stark, & A. Weiser, *You Take the High Node and I'll Take the Low Node* (pp. 397–422). Chicago, Illinois: Chicago Linguistic Society.

Roussou, A., & Tsimpli, I.-M. (2006). On Greek VSO again! *Journal of Greek Linguistics , 42*, 317-354.

Sağlam, E. (2017). *Constitutive Ambiguities. Subjectivities and Memory in the Case of Romeika-Speaking Communities of Trabzon.* Unpublished PhD dissertation, Birkbeck, University of London.

Sapir, E. (1921). *Language: An Introduction to the Study of Speech.* New York.

Schreiber, L. (2018). Romeyka. In G. Haig & G. Khan, *The Languages and Linguistics of Western Asia: An Areal Perspective* (pp. 892–934). Berlin, Boston: De Gruyter.

Schreiber, L., & Sitaridou, I. (2017). Assessing the sociolinguistic vitality of Istanbulite Romeyka: an attitudinal study. *Journal of Multilingual and Multicultural Development* , 1-16.

Şener, S. (2010). *(Non-)Peripheral Matters in Turkish Syntax.* Unpublished PhD dissertation, University of Connecticut.

Sheehan, M. (2010). 'Free' inversion in Romance and the Null Subject Parameter. In T. Biberauer, A. Holmberg, I. Roberts, & M. Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory* (pp. 231-262). Cambridge: Cambridge University Press.

Sitaridou, I. (2016). Reframing the phylogeny of Asia Minor Greek: The view from Pontic Greek. *CHS Research Bulletin*, Center for Hellenic Studies, Harvard University, Volume 4, Issue 1. 1-17.

Sitaridou, I. (2014b). Modality, antiverdicality and complementation: The Romeyka infinitive as a negative polarity item. *Lingua* (148), 118-146.

Sitaridou, I. (2014a). The Romeyka infinitive: continuity, contact and change in the Hellenic varieties of Pontus. *Diachronica , 31* (1), 23-73.

Sitaridou, I. (2013). Greek-speaking enclaves in Pontus today: The documentation and revitalization of Romeyka. In M. C. Jones, & S. Ogilvie, *Keeping languages alive. Language endangerment: documentation, pedagogy and revitalization.* (pp. 98-112). Cambridge: Cambridge University Press.

Sitaridou, I., & Kaltsa, M. (2014). Contrastivity in Pontic Greek. *Lingua* (146), 1-27.

Skopeteas, S. (2012). The empirical investigation of information structure. In M. Krifka, & R. Musan, *The expression of information structure* (pp. 216–246). Berlin: Mouton De Gruyter.

Skopeteas, S., & Franselow, G. (2010). Focus in Georgian and the expression of contrast. *Lingua , 120* (6), 1370–1391.

Skopeteas, S., Féry, C., & Asatiani, R. (2009). Word Order and Intonation in Georgian. *Lingua , 119* (1), 102-127.

Smith, N., & Tsimpli, I.-M. (1995). *The Mind of a Savant: Language Learning and Modularity.* Oxford: Blackwell.

Tsimpli, I.-M. (2003). In I. i. account, & E. Mela-Athanasopoulou (Ed.), *Selected papers on theoretical and applied linguistics* (pp. 214–225). Thessaloniki: Department of Theoretical & Applied Linguistics, School of English, Aristotle University.

Tsimpli, I. M. (1995). Focusing in modern Greek. In K. E. Kiss, *Discourse Configurational Languages* (pp. 176-207). Oxford: Oxford University Press.

Tsimpli, I. M. (1990). The Clause Structure and Word Order of Modern Greek. *UCL Working Papers in Linguistics*, vol. 2, pp. 220-258.

Tsimpli, I.-M., & Dimitrakopoulou, M. (2007). The Interpretability Hypothesis: Evidence from *wh*-interrogatives in second language acquisition. *Second Language Research , 23* (2), 215-242.

Tsimpli, I.-M., & Mastropavlou, M. (2007). Feature Interpretability in L2 Acquisition and SLI: Greek Clitics and Determiners. In H. Goodluck, J. Liceras, & H. Zobl, *The Role of Formal Features in Second Language Acquisition* (pp. 143-183). Routledge.

Underhill, R. (1976). *Turkish Grammar.* Cambridge, MA: Massachusetts Institute of Technology (MIT) Press.

Vallduví, E., & Vilkuna, M. (1998). On rheme and kontrast. *Syntax and Semantics , 29*, 79-108.

van der Wal, J. (2016). Diagnosing focus. *Studies in Language , 40* (2), 259-301.

Vikner, S. (1995). *Verb Movement and Expletive Subjects in the Germanic Languages.* Oxford: Oxford University Press.

Vikner, S. (1997). V-to-I movement and inflection for person in all tenses. In L. Haegeman, *The New Comparative Syntax* (pp. 189-213). London: Longman.

Williams, S. J. (1980). *A Welsh Grammar.* Cardiff: University of Wales Press.

Willis, D. (2011). Reconstructing Last Week's Weather: Syntactic Reconstruction and Brythonic Free Relatives. *Journal of Linguistics , 47*, 407–446.

Zimmermann, M., & Onea, E. (2011). Focus marking and focus interpretation. *Lingua , 121* (11), 1651–1670.

Textual sources

Machairas Leontios, ΕΞΗΓΗΣΙΣ τῆς γλυκείας χώρας Κύπρου, ἡ ποία λέγεται Κρόνακα τουτἔστιν Χρονικ(όν)	http://users.uoa.gr/~nektar/history/2romanity/makhairas_chronicle.htm
New Testament	http://www.ntgateway.com
The Bible in Turkish	http://worldbibles.org/language_detail/eng/tur/Turkish
Ἄχτι Τζετὶτ Καινὴ Διαθήκη, Γιάνι Ραππιμὶζ βὲ χελαστζημὴζ Ἰησοῦς Χριστοζοὺν γενὶ βασιετί, κὶ Ἰντζίλι Σερίφ, Ρεσουλλαρὴν ἀμελλερὶ βὲ μεκτουπλαρή, χὲμ Ἀποκάλυψις τιρ, Χάλια μουτζέττετεν ἀτζὴκ τούρκτζε λισανὰ τερτζουμὲ ὀλουνοὺπ Ἀθήνατα, Γεώργη Πολυμέρηνιν τὰπ᾽ χανεσιντὲ Πασηλμὴς τηρ. 1838.	https://anemi.lib.uoc.gr/search/?dtab=m&search_type=simple&search_help=&display_mode=overview&wf_step=init&show_hidden=0&number=10&keep_number=&cclterm1=καραμανλίδικα&cclterm2=&cclterm3=&cclterm4=&cclterm5=&cclterm6=&cclterm7=&cclterm8=&cclfield1=term&cclfield2=&cclfield3=&cclfield4=&cclfield5=&cclfield6=&cclfield7=&cclfield8=&cclop1=&cclop2=&cclop3=&cclop4=&cclop5=&cclop6=&cclop7=&isp=&display_help=0&offset=11&search_coll[metadata]=1&&stored_cclquery=&skin=&rss=0&store_query=1&show_form=&export_method=none&ioffset=1&display_mode=detail&ioffset=1&offset=14&number=1&keep_number=10&old_offset=11&search_help=detail

Appendix A Information structure questionnaires

1 *wh*-congruence

1.1 Topics

1.1.1 Aboutness topics

(1) Subject (personal pronoun):

esí	do	epítšes?
you.NOM	what.ACC	do.Past.2SG

'What did you do?'

(2) Subject (demonstrative pronoun):

até	do	epítšen?
she	what.ACC	do.Past.3SG

'What did she do?'

(3) Subject (NP):

éna	líkon	do	epítšen?
a.NOM	wolf.NOM	what.ACC	do.Past.3SG

'What did a wolf do?'

(4) Subject (DP):

a. o mehmétis do epítšen?
the.NOM Mehmetis.NOM what.ACC do.Past.3SG
'What did Mehmetis do?'

b. i aišé do epítšen?
the.NOM Ayşe.NOM what.ACC do.Past.3SG
'What did Ayşe do?'

c. o dohtóris do epítšen?
the.NOM doctor.NOM what.ACC do.Past.3SG
'What did the doctor do?'

(5) Object (DP):

a. tin aišén píos epíren?
the.ACC Ayşe.ACC who.NOM marry.Past.3SG
'Who married Ayşe?'

b. ta ɣarðélæ píos epólisen s okʰúlin?
the.ACC children.ACC who.NOM send.Past.3SG to school.ACC
'Who sent the children to school?'

c. to cʰitʰápin píos éndžen?
the.ACC book.ACC who.NOM bring.Past.3SG
'Who brought the book?'

d. to dérsin píos epítšen?
the.ACC homework.ACC who.NOM do.Past.3SG
'Who did the homework?'

e. to boz dolápin píos aɣórasen asín tšáikaran?
the.ACC fridge.ACC who.NOM buy.Past.3SG from.the.ACC Çaykara.ACC
'Who bought the fridge from Çaykara?'

f. to tšáin píos epítšen?
the.ACC tea.ACC who.NOM make.Past.3SG
'Who made the tea?'

1.1.2 Contrastive topics

(6) VP:

o	mehmétis	do	epítšen
the.NOM	Mehmetis.NOM	what.ACC	do.Past.3SG

dž	esí	do.ACC	epítšes?
and	you.NOM	what.ACC	do.Past.2SG

'What did Mehmetis do and what did you do?'

(7) Subject (DP):

a. Context:

alís	éfaen	éna	mílon,
Alis.NOM	eat.Past.3SG	a.ACC	apple.ACC

o	mehmétis	éfaen	énan	ap^{h}íðin.
the.NOM	Mehmetis.NOM	eat.Past.3SG	a.ACC	pear.ACC

'Alis ate an apple and Mehmetis ate a pear.'

Question:

alís	do	éfaen,
Alis.NOM	what.ACC	eat.Past.3SG

dž	o	mehmétis	do	éfaen?
and	the.NOM	Mehmetis.NOM	what.ACC	eat.Past.3SG

'What did Alis eat and what did Mehmetis eat?'

b. Context:

alís	éfaen	to	mílon,
Alis.NOM	eat.Past.3SG	the.ACC	apple.ACC

o	mehmétis	éfaen	t	ap^{h}íðin.
the.NOM	Mehmetis.NOM	eat.Past.3SG	the.ACC	pear.ACC

'Alis ate the apple and Mehmetis ate the pear.'

Question:

píos	éfaen	to	mílon
who.NOM	eat.Past.3SG	the.ACC	apple.ACC

dže	píos	éfaen	t	ap^{h}íðin.
and	who.NOM	eat.Past.3SG	the.ACC	pear.ACC

'Who ate the apple and who ate the pear?'

c. Context:

alís éfaen to χavítsin,
Alis.NOM eat.Past.3SG the.ACCpudding.ACC
o mehmétis éfaen ton tšorbán.
the.NOM Mehmetis.NOM eat.Past.3SG the.ACCsoup.ACC
'Alis ate the pudding and Mehmetis ate the soup.'

Question:

alís do éfaen,
Alis.NOM what.ACC eat.Past.3SG
o mehmétis do éfaen?
the.NOM Mehmetis.NOM what.ACC eat.Past.3SG
'What did Alis eat and what did Mehmetis eat?'

d. Context:

alís epíren énan pont^{h}ólin,
Alis.NOM buy.Past.3SG a.ACC trousers.ACC
o mehmétis epíren énan kazáçin.
the.NOM Mehmetis.NOM buy.Past.3SG a.ACC sweater.ACC
'Alis bought trousers and Mehmeris bought a sweater.'

Question:

píos epíren énan pont^{h}ólin
who.NOM buy.Past.3SG a.ACC trousers.ACC
dže píos epíren énan kazáçin?
and who.NOM buy.Past.3SG a.ACC sweater.ACC
'Who bought trousers and who bought a sweater?'

e. Context:

alís epíren to pont^{h}ólin,

Alis.NOM buy.Past.3SG the.ACCtrousers.ACC

o mehmétis epíren to kazáçin.

the.NOM Mehmetis.NOM buy.Past.3SG the.ACCsweater.ACC

'Alis bought trousers and Mehmetis bought a sweater.'

Question:

alís do epíren

Alis.NOM what.ACC buy.Past.3SG

dž o mehmétis do epíren?

and the.NOM Mehmetis.NOM what.ACC buy.Past.3SG

'What did Alis buy and what did Mehmetis buy?'

(8) Object (DP):

a. Context:

ton alín íðes si mazirán,

the.ACCAlis.ACC see.Past.2SG in.the.ACC Mazira.ACC

to mustafán íðes so χoríon.

the.ACCMustafas.ACC see.Past.2SG in.the.ACC village.ACC

'You saw Alis in Mazira and Mustafas in the village.'

Question:

pútšeka íðes ton alín

where see.Past.2SG the.ACCAlis.ACC

tše pútšeka íðes to mustafán?

and where see.Past.2SG the.ACCMustafas.ACC

'Where did you see Alis and where did you see Mustafas?'

b. Context:

to	líkon	íðes	s	óros,		
the.ACC	wolf.ACC	see.Past.2SG	in.the.ACC	forest.ACC		
ton	árkon	íðes	s	alátæ	pu	ka.
the.ACC	bear.ACC	see.Past.2SG	in	trees.ACC	from	under

'You saw the wolf in the forest and the bear under the trees.'

Question:

pútšeka	íðes	to	líkon
where	see.Past.2SG	the.ACC	wolf.ACC

tše	pútšeka	íðes	ton	árkon?
and	where	see.Past.2SG	the.ACC	bear.ACC

'Where did you see the wolf and where did you see the bear?'

c. Context:

to	pont^{h}ólin	alís	epíren,	
the.ACC	trousers.ACC	Alis.NOM	buy.Past.3SG	
to	kazáçin	o	mehmétis	epíren.
the.ACC	sweater.ACC	the.NOM	Mehmetis.NOM	buy.Past.3SG

'Alis bought the trousers and Mehmetis bought the sweater.'

Question:

to	pont^{h}olin	ts	epíren,	
the.ACC	trousers.ACC	who.NOM	buy.Past.3SG	
to	kazáçin	ts	epíren?	
the.ACC	sweater.ACC	who.NOM	buy.Past.3SG	

'Who bought the trousers and who bought the sweater?'

d. Context:

to	mílon	alís	éfaen,	
the.ACC	apple.ACC	Alis.NOM	eat.Past.3SG	
t	ap^{h}íðin	o	mehmétis	éfaen.
the.ACC	pear.ACC	the.NOM	Mehmetis.NOM	eat.Past.3SG

'Alis ate the apple and Mehmetis ate the pear.'

Question:

to	mílon	ts	éfaen,
the.ACC	apple.ACC	who.NOM	eat.Past.3SG
t	ap^{h}íðin	ts	éfaen?
the.ACC	pear.ACC	who.NOM	eat.Past.3SG

'Who ate the apple and who ate the pear?'

1.2 Foci

1.2.1 Information foci

(9) DP-subjects:

a. píos epíren tin aišén?
who.NOM marry.Past.3SG the.ACC Ayşe.ACC
'Who married Ayşe?'
(alís / o mohal:ímis / o dohtóris)
Alis.NOM the.NOM teacher.NOM the.NOM doctor.NOM
'Alis / the teacher / the doctor'

b. píos epíren tin aišén?
who.NOM marry.Past.3SG the.ACC Ayşe.ACC
'Who married Ayşe?'
(éna ándras / éna mohal:ímis / éna dohtóris)
a.NOM man.NOM a.NOM teacher.NOM a.NOM doctor.NOM
'a man / a teacher / a doctor'

c. píos epíren tin aišén?
who.NOM marry.Past.3SG the.ACC Ayşe.ACC
'Who married Ayşe?'
(is / esí / atós)
someone.NOM you.NOM he.NOM
'someone / you / he'

d. píos epélien ta ɣarðélæ s ok^{h}úl:in?
who.NOM send.Past.3SG the.ACC children.ACC to school.ACC
'Who sent the children to school?'
(o tšíris atuna / o šœféris /
the.NOM father.NOM his the.NOM driver.NOM
éna šœféris / até / eɣó.NOM)
a.NOM driver.NOM she I
'his father / the driver / a driver / she / I'

e. píos aɣórasen to biɭɟísaɟarin so ɣarðélin?
who.NOM buy.Past.3SG the.ACC computer.ACC for.the.ACC child.ACC
'Who bought the computer for the child?'
(i mána t / até)
the.NOM mother.NOM his / she.NOM
'his mother / she'

(10) DP-objects:

a. i aišé tínan epíren?
the.NOM Ayşe.NOM who.ACC marry.Past.3SG
'Who did Ayşe marry?'
(ton alín)
the.ACC Alis.ACC
'Alis'

b. i aišé tínan efílisen?
the.NOM Ayşe.NOM who.ACC kiss.Past.3SG
(ton alín / ton ándran ates / énan ándran /
the.ACC Alis.ACC the.ACC husband.ACC she.POSS a.ACC man.ACC
atón / eménan)
he.ACC I.ACC
'Alis / her husband / a man / him / me'

c. opsé i aišé tínan efáisen?
yesterday the.NOM Ayşe.NOM who.ACC feed.Past.3SG
'Who did Ayşe feed yesterday?'
(ton alín / ta ɣarðélæ tes / ɣarðélæ / eménan)
the.ACC Alis.ACC the.ACC children.ACC she.POSS children.ACC I.ACC
'Alis / her children / children / me'

d. opsé i aišé tínan íðen?
yesterday the.NOM Ayşe.NOM who.ACC see.Past.3SG
(ton mohal:ímin / éna mohal:ímin / eménan / atón)
the.ACC teacher.ACC a.ACC teacher.ACC me he.ACC

e. alís do éfaen?
Alis.NOM what.ACC eat.Past.3SG
'What did Alis eat?'
(to χavítsin / ta χapsíæ / to χašílin /
the.ACCpudding.ACC the.ACCanchovies.ACC the.ACCpudding.ACC
to psomín / éna χavítsin / χapsíæ /
the.ACCbread.ACC a.ACC pudding.ACC anchovies.ACC
éna χašílin / éna psomín)
a.ACC pudding.ACC a.ACC bread.ACC
'the pudding / the anchovies / the pudding / the bread / a pudding / anchovies / a pudding / a bread'

f. alís do éktisen?
Alis.NOM what.ACC build.Past.3SG
'What did Alis build?'
(t ospítin / to rðómon / ton odán /
the.ACChouse.ACC the.ACCstreet.ACC the.ACCroom.ACC
énan ospítin / éna rðómon / énan odán)
a.ACC house.ACC a.ACC street.ACC a.ACC room.ACC
'the house / the street / the room / a house / a street / a room'

g. alís tínan epíren?
Alis.NOM who.ACC marry.Past.3SG
(éna patsí / tšíno to patsí / éna inéka /
a.ACC girl.ACC that.ACC the.ACCgirl.ACC a.ACC woman.ACC
tšínon to inéka / ti mohal:ímena /
that.ACC the.ACCwoman.ACC the.ACCteacher.ACC
éna mohal:ímena / eménan / tin aišén / aténan / tšínin)
the.ACCteacher.ACC I.ACC the.ACCAyşe.ACC she.ACC that.one.ACC
'a girl / that girl / a woman / that woman / the teacher / a teacher / me / Ayše / her / that one'

(11) Predicative complements:

a. alís do en?
Alis.NOM what.NOM be.3SG
'Who is Alis?'
(t emón to ɣarðélin /
the.NOM mine the.NOM son.NOM
t emón o ándras)
the.NOM mine the.NOM husband.NOM
'my son / my husband'

b. alís do en?
Alis.NOM what.NOM be.3SG
'Who is Alis?'
(éna ɣarðélin / éna ándra)
a.NOM child.NOM a.NOM man.NOM
'a child / a man'

c. to šcʰíl:on do en?
the.NOM dog.NOM what.NOM be.3SG
'What is a dog?'
(éna haivánin)
a.NOM animal.NOM
'an animal'

d. ató to šcʰíl:on do en?
this.NOM the.NOM dog.NOM what.NOM be.3SG
'What is this dog?'

e. (t emón to haivánin)
the.NOM mine.NOM the.NOM animal.NOM
'my pet'

(12) PPs:

a. opsé i aišé pútšeka epíen?
yesterday the.NOM Ayşe.NOM where go.Past.3SG
'Where did Ayşe go yesterday?'
(s okʰúlːin / so istambólin / sa parχáræ)
to school.ACC to.the.ACC Istanbul.ACC to.the.ACC pastures.ACC
'to school / to Istanbul / to the pastures'

b. opsé i aišé pútšeka epíen?
yesterday the.NOM Ayşe.NOM where go.Past.3SG
'Where did Ayşe go yesterday?'
(s énan džamín / s éna χoríon)
to a.ACC mosque.ACC to a.ACC village.ACC
'to a mosque / to a village'

(13) Adverbials:

a. alís pútšeka stétši?
Alis.NOM where stay.3SG
'Where does Alis stay?'
(aðátšaka)
here
'here'

b. pútšeka ívres ton dohtórin?
where find.Past.2SG the.ACC doctor.ACC
'Where did you find the doctor?'
(sin tšáikaran)
in.the.ACC Çaykara.ACC
'in Çaykara'

c. póte epíres ton ándra s?
when marry.Past.2SG the.ACC husband.ACC you.POSS
'When did you marry your husband?'

d. katš jašúnda éktises t ospíti s?
how.many old build.Past.2SG the.ACC house.ACC you.POSS
'How old were you when you built the house?'

e. póte epíes son dohtórin?
when go.Past.2SG to.the.ACC doctor.ACC
'When did you go to the doctor?'

f. póte epíes so istambólin?
when go.Past.2SG to.the.ACC Istanbul.ACC
'When did you go to Istanbul?'

(14) Modifier:
i aišé láɣa psomín éfaen?
the.NOM Ayše.NOM what.kind bread.ACC eat.Past.3SG
'What kind of bread did Ayše eat?'
(tsupaðítikon)
corn.ACC
'corn'

(15) Argument focus:
i aišé do éfaen?
the.NOM Ayşe.NOM what.ACC eat.Past3SG
'What did Ayşe eat?'
(tsupaðítikon psomín)
corn.ACC bread.ACC
'cornbread'

(16) Predicate focus:
i aišé do epítšen?
the.NOM Ayşe.NOM what.ACC do.Past.3SG
'What did Ayşe do?'
(i aišé éfaen tsupaðítikon psomín)
the.NOM Ayşe.NOM eat.Past.3SG corn.ACC bread.ACC
'Ayşe ate cornbread.'

1.2.2 Contrastive foci

(17) DP-subject:

a. alís éktisen t ospítin?
Alis.NOM build.Past.3SG the.ACC house.ACC
'Is it Alis that built the house?'
(o mehmétis)
the.NOM Mehmetis.NOM
'Mehmetis'

b. o ramazánis epíren ti zeinép?
the.NOM Ramazanis.NOM marry.Past.3SG the.ACC Zeynep.ACC
'Is it Ramazanis that married Zeynep?'
(œméris)
Ömris.NOM
'Ömeris'

c. o dohtóris epíen so istambólin?
the.NOM doctor.NOM go.Past.3SG to.the.ACC Istanbul.ACC
'Is it the doctor that went to Istanbul?'
(o muftís)
the.NOM imam.NOM

d. i aišé epíen sa parχáræ?
the.NOM Ayşe.NOM go.Past.3SG to.the.ACC fields.ACC
'Is it Ayşe that went to the fields?'
(eɣó)
I.NOM
'I'

e. éna ándra p^{h}ak^{h}laévi t ospítin?
a.NOM man.NOM clean.3SG the.ACChouse
'Is it a man that cleans the house?'
(éna inéka)
a.NOM woman.NOM
'a woman'

f. éna mohal:ímena eχujévi to c^{h}it^{h}ápin?
a.NOM female.teacher.NOM read.3SG the.ACCbook.ACC
'Is it a (female) teacher that reads the book?'
(éna mohal:ímis)
a.NOM male.teacher.NOM
'a (male) teacher'

g. éna kos:ára eséven s ospítin?
a.NOM hen.NOM enter.Past.3SG in.the.ACC house.ACC
'Is it a hen that got into the house?'
(éna furnón)
a.NOM hedgehog.NOM
'a hedgehog'

(18) DP-objects:

a. alís éktisen t ospítin?
Alis.NOM build.Past.3SG the.ACChouse.ACC
'Is it Alis that built the house?'
(to džamín)
the.ACCmosque.ACC
'the mosque'

b. o ramazánis epíren ti zeinép?
the.NOM Ramazanis.NOM marry.Past.3SG the.ACCZeynep.ACC
'Is it Ramazanis that married Zeynep?'
(tin aišén)
'Ayşe'

c. i patsí ep^{h}ak^{h}láepsen t ospítin?
the.NOM girl.NOM clean.Past.3SG the.ACChouse.ACC
'Is it the house that the girl cleaned?'
(to madrín)
the.ACCpen.ACC
'the pen'

d. o dohtóris epíen si stambólin?
the.NOM doctor.NOM go.Past.3SG to.the.ACC Istanbul.ACC
'Is it to Istanbul that the doctor went?'
(sin ingiltéran)
to.the.ACC England.ACC
'to England'

e. i aišé epíen sa parχáræ?
the.NOM Ayşe.NOM go.Past.3SG to.the.ACC pastures.ACC
'Is it to the pastures that Ayşe went?'
(so χoríon / s éna χoríon)
to.the.ACC village.ACC to a.ACC village.ACC
'to the village / to a village'

f. alís epíren tin aišén?
Alis.NOM marry.Past.3SG the.ACC Ayşe.ACC
'Is it Ayşe that Alis married?'
(eménan)
I.ACC
'me'

(19) Verb adjacency test:

a. alís póte írten asín tšáikaran?
Alis.NOM when come.Past.3SG from.the.ACC Çaykara.ACC
'When did Alis come from Çaykara?'
(opsé)
yesterday
'yesterday'

b. alís ap éndžeka írten opsé?
Alis.NOM from where come.Past.3SG yesterday
'From where did Alis come yesterday?'
(asín tšáikaran)
from.the.ACC Çaykara.ACC
'from Çaykara'

c. alís osím:eron írten asín tšáikaran?
Alis.NOM today come.Past.3SG from.the.ACC Çaykara.ACC
'Is it today that Alis came from Çaykara?'
(opsé)
'yesterday'

d. alís opsé írten asín tšáikaran?
Alis.NOM yesterday come.Past.3SG from.the.ACC Çaykara.ACC
'Is it yesterday that Alis came from Çaykara?'
(asín trapezúndan)
from.the.ACC Trabzon.ACC
'from Trabzon'

e. i mána s póte efáisen to musafírin?
the.NOM mother.NOM you.POSS when feed.Past.3SG the.ACCguest.ACC
'When did your mother feed the guest?'
(opsé)
'yesterday'

f. i mána s tínan efáisen opsé?
the.NOM mother.NOM you.POSS who.ACC feed.Past.3SG yesterday
'Who did your mother feed yesterday?'
(to musafírin)
the.ACCguest.ACC
'the guest'

g. i mána s osím:eron efáisen to musafírin?
the.NOM mother.NOM you.POSS today feed.Past.3SG the.ACCguest.ACC
'Is it today that your mother fed the guest?'
(opsé)
'yesterday'

h. i mána s opsé efáisen to musafírin?
the.NOM mother.NOM you.POSS yesterday feed.Past.3SG the.ACCguest.ACC
'Is it yesterday that your mother fed the guest?'
(to mohal:ímin)
the.ACCteacher.ACC
'the teacher'

2 Alternative questions (selective focus)

(20) Verbal phrase:

na eχujévis cʰitʰápæ
PRT.MOD read.2SG books.ACC
ɟóksa na terís tʰelevizɟónin aɣapás?
or PRT.MOD watch.2SG television.ACC love.2SG
'Do you like to read books or watch TV?

(21) Subject (personal pronoun):

a. píos éfaen ton tšorbán?
who.NOM eat.Past.3SG the.ACC soup.ACC?
esí ɟóksa i aišé?
you.NOM or the.NOM Ayşe.NOM
'Who ate the soup? You or Ayşe?'

b. píos en pío tranós? eɣó ɟóksa atós?
who.NOM be.3SG more strong.NOM I.NOM or he.NOM?
'Who is stronger? He or I?'

(22) Subject (NP):

a. o mehmétis ɟóksa alís éfaen to χavítsin?
the.NOM Mehmetis.NOM or Alis.NOM eat.Past.3SG the.ACCpudding.ACC
'Did Mehmetis or Alis eat the pudding?'

b. o dohtóris ɟóksa o mohalːímis
the.NOM doctor.NOM or the.NOM teacher.NOM
epíen sin trapezúndan?
go.Past.3SG to.the.ACC Trabzon.ACC
'Did the doctor or the teacher go to Trabzon?'

c. kahvén ɟóksa tšáin θélis?
coffee.ACC or tea.ACCwant.2SG
'Do you want coffee or tea?'

(23) Subject (DP):

a. o mehmétis míla ɟóksa apʰíðæ aɣórasen?
the.NOM Mehmetis.NOM apples.ACC or pears.ACC buy.Past.3SG
'Did Mehmetis buy apples or pears?'

b. i aišé ton mehmétin
the.NOM Ayşe.NOM the.ACC Mehmetis.ACC
ɟóksa ton alín epíren?
or the.ACC Alis.ACC marry.Past.3SG
'Did Ayşe marry Mehmetis or Alis?'

c. i patsí ton dohtórin
the.NOM girl.NOM the.ACC doctor.ACC
ɟóksa ton mohalːímin aɣapá?
or the.ACC teacher.ACC love.3SG
'Does the girl love the doctor or the teacher?'

(24) Predicative complement:

a. o mehmétis kahvedžís ɟóksa bojadžís en?
the.NOM Mehmetis.NOM coffee.shop.owner.NOM or painter.NOM be.3SG
'Is Mehmetis a coffee shop owner or a painter?'

b. o ščʰílːon éna haivánin ɟóksa énan insánin en?
the.NOM dog.NOM a.NOM animal.NOM or a.NOM human.NOM be.3SG
'Is a dog an animal or a human?'

c. i kátʰa mikútsikon ɟóksa tranón en?
the.NOM cat.NOM small.NOM or big.NOM be.3SG
'Is the cat small or big?'

d. to faín émnoston ɟóksa ánoston éton?
the.NOM food.NOM tasty.NOM or tasteless.NOM be.IMPF.3SG
'Was the food tasty or tasteless?'

(25) Prepositional phrase:

a. láɣa na páɣo son bakʰálin?
how PRT.MOD go.1SG to.the.ACC grocery's.ACC
me t arapán ɟóksa me ta poðáræ m?
by the.ACCcar.ACCor on the.ACCfeet.ACC I.POSS
'How can I get to the supermarket? By car or on foot?'

b. láɣa na páɣo so ɟunanistánin?
how PRT.MOD go.1SG to.the.ACC Greece.ACC
me t arapán ɟóksa me to utšáçin?
by the.ACCcar.ACC or by the.ACCairplane.ACC
'How can I get to Greece? By car or by plane?'

c. láɣa na féris ta χortáræ asá rašíæ?
how PRT.MOD bring.2SG the.ACCgrass.ACC from.the.ACC mountains
me t arapán ɟóksa me ta poðáræ s?
by the.ACCcar.ACCor on the.ACCfeet.ACC you.POSS
'How will you bring the grass from the mountains? By car or on foot?'

d. t aðelfó s sin tšáikaran
the.NOM brother.NOM you.POSS in.the.ACC Çaykara.ACC
ɟóksa sin trapezúndan stétši?
or in.the.ACC Trabzon.ACC stay.3SG

e. to džamín so χoríon
the.NOM mosque.NOM in.the.ACC village.ACC
ɟóksa sa parχáræ en?
or in.the.ACC pastures.ACC be.3SG
'Is the mosque in the village or in the pastures?'

(26) Adverbial phrase:

a. o musafíris osím:eron ɟóksa sapálæ na érθi?
the.NOM guest.NOM today or tomorrow PRT.MOD come.3SG
'Will the guest come today or tomorrow?'

b. i nífe osím:eron ɟóksa opsé
the.NOM daughter-in-law.NOM today or yesterday
efáisen ti mamíka?
feed.Past.3SG the.ACCmother-in-law.ACC
'Did the daughter-in-law feed her mother-in-law today or yesterday?'

3 Mention some questions (non-exhaustive focus)

(27) ap éndžeka na páo
from where PRT.MOD go.1SG
n aɣoráso gazéte s ató to χoríon?
PRT.MOD buy.1SGnewspaper.ACC in this.ACC the.ACCvillage.ACC
'Where will I go to buy a newspaper from this village?'

(28) pútšeka en ta ómorfa ta méræ
where be.PLthe.NOM nice.NOM the.NOM places.NOM
s ató to χoríon na elépo ata?
in this.ACC the.ACCvillage.ACC PRT.MOD see.1SG them
'Where are the most beautiful places to see in this village?'

4 Quantifiers

(29) Numerals become exact:

a. alís ekazánepsen bin lirá s éna mínan.
Alis.NOM earn.Past.3SG thousand lira in a.ACC month.ACC
'Alis earned one thousand lira in a month.'

b. bin lirá alís ekazánepsen (a) s éna mínan.
thousand lira Alis.NOM earn.Past.3SG them in a.ACC month.ACC
'Alis earned one thousand lira in a month.'

c. alís bin lirá ekazánepsen s éna mínan.
Alis.NOM thousand lira earn.Past.3SG in a.ACC month.ACC
'Alis earned one thousand lira in a month.'

(30) Weak quantifiers ('some' and 'few'):

a. alís olíɣon parán ekazánepsen,
Alis.NOM little money earn.Past.3SG
aéts aɣórasen énan arapán.
such.that buy.Past.3SG a.ACC car.ACC

b. alís olíɣon parán ekazánepsen,
Alis.NOM little money.ACC earn.Past.3SG
aéts utš epóresen n aɣorási énan arapán.
such.that NEG can.Past.3SG PRT.MOD buy.3SGa.ACC car.ACC

c. alís [olíɣon]$_{Foc}$ parán ekazánepsen,
Alis.NOM little money.ACC earn.Past.3SG
aéts aɣórasen énan arapán.
such.that buy.Past.3SG a.ACC car.ACC

d. alís [olíɣon]$_{Foc}$ parán ekazánepsen,
Alis.NOM little money.ACC earn.Past.3SG
aéts utš epóresen n aɣorási énan arapán.
such.that NEG can.Past.3SG PRT.MOD buy.3SGa.ACC car.ACC

(31) Unique referent:

a. do eɣvéni asó doɣunin?
what.NOM rise.3SG from.the.ACC east.ACC
'What rises from the east?'
(o ílon)
the.NOM sun.NOM
'the sun'

b. t akšémin do lámbi son uranón?
the.ACC night.ACC what.NOM shine.3SG in.the.ACC sky.ACC
'What shines in the sky at night?'
(o féŋgon)
the.NOM moon.NOM
'the moon'

c. i katsímali dóɣna ékripsen?
the.ACC clouds.ACC what.ACC hide.Past.3SG
'What did the clouds hide?'
(ton ílon)
the.ACC sun.ACC
'the sun'

d. i katsímali dóɣna ékripsen?
the.ACC clouds.ACC what.ACC hide.Past.3SG
'What did the clouds hide?'
(to féŋgon)
the.ACC moon.ACC
'the moon'

(32) Universal quantifiers (‘all’ and ‘every’):

a. píos pái so džamín?
who.NOM go.3SG to.the.ACC mosque.ACC
‘Who goes to the mosque?’
([úl:i]$_{\text{I-Foc}}$)
everyone.NOM
‘everyone’

b. džumá píos pái so džamín?
Friday day who.NOM go.3SG to.the.ACC mosque?
‘Who goes to the mosque on Fridays?’
([úl:i]$_{\text{I-Foc}}$)
everyone.NOM
‘everyone’

c. manaχón i aišé pái so džamín?
only the.NOM Ayşe.NOM go.3SG to.the.ACC mosque.ACC
‘Did only Ayşe go to the mosque?’
([úl:i]$_{\text{C-Foc}}$)
everyone.NOM
‘everyone’

d. píos pái sa parχáræ?
who.NOM go.3SG to.the.ACC pastures.ACC
‘Who goes to the pastures?’
([úl:i]$_{\text{I-Foc}}$)
everyone.NOM
‘everyone’

e. džumartesí píos pái sa parχáræ?
Saturday who.NOM go.3SG to.the.ACC pastures.ACC
‘Who goes to the pastures on Saturdays?’
([úl:i]$_{\text{I-Foc}}$)
everyone.NOM
‘everyone’

f. manaχón i aišé pái sa parχáræ?
only the.NOM Ayşe.NOM go.3SG to.the.ACC pastures.ACC
'Only Ayşe goes to the pastures?'
([úlːi]$_{C\text{-}Foc}$)
everyone.NOM
'everyone'

g. píos epíjen so istambólin?
who.NOM say.Past.3SG to.the.ACC Istanbul.ACC
([kanís]$_{C\text{-}Foc}$)
no-one.NOM
'no-one'

h. ɟetšén sené píos epíjen so istambólin?
last year who.NOM go.Past.3SG to.the.ACC Istanbul.ACC
'Who went to Istanbul last year?'
([kanís]$_{C\text{-}Foc}$)
no-one.NOM
'no-one'

i. úlːi epíɣane so istambólin?
everyone.NOM go.Past.3PL to.the.ACC Istanbul.ACC
'Everyone went to Istanbul.'
([kanís]$_{C\text{-}Foc}$)
no-one.NOM
'no-one'

j. píos epíen sin inɟiltéran?
who.NOM go.Past.3SG to.the.ACC England.ACC
([kanís]$_{C\text{-}Foc}$)
no-one.NOM
'no-one'

k. ɟetšén sené píos epíen sin inɟiltéran?
last year who.NOM go.Past.3SG to.the.ACC England.ACC
'Who went to England last year?'
([kanís]$_{C\text{-}Foc}$)
no-one.NOM
'no-one'

l. úlːi epíɣane sin inɟiltéran?
everyone.NOM go.Past.3PL to.the.ACC England.ACC
'Did everyone go to England?'
([kanís]$_{\text{C-Foc}}$)
no-one.NOM
'no-one'

m. píos eš ospítin so χoríon?
who.NOM have.3SG house.ACC in.the.ACC village.ACC
'Who has a house in the village?'
([úlːi]$_{\text{I-Foc}}$ / [káθa is]$_{\text{I-Foc}}$ / [kanís]$_{\text{C-Foc}}$)
everyone.NOM every one.NOM no-one.NOM
'everyone / everyone / no-one'

n. manaχón alís eš ospítin so χoríon?
only Alis.NOM have.3SG house.ACC in.the.ACC village.ACC
'Does only Alis have a house in the village?'
([úlːi]$_{\text{C-Foc}}$)
everyone.NOM
'everyone'

o. úlːi éχune ospítin so χoríon?
everyone.NOM have.3PL house.ACC to.the.ACC village.ACC
'Does everyone have a house in the village?'
([kanís]$_{\text{C-Foc}}$)
no-one.NOM
'no-one'

p. píos en hastás?
who.NOM be.3SG ill.NOM
'Who is ill?'
([kanís]$_{\text{C-Foc}}$)
no-one.NOM
'no-one'

q. manaχón alís epíen so kurbétin?
only Alis.NOM go.Past.3SG to.the.ACC abroad.ACC
'Did only Alis go abroad?'
([úlːa t andrúðæ]$_{\text{C-Foc}}$)
all.NOM the.NOM men.NOM
'all the men'

r. manaχón o ʝašínis pái s okʰúlːin?
only the.NOM Yasinis.NOM go.3SG to school.ACC
'Does only Yasinis go to school?'
([úlːa ta ɣarðélæ]$_{\text{C-Foc}}$)
all.NOM the.NOM children.NOM
'all the children'

5 The distribution of topics and foci

(33) DP-subject as C-Top and DP-object as C-Foc:

a. Question:

alís	do	epítšen?	
Alis.NOM	what.ACC	do.Past.3SG	
do	éfaen	so	bairámin?
what.ACC	eat.Past.3SG	in.the.ACC	Bayram.ACC

'What did Alis do? What did he eat at Bayram?'

b. Answers:

vál:ahi,	utš	ekséro	alís	do	epítšen,	áma …
frankly	NEG	know.1SG	Alis.NOM	what.ACC	do.Past.3SG	but

'Frankly, I don't know about Alis, but …'

(34) DP-object as C-Top and DP-subject as C-Foc:

a. Question:

o	tšorbás	do	ejéndo?
the.NOM	soup.NOM	what.ACC	happen.Past.3SG
atón	kanís	éfaen	a?
this.ACC	anyone.NOM	eat.Past.3SG	it.ACC

'What about the soup? Has anyone eaten that?'

b. Answer:

vál:ahi,	utš	ekséro	o	tšorbás	do	ejéndo,	áma
frankly	NEG	know.1SG	the.NOM	soup.NOM	what.ACC	happen.Past.3SG	but

'Frankly, I don't know about the soup, but …'

(35) Distrative verb; IO as C-Foc and DO as C-Top:

b. Question:

i	antíka	tše	i	sandáʎa
the.NOM	antique.NOM	and	the.NOM	chair.NOM

do	ejéndo?
what.ACC	happen.Past.3SG

o	pápʰos	tínan	éðocen	a?
the.NOM	grandfather.NOM	who.ACC	give.Past.3SG	it.ACC

'What about the antique table and the chair? Who has your granddad bequeathed that to?'

b. Answer:

vál:ahi	i	antíka	do	ejéndo
frankly	the.NOM	antique.NOM	what.ACC	happen.Past.3SG

utš	ekséro,	áma …
NEG	know.1SG	but

'Frankly, I don't know about the antique table, but …'

(36) Distrative verb; DO as C-Foc and IO as C-Top:

b. Question:

o	tšíri	s	do	ejéndo?
the.NOM	father.NOM	you.POSS	what.ACC	happen.Past.3SG

o	pápʰos	do	éðocen	atón?
the.NOM	grandfather.NOM	what.ACC	give.Past.3SG	he.ACC

'What about your dad? What has granddad bequeathed to him?'

b. Answer:

vál:ahi	o	tšíri	m	do	ejéndo
frankly	the.NOM	father.NOM	I.POSS	what.ACC	happen.Past.3SG

utš	ekséro,	áma …
NEG	know.1SG	but

'Frankly, I don't know about my dad, but …'

(37) DP-object as C-Top and DP-subject as I-Foc:

a. Question:

tsi	birjýlis	t	aðélfæ
the.GEN	Birgül.GEN	the.NOM	siblings.NOM
d	epíkane	so	pártin?
what.ACC	do.Past.3PL	at.the.ACC	party.ACC

'What did Birgül's siblings get to drink at the party?'

b. Answer:

vál:ahi	as	aðélfæ	tes	utš	ekséro,	áma …
frankly	from.the.ACC	siblings.ACC	she.POSS	NEG	know.1SG	but

'Frankly, I do not know about all her siblings but …'

Appendix B Focus interpretation constraint in Romeyka

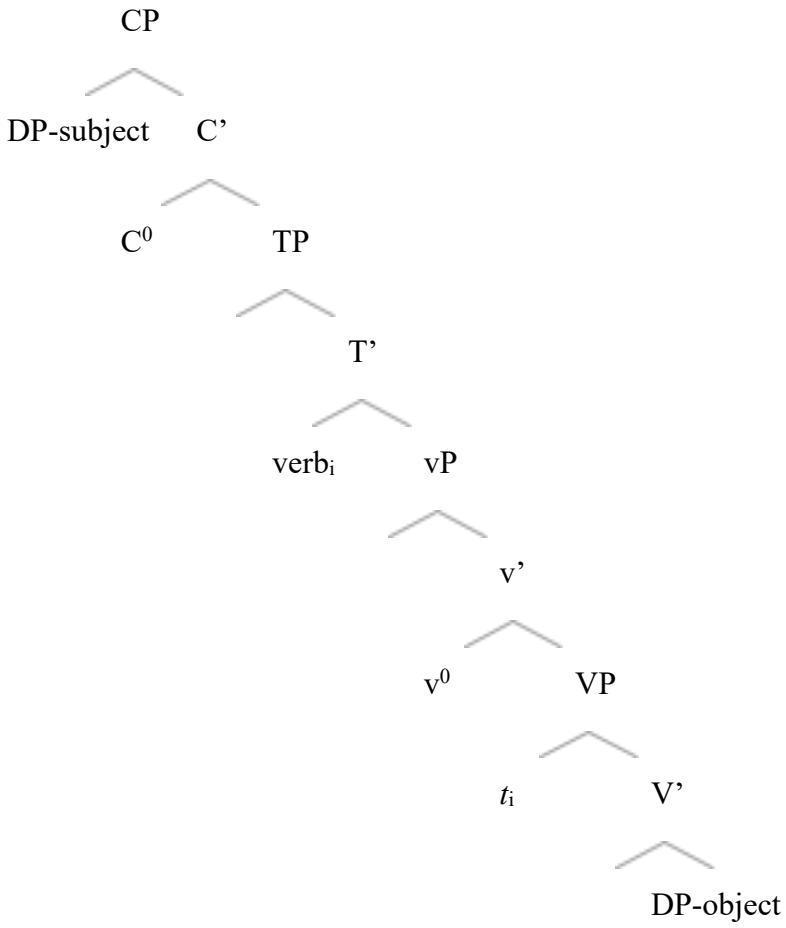
1 Matrix declarative clauses in Romeyka

1.1 SVO

First, in a pragmatically unmarked SVO word order, the verb undergoes V^0-to-T^0 movement, while a ^ feature is associated with the EF of the phase head C^0, resulting in the left dislocation of the DP-subject (see (1) and (2)):

(1) $[_{CP}$ DP-subject C^0 $[_{TP}$ $verb_i$ $[_{vP}$ $[_{VP}$ t_i DP-object]]]]

(2)



For (1) and (2), the Spell-Out rule in (3) is applied:

(3) a. $PF_{\Sigma} = PF_1 \times PF_2 \times PF_3$
b. PF_{Σ} = CP × TP × VP
c. PF_{Σ} = {DP-subject} × {verb} × {DP-object}

For (1) and (2), the Transfer rule in (4) is applied:

(4) a. $LF_{\Sigma} = LF_1 \times LF_2 \times LF_3$
b. LF_{Σ} = CP × TP × VP
c. LF_{Σ} = {DP-subject} × {verb} × {DP-object}

The semantic interpretation of (4) would be like the one in (5):

(5) {∃e (Agent (e, DP-subject) & Theme (e, DP-object)) | e is a verb}

Consider the clause in (6):

(6) Romeyka:

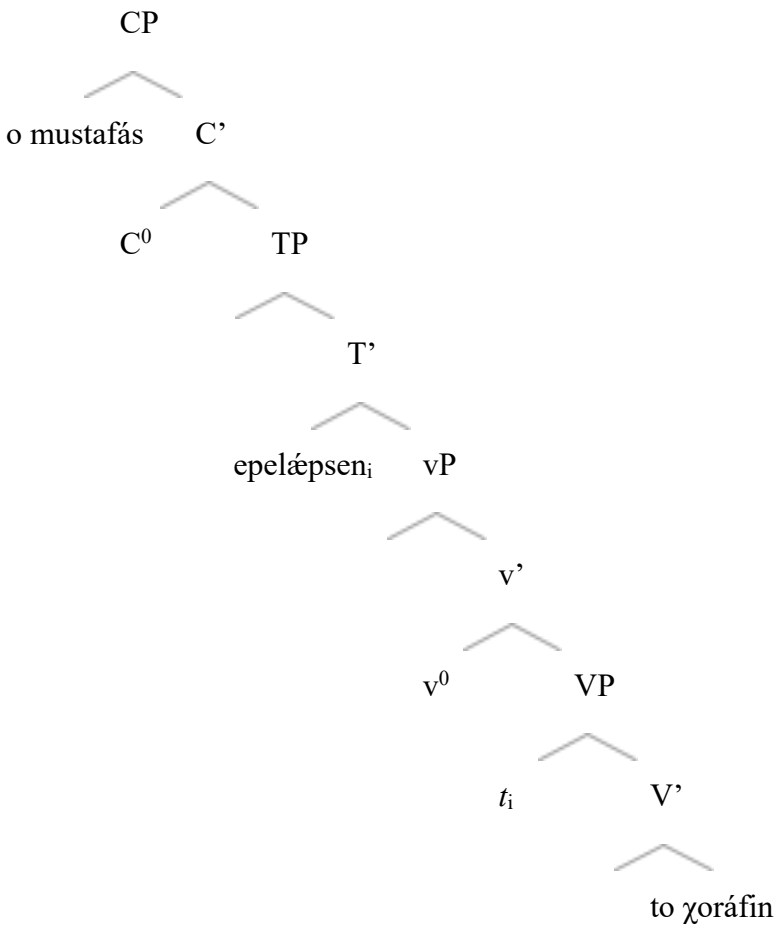
o	mustafás	epeLÆpsen	to	χoráfin.
the.NOM	Mustafas.NOM	put.fertiliser.Past.3SG	the.ACC	field.ACC

'Mustafas put fertiliser on the field.'
(S01; 150703_0040; 02:16)

In the clause in (6), the verb *eplæpsen* 'he put fertiliser' undergoes V^0-to-T^0 movement, while a ^ feature is associated with the EF of the phase head C^0, resulting in the left dislocation of the DP-subject *o mustafás* 'Mustafas' (see (7) and (8)):

(7) $[_{CP}$ o mustafás C^0 $[_{TP}$ epelæpsen$_i$ $[_{vP}$ $[_{VP}$ t_i to χoráfin]]]]

(8)



For (7) and (8), the Spell-Out rule in (9) is applied:

(9) a. $PF_{\Sigma} = PF_1 \times PF_2 \times PF_3$

b. PF_{Σ} = CP × TP × VP

c. PF_{Σ} = {o mustafás} × {epeláepsen} × {to χoráfin}

For (7) and (8), the Transfer rule in (10) is applied:

(10) a. $LF_{\Sigma} = LF_1 \times LF_2 \times LF_3$

b. LF_{Σ} = CP × TP × VP

c. LF_{Σ} = {o mustafás} × {epeláepsen} × {to χoráfin}

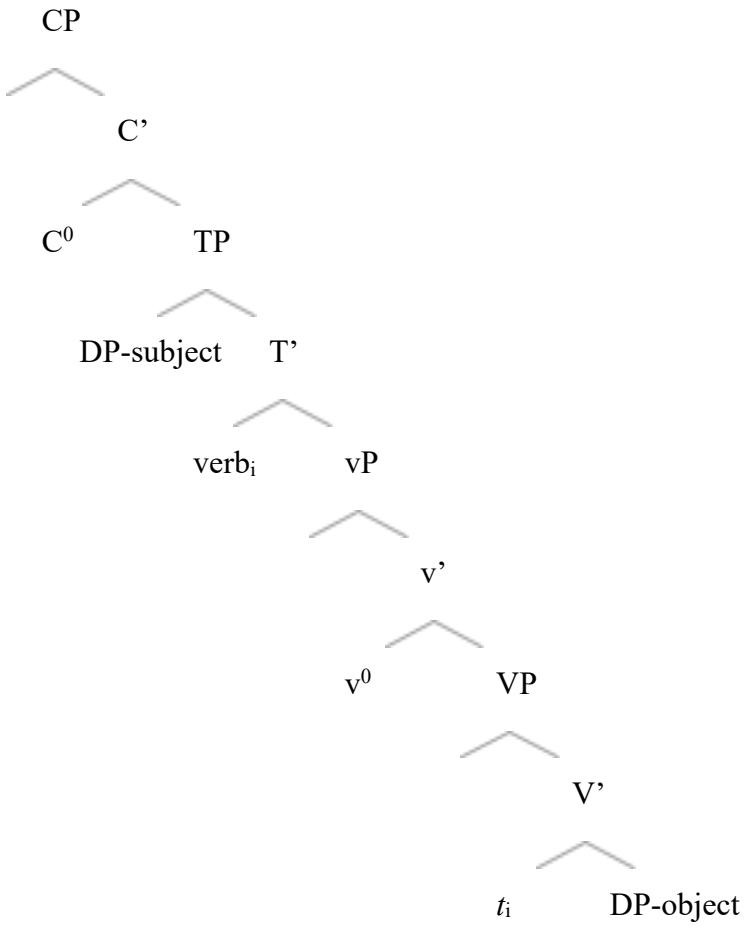
The semantic interpretation of (10) would be like the one in (11):

(11) {∃e (Agent (e, o mustafás) & Theme (e, to χoráfin)) | e is *epeláepsen*}

Second, when the DP-subject is focused in an SVO word order, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the [*u*ϕ] features of T^0, which are inherited by the phase head C^0, resulting in the movement of the DP-subject to the [Spec, TP] (see (12) and (13)):

(12) $[_{CP}\ C^0\ [_{TP}$ DP-subject verb$_i$ $[_{vP}\ [_{VP}\ t_i$ DP-object]]]]

(13)



For (12) and (13), the Spell-Out rule in (14) is applied:

(14) a. $PF_\Sigma = PF_1 \times PF_2$

b. $PF_\Sigma = TP \times VP$

c. $PF_\Sigma = \langle$DP-subject, verb$\rangle \times \{$DP-object$\}$

For (12) and (13), the Transfer rule (15) is applied:

(15) a. $LF_{\Sigma} = LF_1 \times LF_2$
b. LF_{Σ} = TP × VP
c. LF_{Σ} = <DP-subject, verb> × {DP-object}

The semantic interpretation of (15) would be like the one in (16):

(16) a. {∃e ([Agent (e, someone)]$_F$ & Theme (e, DP-object)) | e is a verb}
b. {∃e ([Agent (e, DP-subject)]$_O$ & Theme (e, DP-object)) | e is a verb}

Consider the clause in (17):

(17) Romeyka:

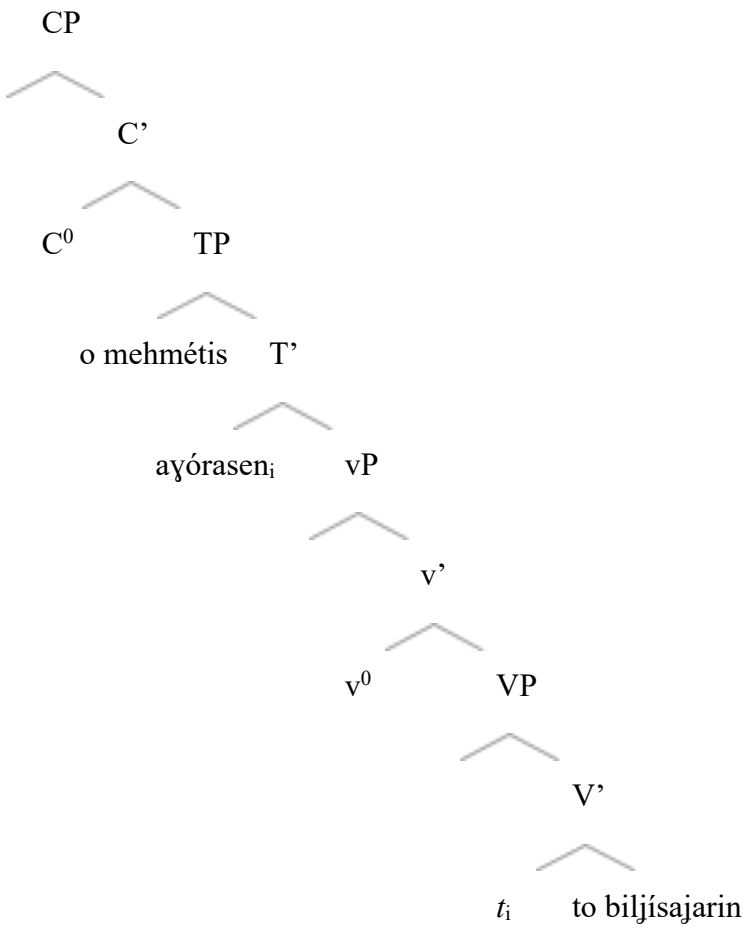
[o	mehMÉtis]$_{Foc}$	aɣórasen	to	biljísajarin.
the.NOM	Mehmetis.NOM	buy.Past.3SG	the.ACC	computer.ACC

'Alis bought the computer.'
(S01; 150703_0040; 05:32)

In the clause in (17), the verb *ayórasen* 'he bought' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the [*u*ϕ] features of T^0, which are inherited by the phase head C^0, resulting in the movement of the DP-subject *o mehmétis* 'Mehmetis' to the [Spec, TP] (see (18) and (19)):

(18) [$_{CP}$ C^0 [$_{TP}$ o mehmétis aɣórasen$_i$ [$_{vP}$ [$_{VP}$ t_i to biljísajarin]]]]

(19)



For (18) and (19), the Spell-Out rule in (20) is applied:

(20) a. $PF_\Sigma = PF_1 \times PF_2$

b. $PF_\Sigma = TP \times VP$

c. PF_Σ = <o mehmétis, aɣórasen> × {to biljísajarin}

For (18) and (19), the Transfer rule (21) is applied:

(21) a. $LF_\Sigma = LF_1 \times LF_2$

b. $LF_\Sigma = TP \times VP$

c. LF_Σ = <o mehmétis, aɣórasen> × {to biljísajarin}

The semantic interpretation of (21) would be like the one in (22):

(22) a. {∃e ([Agent (e, someone)]$_F$ & Theme (e, to biljísajarin)) | e is a *aɣórasen*}

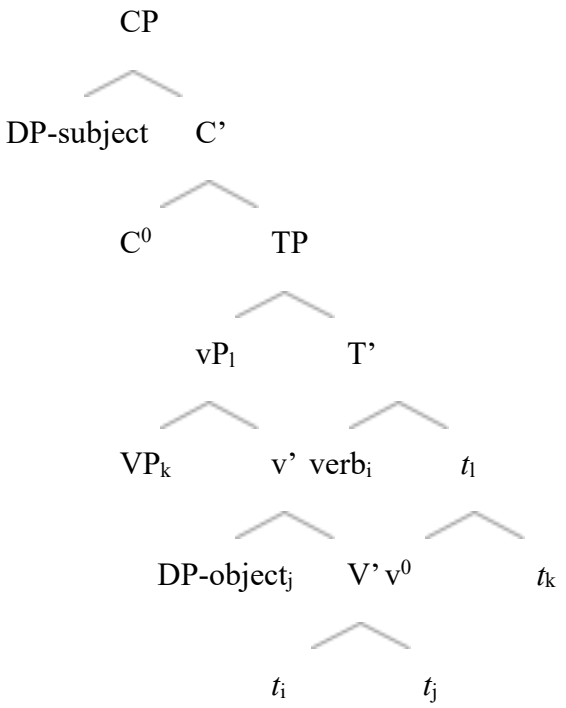
c. {∃e ([Agent (e, o mehmétis)]$_O$ & Theme (e, to biljísajarin)) | e is a *aɣórasen*}

1.2 SOV

First, when the DP-object is focused in an SOV word order, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-object to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (22) and (23)):

(23) $[_{CP}$ DP-subject C^0 $[_{TP}$ $[_{vP}$ $[_{VP}$ DP-object$_j$ t_i $t_j]]$ verb$_i]]$

(24)



For (22) and (23), the Spell-out rule in (25) is applied:

(25) a. $PF_\Sigma = PF_1 \times PF_2$

b. $PF_\Sigma = CP \times TP$

c. PF_Σ = {DP-subject} × <DP-object, verb>

For (22) and (23), the Transfer rule in (26) is applied:

(26) a. $LF_{\Sigma} = LF_1 \times LF_2$
b. LF_{Σ} = CP × TP
c. LF_{Σ} = {DP-subject} × <DP-object, verb>

The semantic interpretation of (26) would be like the one in (27):

(27) a. {∃e (Agent (e, DP-subject) & [Theme (e, someone)]$_F$) | e is a verb}
b. {∃e (Agent (e, DP-subject) & [Theme (e, DP-object)]$_O$) | e is a verb}

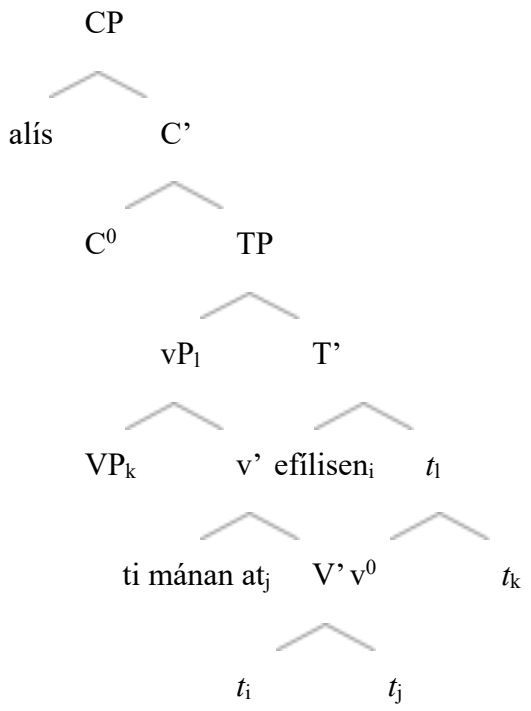
Consider the clause in (28):

(28) Romeyka:
alís [ti MÁnan at]$_{Foc}$ efílisen.
Alis.NOM the.ACCmother.ACC his kiss.Past.3SG
'Alis kissed his mother.'
(S01; 150702_0013; 06:08)

In the clause in (28), the verb *efílisen* 'he kissed' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-object *ti mánan at* 'his mother' to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (29) and (30)):

(29) [$_{CP}$ alís C^0 [$_{TP}$ [$_{vP}$ [$_{VP}$ ti mánan at$_j$ t_i t_j]] efílisen$_i$]]

(30)



For (29) and (30), the Spell-out rule in (31) is applied:

(31) a. $PF_{\Sigma} = PF_1 \times PF_2$
b. $PF_{\Sigma} = CP \times TP$
c. PF_{Σ} = {alís} × <ti mánan at, efílisen>

For (29) and (30), the Transfer rule in (32) is applied:

(32) a. $LF_{\Sigma} = LF_1 \times LF_2$
b. $LF_{\Sigma} = CP \times TP$
c. LF_{Σ} = {alís} × <ti mánan at, efílisen>

The semantic interpretation of (32) would be like the one in (33):

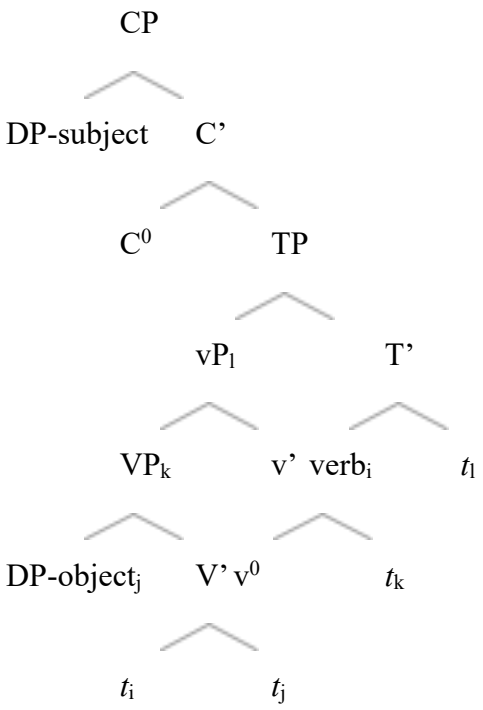
(33) a. {∃e (Agent (e, alís) & [Theme (e, someone)]$_F$) | e is *efílisen*}
b. {∃e (Agent (e, alís) & [Theme (e, ti mánan at)]$_O$) | e is *efílisen*}

Second, when both the DP-subject and the DP-object are focused in an SOV word order, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-

object to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (34) and (35)):

(34) $[_{CP}$ DP-subject C^0 $[_{TP}$ $[_{vP}$ $[_{VP}$ DP-object$_j$ t_i $t_j]]$ verb$_i]]$

(35)



For (34) and (35), the Spell-out rule in (36) is applied:

(36) a. $PF_\Sigma = PF_1 \times PF_2$
b. $PF_\Sigma = CP \times TP$
c. PF_Σ = {DP-subject} × <DP-object, verb>

For (34) and (35), the Transfer rule in (37) is applied:

(37) a. $LF_\Sigma = LF_1 \times LF_2$
b. $LF_\Sigma = CP \times TP$
c. LF_Σ = {DP-subject} × <DP-object, verb>

The semantic interpretation of (37) would be like the one in (38):

(38) a. {∃e [(Agent (e, someone)]$_F$ & [Theme (e, something)]$_F$) | e is a verb}
b. {∃e [(Agent (e, DP-subject)]$_O$ & [Theme (e, DP-object)]$_O$) | e is a verb}

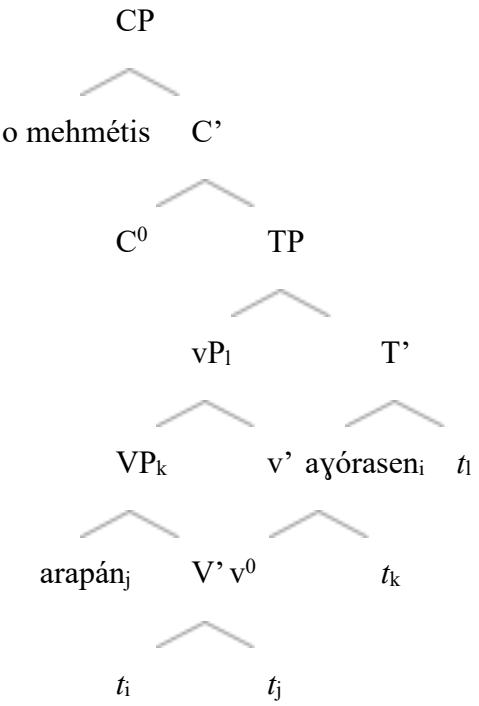
Consider the clause in (39):

(39) Romeyka:
[o mehMÉtis]$_{Foc}$ [araPÁn]$_{Foc}$ aɣórasen.
the.NOM Mehmetis.NOM car.ACC buy.Past.3SG
'Mehmetis bought a car.'
(S01; 150703_0042; 07:39)

In the clause in (39), the verb *aɣórasen* 'he bought' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-object *arapán* 'car' to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (40) and (41)):

(40) [$_{CP}$ o mehmétis C^0 [$_{TP}$ [$_{vP}$ [$_{VP}$ arapán$_j$ t_i t_j]] aɣórasen$_i$]]

(41)



For (40) and (41), the Spell-Out rule in (42) is applied:

(42) a. $PF_{\Sigma} = PF_1 \times PF_2$
b. PF_{Σ} = CP × TP
c. PF_{Σ} = {o mehmétis} × <arapán, aɣórasen>

For (40) and (41), the Transfer rule (43) is applied:

(43) a. $LF_{\Sigma} = LF_1 \times LF_2$
b. LF_{Σ} = CP × TP
c. LF_{Σ} = {o mehmétis} × <arapán, aɣórasen>

The semantic interpretation of (43) would be like the one in (44):

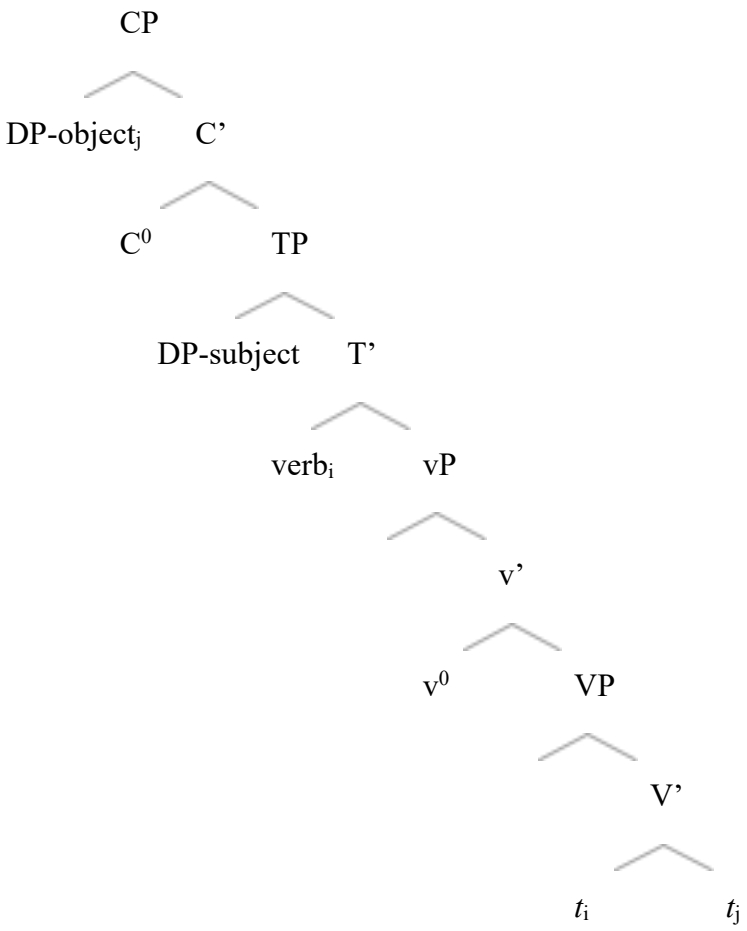
(44) a. {∃e ([Agent (e, someone)]$_F$ & [Theme (e, something)]$_F$) | e is *aɣórasen*}
b. {∃e ([Agent (e, o mehmétis)]$_O$ & [Theme (e, arapán)]$_O$) | e is *aɣórasen*}

1.3 OSV

In an OSV word order, the DP-subject is focused and the DP-object is topicalised. The verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the [*u*ϕ] features of T^0, which are inherited by the phase head C^0, resulting in the movement of the DP-subject to the [Spec, TP]. Moreover, a ^ feature is associated with the EF of the phase head C^0, giving rise to the left dislocation of the DP-object (see (45) and (46)):

(45) [$_{CP}$ DP-object$_j$ C^0 [$_{TP}$ DP-subject verb$_i$ [$_{vP}$ [$_{VP}$ t_i t_j]]]]

(46)



For (45) and (46), the Spell-Out rule in (47) is applied:

(47) a. $PF_\Sigma = PF_1 \times PF_2$

b. PF_Σ = CP × TP

c. PF_Σ = {DP-object} × <DP-subject, verb>

For (45) and (46), the Transfer rule in (48) is applied:

(48) a. $LF_\Sigma = LF_1 \times LF_2$

b. LF_Σ = CP × TP

c. LF_Σ = {DP-object} × <DP-subject, verb>

The semantic interpretation of (48) would be like the one in (49):

(49) a. {∃e ([Agent (e, someone)]$_F$ & Theme (e, DP-object)) | e is a verb}

b. {∃e ([Agent (e, DP-subject)]$_O$ & Theme (e, DP-object)) | e is a verb}

Consider the clause in (50):

(50) Romeyka:

tin	aišén	[o	dohTÓris]$_{Foc}$	epíren.
the.ACC	Ayşe.ACC	the.NOM	doctor.NOM	marry.Past.3SG

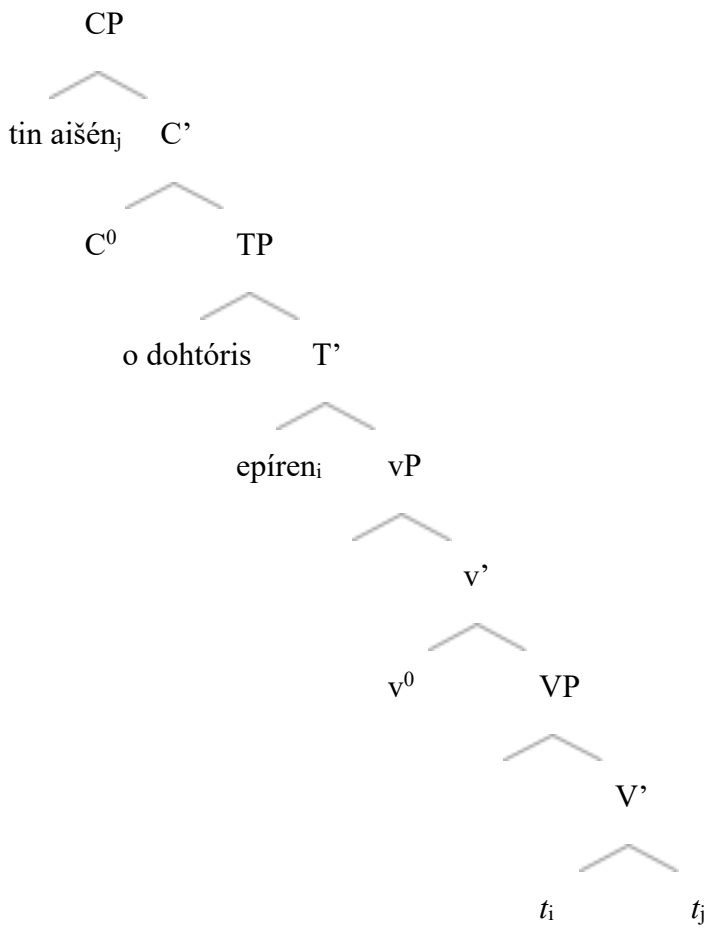
'The doctor married Ayşe.'

(S01; 140102_0008; 01:15)

In the clause in (50), the verb *epíren* 'he married' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the [*u*ϕ] features of T^0, which are inherited by the phase head C^0, resulting in the movement of the DP-subject *o dohtóris* 'the doctor' to the [Spec, TP]. Moreover, a ^ feature is associated with the EF of the phase head C^0, giving rise to the left dislocation of the DP-object DP-object *tin aišén* 'Ayşe' (see (51) and (52)):

(51) [$_{CP}$ tin aišén$_j$ C^0 [$_{TP}$ o dohtóris epíren$_i$ [$_{vP}$ [$_{VP}$ t_i t_j]]]]

(52)



For (51) and (52), the Spell-Out rule in (53) is applied:

(53) a. $PF_{\Sigma} = PF_1 \times PF_2$

b. PF_{Σ} = CP × TP

c. PF_{Σ} = {tin aišén} × <o dohtóris, epíren>

For (51) and (52), the Transfer rule in (54) is applied:

(54) a. $LF_{\Sigma} = LF_1 \times LF_2$

b. LF_{Σ} = CP × TP

c. LF_{Σ} = {tin aišén} × <o dohtóris, epíren>

The semantic interpretation of (54) would be like the one in (55):

(55) a. {∃e ([Agent (e, someone)]$_F$ & (Theme (e, tin aišén)) | e is *epíren*}

c. {∃e ([Agent (e, o dohtóris)]$_O$ & (Theme (e, tin aišén)) | e is *epíren*}

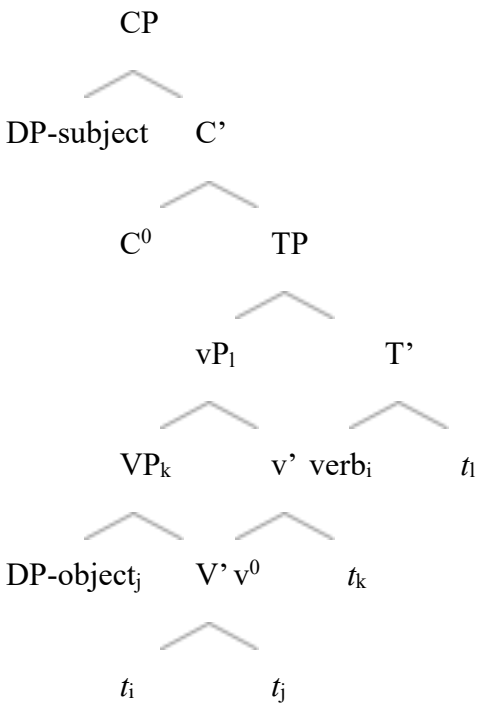
2 Subordinate declarative clauses in Romeyka

2.1 SOV

First, in a pragmatically unmarked SOV word order, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-object to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (56) and (57)). In this case, the ^ feature has not internal semantic structure (see 7.3 for a detailed discussion):

(56) $[_{CP}$ DP-subject C^0 $[_{TP}$ $[_{vP}$ $[_{VP}$ DP-object$_j$ t_i $t_j]]$ verb$_i]]$

(57)



For (56) and (57), the Spell-out rule in (58) is applied:

(58) a. $PF_\Sigma = PF_1 \times PF_2$

b. $PF_\Sigma = CP \times TP$

c. PF_Σ = {DP-subject} × <DP-object, verb>

For (56) and (57), the Transfer rule (59) is applied:

(59) a. $LF_{\Sigma} = LF_1 \times LF_2$
b. LF_{Σ} = CP × TP
c. LF_{Σ} = {DP-subject} × <DP-object, verb>

The semantic interpretation of (59) would be like the one in (60):

(60) {∃e (Agent (e, DP-subject) & Theme (e, DP-object)) | e is a verb}

Consider the clause in (61):

(61) Romeyka:

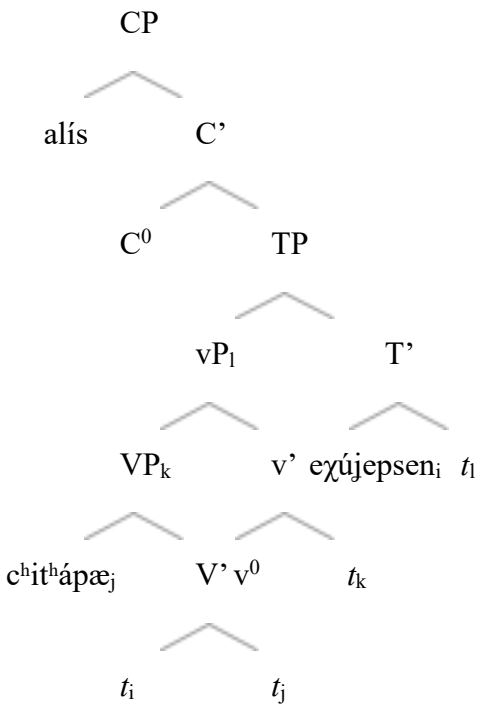
eɣó	θaRÓ,	alís	pol:á	ómorfa	cʰitʰápæ	eχÚjepsen.
I	think.1SG	Alis.NOM	many.ACC	nice.ACC	books.ACC	read.Past.3SG

'I think that Alis read many nice books.'
(S01; 812_0059; 00:20)

In the clause in (61), the verb *eχújepsen* 'he read' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-object *cʰitʰápæ* 'books' to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (62) and (63)). In this case, the ^ feature has not internal semantic structure (see 7.3 for a detailed discussion):

(62) $[_{CP}$ alís C^0 $[_{TP}$ $[_{vP}$ $[_{VP}$ cʰitʰápæ$_j$ t_i $t_j]]$ eχújepsen$_i]]$

(63)



For (62) and (63), the Spell-out rule in (64) is applied:

(64) a. $PF_{\Sigma} = PF_1 \times PF_2$
b. $PF_{\Sigma} = CP \times TP$
c. PF_{Σ} = {alís} × <c^{h}it^{h}ápæ, eχújepsen>

For (62) and (63), the Transfer rule (65) is applied:

(65) a. $LF_{\Sigma} = LF_1 \times LF_2$
b. $LF_{\Sigma} = CP \times TP$
c. LF_{Σ} = {alís} × <c^{h}it^{h}ápæ, eχújepsen>

The semantic interpretation of (65) would be like the one in (66):

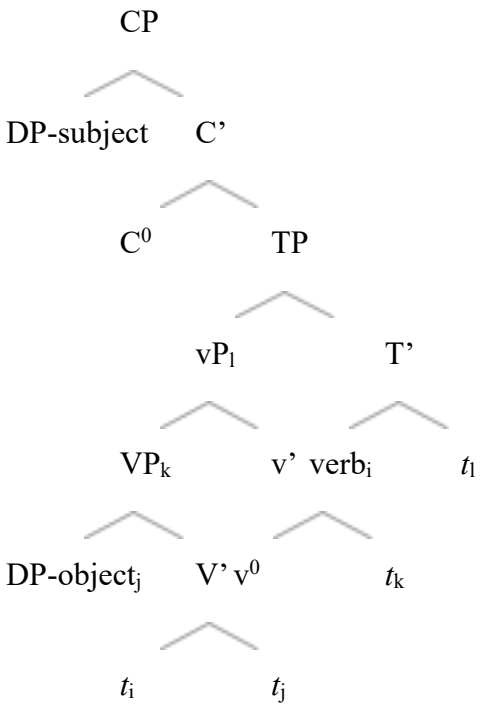
(66) {∃e (Agent (e, alís) & Theme (e, c^{h}it^{h}ápæ)) | e is *eχújepsen*}

Second, when the DP-object is focused in an SOV word order, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-object to the [Spec, VP]. The ^

feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (67) and (68)):

(67) $[_{CP}$ DP-subject C^0 $[_{TP}$ $[_{vP}$ $[_{VP}$ DP-object$_j$ t_i $t_j]]$ verb$_i]]$

(68)



For (67) and (68), the Spell-out rule (69) is applied:

(69) a. $PF_\Sigma = PF_1 \times PF_2$
b. $PF_\Sigma = CP \times TP$
c. PF_Σ = {DP-subject} × <DP-object, verb>

For (67) and (68), the Transfer rule (70) is applied:

(70) a. $LF_\Sigma = LF_1 \times LF_2$
b. $LF_\Sigma = CP \times TP$
c. LF_Σ = {DP-subject} × <DP-object, verb>

The semantic interpretation in (70) would be like the one in (71):

(71) a. (∃e (Agent (e, DP-subject) & [Theme (e, someone)]$_F$) | e is a *verb*}

b. (∃e (Agent (e, DP-subject) & [Theme (e, DP-object)]$_O$) | e is a *verb*}

Consider the clause in (72):

(72) Romeyka:

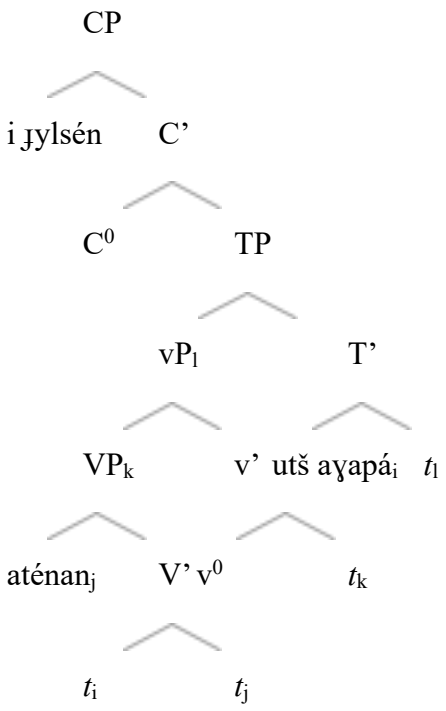
o	mohal:ímis	Ípen,		
the.NOM	teacher.NOM	say.Past.3SG		
i	ɟylsén	[aTÉnan]$_{Foc}$	utš	aɣapá.
the.NOM	Gülsen.NOM	she.ACC	NEG	love.3SG

'The teacher said that Gülsen doesn't like her.'

In the clause in (72), when the DP-object *aténan* 'her' is focused in an SOV word order, the verb *utš aɣapá* 'she doesn't like' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-object *aténan* 'her' to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (73) and (74)):

(73) [$_{CP}$ i ɟylsén C^0 [$_{TP}$ [$_{vP}$ [$_{VP}$ aténan$_j$ t_i t_j]] utš aɣapá$_i$]]

(74)



For (73) and (74), the Spell-out rule (75) is applied:

(75) a. $PF_{\Sigma} = PF_1 \times PF_2$
b. $PF_{\Sigma} = CP \times TP$
c. PF_{Σ} = {i ɟylsén} × <aténan, utš aɣapá>

For (73) and (74), the Transfer rule (76) is applied:

(76) a. $LF_{\Sigma} = LF_1 \times LF_2$
b. $LF_{\Sigma} = CP \times TP$
c. LF_{Σ} = {i ɟylsén} × <aténan, utš aɣapá>

The semantic interpretation in (76) would be like the one in (77):

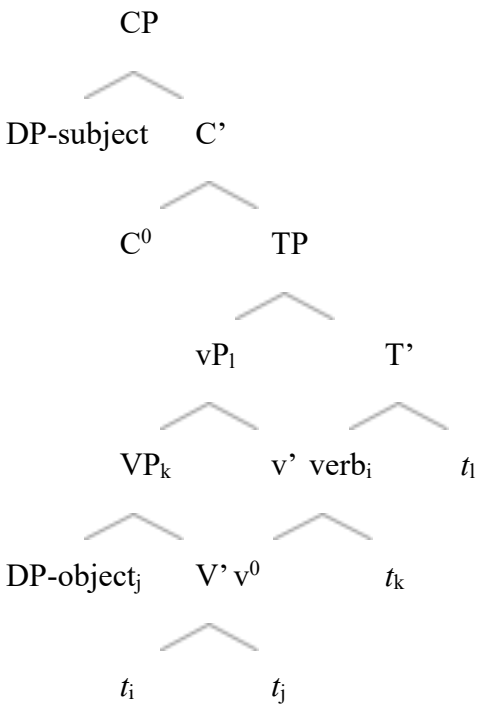
(77) a. (∃e (Agent (e, i ɟylsén) & [Theme (e, someone)]$_F$) | e is *utš aɣapá*}
b. (∃e (Agent (e, i ɟylsén) & [Theme (e, aténan)]$_O$) | e is *utš aɣapá*}

Third, when both the DP-subject and the DP-object are focused in an SOV word order, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-object to the

[Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (78) and (79)):

(78) $[_{CP}$ DP-subject C^0 $[_{TP}$ $[_{vP}$ $[_{VP}$ DP-object$_j$ t_i $t_j]]$ verb$_i]]$

(79)



For (78) and (79), the Spell-out rule (80) is applied:

(80) a. $PF_\Sigma = PF_1 \times PF_2$
b. $PF_\Sigma = CP \times TP$
c. PF_Σ = {DP-subject} × <DP-object, verb>

For (78) and (79), the Transfer rule (81) is applied:

(81) a. $LF_\Sigma = LF_1 \times LF_2$
b. $LF_\Sigma = CP \times TP$
c. LF_Σ = {DP-subject} × <DP-object, verb>

The semantic interpretation in (81) would be like the one in (82):

(82) a. {∃e [(Agent (e, someone)]$_F$ & [Theme (e, something)]$_F$) | e is a verb}
b. {∃e [(Agent (e, DP-subject)]$_O$ & [Theme (e, DP-object)]$_O$) | e is a verb}

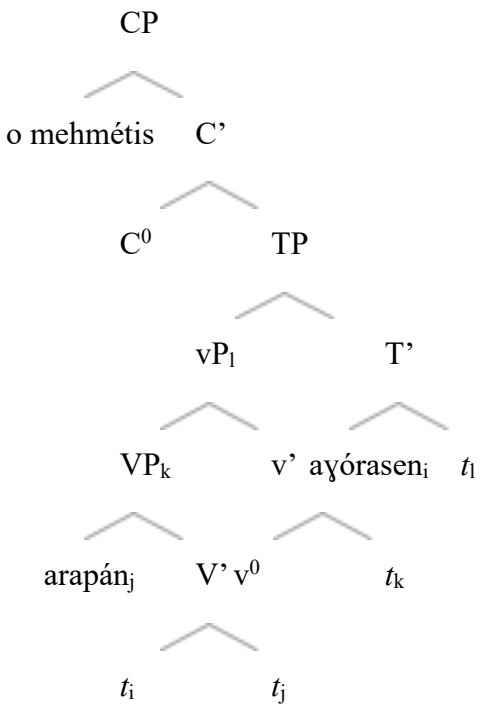
Consider the clause in (83):

(83) Romeyka:
eɣó pa léɣo se,
I.NOM PRT say.1SG you.ACC
[o mehMÉtis]$_{Foc}$ [araPÁN]$_{Foc}$ aɣórasen.
the.NOM Mehmetis.NOM car.ACC buy.Past.3SG
'I tell you that Mehmetis bought a car.'
(S01; 150703_0042; 08:20)

In the clause in (83), the verb *aɣórasen* 'he bought' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-object *arapán* 'car' to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (84) and (85)):

(84) [$_{CP}$ o mehmétis C^0 [$_{TP}$ [$_{vP}$ [$_{VP}$ arapán$_j$ t_i t_j]] aɣórasen$_i$]]

(85)



For (84) and (85), the Spell-out rule (86) is applied:

(86) a. $PF_{\Sigma} = PF_1 \times PF_2$

b. $PF_{\Sigma} = CP \times TP$

c. PF_{Σ} = {o mehmétis} × <arapán, aɣórasen>

For (84) and (85), the Transfer rule (87) is applied:

(87) a. $LF_{\Sigma} = LF_1 \times LF_2$

b. $LF_{\Sigma} = CP \times TP$

c. LF_{Σ} = {o mehmétis} × <arapán, aɣórasen>

The semantic interpretation in (87) would be like the one in (88):

(88) a. {∃e [(Agent (e, someone)]$_F$ & [Theme (e, something)]$_F$) | e is *aɣórasen*}

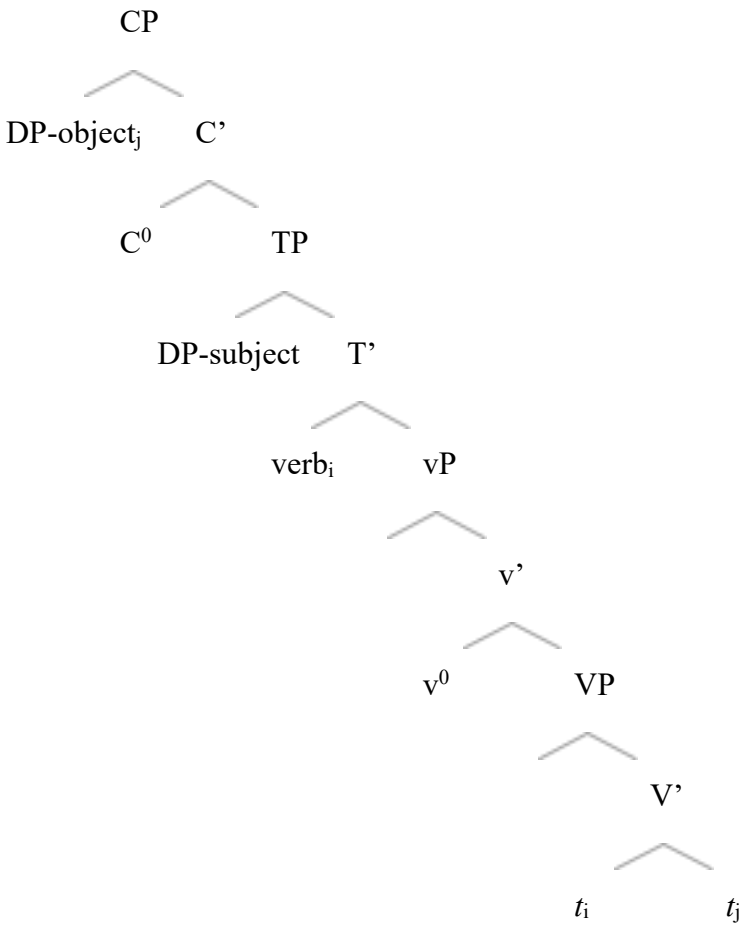
b. {∃e [(Agent (e, o mehmétis)]$_O$ & [Theme (e, arapán)]$_O$) | e is *aɣórasen*}

2.2 OSV

When in an OSV word order the DP-subject is focused and the DP-object is topicalised, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the [*u*ϕ] features of T^0, which are inherited by the phase head C^0, resulting in the movement of the DP-subject to the [Spec, TP]. Moreover, a ^ feature is associated with the EF of the phase head C^0, giving rise to the left dislocation of the DP-object (see (89) and (90)):

(89) $[_{CP}$ DP-object$_j$ C^0 $[_{TP}$ DP-subject verb$_i$ $[_{vP}$ $[_{VP}$ t_i $t_j]]]]$

(90)



For (89) and (90), the Spell-Out rule in (91) is applied:

(91) a. $PF_\Sigma = PF_1 \times PF_2$

b. PF_Σ = CP × TP

c. PF_Σ = {DP-subject} × <DP-subject, verb>

For (89) and (90), the Transfer rule in (92) is applied:

(92) a. $LF_{\Sigma} = LF_1 \times LF_2$
b. LF_{Σ} = CP × TP
c. LF_{Σ} = {DP-subject} × <DP-subject, verb>

The semantic interpretation of (92) would be like the one in (93):

(93) a. {∃e (Theme (e, DP-object) & [Agent (e, someone)]$_F$) | e is a verb}
b. {∃e (Theme (e, DP-object) & [Agent (e, DP-subject)]$_O$) | e is a verb}

Consider the clause in (94):

(94) Romeyka:

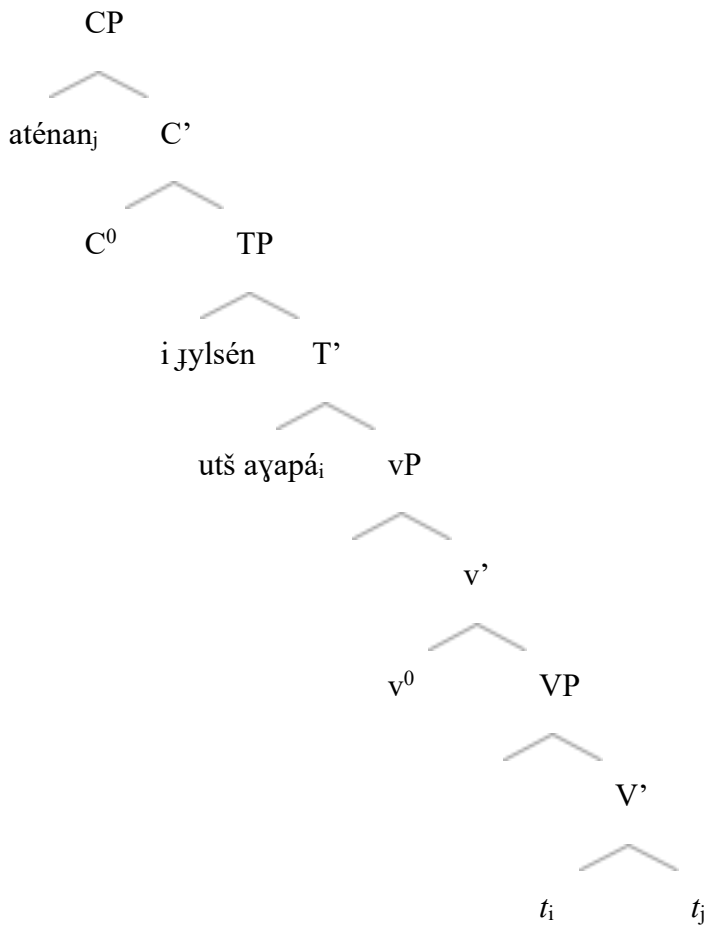
o	mohal:ímis	ípen,		
the.NOM	teacher.NOM	say.Past.3SG		
aténan	[i	ɟylSÉN]$_{Foc}$	utš	aɣapá.
she.ACC	the.NOM	Gülsen.NOM	NEG	love.3SG

'The teacher said that Gülsen doesn't like her.'
(S07; 812_0065; 05:01)

In the clause in (94), the verb *utš aɣapá* 'she doesn't like' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the [*u*ϕ] features of T^0, which are inherited by the phase head C^0, resulting in the movement of the DP-subject *i ɟylsén* 'Gülsen' to the [Spec, TP]. Moreover, a ^ feature is associated with the EF of the phase head C^0, giving rise to the left dislocation of the DP-object *aténan* 'her' (see (95) and (96)):

(95) [$_{CP}$ aténan$_j$ C^0 [$_{TP}$ i ɟylsén utš aɣapá$_i$ [$_{vP}$ [$_{VP}$ t_i t_j]]]]

(96)



For (95) and (96), the Spell-Out rule in (97) is applied:

(97) a. $PF_\Sigma = PF_1 \times PF_2$

b. $PF_\Sigma = CP \times TP$

c. PF_Σ = {aténan} × <i ɟylsén, utš aɣapá>

For (95) and (96), the Transfer rule in (98) is applied:

(98) a. $LF_\Sigma = LF_1 \times LF_2$

b. $LF_\Sigma = CP \times TP$

c. LF_Σ = {aténan} × <i ɟylsén, utš aɣapá>

The semantic interpretation of (98) would be like the one in (99):

(99) a. {∃e [Agent (e, someone)]$_F$ & (Theme (e, aténan)) | e is *utš aɣapá*}

b. {∃e [Agent (e, i ɟylsén)]$_O$ & (Theme (e, aténan)) | e is *utš aɣapá*}

3 Direct questions in Romeyka

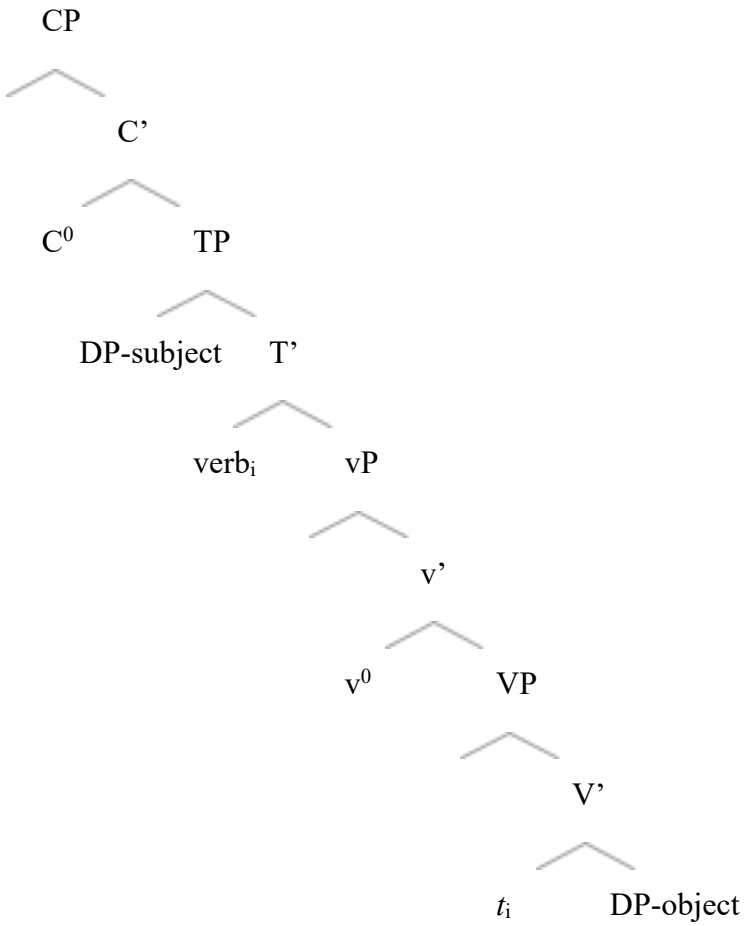
3.1 Direct yes/no questions in Romeyka

3.1.1 SVO

When the DP-subject is focused in an SVO yes/no question, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the [*u*ϕ] features of T^0, which are inherited by the phase head C^0, resulting in the movement of the DP-subject to the [Spec, TP] (see (100) and (101)):

(100) $[_{CP}\ C^0\ [_{TP}$ DP-subject verb$_i$ $[_{vP}\ [_{VP}\ t_i$ DP-object]]]]

(101)



For (100) and (101), the Spell-Out rule in (102) is applied:

(102) a. $PF_{\Sigma} = PF_1 \times PF_2$
b. PF_{Σ} = TP × VP
c. PF_{Σ} = <DP-subject, verb> × {DP-object}

For (100) and (101), the Transfer rule in (103) is applied:

(103) a. $LF_{\Sigma} = LF_1 \times LF_2$
b. LF_{Σ} = TP × VP
c. LF_{Σ} = <DP-subject, verb> × {DP-object}

The semantic interpretation of (103) would be like the one in (104):

(104) a. {∃e ([Agent (e, someone)]$_F$ & Theme (e, object)) | e is a verb}
b. {∃e ([Agent (e, DP-subject)]$_O$ & Theme (e, object)) | e is a verb}

Consider the clause in (105):

(105) Romeyka:

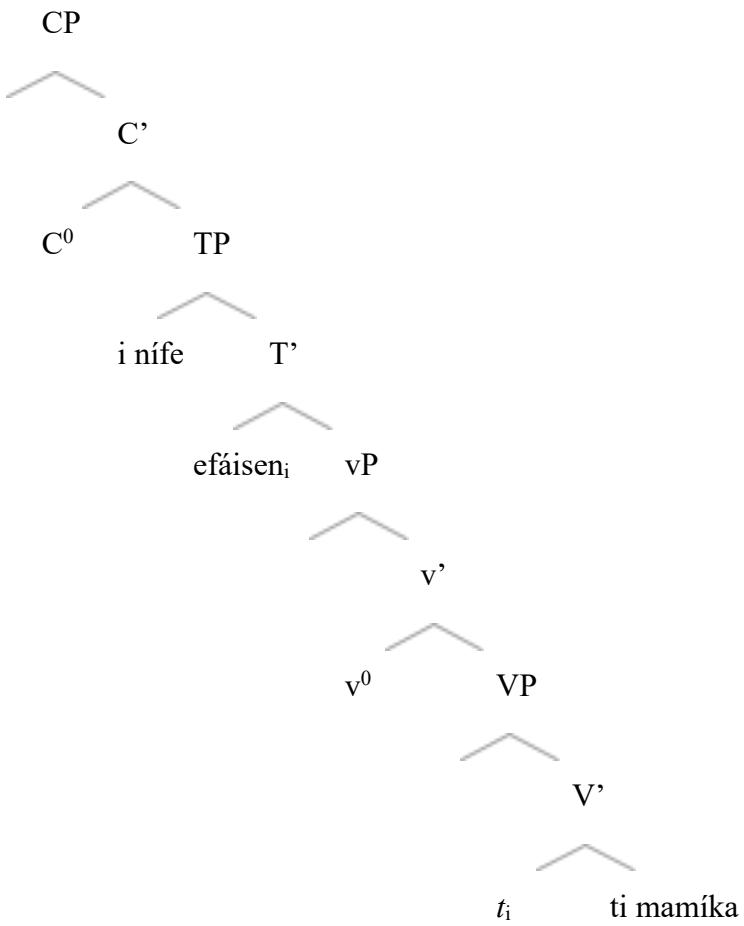
[i	NÍfe]$_{Foc}$	efáisen	ti	mamíka?
the.NOM	daughter-in-law.NOM	feed.Past.3SG	the.ACC	mother-in-law.ACC

'Did the mother-in-law feed her mother-in-law?'
(S01; 150702_0013; 13:53)

In the clause in (105), the verb *efáisen* 'she fed' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the [*u*ϕ] features of T^0, which are inherited by the phase head C^0, resulting in the movement of the DP-subject *i nífe* 'the daughter-in-law' to the [Spec, TP] (see (106) and (107)):

(106) [$_{CP}$ C^0 [$_{TP}$ i nífe efáisen$_i$ [$_{vP}$ [$_{VP}$ t_i ti mamíka]]]]

(107)



For (106) and (107), the Spell-Out rule in (108) is applied:

(108) a. $PF_\Sigma = PF_1 \times PF_2$

b. PF_Σ = TP × VP

c. PF_Σ = <i nífe, efáisen> × {ti mamíka}

For (106) and (107), the Transfer rule in (109) is applied:

(109) a. $LF_\Sigma = LF_1 \times LF_2$

b. LF_Σ = TP × VP

c. LF_Σ = <i nífe, efáisen> × {ti mamíka}

The semantic interpretation of (109) would be like the one in (110):

(110) a. {∃e ([Agent (e, someone)]$_F$ & Theme (e, ti mamíka)) | e is *efáisen*}

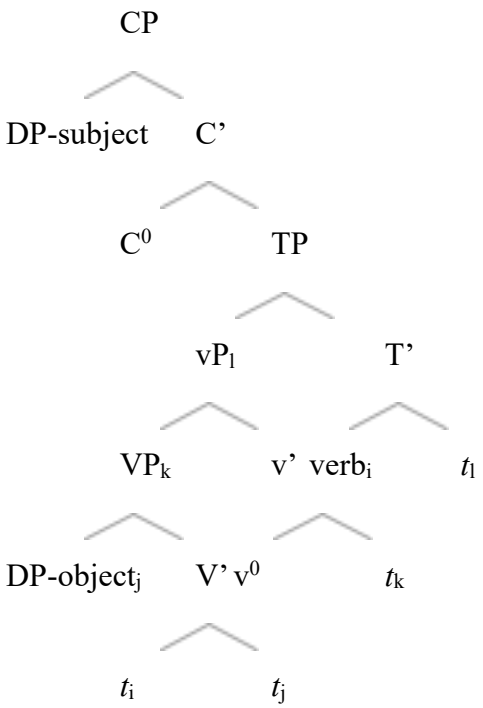
b. {∃e ([Agent (e, i nífe)]$_O$ & Theme (e, ti mamíka)) | e is *efáisen*}

3.1.2 SOV

First, when the DP-object is focused in an SOV yes/no question, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-object to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (111) and (112)):

(111) $[_{CP}$ DP-subject C^0 $[_{TP}$ $[_{vP}$ $[_{VP}$ DP-object$_j$ t_i $t_j]]$ verb$_i]]$

(112)



For (111) and (112), the Spell-out rule in (113) is applied:

(113) a. $PF_\Sigma = PF_1 \times PF_2$

b. $PF_\Sigma = CP \times TP$

c. PF_Σ = {DP-subject} × <DP-object, verb>

For (111) and (112), the Transfer rule in (114) is applied:

(114) a. $LF_{\Sigma} = LF_1 \times LF_2$

b. LF_{Σ} = CP × TP

c. LF_{Σ} = {DP-subject} × <DP-object, verb>

The semantic interpretation of (114) would be like the one in (115):

(115) a. {∃e (Agent (e, DP-subject) & [Theme (e, something)]$_F$) | e is a verb}

b. {∃e (Agent (e, DP-subject) & [Theme (e, object)]$_O$) | e is a verb}

Consider the clause in (116):

(116) Romeyka:

esís	[ta	tsuPÁðæ]$_{Foc}$	θerízete?
you.NOM	the.ACC	corn.ACC	harvest.2PL

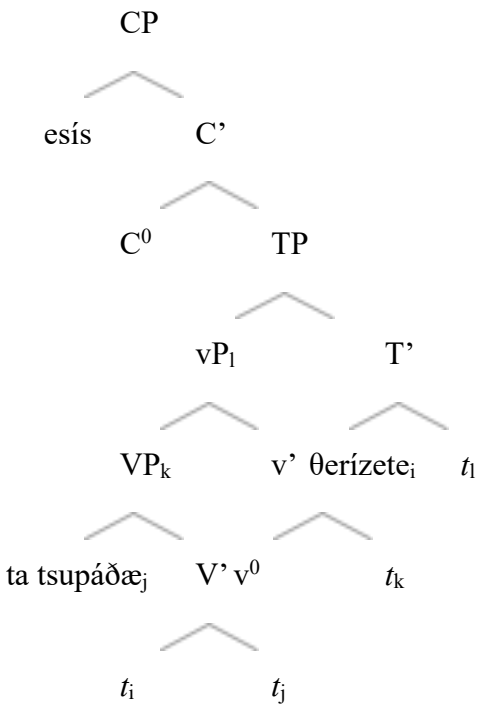
'Do you harvest the corn?'

(S07; 812_0067; 01:58)

In the clause in (116), the verb *θerízete* 'you harvest' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-object *ta tsupáðæ* 'the corn' to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (117) and (118)):

(117) [$_{CP}$ esís C^0 [$_{TP}$ [$_{vP}$ [$_{VP}$ ta tsupáðæ$_j$ t_i t_j]] θerízete$_i$]]

(118)



For (117) and (118), the Spell-out rule in (119) is applied:

(119) a. $PF_{\Sigma} = PF_1 \times PF_2$

b. PF_{Σ} = CP × TP

c. PF_{Σ} = {esís} × <ta tsupáðæ, θerízete>

For (117) and (118), the Transfer rule in (120) is applied:

(120) a. $LF_{\Sigma} = LF_1 \times LF_2$

b. LF_{Σ} = CP × TP

c. LF_{Σ} = {esís} × <ta tsupáðæ, θerízete>

The semantic interpretation of (120) would be like the one in (121):

(121) a, {∃e (Agent (e, esís) & [Theme (e, something)]$_F$) | e is *θerízete*}

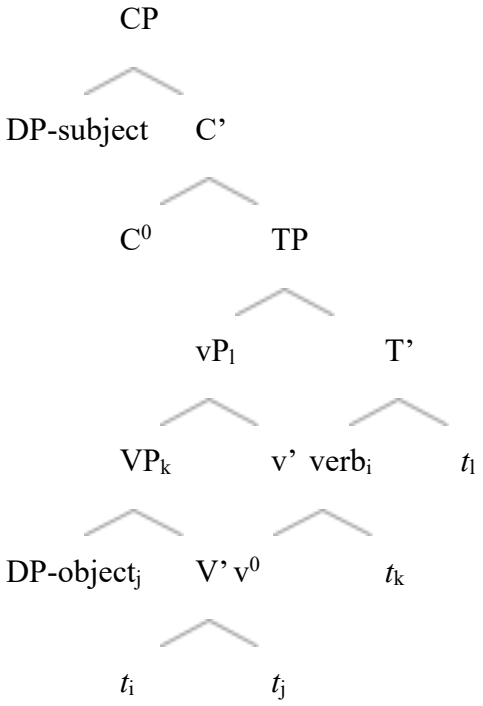
b. {∃e (Agent (e, esís) & [Theme (e, ta tsupáðæ)]$_O$) | e is *θerízete*}

Second, when both the DP-subject and the DP-object are focused in an SOV yes/no question, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-

object to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (122) and (123)):

(122) $[_{CP}$ DP-subject C^0 $[_{TP}$ $[_{vP}$ $[_{VP}$ DP-object$_j$ t_i $t_j]]$ verb$_i]]$

(123)



For (122) and (123), the Spell-out rule in (124) is applied:

(124) a. $PF_\Sigma = PF_1 \times PF_2$

b. $PF_\Sigma = CP \times TP$

c. PF_Σ = {DP-subject} × <DP-object, verb>

For (122) and (123), the Transfer rule in (125) is applied:

(125) a. $LF_\Sigma = LF_1 \times LF_2$

b. $LF_\Sigma = CP \times TP$

c. LF_Σ = {DP-subject} × <DP-object, verb>

The semantic interpretation in (125) would be like the one in (126):

(126) a. {∃e [(Agent (e, someone)]$_F$ & [Theme (e, something)]$_F$) | e is a verb}

b. {∃e [(Agent (e, DP-subject)]$_O$ & [Theme (e, DP-object)]$_O$) | e is a verb}

Consider the clause in (127):

(127) Romeyka:

[eSÍS]$_{Foc}$ [ta tsuPÁðæ]$_{Foc}$ θerízete?

you.NOM the.ACCcorn.ACC harvest.2PL

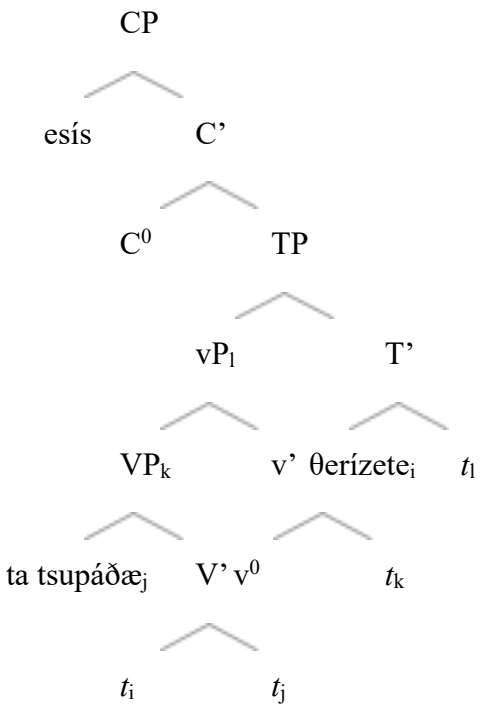
'Do you harvest the corn?'

(S07; 812_0067; 02:14)

In the clause in (127), the verb *θerízete* 'you harvest' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-object *ta tsupáðæ* 'the corn' to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (128) and (129)):

(128) [$_{CP}$ esís C^0 [$_{TP}$ [$_{vP}$ [$_{VP}$ ta tsupáðæ$_j$ t_i t_j]] θerízete$_i$]]

(129)



For (128) and (129), the Spell-out rule in (130) is applied:

(130) a. $PF_{\Sigma} = PF_1 \times PF_2$

b. $PF_{\Sigma} = CP \times TP$

c. PF_{Σ} = {esís} × <ta tsupáðæ, θerízete>

For (128) and (129), the Transfer rule in (131) is applied:

(131) a. $LF_{\Sigma} = LF_1 \times LF_2$

b. $LF_{\Sigma} = CP \times TP$

c. LF_{Σ} = {esís} × <ta tsupáðæ, θerízete>

The semantic interpretation in (131) would be like the one in (132):

(132) a. {∃e [(Agent (e, someone)] & [Theme (e, something)]$_F$) | e is *θerízete*}

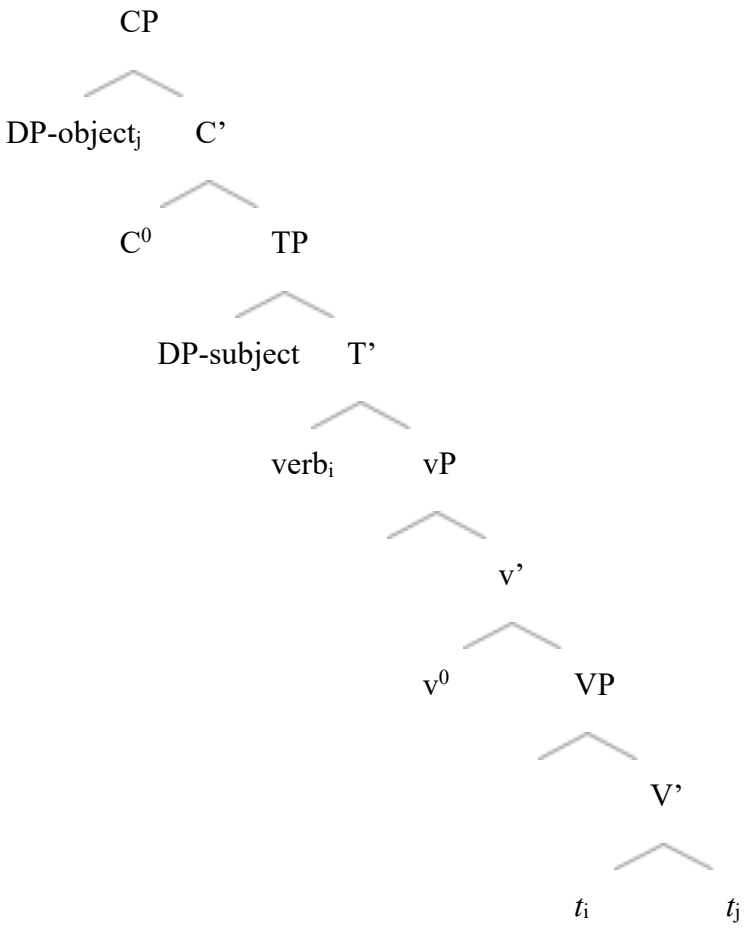
b. {∃e [(Agent (e, esís)] & [Theme (e, ta tsupáðæ)]$_O$) | e is *θerízete*}

3.1.3 OSV

When the DP-subject is focused and the DP-object is topicalised in an OSV yes/no question, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the [*u*ϕ] features of T^0, which are inherited by the phase head C^0, resulting in the movement of the DP-subject to the [Spec, TP]. Moreover, a ^ feature is associated with the EF of the phase head C^0, giving rise to the left dislocation of the DP-object (see (133) and (134)):

(133) $[_{CP}$ DP-object$_j$ C^0 $[_{TP}$ DP-subject verb$_i$ $[_{vP}$ $[_{VP}$ t_i $t_j]]]]$

(134)



For the (133) and (134), the Spell-Out rule (135) is applied:

(135) a. $PF_\Sigma = PF_1 \times PF_2$

b. PF_Σ = CP × TP

c. PF_Σ = {DP-object} × <DP-subject, verb>

For the (133) and (134), the Transfer rule (136) is applied:

(136) a. $LF_{\Sigma} = LF_1 \times LF_2$

b. LF_{Σ} = CP × TP

c. LF_{Σ} = {DP-object} × <DP-subject, verb>

The semantic interpretation of (136) would be like the one in (137):

(137) a. {∃e (Theme (e, DP-object) & [Agent (e, someone)]$_F$) | e is a verb}

b. {∃e (Theme (e, DP-object) & [Agent (e, DP-subject)]$_O$) | e is a verb}

Consider the clause in (138):

(138) Romeyka:

ató	[o	meHMÉtis]$_{Foc}$	éndžen	a?
this.ACC	the.NOM	Mehmetis.NOM	bring.Past.3SG	it.ACC

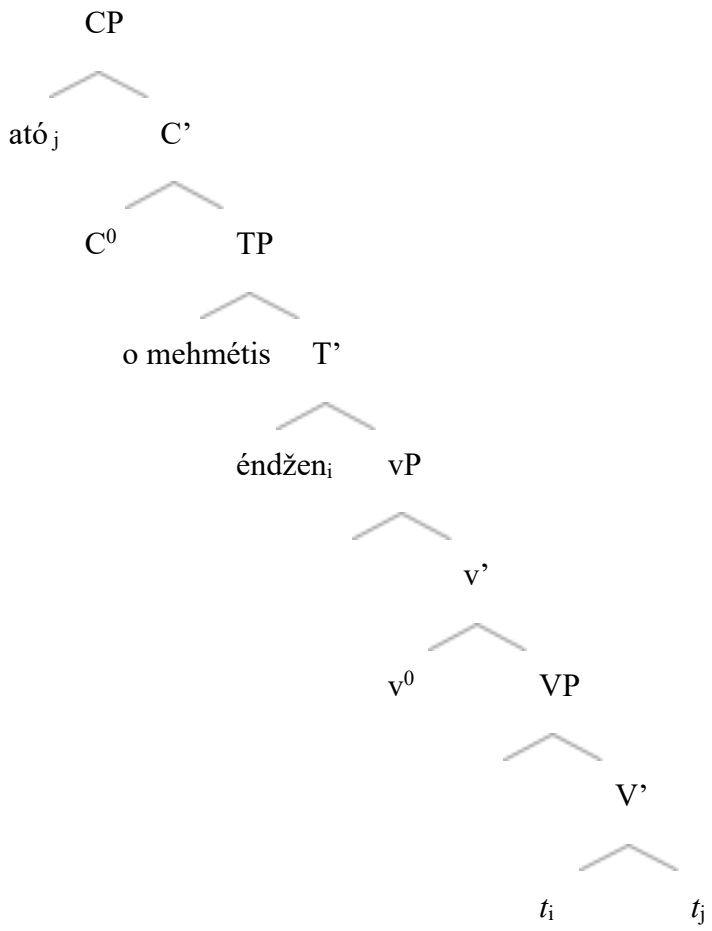
'Did Mehmetis bring that?'

(S01; 150703_0042; 06:36)

In the clause in (138), the verb *éndžen* 'he brought' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the [*u*ϕ] features of T^0, which are inherited by the phase head C^0, resulting in the movement of the DP-subject *o mehmétis* 'Mehmetis' to the [Spec, TP]. Moreover, a ^ feature is associated with the EF of the phase head C^0, giving rise to the left dislocation of the DP-object *ató* 'this' (see (139) and (140)):

(139) [$_{CP}$ ató$_j$ C^0 [$_{TP}$ kanís éndžen$_i$ [$_{vP}$ [$_{VP}$ t_i t_j]]]]

(140)



For the (139) and (140), the Spell-Out rule (141) is applied:

(141) a. $PF_{\Sigma} = PF_1 \times PF_2$

b. PF_{Σ} = CP × TP

c. PF_{Σ} = {ató} × <o mehmétis, éndžen>

For the (139) and (140), the Transfer rule (142) is applied:

(142) a. $LF_{\Sigma} = LF_1 \times LF_2$

b. LF_{Σ} = CP × TP

c. LF_{Σ} = {ató} × <o mehmétis, éndžen>

The semantic interpretation of (142) would be like the one in (143):

(143) a. {∃e (Theme (e, ató) & [Agent (e, someone)]$_F$) | e is *éndžen*}

b. {∃e (Theme (e, ató) & [Agent (e, o mehmétis)]$_O$) | e is *éndžen*}

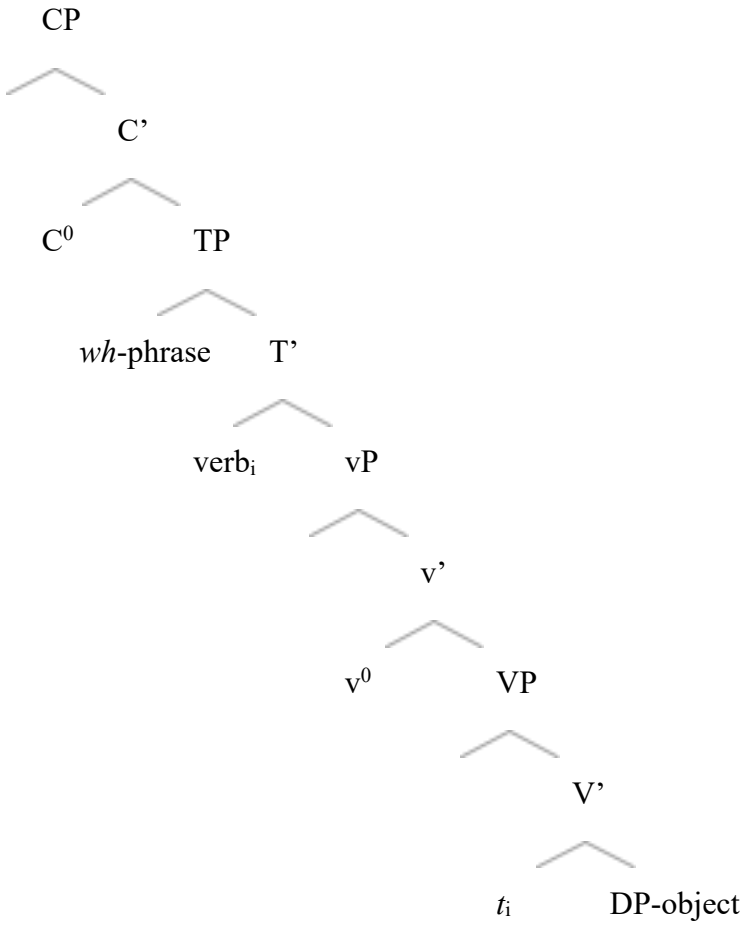
3.2 Direct *wh*-questions in Romeyka

3.2.1 SVO

When the subject is a *wh*-phrase in an SVO *wh*-question, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the [*u*ϕ] features of T^0, which are inherited by the phase head C^0, resulting in the movement of the *wh*-phrase to the [Spec, TP] (see (144) and (145)):

(144) $[_{CP}\ C^0\ [_{TP}$ *wh*-phrase verb$_i$ $[_{vP}\ [_{VP}\ t_i$ DP-object]]]]

(145)



For (144) and (145), the Spell-Out rule in (146) is applied:

(146) a. $PF_\Sigma = PF_1 \times PF_2$

b. $PF_\Sigma = TP \times VP$

c. PF_Σ = <*wh*-phrase, verb> × {DP-object}

For (144) and (145), the Transfer rule in (147) is applied:

(147) a. $LF_{\Sigma} = LF_1 \times LF_2$

b. LF_{Σ} = TP × VP

c. LF_{Σ} = <*wh*-phrase, verb> × {DP-object}

The semantic interpretation of (147) would be like the one in (148):

(148) {∃e ([Agent (e, *wh*-phrase)]$_F$ & Theme (e, DP-object)) | e is a verb}

Consider the clause in (149):

(149) Romeyka:

[PÍos]$_{wh}$ eðótšen tin kos:áran?

who.NOM give.Past.3SG the.ACChen.ACC

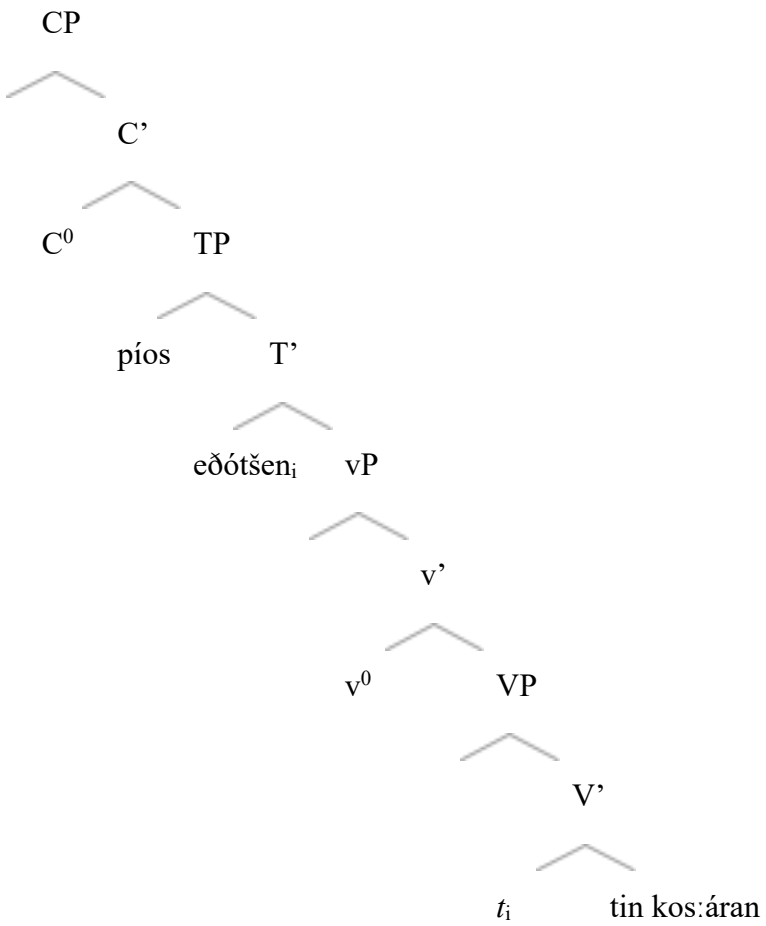
'Who gave the hen?

(S01; 812_0093; 00:03)

In the clause in (149), the verb *eðótšen* 's/he gave' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the [*u*ϕ] features of T^0, which are inherited by the phase head C^0, resulting in the movement of the *wh*-phrase *píos* 'who' to the [Spec, TP] (see (150) and (151)):

(150) [$_{CP}$ C^0 [$_{TP}$ píos eðótšen$_i$ [$_{vP}$ [$_{VP}$ t_i tin kos:áran]]]]

(151)



For (150) and (151), the Spell-Out rule in (152) is applied:

(152) a. $PF_\Sigma = PF_1 \times PF_2$

b. $PF_\Sigma = TP \times VP$

c. PF_Σ = <píos, eðótšen> × {tin kos:áran}

For (150) and (151), the Transfer rule in (153) is applied:

(153) a. $LF_\Sigma = LF_1 \times LF_2$

b. $LF_\Sigma = TP \times VP$

c. LF_Σ = <píos, eðótšen> × {tin kos:áran}

The semantic interpretation of (153) would be like the one in (154):

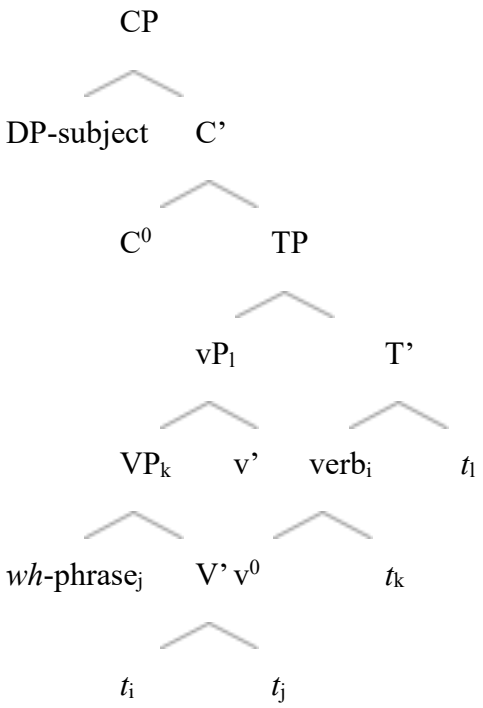
(154) {∃e ([Agent (e, píos)]$_F$ & Theme (e, tin kos:áran)) | e is *eðótšen*}

3.2.2 SOV

First, when the object is a *wh*-phrase in an SOV *wh*-question, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the *wh*-phrase to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (155) and (156)):

(155) $[_{CP}$ DP-subject C^0 $[_{TP}$ $[_{vP}$ $[_{VP}$ *wh*-phrase$_j$ t_i $t_j]]$ verb$_i]]$

(156)



For (155) and (156), the Spell-out rule in (157) is applied:

(157) a. $PF_\Sigma = PF_1 \times PF_2$

b. $PF_\Sigma = CP \times TP$

c. PF_Σ = {DP-subject} × <*wh*-phrase, verb>

For (155) and (156), the Transfer rule in (158) is applied:

(158) a. $LF_{\Sigma} = LF_1 \times LF_2$

b. LF_{Σ} = CP × TP

c. LF_{Σ} = {DP-subject} × <*wh*-phrase, verb>

The semantic interpretation of (158) would be like the one in (159):

(159) {∃e (Agent (e, DP-subject) & [Theme (e, *wh*-phrase)]$_F$) | e is a verb}

Consider the clause in (160):

(160) Romeyka:

alís [DÓɣna]$_{wh}$ aɣórasen?

Alis.NOM what.ACC buy.Past.3SG

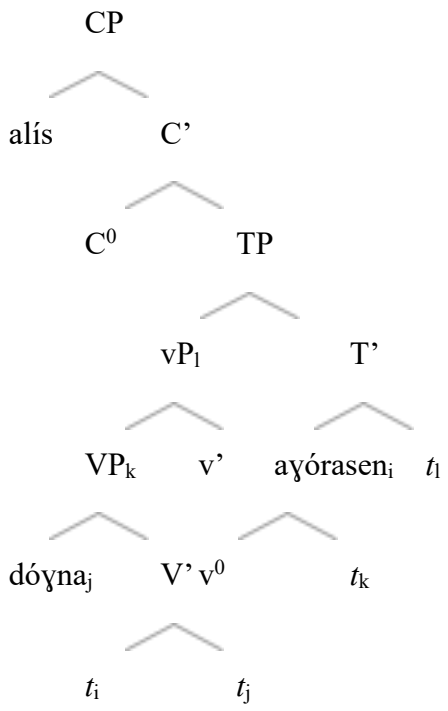
'What did Alis buy?'

(S01; 812_0056; 04:13)

In the clause in (160), the verb *aɣórasen* 'he bought' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the *wh*-phrase *dóɣna* 'what' to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (161) and (162)):

(161) [$_{CP}$ alís C^0 [$_{TP}$ [$_{vP}$ [$_{VP}$ dóɣna$_j$ t_i t_j]] aɣórasen$_i$]]

(162)



For (161) and (162), the Spell-out rule in (163) is applied:

(163) a. $PF_{\Sigma} = PF_1 \times PF_2$

b. $PF_{\Sigma} = CP \times TP$

c. PF_{Σ} = {alís} × <dóɣna, aɣórasen>

For (161) and (162), the Transfer rule in (164) is applied:

(164) a. $LF_{\Sigma} = LF_1 \times LF_2$

b. $LF_{\Sigma} = CP \times TP$

c. LF_{Σ} = {alís} × <dóɣna, aɣórasen>

The semantic interpretation of (164) would be like the one in (165):

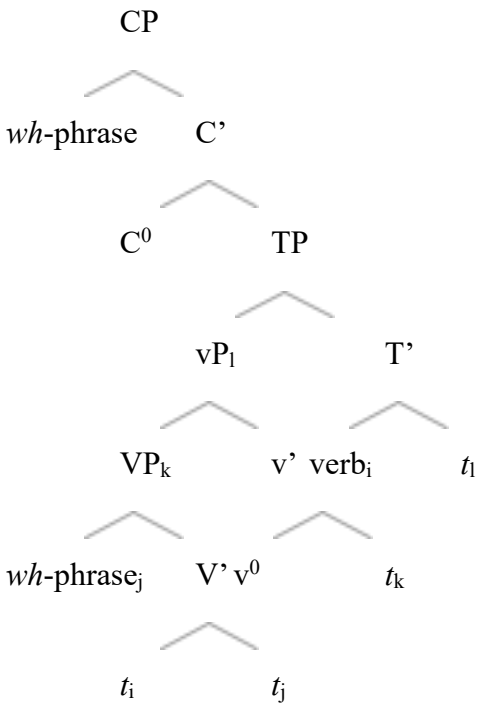
(165) {∃e (Agent (e, alís) & [Theme (e, dóɣna)]$_F$) | e is *aɣórasen*}

Second, when both the subject and the object are *wh*-phrases in an SOV *wh*-question, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the *wh*-phrase ‘whom’ to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the

[Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (166) and (167)):

(166) $[_{CP}$ *wh*-phrase C^0 $[_{TP}$ $[_{vP}$ $[_{VP}$ *wh*-phrase$_j$ t_i $t_j]]$ verb$_i]]$

(167)



For (166) and (167), the Spell-out rule in (168) is applied:

(168) a. $PF_\Sigma = PF_1 \times PF_2$

b. $PF_\Sigma = CP \times DP$

c. PF_Σ = {*wh*-phrase} × <*wh*-phrase, verb>

For (166) and (167), the Transfer rule in (169) is applied:

(169) a. $LF_\Sigma = LF_1 \times LF_2$

b. $LF_\Sigma = CP \times TP$

c. LF_Σ = {*wh*-phrase} × <*wh*-phrase, verb>

The semantic interpretation of (169) would be like the one in (170):

(170) {∃e ([Agent (e, *wh*-phrase)]$_F$ & [Theme (e, *wh*-phrase)]$_F$) | e is a verb}

Consider the clause in (171):

(171) Romeyka:

[PÍos]$_{wh}$ [DÓɣna]$_{wh}$ aɣórasen?

who.NOM what.ACC buy.Past.3SG

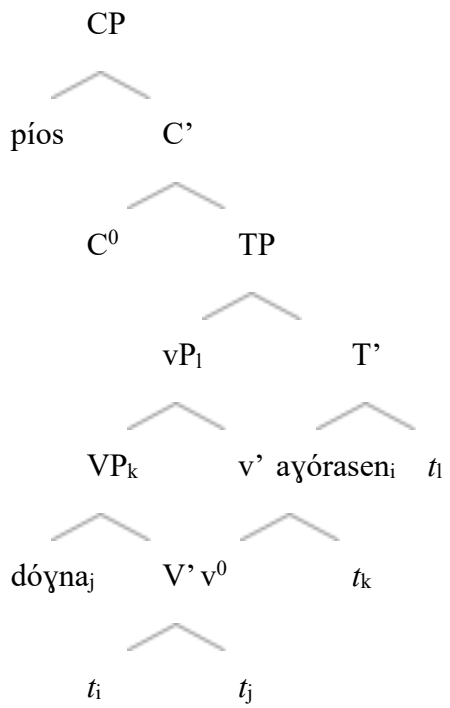
'Who bought what?'

(S01; 150703_0042; 07:37)

In the clause in (171), the verb *aɣórasen* 'bought' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the *wh*-phrase *dóɣna* 'what' to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (172) and (173)):

(172) [$_{CP}$ píos C^0 [$_{TP}$ [$_{vP}$ [$_{VP}$ dóɣna$_j$ t_i t_j]] aɣórasen$_i$]]

(173)



For (172) and (173), the Spell-out rule in (174) is applied:

(174) a. $PF_{\Sigma} = PF_1 \times PF_2$

b. PF_{Σ} = CP × DP

c. PF_{Σ} = {píos} × <dóɣna, aɣórasen>

For (172) and (173), the Transfer rule in (175) is applied:

(175) a. $LF_{\Sigma} = LF_1 \times LF_2$

b. LF_{Σ} = CP × TP

c. LF_{Σ} = {píos} × <dóɣna, aɣórasen>

The semantic interpretation of (175) would be like the one in (176):

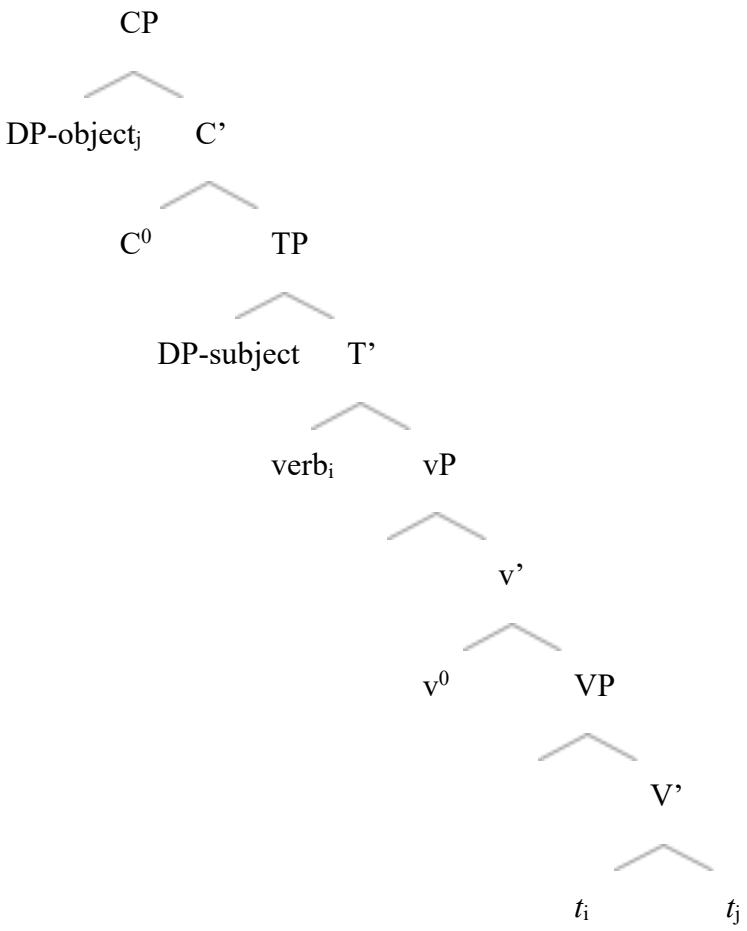
(176) $\{\exists e\ ([\text{Agent}\ (e, \text{píos})]_F\ \&\ [\text{Theme}\ (e, \text{dóɣna})]_F) \mid e \text{ is } \textit{aɣórasen}\}$

3.2.3 OSV

When the subject is a *wh*-phrase and the DP-object is tropicalised in an OSV *wh*-question, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the [*u*ϕ] features of T^0, which are inherited by the phase head C^0, resulting in the movement of the *wh*-phrase to the [Spec, TP]. Moreover, a ^ feature is associated with the EF of the phase head C^0, giving rise to the left dislocation of the DP-object (see (177) and (178)):

(177) $[_{CP}$ DP-object$_j$ C^0 $[_{TP}$ *wh*-phrase verb$_i$ $[_{vP}$ $[_{VP}$ t_i $t_j]]]]$

(178)



For (177) and (178), the Spell-Out rule in (179) is applied:

(179) a. $PF_\Sigma = PF_1 \times PF_2$

b. PF_Σ = CP × TP

c. PF_Σ = {DP-object} × <*wh*-phrase, verb>

For (177) and (178), the Transfer rule in (180) is applied:

(180) a. $LF_\Sigma = LF_1 \times LF_2$

b. LF_Σ = CP × TP

c. LF_Σ = {DP-object} × <*wh*-phrase, verb>

The semantic interpretation of (180) would be like the one in (181):

(181) {∃e ([Agent (e, *wh*-phrase)]$_F$ & Theme (e, DP-object)) | e is a verb}

Consider the clause in (182):

(182) Romeyka:

χavítsæ [PÍos]$_{wh}$ éfaen?

anchovies.ACC who.NOM eat.Past.3SG

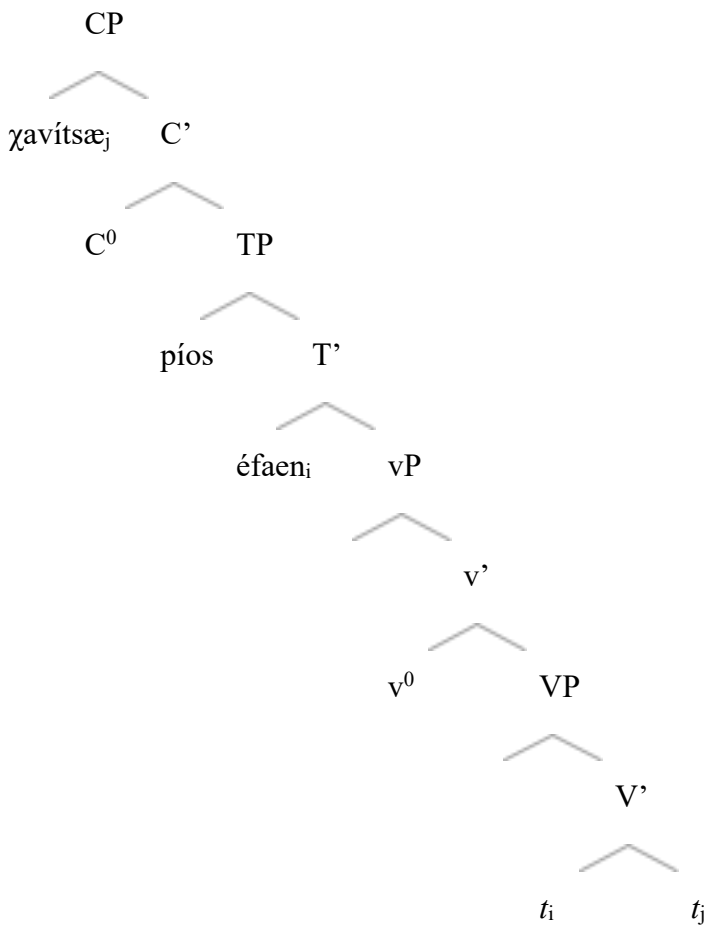
'Who ate anchovies?'

(S01; 812_0057; 04:06)

In the clause in (182), the verb *éfaen* 'ate' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the [*u*ϕ] features of T^0, which are inherited by the phase head C^0, resulting in the movement of the *wh*-phrase to the [Spec, TP]. Moreover, a ^ feature is associated with the EF of the phase head C^0, giving rise to the left dislocation of the DP-object (see (183) and (184)):

(183) [$_{CP}$ χavítsæ$_j$ C^0 [$_{TP}$ píos éfaen$_i$ [$_{vP}$ [$_{VP}$ t_i t_j]]]]

(184)



For (183) and (184), the Spell-Out rule in (185) is applied:

(185) a. $PF_{\Sigma} = PF_1 \times PF_2$

b. PF_{Σ} = CP × TP

c. PF_{Σ} = {χavítsæ} × <píos, éfaen>

For (183) and (184), the Transfer rule in (186) is applied:

(186) a. $LF_{\Sigma} = LF_1 \times LF_2$

b. LF_{Σ} = CP × TP

c. LF_{Σ} = {χavítsæ} × <píos, éfaen>

The semantic interpretation of (186) would be like the one in (187):

(187) {∃e ([Agent (e, píos)]$_F$ & Theme (e, χavítsæ)) | e is *éfaen*}

4 Indirect questions in Romeyka

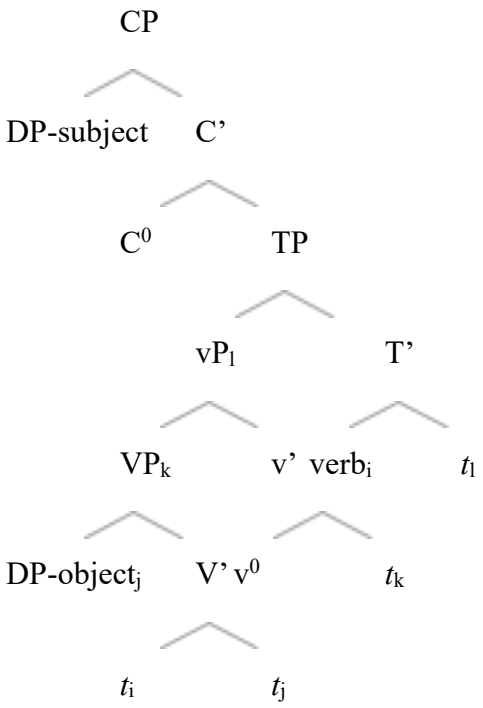
4.1 Indirect yes/no questions in Romeyka

4.1.1 SOV

First, when the DP-object is focused in an SOV yes/no question, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-object to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (188) and (189)):

(188) $[_{CP}$ DP-subject C^0 $[_{TP}$ $[_{vP}$ $[_{VP}$ DP-object$_j$ t_i $t_j]]$ verb$_i]]$

(189)



For (188) and (189), the Spell-out rule in (190) is applied:

(190) a. $PF_\Sigma = PF_1 \times PF_2$

b. $PF_\Sigma = CP \times TP$

c. PF_Σ = {DP-subject} × <DP-object, verb>

For (188) and (189), the Transfer rule in (191) is applied:

(191) a. $LF_{\Sigma} = LF_1 \times LF_2$

b. LF_{Σ} = CP × TP

c. LF_{Σ} = {DP-subject} × <DP-object, verb>

The semantic interpretation of (191) would be like the one in (192):

(192) a. {∃e (Agent (e, DP-subject) & [Theme (e, DP-object)]$_F$) | e is a verb}

b. {∃e (Agent (e, DP-subject) & [Theme (e, DP-object)]$_O$) | e is a verb}

Consider the clause in (193):

(193) Romeyka:

rotás	me,	alís	[tin	aiŠÉN]$_{Foc}$	efílisen?
ask.NOM.2SG	I.ACC	Alis.NOM	the.ACC	Ayşe	kiss.Past.3SG

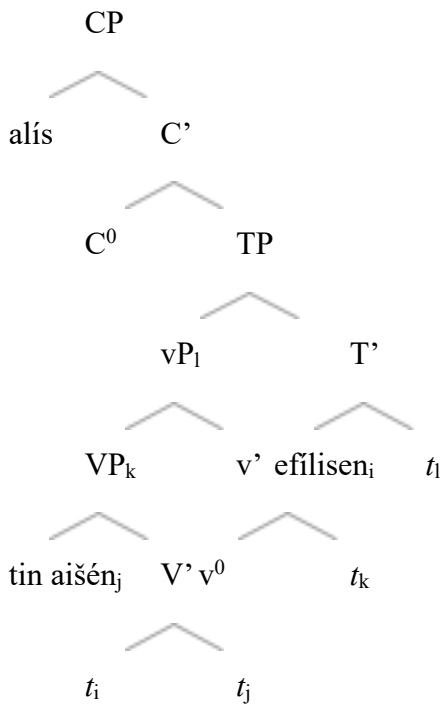
'You ask me, did Alis kiss Ayşe?'

(S01; 150702_0022; 03:16)

In the clause in (193), the verb *efílisen* 'he kissed' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-object *tin aišén* 'Ayşe' to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (194) and (195)):

(194) [$_{CP}$ alís C^0 [$_{TP}$ [$_{vP}$ [$_{VP}$ tin aišén$_j$ t_i t_j]] efílisen$_i$]]

(195)



For (194) and (195), the Spell-out rule in (196) is applied:

(196) a. $PF_{\Sigma} = PF_1 \times PF_2$

b. $PF_{\Sigma} = CP \times TP$

c. PF_{Σ} = {alís} × <tin aišén, efílisen>

For (194) and (195), the Transfer rule in (197) is applied:

(197) a. $LF_{\Sigma} = LF_1 \times LF_2$

b. $LF_{\Sigma} = CP \times TP$

c. LF_{Σ} = {alís} × <tin aišén, efílisen>

The semantic interpretation of (197) would be like the one in (198):

(198) a. {∃e (Agent (e, alís) & [Theme (e, someone)]$_F$) | e is *efílisen*}

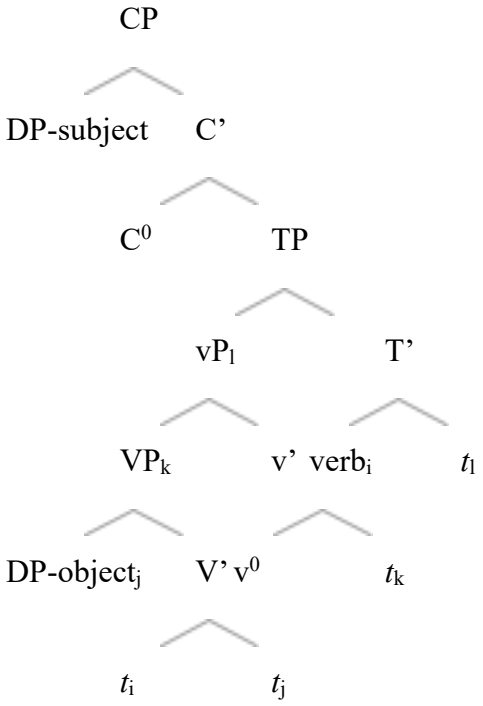
b. {∃e (Agent (e, alís) & [Theme (e, tin aišén)]$_O$) | e is *efílisen*}

Second, when both the DP-subject and the DP-object are focused in an SOV yes/no question, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-

object to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (199) and (200)):

(199) $[_{CP}$ DP-subject C^0 $[_{TP}$ $[_{vP}$ $[_{VP}$ DP-object$_j$ t_i $t_j]]$ verb$_i]]$

(200)



For (199) and (200), the Spell-out rule in (201) is applied:

(201) a. $PF_\Sigma = PF_1 \times PF_2$

b. $PF_\Sigma = CP \times TP$

c. PF_Σ = {DP-subject} × <DP-object, verb>

For (199) and (200), the Transfer rule in (202) is applied:

(202) a. $LF_\Sigma = LF_1 \times LF_2$

b. $LF_\Sigma = CP \times TP$

c. LF_Σ = {DP-subject} × <DP-object, verb>

The semantic interpretation of (202) would be like the one in (203):

(203) a. {∃e [(Theme (e, someone)]$_F$ & [Agent (e, someone)]$_F$) | e is a verb}

b. {∃e [(Theme (e, DP-subject)]$_O$ & [Agent (e, DP-object)]$_O$) | e is a verb}

Consider the clause in (204):

(204) Romeyka:

rotás	me,	[aLÍS]$_{Foc}$	[tin	aiŠÉN]$_{Foc}$	efílisen?
ask.NOM.2SG	I.ACC	Alis.NOM	the.ACC	Ayşe	kiss.Past.3SG

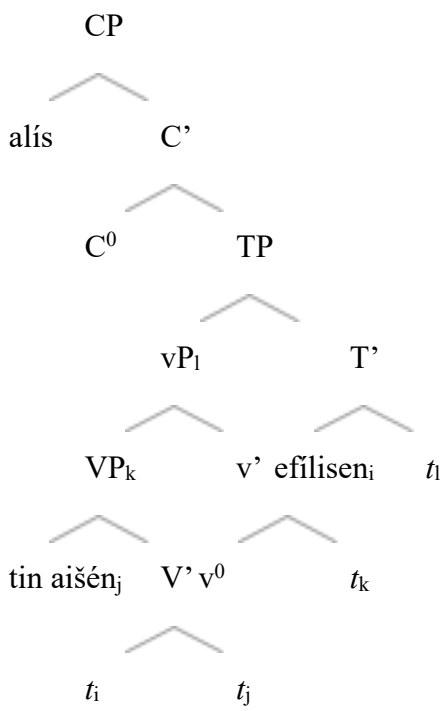
'You ask me, did Alis kiss Ayşe?'

(S01; 150702_0022; 03:52)

In the clause in (204), the verb *efílisen* 'he kissed' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-object *tin aišén* 'Ayşe' to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (205) and (206)):

(205) [$_{CP}$ alís C^0 [$_{TP}$ [$_{vP}$ [$_{VP}$ tin aišén$_j$ t_i t_j]] efílisen$_i$]]

(206)



For (205) and (206), the Spell-out rule in (207) is applied:

(207) a. $PF_{\Sigma} = PF_1 \times PF_2$
b. PF_{Σ} = CP × TP
c. PF_{Σ} = {alís} × <tin aišén, efílisen>

For (205) and (206), the Transfer rule in (208) is applied:

(208) a. $LF_{\Sigma} = LF_1 \times LF_2$
b. LF_{Σ} = CP × TP
c. LF_{Σ} = {alís} × <tin aišén, efílisen>

The semantic interpretation of (208) would be like the one in (209):

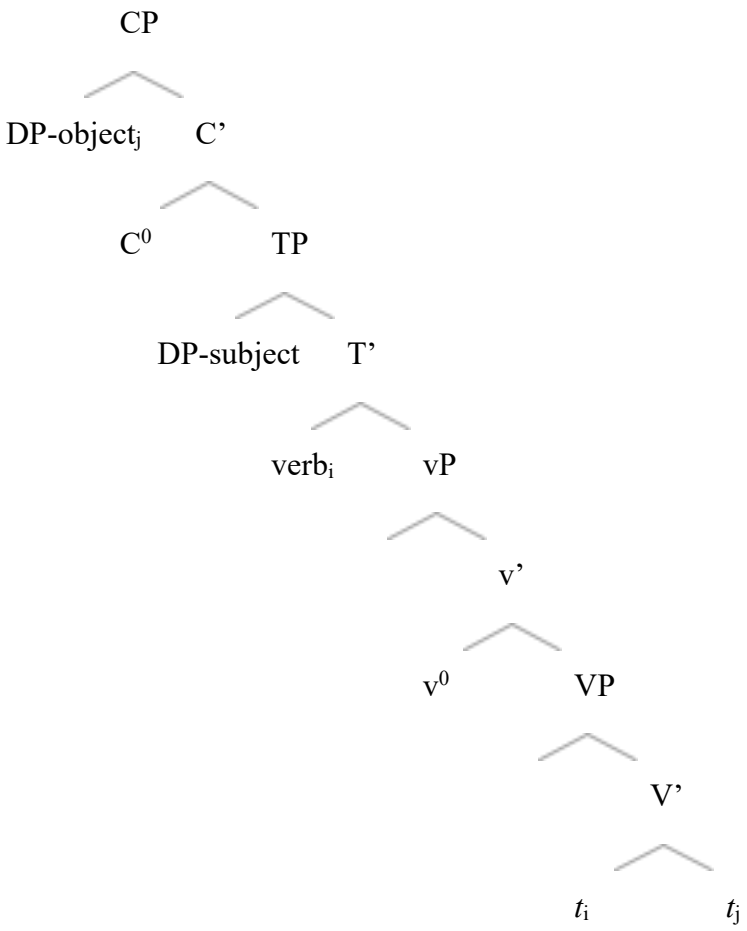
(209) a. {∃e [(Theme (e, someone)]$_F$ & [Agent (e, someone)]$_F$) | e is *efílisen*}
b. {∃e [(Theme (e, alís)]$_O$ & [Agent (e, tin aišén)]$_O$) | e is *efílisen*}

4.1.2 OSV

When the DP-subject is focused and the DP-object is topicalised in an OSV yes/no question, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the [*u*ϕ] features of T^0, which are inherited by the phase head C^0, resulting in the movement of the DP-subject to the [Spec, TP]. Moreover, a ^ feature is associated with the EF of the phase head C^0, giving rise to the left dislocation of the DP-object (see (210) and (211)):

(210) [$_{CP}$ DP-object$_j$ C^0 [$_{TP}$ DP-subject verb$_i$ [$_{vP}$ [$_{VP}$ t_i t_j]]]]

(211)



For (210) and (211), the Spell-Out rule in (212) is applied:

(212) a. $PF_\Sigma = PF_1 \times PF_2$

b. PF_Σ = CP × TP

c. PF_Σ = {DP-object} × <DP-subject, verb>

For (210) and (211), the Transfer rule in (213) is applied:

(213) a. $LF_\Sigma = LF_1 \times LF_2$

b. LF_Σ = CP × TP

c. LF_Σ = {DP-object} × <DP-subject, verb>

The semantic interpretation of (213) would be like the one in (214):

(214) a. {∃e ([Agent (e, someone)]$_F$ & Theme (e, DP-object)) | e is a verb}

b. {∃e ([Agent (e, DP-subject)]$_O$ & Theme (e, DP-object)) | e is a verb}

Consider the clause in (215):

(215) Romeyka:

rotás	me,	to	$c^h it^h$ápin	[aLÍS]$_{Foc}$	eχújepsen?
ask.2SG	I.ACC	the.ACC	book.ACC	Alis.NOM	read.Past.3SG

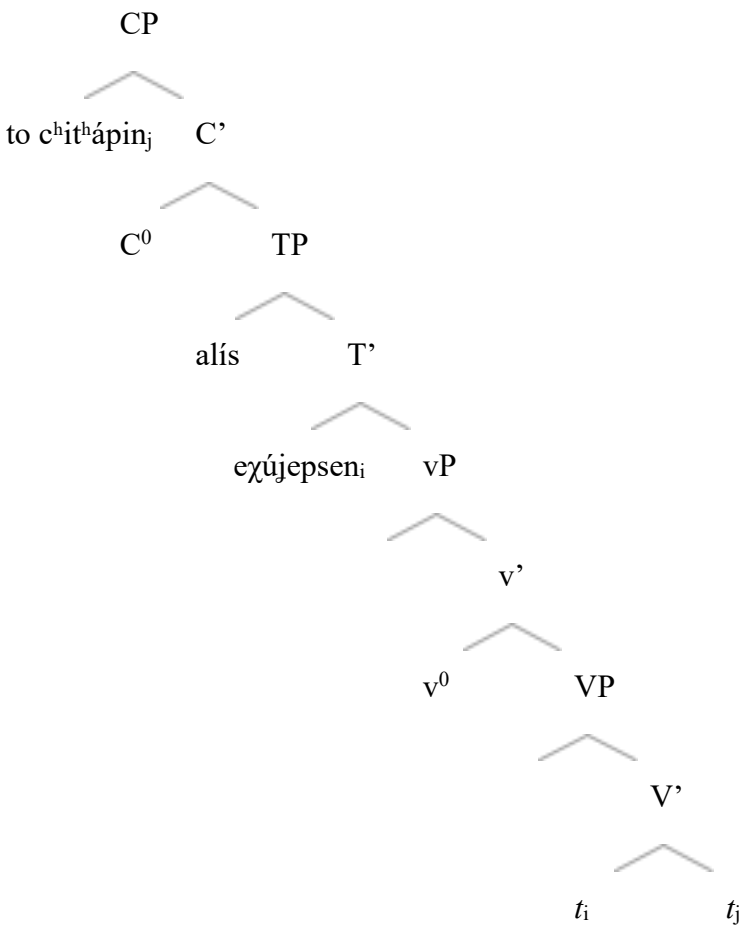
'You ask me, did Alis read the book?'

(S01; 150702_0022; 06:13)

In the clause in (215), the verb *eχújepsen* 'he read' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the [*u*ϕ] features of T^0, which are inherited by the phase head C^0, resulting in the movement of the DP-subject *alís* 'Alis' to the [Spec, TP]. Moreover, a ^ feature is associated with the EF of the phase head C^0, giving rise to the left dislocation of the DP-object *to $c^h it^h$ápin* 'the book' (see (216) and (217)):

(216) [$_{CP}$ to $c^h it^h$ápin$_j$ C^0 [$_{TP}$ alís eχújepsen$_i$ [$_{vP}$ [$_{VP}$ t_i t_j]]]]

(217)



For (216) and (217), the Spell-Out rule in (218) is applied:

(218) a. $PF_{\Sigma} = PF_1 \times PF_2$
b. PF_{Σ} = CP × TP
c. PF_{Σ} = {to c^{h}it^{h}ápin} × <alís, eχújepsen>

For (216) and (217), the Transfer rule in (219) is applied:

(219) a. $LF_{\Sigma} = LF_1 \times LF_2$
b. LF_{Σ} = CP × TP
c. LF_{Σ} = {to c^{h}it^{h}ápin} × <alís, eχújepsen>

The semantic interpretation of (219) would be like the one in (220):

(220) a. {∃e [Agent (e, someone)]$_F$ & (Theme (e, to c^{h}it^{h}ápin)) | e is *eχújepsen*}
b. {∃e [Agent (e, alís)]$_O$ & (Theme (e, to c^{h}it^{h}ápin)) | e is *eχújepsen*}

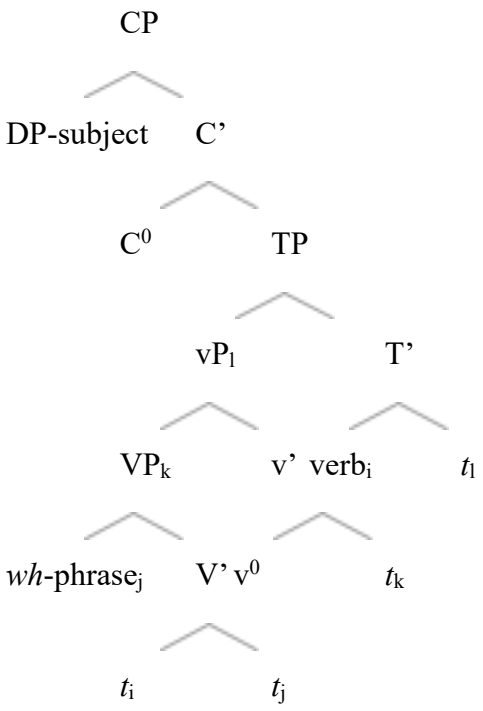
4.2 Indirect *wh*-questions in Romeyka

4.2.1 SOV

First, when the object in an SOV *wh*-question is a *wh*-phrase, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the *wh*-phrase to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (221) and (222)):

(221) [$_{CP}$ DP-subject C^0 [$_{TP}$ [$_{vP}$ [$_{VP}$ *wh*-phrase$_j$ t_i t_j]] verb$_i$]]

(222)



For (221) and (222), the Spell-out rule in (223) is applied:

(223) a. $PF_{\Sigma} = PF_1 \times PF_2$

b. PF_{Σ} = CP × TP

c. PF_{Σ} = {DP-subject} × <*wh*-phrase, verb>

For (221) and (222), the Transfer rule in (224) is applied:

(224) a. $LF_{\Sigma} = LF_1 \times LF_2$

b. LF_{Σ} = CP × TP

c. LF_{Σ} = {DP-subject} × <*wh*-phrase, verb>

The semantic interpretation of the LF_{Σ} in (224) would be like the one in (225):

(225) {∃e (Agent (e, subject) & [Theme (e, *wh*-phrase)]$_{F}$,) | e is a verb}

Consider the clause in (226):

(226) Romeyka:

do	θarís,	alís	[TÍnan]$_{wh}$	efílisen?
what.ACC	think.2SG	Alis.NOM	who.ACC	kiss.Past.3SG

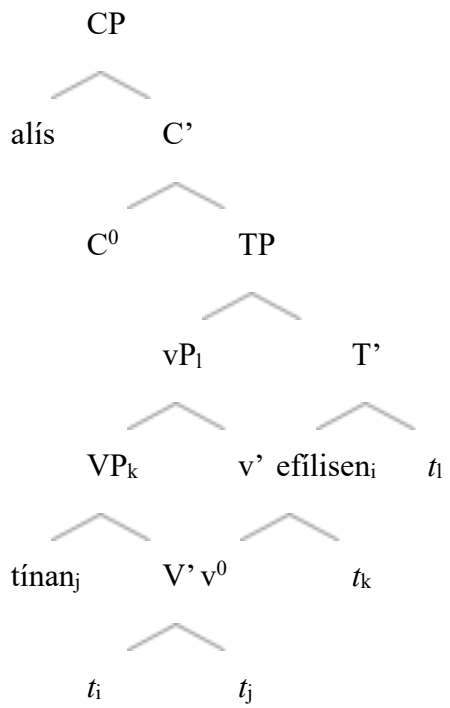
'What do you think, who did Alis kiss?'

(S01; 150703_0040; 19:24)

In the clause in (226), the verb *efílisen* 'he kissed' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the *wh*-phrase *tínan* 'whom' to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (227) and (228)):

(227) [$_{CP}$ alís C^0 [$_{TP}$ [$_{vP}$ [$_{VP}$ tínan$_j$ t_i t_j]] efílisen$_i$]]

(228)



For (227) and (228), the Spell-out rule in (229) is applied:

(229) a. $PF_{\Sigma} = PF_1 \times PF_2$
b. PF_{Σ} = CP × TP
c. PF_{Σ} = {alís} × <tínan, efílisen>

For (227) and (228), the Transfer rule in (230) is applied:

(230) a. $LF_{\Sigma} = LF_1 \times LF_2$
b. LF_{Σ} = CP × TP
c. LF_{Σ} = {alís} × <tínan, efílisen>

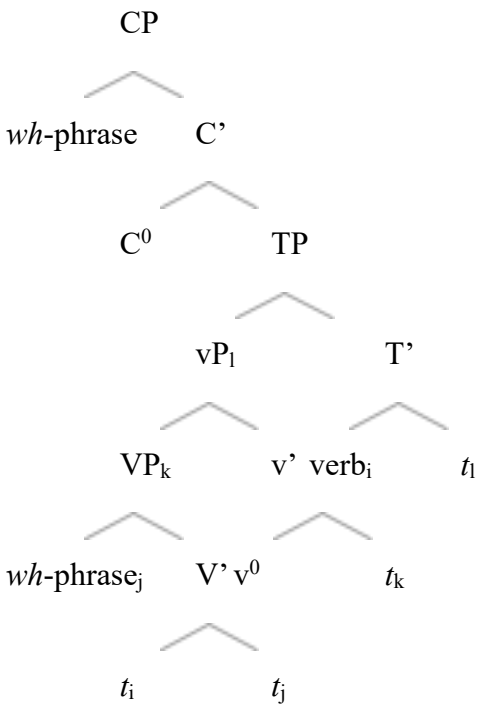
The semantic interpretation of the LF_{Σ} in (230) would be like the one in (231):

(231) {∃e (Agent (e, alís) & [Theme (e, tínan)]$_F$) | e is *efílisen*}

Second, when both the subject and the object are *wh*-phrases in an SOV *wh*-question, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-object to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (232) and (233)):

(232) [$_{CP}$ *wh*-phrase C^0 [$_{TP}$ [$_{vP}$ [$_{VP}$ *wh*-phrase$_j$ t_i t_j]] verb$_i$]]

(233)



For (232) and (233), the Spell-out rule in (234) is applied:

(234) a. $PF_{\Sigma} = PF_1 \times PF_2$

b. $PF_{\Sigma} = CP \times TP$

c. PF_{Σ} = {*wh*-phrase} × <*wh*-phrase, verb>

For (232) and (233), the Transfer rule in (235) is applied:

(235) a. $LF_{\Sigma} = LF_1 \times LF_2$

b. $LF_{\Sigma} = CP \times TP$

c. LF_{Σ} = {*wh*-phrase} × <*wh*-phrase, verb>

The semantic interpretation of (235) would be like the one in (236):

(236) $\{\exists e\ [(\text{Agent}\ (e, \textit{wh}\text{-phrase})]_F\ \&\ [\text{Theme}\ (e, \textit{wh}\text{-phrase})]_F) \mid e \text{ is a verb}\}$

Consider the clause in (237):

(237) Romeyka:

esí	erotás	me,	[PÍos]$_{wh}$	[DÓɣna]$_{wh}$	aɣórasen?
you.NOM	ask.2SG	I.ACC	who.NOM	what.ACC	buy.Past.3SG

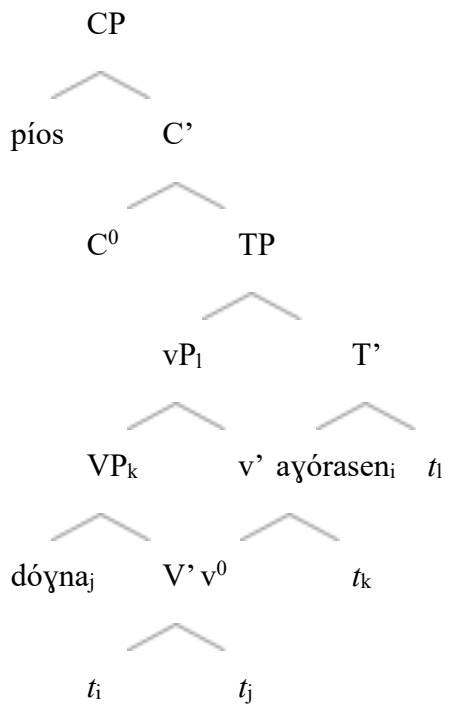
'You ask me, who bought what?'

(S01; 150703_0042; 08:16)

In the clause in (237), the verb *ayórasen* 's/he bought' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the c-selection feature of V^0, which is inherited by the phase head v^0, resulting in the movement of the DP-object *dóyna* 'what' to the [Spec, VP]. The ^ feature is then spread to v^0, triggering movement of the VP to the [Spec, vP]. Finally, the ^ feature is spread to T^0, triggering movement of the vP to the [Spec, TP] (see (238) and (239)):

(238) [$_{CP}$ píos C^0 [$_{TP}$ [$_{vP}$ [$_{VP}$ dóɣna$_j$ t_i t_j]] aɣórasen$_i$]]

(239)



For (238) and (239), the Spell-out rule in (240) is applied:

(240) a. $PF_{\Sigma} = PF_1 \times PF_2$
b. PF_{Σ} = CP × TP
c. PF_{Σ} = {píos} × <dóɣna, aɣórasen>

For (238) and (239), the Transfer rule in (241) is applied:

(241) a. $LF_{\Sigma} = LF_1 \times LF_2$
b. LF_{Σ} = CP × TP
c. LF_{Σ} = {píos} × <dóɣna, aɣórasen>

The semantic interpretation of (241) would be like the one in (242):

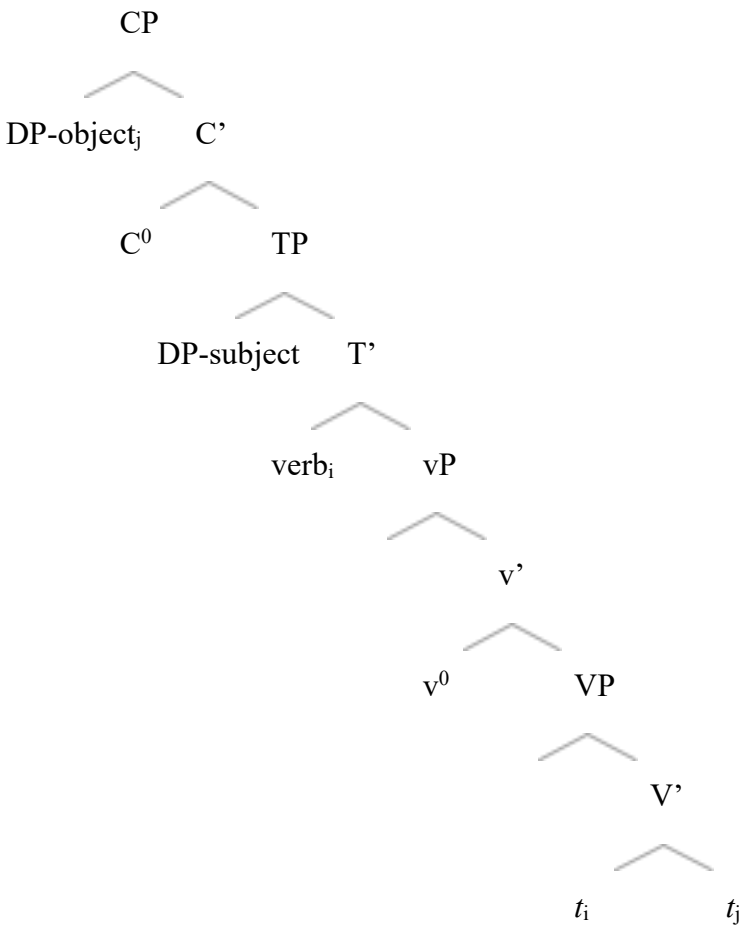
(242) {∃e $[(\text{Agent (e, píos)})]_F$ & $[\text{Theme (e, dóɣna)}]_F$) | e is *aɣórasen*}

4.2.2 OSV

When the DP-subject is focused and the DP-object is topicalised in an OSV *wh*-question, the verb undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the [*u*ϕ] features of T^0, which are inherited by the phase head C^0, resulting in the movement of the DP-subject to the [Spec, TP]. Moreover, a ^ feature is associated with the EF of the phase head C^0, giving rise to the left dislocation of the DP-object (see (243) and (244)):

(243) $[_{CP}$ DP-object$_j$ C^0 $[_{TP}$ DP-subject verb$_i$ $[_{vP}$ $[_{VP}$ t_i $t_j]]]]$

(244)



For (243) and (244), the Spell-Out rule in (245) is applied:

(245) a. $PF_\Sigma = PF_1 \times PF_2$

b. PF_Σ = CP × TP

c. PF_Σ = {DP-object} × <DP-subject, verb>

For (243) and (244), the Transfer rule in (246) is applied:

(246) a. $LF_\Sigma = LF_1 \times LF_2$

b. LF_Σ = CP × TP

c. LF_Σ = {DP-object} × <DP-subject, verb>

The semantic interpretation of (246) would be like the one in (247):

(247) {∃e [Agent (e, *wh*-phrase)]$_F$ & (Theme (e, DP-object)) | e is a verb}

Consider the clause in (248):

(248) Romeyka:

rotás	me,		tin	aišén	[TS]$_{wh}$	epíren?
ask.2SG	I.ACC	the.NOM	the.ACC	Ayşe.ACC	who.NOM	marry.Past.3SG

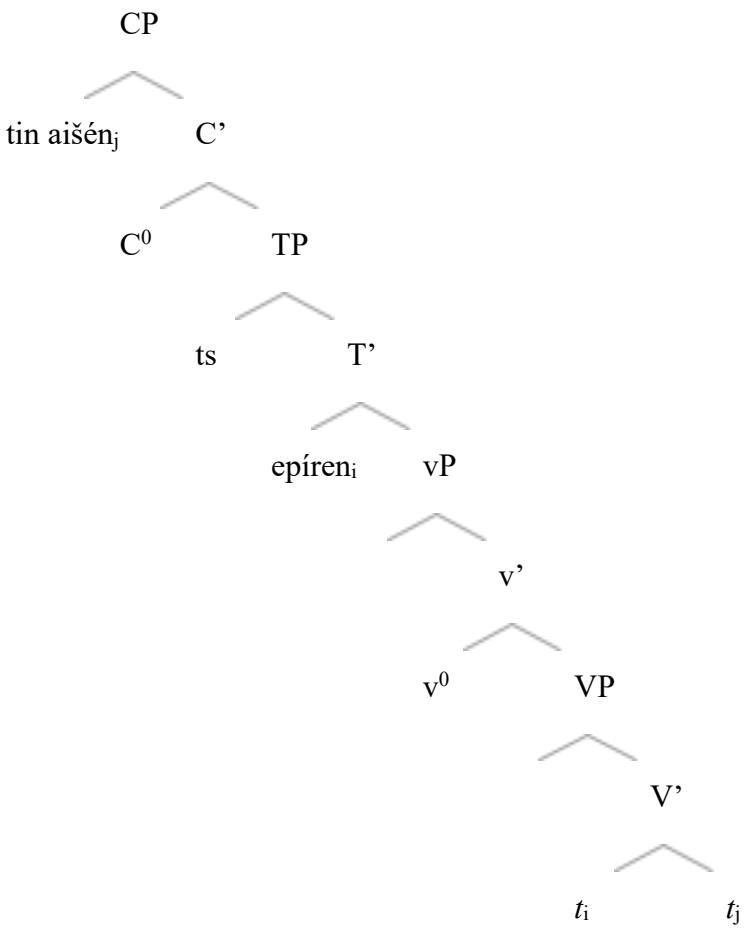
'You ask me, who married Ayşe?'

(S01; 140102_0008; 01:41)

In the clause in (248), the verb *epíren* 'he married' undergoes a V^0-to-T^0 movement, while a ^ feature is associated with the [*u*ϕ] features of T^0, which are inherited by the phase head C^0, resulting in the movement of the DP-subject *ts* 'who' to the [Spec, TP]. Moreover, a ^ feature is associated with the EF of the phase head C^0, giving rise to the left dislocation of the DP-object *tin aišén* 'Ayşe' (see (249) and (250)):

(249) [$_{CP}$ tin aišén$_j$ C^0 [$_{TP}$ ts epíren$_i$ [$_{vP}$ t_i t_j]]]]

(250)



For (249) and (250), the Spell-Out rule in (251) is applied:

(251) a. $PF_{\Sigma} = PF_1 \times PF_2$

b. PF_{Σ} = CP × TP

c. PF_{Σ} = {tin aišén} × <ts, epíren>

For (249) and (250), the Transfer rule in (252) is applied:

(252) a. $LF_{\Sigma} = LF_1 \times LF_2$

b. LF_{Σ} = CP × TP

c. LF_{Σ} = {tin aišén} × <ts, epíren>

The semantic interpretation of (252) would be like the one in (253):

(253) $\{\exists e\ [\text{Agent}\ (e, ts)]_F)$ & (Theme (e, tin aišén) | e is *epíren*}