Chapter 12

Word order variation in Romeyka

Laurentia Schreiber[a] & Mark Janse[b]
[a]University of Bamberg [b]University of Cambridge

The present chapter describes word order variation in Romeyka based on the multilingual spoken language dataset of the Word Order in Western Asia (WOWA) corpus. Descending from VO ancestry, Romeyka shows under contact from Turkish increasingly head-final syntax. While cross-linguistically relevant factors such as semantic role, flagging and phonological weight do not offer much explanation for the flexibility between pre- and post-verbal constituents in Romeyka, information structure and phrase type do seem to be relevant. In addition, inter-speaker variation has been found significant to account for word order variation in Romeyka, suggesting that in a setting of language shift, individual forms of bilingualism affect word order.

1 Introduction

Romeyka is a variety of Pontic Greek (henceforth PG) that is at present still spoken by Muslims in Trabzon Province in northeastern Turkey, although its status can be characterized as endangered (Schreiber 2016; Schreiber & Sitaridou 2017). Romeyka belongs to the Hellenic branch of Indo-European. Like most modern Greek dialects, it is descended from postclassical (Koine) Greek, with an undisputable Ionic substrate dating back to the Milesian colonization of the Black Sea coast in the 7th–6th c. BCE. The closest relatives of Pontic are Pharasiot and Cappadocian, collectively known as (East) Asia Minor Greek (henceforth AMG).[1]

[1]The designation "East Asia Minor Greek" is due to Janse (e.g., 2008: 191–192, 20202020: 202–203) and implied in Dawkins' "Greek of eastern Asia Minor" (1916: 213). The qualification "East" is necessary in light of Ralli's (2020) broader acceptation of the geographical designation "Asia Minor Greek" (Ralli 2020). For a different view on the internal relationship between Pontic, Pharasiot and Cappadocian see Karatsareas (2016: 40–55).

Laurentia Schreiber & Mark Janse. 2024. Word order variation in Romeyka. In Geoffrey Haig, Mohammad Rasekh-Mahand, Donald Stilo, Laurentia Schreiber & Nils N. Schiborr (eds.), *Post-predicate elements in the Western Asian Transition Zone: A corpus-based approach to areal typology*, 337–373. Berlin: Language Science Press. DOI: 10.5281/zenodo.14266353

The location of Romeyka, which is nestled in the remote, mountainous area of the Pontic Alps, has probably contributed to the preservation of the language and of some archaic features and facilitated the development of three main dialect areas around the townships of Of/Çaykara (abbreviated ROf), Sürmene (abbreviated RSür), and Tonya (see Schreiber 2024). Romeyka has been in contact with Turkish varieties at least since the 17th century, although early contacts may date back till the 11th c. (Drettas 1997: 5–6) and it is difficult to assess the intensity of contact with Turkish throughout the Middle Ages and the Modern period. At least since intensified labour migration to larger cities in Turkey (as well as abroad) since the 1960s, the influence of Turkish has significantly increased, causing cultural assimilation and language shift at least in the urban speech communities (Schreiber & Sitaridou 2017). Today, the majority of Romeyka speakers are recessive bilinguals, with Turkish as dominant language, which affects the linguistic structure of Romeyka and facilitates contact-induced changes (Schreiber 2024).

The density of documentation and grammatical description of Romeyka is moderate, with increased interest in the field in the last years. After Mackridge's (1987) work on the Muslim Pontic Greek variety still spoken in Turkey, Sitaridou (2014b, 2013, 2014a, 2016, 2021) has contributed research mainly on the syntactic domain of Romeyka, followed by Neocleous' 2017 doctoral thesis on word order and information structure in Romeyka (Neocleous 2020, cf. Neocleous 2022). Recently, Schreiber (2024) has presented the first comprehensive grammatical description of Romeyka based on a naturalistic spoken language corpus. A considerably larger body of literature is available on Pontic Greek as spoken by Christian speakers in Turkey before the Greek-Turkish population exchange in 1923 (Deffner 1878, Parcharidis 1880, 1888) and in Greece after 1923 (Dawkins 1931, 1937, Papadopoulos 1933, 1955, 1958–1961, Tombaidis 1992, 1996, Drettas 1997, 1999, Revithiadou & Spyropoulos 2009) as well as on its closest relatives Cappadocian (Janse In press and references therein) and Pharasiot (Bağrıaçık 2018 and references therein).

The aim of the present chapter is to analyze the WOWA dataset of Romeyka (Schreiber 2021) with regard to the grammatical entities that occur in the postverbal domain and to word order in Romeyka in general. Since Romeyka is a shifting language under strong contact with Turkish (see Schreiber 2016 and Schreiber & Sitaridou 2017 for a sociolinguistic assessment of language vitality), which displays fundamentally different word orders, the data reveal a lot of variation which needs to be accounted for. Thereby, theoretically relevant topics from Greek linguistics such as word order properties of strong and weak pronouns and the question about a potential shift in Romeyka word order directionality will be

touched upon. Most importantly, it will be argued that basic word order properties (see Section 2 for a discussion of terminology) in Romeyka as a shifting language varies significantly between speakers and seems to be crucially determined by the individual multilingual profiles of the speakers.

The WOWA dataset of Romeyka has been compiled and coded according to the methodology of the WOWA corpus outlined in Haig et al. 2024 [this volume]; basic descriptive statistics have been carried out in R (logistic regression models). The data consist of five coherent texts produced by three speakers, extracted from the larger naturalistic spoken language corpus of Romeyka compiled by Schreiber (In prep).[2] The texts were recorded during fieldwork in Turkey in June/July 2019. Romeyka examples in the present chapter are referenced as follows: examples that stem from the WOWA dataset are referenced as Schreiber (2021: text ID, token ID), and examples from the Romeyka corpus by Schreiber (In prep) are referenced by the respective code in the corpus (as explained in Schreiber 2024).

In the following, Section 2 summarizes the present stage of research on word order in Romeyka. Section 3 sketches the general impact of information structure. In Section 4, the word order profile of Romeyka is characterized based on the WOWA data by focusing on different clause types (the NP in Section 4.1, the PP in Section 4.2), semantic and grammatical roles (Sections 4.3-4.5), auxiliaries in Section 4.6, and complex clauses in Section 4.7. Section 5 on areal aspects of language contact and Section 6 on the impact of recessive bilingualism discuss further important factors in the variability of Romeyka word order patterns.

2 Romeyka word order: Background and previous analyses

It is not straightforward to determine the present word order profile of Romeyka (a) due to a high pragmatically conditioned variability (see Section 3), and (b) due to ongoing language shift to Turkish and consequently high variation. It is generally agreed that Greek word order is determined by information structure. This was the case in Ancient Greek (van Emde Boas et al. 2019: 702–721), in Medieval Greek[3](Holton et al. 2019: 2022–2024) as well as in Standard Modern Greek (Holton et al. 2012: 518–520). As far as PG is concerned, it has been

[2]For an overview of speakers and data in the WOWA dataset of Romeyka, see Table 1 in Section 6. For a description of data collection and metadata see multicast.aspra.uni-bamberg.de/resources/wowa/data/hellenic/ponticgreek_romeyka/wowa_hell_ponticgreek_romeyka__metadata.pdf.

asserted that there is no basic worder (Drettas 1997: 277–280). On the other hand, Neocleous (2022, 2020), following Sitaridou (2016), investigated the diachronic development of Romeyka word order within the Minimalist framework and concluded that Romeyka is a "mixed directionality language" (Neocleous & Sitaridou 2022) with inherited VO order in pragmatically unmarked main clauses and OV order in subordinate clauses reinforced by contact with Turkish.

Neocleous (2022, 2020) claims that Romeyka has in main clauses an underlying VO word order, which is evident when all possible information-structural complications are set aside. Deviations from VO are explained through information-structural conditions holding in certain contexts (for a similar approach for Pharasiot, see Bağrıaçık 2018). In subordinate declarative clauses, unmarked word order is, according to Neocleous, OV with finite verbs and VO with infinitives. Although Neocleous seems not to distinguish between complement clauses preceding and following the matrix clause, most of his examples of subordinate clauses follow the matrix clause.

As has been addressed by Haig et al. (2024 [this volume]), it is by no means straightforward to determine unmarked or "basic" word order in a language and there are different accounts on how to establish this (see Dryer 2013 for a "rule of thumb" in determining dominant word order based on relative frequency, cf. also Dryer 1995). Neocleous (2022, 2020) defines a basic word order for both matrix and subordinate clauses in Romeyka based on the pragmatically unmarked word order, which he defines as an "'all-focus sentence', aka 'a presentational focus sentence', containing neither old information nor any presuppositions" (Neocleous 2020: 143). He elicits such clauses in response to the question 'What happened?', departing from the assumption that in response to this question all information is new and thus of equal discourse-pragmatic status.[4] However, departing from a naturalistic spoken language corpus of multilingual speakers as in Schreiber (2024), frequencies of word order patterns show a different picture, although certain limitations apply as well, such as the size of the corpus. While many observations of Neocleous are confirmed by the present data (e.g., the impact of information structure), the analysis cannot be adopted wholesale (see especially Section 6 on inter-speaker variation). Furthermore, the present analysis follows largely the WOWA approach (see Haig et al. 2024 [this volume])

[3]See also Horrocks (1990: 45) and Rafiyenko & Seržant (2020: 11) for Postclassical Greek. It should be noted that in later publications, Horrocks (2007) assumes VSO as the informationally most neutral word order in both Postclassical (Horrocks 2007: 623) and Late Medieval and Early Modern Greek (Horrocks 2019: 2022–2023).

[4]Certainly, Neocleous (2020) is not the first to use this approach; for a description of how neutral word order has been defined in the previous literature, see Bağrıaçık (2018: 146–151).

which is role-specific; the existing literature on word order in Romeyka does, to our knowledge, not include claims specific to Goals, Locations, etc. So, our prominent question should not be whether Greek had at any stage (Postclassical, Medieval, Modern) an unmarked VO word order, but rather whether OV has become more prominent in Romeyka — regardless of information structure — under contact influence from Turkish. Therefore, our analysis shall be based on frequencies of head-final orders. Note that this is still not straightforward to deal with, as Turkish allows for word order variation as well (Göksel & Kerslake 2005: 343–349), at least in informal spoken language — and certainly in the Trabzon Turkish dialect (see Schreiber Submitted).

3 The role of information structure

This section summarizes the role of information structure in Romeyka word order as proposed by Neocleous (2022, 2020), although the naturalistic corpus data do not always conform to these predictions and, for example, inter-speaker differences will need to be kept in mind (see also Janse & Schreiber In prep). Word order in Romeyka is largely determined by information structure, defined in terms of the concepts of topic and focus. A topic of a clause is defined here broadly as old/given information, that is, "an entity that has usually already been introduced into the discourse and is taken up again" (Bağrıaçık 2018: 114) and which is, if not already familiar to the hearer (or at least to the speaker), "agreed on by the speakers" (Soltic 2015: 48, following Gundel & Fretheim 2004). A constituent is in focus if it contains emphasized information which is generally assumed to be in this context new to the hearer (cf. "information focus" in Bağrıaçık 2018: 115). Both topic and focus can also yield contrastive information, thus called contrastive topic and contrastive focus. A contrastive topic is "an element that induces alternatives which have no impact on the focus value and creates oppositional pairs with respect to other topics" (Frascarelli & Hinterhölzl 2007: 88; cf. Bağrıaçık 2018: 267).

Neocleous (2020: 105) argues that there is a single subject position in Romeyka main and subordinate clauses and that all subjects (no differentiation regarding specificity) in pragmatically unmarked orders in Romeyka are left-dislocated topics,[5] appearing in the left-most clause position (1).

(1) Romeyka (Schreiber 2021: B, 0238)
[ena peðas]$_{TOP}$ ebidže havus
a boy make.AOR.3SG pool
'A boy made a pool.'

Both the pre-verbal and the post-verbal domain can host topics, but contrastive topics are only possible in pre-verbal position (example (2); Neocleous 2020: 186). Definite object topics, both pre-verbal (Neocleous 2020: 128) and post-verbal (Schreiber 2021), optionally trigger clitic doubling, that is, the coindexation of a (pro)nominal object topic by a clitic pronoun. Clitic doubling of definite object topics is an inherited Greek feature, which is obligatory in PG and other East Asia Minor Greek varieties such as Cappadocian and Pharasiot (Janse 2008), but not systematically attested in the Romeyka corpus. In PG (Drettas 1997: 276–280), clitic doubling occurs for nominal object topics in left-dislocated position (but not when the topicalizer *=ba(l)* occurs) by means of a referential resumptive pronoun to distinguish a topic from pre-verbal focus since focalization never triggers clitic doubling.[6]

Romeyka displays, like PG, an (arguably)[7] clitic topicalization particle *=ba(l)* (3a/b), which according to Neocleous (2020: 120) assigns contrastive topic to the marked constituent, although the manifold functions of this particle outlined in Schreiber (2024: 141–143) require more detailed investigation. The focus position in Romeyka (for both information and contrastive focus) is immediately to the left of the verb (4) and can be filled by several constituents (Neocleous 2020: 129); multiple focus is possible resulting in movement of all focused constituents to pre-verbal position (Neocleous 2020: 181).

In sum, according to Neocleous (2020, 2022), the pragmatically neutral basic word order in Romeyka main clauses is (S)VO, as illustrated in (1) above; if OV order occurs in main clauses, this is argued to be due to either focalization or topicalization.

[5]We are aware that the term left-dislocation is associated with specific mechanisms in the generative literature (Kaltsa & Sitaridou 2010, Neocleous 2020, 2022), but since we are not working within a generative framework, we prefer to use "topicalization" for preverbal topics which are still within the clause and "topic left-dislocation" for preverbal topics which are prosodically detached from the nuclear clause and constitute a separate intonation unit (Janse 2008), "backgrounding" for postverbal topics within the clause and "topic right-dislocation" for postverbal topics which constitute a separate intonation unit (in accordance with Janse 2008: 167–168).

[6]Compare Horrocks' 2019 rule [92] for the obligatory co-occurrence of clitic-doubling with dislocated and non-dislocated preverbal topics versus rule [93] for the absence of clitic-doubling with preverbal foci in Late Medieval and Early Modern Greek (Horrocks 2019: 2024–2025).

[7]On the enclitic status of *pa(l)* in PG, see Papadopoulos (1955: 119, 1958–1961: volume 2: 138), Setatos (1994), Drettas (1997: 46, 434), Janse (2002: 225–226), Ralli (2006: 131–132), Kaltsa & Sitaridou (2010: 263). Compare the use of *πάλιν* as an enclitic topic marker in Late Medieval Greek (Soltic 2013).

(2) Romeyka (Schreiber 2024: 247, ex. 75a, constructed example)
[avudo to saχan]$_{TOPi}$ epero to$_i$
this the plate take.PRS.1SG OPN.CL.3SG
'I take this plate.'

(3) a. Romeyka (Schreiber 2021: D, 0390)
ulin efteme
all make.PRS.1PL
'We make everything.'
b. Romeyka (Schreiber 2021: D, 0391)
pikniki ba efteme
picnic TOP make.PRS.1PL
'We make a picnic.'

(4) Romeyka (Schreiber 2021: B, 0243)
[ta is]$_{foc}$ ebidže
the footprints make.AOR.3SG
'He left footprints.'

4 Word order profile

4.1 Word order in the nominal phrase (NP)

Word order in the NP is head-final in Romeyka and nominal modification is prenominal with the exception of enclitic genitive pronouns. Attributive adjectives precede the head noun (5),[8] as do demonstratives (6) and numerals (7). Definite NPs trigger determiner spreading on attributive adjectives and in principle also on numerals,[9] that is, the definite article occurs before each modifying element as well as before the head noun (e.g., examples 6, 8, 10; for details on determiner spreading and nominal agreement in Romeyka see Schreiber 2024).

[8]While we apply in this chapter generally a very simplified glossing system that ignores some morphological information (and we also abstained from indicating word accent in the examples), we indicate case information only in Section 4.1 on NP word order, but ignore nominal number and gender. It has to be noted that case flagging vs. bare marking, as it is differentiated in the WOWA coding strategy, cannot be considered a reliable factor in Romeyka, since nominative and accusative(/oblique) case endings are often reduced and it can be partly only inferred from the syntactic context which case is expected in a certain example. In the coding of the WOWA dataset, only those tokens are tagged as 'case' which show a clear case ending, like MASC.SG.NOM. *-os* and MASC.SG.ACC. *-on*.

[9]Although the Romeyka corpus shows several deviations (see Schreiber 2024).

(5) Romeyka (Schreiber 2021: A, 0042)
geniše ðromo utš en
broad road.NOM NEG be.PRS.3SG
'There is no broad road.'

(6) Romeyka (Schreiber In prep: 08_04072019M_3; 161)
hatšino d omorfo don dobo
this.ACC the nice.ACC the.ACC place.ACC
'this nice place'

(7) Romeyka (Schreiber 2021: C, 0285)
ðio nomade
two persons.NOM
'two people'

(8) Romeyka (Schreiber In prep: 04_01072019F_2; 087)
mo ta dœrt tane ta za
with the.ACC four piece the.ACC cows.ACC
'with the four cows'

Romeyka has pre-nominal nominal genitives (both nouns and NPs), i.e., the possessor precedes the possessed (9, 10).

(9) Romeyka (Schreiber 2021: C, 0303)
tu spid i arθob
the.GEN house.GEN the.NOM people.NOM
'the people of the house'

(10) Romeyka (Schreiber In prep: 08_04072019M_2; 090–092)
du dünja olon da tehlikelija da dobe
the.GEN world.GEN all.GEN the.NOM dangerous.NOM the.NOM places.NOM
'the world's most dangerous places'

As for pronominal possession, Romeyka has pre-nominal full possessive pronouns (11) but also post-nominal weak/enclitic possessive(/genitive) pronouns (12).[10] As with other pre-nominal modifiers in definite NPs, the head noun keeps its definite article when combined with a full (pre-nominal) possessive pronoun (11), which historically includes an incorporated definite article as well.

(11) Romeyka (Schreiber 2021: A, 0129)
temetero do barχari
POSS.1PL the pasture
'our pasture'

(12) Romeyka (Schreiber 2021: E, 0564)
andras ades
husband.NOM POSS.CL.3SG
'her husband'

Relative clauses in Romeyka are in principle pre-nominal (13), although post-nominal relative clauses exist as well (14). It is argued that pre-nominal relative clauses in Romeyka have evolved under contact influence from Turkish, while post-nominal relative clauses are a Hellenic relic (Neocleous 2020; Schreiber 2024).

(13) Romeyka (Gandon 2016: 222, ex. 517, glosses modified)
opse iða [alis p epiren] ineka
yesterday see.AOR.1SG Alis.NOM REL take.AOR.3SG woman.ACC
'Yesterday I saw the woman who Ali married.'

(14) Romeyka (Neocleous 2020: 71, ex. 87, presentation/glosses modified)
o peðas [op erθen aso cicenin] temon t anepsin
the child REL come.AOR.3SG from.the grocery my the nephew
en
be.PRS.3SG
'The child who came from the grocery's is my nephew.'

4.2 Adpositional phrases

Romeyka is a prepositional language. Prepositional phrases express in Romeyka semantic roles of location, Goal, source/origin, instrument or benefactive, as well as some temporal (and other) adjuncts. However, these semantic categories are not exclusively flagged with prepositions. The overall rate of prepositional marking in the WOWA dataset for Romeyka is 61% (percentage calculated based on

[10]Throughout this chapter, we indicate clitic status only for weak possessive and object pronouns, and refrain from doing so for all other parts of speech that are traditionally considered clitics in Greek linguistics, such as definite articles, prepositions, relativizer, auxiliaries, negation and modal particles, and the coordinating conjunction *tše* 'and'.

the occurrence of prepositions in the functions ablative, addressee, benefactive, comitative, (caused) Goal, instrumental, locative, recipient and beneficiary). For differences in the frequency with which adpositional phrases occur after the verb, see the sections on the respective semantic roles below.

Finally, it must be noted that Romeyka has a complex system of spatial orientation (see Schreiber 2024) and some spatial adverbs could be potentially considered circumpositions (see also Karatsareas 2016 work on circumpositions in Cappadocian), if they co-occur with a preposition, although their status as bound elements is not clear (examples 15, 16 vs. 17).

(15) Romeyka (Schreiber In prep: 05_03072019M_3; 29)
s ena sergi eban
on a blanket above
'on (top of) a blanket'

(16) Romeyka (Schreiber In prep: 01_28062019F_3; 42–43)
s oros apes-merea
at.the forest inside-somewhere
'inside the forest'

(17) Romeyka (Schreiber In prep: 01_04022016F_1; 052)
eb-ebuka asi ɣoryoran
from-under from.the Gorgoras.ACC
'from lower Gorgoras' [as differentiated from upper Gorgoras]

4.3 Ordering of spatial expressions relative to the verb

4.3.1 Locations

The preposition indicating location (as well as Goal/direction) is *s* 'to, at', which can merge with determiners and object pronouns, e.g., *s* 'to, at' + *to*.DET.ACC.SG > *so* 'to the' (18), *s* 'to, at' + *emasuna*.OPN.1PL > *semasuna* 'to us'. The overall frequency of post-verbal locations (including PPs and spatial adverbials) is 42%, as in (19), the majority of locations is pre-verbal. There is no significant statistical correlation found between the position and the independent variables animacy, weight or flagging (i.e., any overt phonological marking of case). However, there is some inter-speaker variation to be observed (see Table 3 in Section 6 below). The position of spatial adjuncts is sensitive to discourse. The immediate pre-verbal position is argued to be the (information) focus position (Neocleous 2020: 132, 148). In this vein, as exemplified in (19), multiple dislocation of (place) constituents is possible in Romeyka, as is the case for SMG (Alexiadou 1997: 58).

(18) Romeyka (Schreiber 2021: A, 0187)
eskiden gœl$_{FOC}$ *ebiname so bodami*$_{TOP}$
formerly lake make.IPF.1PL at.the valley
'In earlier times, we made a lake at the valley.'

(19) Romeyka (Schreiber 2021: A, 0001)
emistine$_{TOP}$ *[aða sin otšena]*$_{FOC}$ *jašaevume*
we here at.the Ogene live.PRS.1PL
'We live here at Ogene.'

4.3.2 Goals

For the discussion of goals in Romeyka, only goals of motion verbs and verbs of caused motion are considered. Goals in the WOWA dataset are predominantly expressed by prepositional phrases headed by the prepositions *s* 'to' and *os* 'until', but also adverbially. 78% of goals (both prepositional and adverbial) are post-verbal (20), which shows that goals are much more likely to be post-verbal than locations. While information structure affects word order variation in spatial adjuncts in Romeyka, this is not likely to be the only factor determining word order in light of the significantly higher number of post-verbal Goals; see Rasekh-Mahand et al. 2024 [this volume] on spoken Persian for a similar observation. However, in light of an assumed VO/VX order in declarative clauses in Romeyka (Neocleous 2020) — as in other dialects of modern Greek — it is rather the high number of pre-verbal locations and sources (see Sections 4.3.1 and 4.3.3) that is striking and requires explanation. Comparing (20) with (21), information structure accounts for the pre-verbal position of the PP in (21). Otherwise, no statistically significant correlation is found between position and the dependent variables animacy, weight, and flagging, although adverbial goals have a higher likelihood to be post-verbal than PPs. Finally, it has to be noted that there is high inter-speaker variation (see Table 4 in Section 6 below).

(20) Romeyka (Schreiber 2021: A, 0175)
pao so raši
go.PRS.1SG to.the mountain
'I go to the mountain.'

(21) Romeyka (Schreiber 2021: A, 0136)
[son barχari muna]$_{TOP}$ *[direk araba]*$_{FOC}$ *bai*
to.the pasture POSS.CL.1PL direct car go.PRS.3SG
'A bus goes directly to our pasture.'

As for caused goals in the WOWA dataset, 100% are post-verbal (22), but note that due to very low token numbers (N=4) in the generally small WOWA dataset for Romeyka this information is only tentative and needs further investigation.

(22) Romeyka (Schreiber 2021: D, 0488)
eferenam son barχari ksila
bring.IPF.1PL to.the pasture wood
'We brought wood to the pasture.'

4.3.3 Sources

The majority of sources in the WOWA dataset, i.e., 80%, are pre-verbal, including PPs with the preposition *as* 'from' and adverbials. This result needs to be taken with some caution, though, since nearly all examples stem from a single speaker. Information structure is likely to account for the position of the source, cf. ex. (23a) in assumed historically unmarked VX position vs. (23b) in pre-verbal contrastive focus position where the PP contrasts with the PP in (23a), i.e., the previous sentence in the same recording.

(23) a. Romeyka (Schreiber 2021: A, 0024–0025)
bazen para berename edroɣame asa bakale
sometimes money take.IPF.1PL eat.IPF.1PL from.the shops
'Sometimes we took money, we ate [food] from the shops.'

b. Romeyka (Schreiber 2021: A, 0026)
bazen aso spidi eberenam tš ebejname
sometimes from.the house take.IPF.1PL and go.IPF.1PL
'Sometimes we took [food] from the house.'

4.4 Ordering of direct objects relative to the verb

4.4.1 Nominal direct object

This section focuses on direct objects (DOs) of transitive verbs. In the WOWA dataset, 66% of nominal direct objects are post-verbal (arguments of 'have'/existentials, which are predominantly pre-verbal, not included in this count). The percentage appears to be the same for definite and indefinite nominal DOs, which suggests that definiteness does not play a role. Note, however, that in Cappadocian, word order is sensitive to definiteness, that is, indefinite object NPs tend to occur in post-verbal position (Janse 2006). The difference in word order between examples (24) and (25) is probably rather than in definiteness to

be found in information structure, with the object NP in (24) in focus position and the object NP in (25) in topic position. However, there may be also an interference from the frequent occurrence of post-verbal nominal object NPs in Trabzon Turkish (Schreiber Submitted), a phenomenon described for informal spoken Standard Turkish as "backgrounding", and which applies usually to definite NPs, although a non-definite NP can be placed in spoken Turkish in the post-verbal position "if it refers to an entity or category that has been mentioned (or implied) in the immediately preceding discourse" (Göksel & Kerslake 2005: 346).

Numerically, more complex NPs/PPs in the WOWA dataset, i.e., those with more than two words usually involving some kind of nominal modification or genitives, tend to be pre-verbal (26). Although this is contrary to the traditional assumption that very complex NPs/PPs preferably appear at the end of an utterance (see, e.g., Behaghel 1909 departing from German), it does align with similar findings from the spoken-language corpora investigated in this volume (see, e.g., Rasekh-Mahand et al. 2024 [this volume] and Leitner 2024 [this volume]). However, it is likely that information structure is the more decisive factor, which might just happen to overlap with the factor of "weight" (see, e.g., 26).

A significant correlation was found between the dependent variable position and the independent variable animacy: human and animate nominal DOs are slightly more likely to be pre-verbal than inanimate nominal DOs. Interestingly, a reverse significant effect was found for pronominal human and animate DOs, which tend to be more likely post-verbal than inanimate pronominal DOs (see Section 4.4.2). As is the case for all semantic categories, the order of nominal DOs with regard to the verb is sensitive to information structure (see Section 3) and inter-speaker variation is relevant here as well (see Table 2 in Section 6). Finally, it also needs to be stated that the linear order of nominal objects with regard to the verb seems to be at times within the same speaker in free variation, even if other variables remain stable (e.g., 27 vs. 28), i.e., there is intra-speaker variation. This needs to be most likely attributed to ongoing language shift under strong influence from Turkish OV orders, as information structure cannot always be convincingly invoked as explanation for all OV orders.

(24) Romeyka (Schreiber 2021: E, 0564)
andras ades ba ksila gofdi
husband POSS.CL.3SG TOP wood cut.PRS.3SG
'Her husband cuts wood.'

(25) Romeyka (Schreiber 2021: E, 0572)
efidže so ðormo ban ena ksilo
leave.AOR.3SG at.the road above a wood
'He lost a log on the street.'

(26) Romeyka (Schreiber 2021: A, 0144–0145)
ifedi eɣo ebiɣa išde muskaræ utš dane muskar ebiɣa
last_year I buy.AOR.1SG DP calves three piece calve buy.AOR.1SG
'Last year I bought calves. I bought three calves.'

(27) Romeyka (Schreiber 2021: A, 0073)
efteme fasulijas
do.PRS.1PL beans
'We do beans.' [i.e., 'We grow beans.']

(28) Romeyka (Schreiber 2021: A, 0099)
opsar ebsame
fish catch.AOR.1PL
'We fished.'

4.4.2 Pronominal direct objects

The percentage of post-verbal pronominal direct objects in the WOWA dataset is with 58% slightly lower than that of nominal DOs. Importantly, (en)clitic object pronouns are not coded in the corpus since they are bound. Romeyka has both clitic object pronouns following the predicate and free object pronouns preceding the predicate. The two are mainly differentiated based on the criterion of stress, whereby enclitic object pronouns have a reduced phonological form and do not impact the stress pattern of the verb. However, in the Romeyka corpus (Schreiber In prep) there appear to be also post-verbal full object pronouns, which are sometimes difficult to differentiate from weak ones and which are probably reinforced by contact with Turkish (see below; see also Schreiber 2024: 103). In the WOWA dataset, the percentage of 58% post-predicate pronominal DOs only refers to full pronominal forms, the vast majority being third person pronouns. While none of the coded categories like weight or flagging is significant for the position of pronominal DOs, there is a significant tendency for pronominal DOs denoting human or animate entities to appear post-verbally as opposed to inanimate pronominal DOs (29 vs. 30). However, it is obvious from the examples, that information structure plays a significant role here as well (see 29). Still, it is unclear whether 42% of pre-verbal pronominal DOs are largely due to information

structure since pronouns are usually given information which can occur in pre- and post-verbal position (see Section 3).

(29) Romeyka (Schreiber 2021: D, 0408–0409)
laɣo dune ado utš eksero
how be.IPF.3SG this NEG know.PRS.1SG
'How it was? I don't know this.'

(30) Romeyka (Schreiber 2021: B, 0246)
utš iðan adona
NEG see.AOR.3PL OPN.3SG
'They did not see him [=the bear].'

An important issue is the post-verbal placement of full pronominal objects (often third person pronouns) versus weak enclitic object pronouns. In the WOWA dataset, 60% of pronominal DOs are weak enclitic pronouns (thus not coded for position), 40% are full pronominal forms out of which 58% are post-verbal. In comparison, in the larger Romeyka corpus, 63% of all pronominal DOs are clitic pronouns (Schreiber 2024: 100, Table 12). In (30), a weak object pronoun would be likely to occur since 'the bear' is a topic which has been mentioned several times in the preceding context. Instead, the full pronominal form is placed in post-predicate (enclitic) position. As another example, in (31), the full pronominal form in post-verbal position is preferred over the weak pronominal form *=(a)ta*. It is possible that a contact explanation can account for the preference of post-posed full object pronouns in Romeyka, as (Trabzon) Turkish has no clitic object pronouns and object pronouns can appear post-verbally in informal spoken Turkish (also cf. only 13% post-verbal pronominal DOs in Hodgson 2024). According to Brendemoen (2005: 30), post-verbal pronominal DOs in Trabzon Turkish have arisen due to contact with PG. The potential mutual influence suggests a convergence-type of change between Romeyka and Trabzon Turkish.

(31) Romeyka (Schreiber 2021: B, 0234)
ama utš eboresane dosin adonusine
but NEG can.AOR.3PL give.INF OPN.3PL
'But they could not hit them.'

In sum, VO is assumed to be the unmarked order of pronominal DO and verb in Romeyka, whereby OV orders are motivated by information structure. The remaining question is what motivates the post-verbal position of full object pronouns. In terms of information structure, pronominal VO with full object pronouns can only be explained, if their information structural value is comparable

to that of weak clitic pronouns, that is "familiar topics". If the post-verbal position of full pronominal DOs in the Romeyka corpus is split up according to person and number, it is evident that third person singular full object pronouns are most likely to be post-verbal, although strikingly third person plural full object pronouns have the least chance to be post-predicate (unlike ex. 31).[11] So probably, gender could play a role in the sense that gender cannot be differentiated in weak third person singular object pronouns but does reflect in strong third person object pronouns. Furthermore, it is not clear whether phonological weight could also play a role, although 3SG strong pronominal forms are not significantly shorter (i.e., more "clitic-like") than 3PL object pronouns.

4.5 Ordering of other obliques relative to the verb

Other obliques refer to the semantic roles of recipients, addressees, comitatives, instruments, benefactives, and others. In the WOWA dataset, these categories do not figure prominently, so any quantitative analysis is pointless. Instead, the present section discusses some examples for each semantic category.

The WOWA dataset contains only two tokens for recipients, both of which are pronominal. Both pronominal recipients occur pre-verbally (32; but cf. 35). In general, the unmarked word order for nominal recipients is VO (33) with OV orders triggered by information structure, namely focus (34; see Schreiber 2024: 249–250). In ditransitive constructions, the unmarked word order is V–IO–DO (Schreiber 2024: 249). This applies also to pronominal recipients (see (35) and unlike (32) with a topicalized object pronoun).

(32) Romeyka (Schreiber 2021: A, 0182)
emenan ndona na ðiɣune
OPN.1SG what PRT give.PRS.3PL
'What do they give to me?'

(33) Romeyka (Schreiber 2024: 250, ex. 93, questionnaire data)
ta mila ðokan ti mana tuna
the apples give.AOR.3PL the mother POSS.CL.3PL
'They gave the apples to their mother.'

[11] Distribution of bound (weak) pronominal object pronouns in the Romeyka corpus (Schreiber In prep) per person and number: 1SG: 41 bound out of 67, i.e., 61% bound; 2SG: 18 out of 28, i.e., 64% bound; 3SG: 72 out of 133; i.e., 54% bound; 1PL: 17 out of 26, i.e., 65% bound; 2PL: 6 out of 10, i.e., 60% bound; 3PL: 63 out of 79, i.e., 80% bound.

(34) Romeyka (Schreiber 2024: 250, ex. 96, questionnaire data)
din batsi eðotše ena ido ..sturatši
the girl give.AOR.3SG a DEM ..stick
'He gave a stick to the girl.'

(35) Romeyka (Schreiber 2024: 251, ex. 102, questionnaire data)
etšine bal ðotš emena milo
he TOP give.AOR.3SG OPN.1SG apple
'He gave me an apple.'

In the WOWA dataset, all three tokens of addressees are pronominal and appear post-verbally (36, the stress pattern *ípen ádona* reveals that the pronoun must be indeed the post-posed strong form and not a clitic). In general, like with other obliques, nominal addressees are expected to follow the verb in unmarked word order, although they can move to pre-verbal focus position (note that the Romeyka corpus does not feature an example of a nominal addressee, but cf. ex. (37) with *ta patsiðes* 'the girls' as contrastive topic). Pronominal addressees, seem to occur nearly in all cases (of the Romeyka corpus) post-verbally.

(36) Romeyka (Schreiber 2021: E, 0582)
iben adona
say.AOR.3SG OPN.3SG
'He told her […].'

(37) Romeyka (Schreiber In prep: 02_02022015F_1; 073–074)
eleɣane ištera ta peðia kopela ta patsiðes eleɣane ɣospiðes
say.IPF.3PL later the boys girl the girls say.IPF.3PL prostitutes
'Then they said to the boys girl, to the girls they said prostitutes.'

Comitatives, i.e., referents denoting accompanying persons (or at least animate entities), occur in the WOWA dataset predominantly in post-verbal position (38a), although pre-verbal placement is possible as well (see the PP as pre-posed given topic in (38b), although stylistic variation seems to play a role here as well).

(38) a. Romeyka (Schreiber 2021: A, 0060–0061)
jaja ebejnane me ta za
on_foot go.IPF.3PL with the cows

b. *me ta za jaja ebejnane*
with the cows on_foot go.IPF.3PL
'They went on foot with the cows.'

Instruments appear in the WOWA dataset in the vast majority pre-verbally, which may be explained by focalization of new information (39) or contrastive focus. However, based on the data of the Romeyka corpus, it seems that instruments do not differ from other obliques in terms of unmarked VO order (40). Ex. (41) shows that for information structural reasons both nominal obliques and objects can occur pre-verbally, i.e., Romeyka allows for multiple focus, which is argued to be "order-preserving" (Neocleous 2020: 181–182). However, then it is not clear why the oblique precedes the DO in (29).

(39) Romeyka (Schreiber 2021: E, 0570)
mo d aksinari ešgise da ksila
with the axe split.AOR.3SG the woods
'He split the wood with the axe.'

(40) Romeyka (Schreiber In prep: 09_04072019_7; 11)
ekoftame me ti kerenti
cut.IPF.1PL with the scythe
'We cut with the scythe.'

(41) Romeyka (Schreiber 2021: A, 0036)
ula me ta rašes ta ɣomare ekovalename
all with the pannier the loads carry.IPF.1PL
'We always carried the loads with the panniers.'

Benefactives refer to situations where X does something in the interest of Y, generally implying that Y is a sentient being. Although no benefactives appear in the WOWA dataset, they are realized in Romeyka by means of the preposition *ja(t)* 'for' (42). Their position seems not to differ from that of comitatives, which are prepositional phrases as well, and which appear often in the pre-verbal domain for information structural reasons, although their unmarked word order is post-verbal.

(42) Romeyka (Schreiber 2024: 147, ex. 565, questionnaire data)
sade jad emena faji utš eθelisa pseθinimo
only for OPN.1SG food NEG want.AOR.1SG cook.NMZ
'I did not want to cook just for myself.'

Among other obliques in the WOWA dataset, i.e., those which are none of the semantic roles above, 57% are in post-verbal position, which equals roughly the overall percentage of post-verbal placement in the WOWA dataset which is 55%. Since nearly all of these other obliques are prepositional phrases, they follow the constraints outlined for other PPs above.

4.6 Auxiliaries

Neocleous & Sitaridou (2022) discuss the following auxiliaries in Romeyka: (i) *ime* (+ particle), (ii) *iχa* (+ infinitive), (iii) *eš(i)* (+ finite verb in present tense or imperfective past). They always precede the main verb, irrespective of whether it is finite, infinitive or particle. Additionally, there is a periphrastic progressive construction with *steko/stekome* 'stand' or *kahome* 'sit.'

The invariable form *eš(i)* plus finite verb is used as a periphrastic progressive denoting processes that are close to completion and goes most likely back to the 3SG present tense form of the verb *eχo* > *eš(i)* (ex. 43; cf. Drettas 1997: 334 on a related progressive form in PG).[12] The inflected imperfective forms of the verb *eχo* 'have' are used as an auxiliary in the formation of counterfactual conditional clauses (44), which are formed by one of the modal particles *na/an/as* + inflected imperfective of *eχo* + non-finite verb/inflected infinitive (Schreiber 2024: 298, Table 37; see also Sitaridou 2014a).

(43) Romeyka (Schreiber 2021: B, 0241)
eš erde argo
AUX come.AOR.3SG bear
'The bear is/was? coming.'

(44) Romeyka (Schreiber In prep: 04_01072019F_13; 53)
eɣo na m iχa škisen da[13] *da ksila […]*
I PRT NEG AUX split.INF? OPN.CL.3PL the woods
'If I had not chopped the wood, […].'

With the limited use of the auxiliary *iχa* (morphologically homonymous with imperfective 1SG form of *eχo* suggesting two functions of *eχo* as lexical verb and auxiliary), Romeyka stands out from other varieties of modern Greek like SMG. Its closest relative Cappadocian has a pluperfect with impersonal *iton/itan* (homonymous with the 3rd person aorist form of *ime*) preceded by a finite main verb. Diachronically AMG also uses the auxiliary *ime*, morphologically homonymous with the copula *ime* 'be' (i.e., *ime* can thus function both as existential and auxiliary). For the word order properties of copula and 'become' complements in Romeyka, see Section 4.7.2.

[12] According to M. Bagriacik (p.c.), the form *eš* goes probably back to a homonymous existential auxiliary already existing in older stages of Greek, rather than to the 'have'-auxiliary, as the former presents a more likely grammaticalization pathway for a progressive form.

[13] The analysis of the verb form *škis-en=da* is not clear, we analyse *=ta* here as a weak object pronoun but it also resembles the suffix *-ta* marking gerunds like *jelaχ-ta* 'laughing'. The form *škisen-* could be potentially a reduced infinitive, cf. *škisini*.INF.

In both cases where forms of *eχo* are used as an auxiliary in Romeyka, the auxiliary precedes the main verb, although importantly, in the counterfactual conditionals the verb following the auxiliary is not (i.e., in the case of the infinitive) or only partly (i.e., in the case of the "inflected infinitive", see Schreiber 2024: 229–233 going back to Sitaridou 2014a) inflected. Neocleous (2020: 269) confirms AuxV order in Romeyka main and subordinate clauses. The construction Vfin + *iton* in pluperfects which exists in Cappadocian is not attested for Romeyka.

4.7 Complementation

4.7.1 Complement and adjunct clauses

In Romeyka both finite and non-finite complementation exist. Depending on the type of predicate, a complementizer is used to introduce the complement clause (CC), while other clause types do not require a complementizer. In Schreiber (2024: 278), it has been argued that more than one complementation strategy is available for some clause types due to contact influence from Turkish, resulting in an increase of non-finite complementation strategies. While the non-finite strategy of using infinitives in some CCs is an archaic trait of Romeyka, non-finite deverbal nouns as a complementation strategy have increased under contact with Turkish. Within the finite complementation strategies, complementation by means of the complementizer *na* is more restricted in Romeyka compared to SMG, while juxtaposition with paratactic syntax and without complementizer is widespread (45), especially with verbs of saying and in (in)direct speech.

(45) Romeyka (Schreiber In prep: 02_2906019F_1; 02)
egusame [o jaja evren arkon]
hear.AOR.1PL the Yahya find.AOR.3SG bear
'We heard that Yahya has found a bear.'

In Romeyka CCs, the complement clause predominantly follows the matrix verb (45). For some predicate types, among which are verbs of saying, the reverse order is possible as well, see for example the preverbal headless relative clause in (46); even circum-positions exist (47).

(46) Romeyka (Schreiber In prep: 01_28062019F_3; 24)
[to leɣo] utš eɣrigo
what say.PRS.1SG NEG understand.PRS.1SG
'I don't understand what I say.'

(47) Romeyka (Schreiber In prep: 08_04072019M_2; 066)
[t aleɣo] kseris [dohna e]
the horse know.PRS.2SG what be.PRS.3SG
'Do you know what "aleɣo" is?'

Na-clauses and infinitives obligatorily follow the main verb (48), (49). However, deverbal nouns calquing Turkish complement clauses appear before the main verb but they are strictly speaking NPs and thus no actual CCs (50). Nominalizations selected by some aspectual verbs like *bašlaevo* 'start' which requires a PP follow the main verb (51). If a complementizer is used, it appears at the beginning of the CC.

(48) Romeyka (Schreiber In prep: 02_02022015F_1; 014)
utš eθelena [n andriza]
NEG want.IPF.1SG PRT marry.IPF.1SG
'I didn't want to marry.'

(49) Romeyka (Schreiber 2024: 285, ex. 376, questionnaire data)
utš eboresa [tšimeθina]
NEG could.AOR.1SG sleep.INF.1SG
'I could not sleep.'

(50) Romeyka (Schreiber 2021: A, 0193)
[hab-aðadžega to panimo] eɣo utš eɣabo
from-here the go.NMZ I NEG like.PRS.1SG
'I don't want to go from here.'

(51) Romeyka (Schreiber In prep: 04_01072019F_13; 30)
ebašlaepse [so borbatima]
start.AOR.3SG at.the walking
'She started to walk.'

Subordination of adjunct and relative clauses exhibits pre-dominantly head-final syntax, although there is a lot of variation only partly dependent on the clause type. Romeyka adjunct and relative clauses are predominantly finite, although some non-finite strategies exist. The dependent clause is predominantly pre-posed and generally introduced by pre-verbal adverbial subordinators or relative markers (ex. 52; for the syntax of relative clauses, see also Section 4.1. above).

(52) Romeyka (Schreiber In prep: 04_01072019F_13; 45)
[omon d eruise s ormi] ejendune natsurula
when REL fall.AOR.3SG to.the river become.IPF.3SG wet
'When she fell into the river, she got wet.'

With regard to head-order directionality within the complement clause, according to Neocleous (2020: 118), the unmarked word order in subordinate/complement clauses is head-final (53) as a consequence of contact with Turkish head-final syntax. However, the example in (45) above seems to form a counter-example to this generalization, since the CC in (45) is a pragmatically neutral, information-structurally unmarked kind of statement (the object is new, not topical nor contrastive) and following Neocleous (2020), we would expect OV order here. Thus, both OV and VO orders seem to be possible in unmarked CCs; potential restrictions with regard to predicate types need further research.

(53) Romeyka (Schreiber 2021: C, 0332)
eterezen [o argo erθen]
look.AOR.3SG the bear come.AOR.3SG
'He saw that the bear came.'

4.7.2 Copula and 'become'-complements

In the WOWA dataset, only 8% of copula complements appear after the copula, which means that the copula *ime* 'be' appears predominantly in clause-final position. Neocleous (2020: 117) also confirms that copula clauses with the 3rd person form *en/ine* are always head-final. However, when it comes to an explanation of the very low number of copula complements that appear post-verbally in the WOWA dataset (N=2), it is not straightforward to determine any decisive factor. In any case, predicate nominals (54) and predicate adjectives (55) seem to behave alike. Information structure can also account for some of the pre-predicate copula complements (56), while interrogative copula clauses are head-final due to the focus position of the *wh*-element (57). Finally, since the overwhelming majority of clause-final copulas occur in the text of a single speaker (see Table 5 in Section 6 below), the speaker variable may have an effect as well. Still, it is not clear what determines post-predicate copula complements as in (58), although information structure may serve as an explanation here as well.

(54) Romeyka (Schreiber 2021: A, 0128)
ta mandria muna boš en
the stables POSS.CL.1PL empty be.PRS.3SG
'Our stables are empty.'

(55) Romeyka (Schreiber 2021: A, 0020)
džumartesi tatili en
saturday holiday be.PRS.3SG
'Saturday is a free day.'

(56) a. Romeyka (Schreiber 2021: D, 0504–0506)
ama do vutero poli
but the butter much
'But the butter [was] plentiful.'

b. *do vutero bolin adone*
the butter much be.IPF.3SG
'The butter was plentiful.'

c. *eliɣo utš en do diri*
little NEG be.PRS.3SG the cheese
'The cheese was not scarce.'

(57) Romeyka (Schreiber 2021: D, 0408)
laɣo dune
how be.IPF.3SG
'How was it?'

(58) Romeyka (Schreiber 2021: A, 0049)
en ja ta ðorma muna patikas
be.PRS.3SG DP the roads POSS.CL.1PL unpaved
'Our roads are unpaved (as you know).'

Finally, it has to be noted that the copula — and especially its 3SG present tense form *en* — is often omitted (59, also 56a above; see also the last column in Table 5 in Section 6). Due to clause-final null-copulas in Turkish, it is compelling to assume a contact influence here, although an internal explanation may also play a role due to the particular nature of the verb *ime*. In Cappadocian, the full verb *ime* 'be' is used in existentials, where it is never dropped; besides, a copula *ime* 'be' exists which behaves as a clitic (*=me*) since MedGr times and thus results always in verb-final orders, although it is never left out in third person singular present tense.

(59) Romeyka (Schreiber In prep: 02_9062019F_1; 24)
etšinos χaremenos
he happy
'He is happy.'

Interestingly, 'become'-complements in the WOWA dataset are 69% post-verbal, suggesting a different behaviour compared to copula clauses. However, it must be noted that nearly all 'become'-complements are produced by the same speaker (i.e., Speaker 2, see Table 1 in Section 6), so this result is to be treated with caution. Although the number of coded tokens is too small for any statistical analysis, there is in the dataset a tendency for adjectives as 'become'-complements to be more likely post-predicate than nominal complements (but cf. *ejendune gedže* 'it became night' (Schreiber 2021: B, 0262; E, 0529), also *inete akšemis* 'it became evening' (Schreiber 2021: D, 0437) vs. *gedže ejendune* (Schreiber 2021: B, 0242)). Again, information structure has an influence here; see (60) vs. (61) in topic position.

(60) Romeyka (Schreiber 2021: D, 0445)
ula inumunesten annera
all become.IPF.1PL soakingly_wet
'We became allover soakingly wet.'

(61) Romeyka (Schreiber 2021: A, 0210)
emeklis ba na inese na stetšis aðatšeka
retired TOP PRT become.PRS.2SG PRT stay.PRS.2SG here
'When you become retired, you will stay here.'

5 Areal issues & language contact

Similarly to the well-documented influences of Anatolian Turkish on AMG described by Dawkins (1916), who inspired much of the literature on language contact, long-standing language contact with Turkish in the Trabzon area has evidently led to contact-induced changes in Romeyka in several domains of the language including the lexicon and grammar (for a tentative overview see Chapter 6 in Schreiber 2024; see also Brendemoen 1998, 1999, 2002, 2005, 2006, 2010, 2019 for Greek influences on Turkish). Importantly, similarly strong influences of Turkish are reported from Laz (a Kartvelian language, see Kutscher 2008; Lacroix 2019, 2009, Öztürk & Pöchtrager 2011) and Homshetsma/Hemshinli (Western Armenian; Vaux 2007) which are both minority languages in the neighbouring provinces of Rize and Artvin in northeastern Turkey and which share a sociolinguistically and historically similar contact setting with Turkish as dominant language of the area. Although a comparison of contact-induced influences from Turkish on the morphosyntax of Laz and Hemshinli is beyond the scope of this chapter, Neocleous (2020: 282–284) reports a clear restructuring of Laz

word order patterns based on Turkish to pragmatically unmarked word order OV (see also the extremely low percentage of post-predicate elements in the WOWA dataset of Arhavi Laz, Stilo & Lacroix 2021) and unlike its Kartvelian relative Georgian (Neocleous 2020: 243–247). As far as what has been deducible for Hemshinli, it also shares similarities with Turkish, for example in non-finite subordination which is otherwise less common in related varieties outside the specific contact setting (Gandon 2016: 210–212). Finally, comparing the word order properties of PG as spoken in Armenia (Hodgson 2024) with that of Romeyka, Armenian PG has only an overall score of 32% post-posed elements, as opposed to 55% in Romeyka.

Another aspect of language contact in the area is the contact influence PG exerted on the Turkish Eastern Black Sea dialect in several domains including word order (Brendemoen 1998, 1999, 2005, 2006, 2019). For example, post-verbal pronominal DOs in Trabzon Turkish (62) are argued by Brendemoen (2005: 30) to arise due to contact with PG (cf. the percentage of post-verbal pronominal DOs in colloquial spoken Turkish of Ankara (Iefremenko 2021) at 7%). When considering the areal picture of contact-induced changes and convergence, the influence of indigenous minority languages on the majority language should not be neglected, although this seems to apply to a lesser degree to the influence of Armenian (and Laz?) on regional Trabzon Turkish (Brendemoen 2005: 29). In any case, potential mutual contact influences highlight the point that for the purpose of inter-language comparison, not primarily the respective standard variety (such as Istanbul Turkish) should be considered, but rather regional varieties (as is done by Neocleous 2020).

(62) Trabzon Turkish (Brendemoen 2005: 30, ex. 2, presentation adapted)
yedi *oni*
eat.PST.3SG OPN.3SG
'He ate it.'

6 The role of inter-speaker variation

In order to explain (some of) the variability in word order patterns in Romeyka that appear in the quantitative data and can only partly be explained by linguistic factors, it is crucial to consider the nature of Romeyka which is currently spoken in Turkey as shifting variety, and the composition of the present sample. As Schreiber (2024) has shown, the character of Romeyka as shifting language is not (yet) characterized by language attrition as defined by Thomason (2001: 12)

as including structural simplification and loss without compensation but rather with a high inter- (and intra-)speaker variation. Indeed, it is the idiolectal variation based on the individual multilingual profiles of the speakers which explains some of the variability in word order patterns. In other words, the overall figure of 55% post-verbal placement (including 66% post-verbal placement of nominal DOs) in the WOWA dataset of Romeyka does not reflect a stable norm in "the speech community", but a mean value aggregating over very different individuals. As Craevschi (2022) has shown, the Romeyka data set exhibits the greatest degree of inter-speaker variation among the twenty-four WOWA data sets analysed by him. In fact, the influence of the independent variable "Speaker" outweighs significantly the influence of all other variables which were controlled for in the analysis, even when the imbalanced contribution of the three speakers to the overall data set is taken into consideration (Speaker 3 only contributes around 10% of the total tokens). This is shown in Table 1.

Table 1: Overview of speakers and their corresponding texts in the WOWA dataset

Speaker	Speaker characteristics	Texts provided	Total tokens
1	male, middle-aged, ROf, Karaçam	A	198
2	female, middle-aged, ROf, Karaçam	B, D, E	251
3	male, middle-aged, RSür, Beşköy	C	52

The influence of the speaker variable is not the same for all constituent types, and due to the low absolute numbers of tokens for some constituent types, cannot be readily statistically validated for all roles. It does turn out to be significant for predicting the placement of nominal DOs (see also Craevschi 2022). As shown in Table 2, Speaker 1 has with nominal DOs predominantly OV, while Speaker 2 has dominantly OV order. Speaker 3 appears to be largely balanced but note the smaller absolute number of tokens in the data of Speaker 3. A similar inter-

Table 2: Percentage of post-verbal nominal direct objects per speaker

	Total nominal DOs ('do'+'do-def')	Total VO	% VO
Speaker 1 (=text A)	51	18	35%
Speaker 2 (=text B, D, E)	103	86	83%
Speaker 3 (=text C)	21	12	57%

speaker difference is visible for free pronominal DOs, although the percentages for post-verbal placement are there lower in general (see also Section 4.4.2).

A similar picture of inter-speaker variation arises for the semantic roles of locations (Table 3) and goals (Table 4). While Speaker 1 shows predominantly pre-verbal locations (and more or less balanced pre- and post-verbal Goals), Speaker 2 uses predominantly post-verbal locations and especially goals. Speaker 3 shows clearly preverbal locations — other than with nominal DOs — although the amount of data provided by Speaker 3 is too little to get a clear picture.

Table 3: Percentage of post-verbal locations per speaker

	Total locations	Total VX	%VX
Speaker 1 (=text A)	32	11	34%
Speaker 2 (=text B, D, E)	35	23	66%
Speaker 3 (=text C)	16	1	6%

Table 4: Percentage of post-verbal Goals per speaker

	Total goals (no pronouns)	Total VX	%VX
Speaker 1 (=text A)	32	17	53%
Speaker 2 (=text B, D, E)	44	43	98%
Speaker 3 (=text C)	1	1	nc

Ignoring for a moment Speaker 3 (due to low absolute token numbers), it is evident that Speakers 1 and 2 differ rather consistently: Speaker 2 postposes these three argument types approximately twice as often as Speaker 1. With regard to nominal direct objects, locations, and goals, then, an account of word order purely in terms of language-internal features such as information structure, animacy, or weight, is clearly missing a very important dimension of variation.

To account for the inter-speaker variation just described, one would need to consider the individual multilingual profiles and language competences of the speakers. Speaker 1 and 2 stem from the same municipality, which is located remotely and is reported to have high language vitality (Schreiber 2016); both speakers are roughly of the same age group. Still, Speaker 1 shows more head-final syntax than Speaker 2. This can be explained by interference from Turkish head-final word orders due to a multilingual profile where Turkish is (at least at

the time of data collection) the dominant language. Gender is likely to account for the dominance of Turkish in the multilingual repertoire of (male) Speaker 1 as opposed to (female) Speaker 2. Usually, men are more exposed to Turkish than woman due to higher mobility for various reasons (like labour and military service, but also differences in education, see Schreiber 2016). However, it should be noted that individual language biographies can easily override this gender bias. Since unfortunately no detailed biographical information is available for Speaker 1 and 2, gender is initially assumed here as the decisive variable causing the different word order patterns.

As for the individual multilingual profile of (male) Speaker 3 who shows a more balanced word order with regard to pre-predicate and post-predicate elements, it has to be mentioned that apart from Turkish and Romeyka he is also competent in Modern Greek, which is likely to have in turn an influence on his data.

In sum, inter-speaker differences in the individual multilingual repertoires of the speakers are able to account for at least some of the variation in word order patterns found in Romeyka.

However, when it comes to copula complements (Table 5), the picture of inter-speaker variation is different with overall low numbers of post-predicate complements irrespective of the percentage of post-predicate other semantic roles. This suggests that there is indeed a change in word-order patterns in copula complements, which is reflected in the data by all three speakers. Moreover, all three speakers tend to omit some copulas; the explanation of this phenomenon, however, requires further research and could be potentially affected by the method of data elicitation for the WOWA dataset (see below).

Table 5: Copula complements in the WOWA dataset per speaker

	Total copula complements	Total VX	% VX	Omitted copulas
Speaker 1 (=text A)	19	2	10%	2
Speaker 2 (=text B, D, E)	4	0	0%	6
Speaker 3 (=text C)	3	1	33%	4

In order to be able to integrate the findings on word order variation in the present WOWA dataset correctly, some critical notes on the dataset and methodology of data collection are in order. Firstly, with 500 tokens, the present dataset is very small and on the lower edge of what can meaningfully be analysed by

quantitative means. It has also to be noted that the WOWA Romeyka dataset has a relatively high number of tokens (N=98) that could not be coded due to non-canonical constructions such as mixing with Turkish or elliptical constructions which also add to the picture of Romeyka as a shifting variety. Secondly, part of the five texts in the WOWA dataset were elicited by means of a storytelling task prompted by picture cards which could have had an influence on the data as well, especially with regard to the omission of copulas which was striking in the narratives elicited with the help of the picture cards. Finally, the fact that the five texts stem from three different speakers which are not directly comparable in terms of gender, age, speech community and their multilingual competences further complicates the analysis. It has to be mentioned, though, that with regard to diatopic variation, the three main dialect areas (Schreiber 2019) are not expected to show significant differences in the domain of word order, except for the potentially different amount of exposure to and use of Turkish, which is considered to be higher in the villages closer to the sea and lower in the more remote mountain villages (see Schreiber 2019).

7 Conclusion

The aim of the present chapter was to analyze the WOWA dataset of Romeyka (Schreiber 2021) with regard to word order in Romeyka in general and the grammatical entities that occur in the post-verbal domain. Romeyka has been described as having inherited VO word order, which has developed mixed directionality under contact with Turkish, which is visible especially in unmarked OV order in subordinate clauses (Neocleous 2020, 2022) but also in ongoing change in certain domains such as copula complements, although the diachronic picture is complex here. Furthermore, the analysis of the present WOWA dataset has revealed that inter-speaker variation mirrors ongoing language shift to Turkish, which complicates the attempt to define default word order patterns for the language and requires a nuanced methodology in assessing word order as if to allow for any meaningful conclusion. In general, information structure accounts for much of the variability in word order patterns that has been described above. Furthermore, PPs seem to behave differently than other obliques/objects, as well in locations and goals as in benefactives and comitatives.

To conclude, the picture of word order in Romeyka is by no means clearly deducible from quantitative data, as is indicated by the number of 55% overall frequency for post-verbal placement. There are several aspects accounting for the significant variation found in the present WOWA dataset: (i) the role of pragmatics, namely information structure, is highly relevant in Romeyka with at the

same time hardly a significant correlation of other factors such as semantic role, flagging or weight, although goals show a tendency to be post-predicate which seems to qualify as an areal (or otherwise for certain reasons universal?) pattern in other languages of the area as well; (ii) the high amount of inter-speaker variation which can be traced back to different levels of Turkish influence in the individual multilingual profiles of the speakers and reflects the status of Romeyka as a shifting variety; but also (iii) considerable intra-speaker variation as a characteristic of language shift; and (iv) ongoing language change as in the case of copula complements, be it internally caused and reinforced by language contact, which requires a very fine-grained and domain-specific investigation not only on the synchronic feature but involving diachronic developments as well as potential contact influences.

Abbreviations

ACC	accusative	NOM	nominative
AOR	aorist	OPN	object pronoun
AUX	auxiliary	PL	plural
CL	clitic	POSS	possessive pronoun
DEM	demonstrative	PRS	present tense
DP	discourse particle	PRT	aspectual particle
GEN	genitive	PST	past tense
INF	infinitive	REL	relativizer
IPF	imperfective	SG	singular
NEG	negation	TOP	topicalizer
NMZ	nominalization	1, 2, 3	1st, 2nd, 3rd person

Acknowledgements

The authors thank the Romeyka speakers and acknowledge financial support from the Alexander-von-Humboldt Foundation within the framework of the project on "Post-predicate Elements in Iranian: Inheritance, Contact, and Information Structure" (PI's: Geoffrey Haig, Mohammad Rasekh-Mahand) for field research in Trabzon/Çaykara in June/July 2019. Furthermore, our gratitude is due to an anonymous reviewer for their very helpful, detailed remarks.

References

Alexiadou, Artemis. 1997. *Adverb placement. A case study in antisymmetric syntax*. Amsterdam/Philadelphia: Benjamins. DOI: 10.1075/la.18.

Bağrıaçık, Metin. 2018. *Pharasiot Greek: Word order and clause structure*. Ghent: Ghent University. (Doctoral dissertation).

Behaghel, Otto. 1909. Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen* 25. 110–142.

Brendemoen, Bernt. 1998. Einige Bemerkungen über die Wortstellung in anatolischen Dialekten. In Nurettin Demir & Erika Taube (eds.), *Turkologie heute – Tradition und Perspektive: Materialien der 3. Deutschen Turkologen-Konferenz*, 27–46. Wiesbaden: Harrassowitz.

Brendemoen, Bernt. 1999. Greek and Turkish language encounters in Anatolia. In Bernt Brendemoen, Elizabeth Lanza & Else Ryen (eds.), *Language encounters across time and space: Studies in language contact*, 353–378. Oslo: Novus.

Brendemoen, Bernt. 2002. *The Turkish dialects of Trabzon: Their phonology and historical development*. Wiesbaden: Harrassowitz.

Brendemoen, Bernt. 2005. Sentence subordination in the Turkish Eastern Black Sea dialect: Another example of copying from Greek? In Dag Haug & Eirik Welo (eds.), *Haptačahaptāitiš: Festschrift for Fridrik Thordarson on the occasion of his 77th birthday*, 29–47. Oslo: Novus.

Brendemoen, Bernt. 2006. Aspects of Greek-Turkish language contact in Trabzon. In Hendrik Boeschoten & Lars Johanson (eds.), *Turkic languages in contact*, 63–73. Wiesbaden: Harrassowitz.

Brendemoen, Bernt. 2010. Some remarks on the copula in the Eastern Black Sea dialects. In Hendrik Boeschoten & Julian Rentzsch (eds.), *Turkology in Mainz/Turkologie in Mainz*, 47–54. Wiesbaden: Harrassowitz.

Brendemoen, Bernt. 2019. The Eastern Black Sea dialects of Turkey from the point of view of language contact. In Christos Tzitzilis & Georgos Papanastassiou (eds.), *Language contact in the Balkans and Asia Minor*, 248–261. Thessaloniki: Aristotle University of Thessaloniki.

Craevschi, Alexandru. 2022. *Historical contingency and typological tendencies in languages of Western Asia: A quantitative study of word order of non-subject constituents*. Bamberg: University of Bamberg. (MA thesis).

Dawkins, Richard M. 1916. *Modern Greek in Asia Minor*. Cambridge: Cambridge University Press.

Dawkins, Richard M. 1931. Notes on the study of the Modern Greek of Pontos. *Byzantion* 6. 389–400.

Dawkins, Richard M. 1937. The Pontic dialect of Modern Greek in Asia Minor and Russia. *Transactions of the Philological Society* 36. 15–52.

Deffner, Michael. 1878. Die Infinitive in den pontischen Dialekten und die zusammengesetzten Zeiten im Neugriechischen. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin* 1878. 191–230.

Drettas, Georges. 1997. *Aspects pontiques*. Paris: Association de Recherches Pluridisciplinaires.

Drettas, Georges. 1999. Το ελληνο-ποντιακό διαλεκτικό σύνολο [The Greek-Pontic dialectal set]. In Anastasios-Foivos Christidis, Maria Arapopoulou & Giannoula Giannoulopoulou (eds.), *Διαλεκτικοί θύλακοι της ελληνικής γλώσσας*, 91–100. Athens: Ministry of National Education & Religious Affairs/Centre for the Greek Language.

Dryer, Matthew S. 1995. Frequency and pragmatically unmarked word order. In Pamela A. Downing & Michael Noonan (eds.), *Word order in discourse*, 105–135. Amsterdam: Benjamins.

Dryer, Matthew S. 2013. Determining dominant word order. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/s6 (15 November, 2022).

Frascarelli, Mara & Roland Hinterhölzl. 2007. Types of topics in German and Italian. In Kerstin Schwabe & Susanne Winkler (eds.), *On information structure, meaning and form: Generalizations across languages*, 87–116. Amsterdam: Benjamins.

Gandon, Ophélie. 2016. *La relativisation dans une perspective aréale: l'aire Caucase – Anatolie de l'Est – Iran de l'Ouest*. Université Sorbonne Nouvelle. (Doctoral dissertation).

Göksel, Aslı & Celia Kerslake. 2005. *Turkish: A comprehensive grammar*. London/New York: Routledge.

Gundel, Jeanette K. & Thorstein Fretheim. 2004. Topic and focus. In Laurence Horn & Gregory Ward (eds.), *The handbook of pragmatics*, 175–196. Oxford: Blackwell.

Haig, Geoffrey, Mohammad Rasekh-Mahand, Donald Stilo, Laurentia Schreiber & Nils Schiborr. 2024. Post-predicate elements in the Western Asian Transition Zone: Data, theory, and methods. In Geoffrey Haig, Mohammad Rasekh-Mahand, Donald Stilo, Laurentia Schreiber & Nils N. Schiborr (eds.), *Post-predicate elements in the Western Asian Transition Zone: A corpus-based approach to areal typology*, 3–54. Berlin: Language Science Press. DOI: 10.5281/zenodo.14266331.

Hodgson, Katherine. 2024. Pontic Greek (Armenia). In Geoffrey Haig, Donald Stilo, Mahîr C. Doğan & Nils N. Schiborr (eds.), *WOWA — Word Order in Western Asia: A spoken-language-based corpus for investigating areal effects in word order variation*. Bamberg: University of Bamberg. https://multicast.aspra.uni-bamberg.de/resources/wowa/.

Holton, David, Geoffrey Horrocks, Marjolijne Janssen, Tina Lendari, Io Manolessou & Notis Toufexis. 2019. *The Cambridge grammar of medieval and early Modern Greek*. Cambridge: Cambridge University Press.

Holton, David, Peter Mackridge, Irene Phillipaki-Warburton & Vassilios Spyropoulos. 2012. *Greek: A comprehensive grammar of the modern language*. London/New York: Routledge.

Horrocks, Geoffrey. 1990. Clitics in Greek: A diachronic review. In Maria Roussou & Stavros Panteli (eds.), *Greek outside Greece, vol. 2*, 35–52. Athens: Diaspora Books.

Horrocks, Geoffrey. 2007. Syntax: From Classical Greek to the Koiné. In Anastasios F. Christidis & Maria Arapopoulou (eds.), *A history of Ancient Greek: From the beginnings to Late Antiquity*, 618–631. Cambridge: Cambridge University Press.

Horrocks, Geoffrey. 2019. Syntax. In David Holton, Geoffrey C. Horrocks, Marjolijne Janssen, Tina Lendari, Io Manolessou & Notis Toufexis (eds.), *The Cambridge grammar of medieval and early Modern Greek, vol. 4*, 1861–2040. Cambridge: Cambridge University Press.

Iefremenko, Kateryna. 2021. Oghuz (Ankara). In Geoffrey Haig, Donald Stilo, Mahîr C. Doğan & Nils N. Schiborr (eds.), *WOWA — Word Order in Western Asia: A spoken-language-based corpus for investigating areal effects in word order variation*. Bamberg: University of Bamberg. multicast.aspra.uni-bamberg.de/resources/wowa/#turkic.

Janse, Mark. 2002. Aspects of Pontic grammar. *Journal of Greek Linguistics* 3. 203–231.

Janse, Mark. 2006. Object position in Cappadocian and other Asia Minor Greek dialects. *Modern Greek Dialects and Linguistic Theory* 2. 115–129.

Janse, Mark. 2008. Clitic doubling from ancient to Asia Minor Greek. In Dalina Kallulli & Leon Tasmowski (eds.), *Clitic doubling in the Balkan languages*, 165–202. Amsterdam: Benjamins.

Janse, Mark. 2020. Back to the future: Akritic light on diachronic variation in Cappadocian (East Asia Minor Greek). In Klaas Bentein & Mark Janse (eds.), *Varieties of Postclassical and Byzantine Greek*, 201–239. Berlin: De Gruyter Mouton.

Janse, Mark. In press. Η καππαδοκική διάλεκτος [The Cappadocian dialect]. In Christos Tzitzilis (ed.), *Neoellinikes ðialekti [The Modern Greek dialects]*. Thessaloniki: Instituto Neoellinikon Spuðon.

Janse, Mark & Laurentia Schreiber. In prep. Word order and information structure in Romeyka.

Kaltsa, Maria & Ioanna Sitaridou. 2010. Topicalisation in Pontic Greek. *Modern Greek Dialects and Linguistic Theory* 4. 262–279.

Karatsareas, Petros. 2016. The Asia Minor Greek adpositional cycle. A tale of multiple causation. *Journal of Greek Linguistics* 16. 47–86.

Kutscher, Silvia. 2008. The language of the Laz in Turkey: Contact-induced change or gradual language loss. *Turkic Languages* 12. 82–102.

Lacroix, René. 2009. *Description du dialecte Laze d'Arhavi (Caucasique du Sud, Turquie)*. Lyon: Université Lumière Lyon 2. (Doctoral dissertation).

Lacroix, René. 2019. Laz. In Geoffrey Haig & Geoffrey Khan (eds.), *The languages and linguistics of Western Asia: An areal perspective*, 830–860. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110421682-021.

Leitner, Bettina. 2024. Khuzestani Arabic. In Geoffrey Haig, Mohammad Rasekh-Mahand, Donald Stilo, Laurentia Schreiber & Nils N. Schiborr (eds.), *Post-predicate elements in the Western Asian Transition Zone: A corpus-based approach to areal typology*, 413–430. Berlin: Language Science Press. DOI: 10.5281/zenodo.14266357.

Mackridge, Peter. 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.

Neocleous, Nicolaos. 2020. *Word order and information structure in Romeyka: A syntax and semantics interface account of order in a minimalist system*. Cambridge: University of Cambridge. (Doctoral dissertation).

Neocleous, Nicolaos. 2022. The evolution of VO and OV alternation in Romeyka. In Hiwa Asadpour & Thomas Jügel (eds.), *Word order variation: Semitic, Turkic and Indo-European languages in contact*, 141–161. Berlin: De Gruyter Mouton.

Neocleous, Nicolaos & Ioanna Sitaridou. 2022. Never just contact: The rise of final auxiliaries in Asia Minor Greek. *Diachronica* 39. 369–408.

Öztürk, Balkız & Markus A. Pöchtrager. 2011. *Pazar Laz*. München: Lincom Europa.

Papadopoulos, Anthimos A. 1933. Συμβολὴ εἰς τὴν ἔρευναν τῆς ποντικῆς διαλέκτου (ἤτοι γραμματικὰ καὶ συντακτικὰ ζητήματα) [A contribution to the study of the Pontic dialect: Grammatical and syntactic issues]. *Ἀθηνᾶ* 45. 15–48.

Papadopoulos, Anthimos A. 1955. *Ἱστορικὴ γραμματικὴ τῆς ποντικῆς διαλέκτου [Historical grammar of the Pontic dialect]*. Athens: Ἐπιτροπὴ Ποντιακῶν Μελετῶν [Commission of Pontic Studies].

Papadopoulos, Anthimos A. 1958–1961. *Ἱστορικὸν λεξικὸν τῆς ποντικῆς διαλέκτου [Historical dictionary of the Pontic dialect]*, vol. 1–2. Athens: Myrtidis.

Parcharidis, Ioannis. 1880. *Γραμματική της διαλέκτου Τραπεζουντος. Περί επιθέτων, αντωνυμιών και βαρυτόνων ρημάτων [Grammar of the Trabzon dialect: On adjectives, pronouns and baritone verbs]*, Manuscript No. 335. Athens: Κέντρο Ερεύνης των Νεοελληνικών Διαλέκτων και Ιδιωμάτων [Research Centre of Modern Greek Dialects & Idioms], Academy of Athens.

Parcharidis, Ioannis. 1888. Συλλογή ζώντων μνημείων της αρχαίας ελληνικής γλώσσας εν Όφει' [Collection of living monuments of the ancient Greek language in Ofi']. *Ο εν Κωνσταντινουπόλει Ελληνικός Φιλολογικός Σύλλογος [The Greek Philological Association in Constantinople]* 18. 120–178.

Rafiyenko, Dariya & Ilja A. Seržant. 2020. Postclassical Greek: An overview. In Dariya Rafiyenko & Ilja A. Seržant (eds.), *Postclassical Greek: Contemporary approaches to philology and linguistics*, 1–15. Berlin: De Gruyter Mouton.

Ralli, Angela. 2006. Syntactic and morphosyntactic phenomena in Modern Greek dialects: The state of the art. *Journal of Greek Linguistics* 7. 121–159.

Ralli, Angela. 2020. Selected topics of the morphology of Asia Minor Greek: An introduction. In Angela Ralli (ed.), *The morphology of Asia Minor Greek*, 1–19. Leiden: Brill.

Rasekh-Mahand, Mohammad, Elham Izadi, Mehdi Parizadeh, Geoffrey Haig & Nils Schiborr. 2024. Post-predicate elements in modern colloquial Persian: A multifactorial analysis. In Geoffrey Haig, Mohammad Rasekh-Mahand, Donald Stilo, Laurentia Schreiber & Nils N. Schiborr (eds.), *Post-predicate elements in the Western Asian Transition Zone: A corpus-based approach to areal typology*, 197–230. Berlin: Language Science Press. DOI: 10.5281/zenodo.14266343.

Revithiadou, Anthoula & Vasilis Spyropoulos. 2009. *Οφίτικη ποντιακή: έρευνα γλωσσικής καταγραφής με έμφαση στη διαχρονία και συγχρονία της διαλέκτου [Ophitic Pontic: Language inventory survey with emphasis on the diachrony and synchrony of the dialect]*. Athens: John S. Latsis Public Benefit Foundation.

Schreiber, Laurentia. 2016. *Assessing sociolinguistic vitality: An attitudinal study of Rumca (Romeyka)*. Heike Wiese, Heiko F. Marten & Oliver Bunk (eds.) (Arbeitspapiere "Sprache, Variation und Migration": Studentische Arbeiten 4). Potsdam: Universität Potsdam. http://www.uni-potsdam.de/fileadmin01/projects/svm/Arbeitspapiere/MA_Schreiber_May.pdf (17 November, 2017).

Schreiber, Laurentia. 2019. Romeyka. In Geoffrey Haig & Geoffrey Khan (eds.), *The languages and linguistics of Western Asia: An areal perspective*, 892–933. Berlin: De Gruyter Mouton.

Schreiber, Laurentia. 2021. Pontic Greek (Romeyka). In Geoffrey Haig, Donald Stilo, Nils N. Schiborr & Mahîr Dogan (eds.), *WOWA — Word Order in Western Asia: A spoken-language-based corpus for investigating areal effects in word order variation*. Bamberg: University of Bamberg. multicast.aspra.uni-bamberg.de/resources/wowa/.

Schreiber, Laurentia. 2024. *A (Contact-)Grammar of Romeyka*. Bamberg/Ghent: University of Bamberg / Ghent University. (Doctoral dissertation). DOI: 10.20378/irb-95115.

Schreiber, Laurentia. Submitted. Muslim Pontic Greek in Turkey: Language shift and bilingual language use. In Konstantinos Sampanis, Matthias Kappler & Eleni Karantzola (eds.), *Sociocultural aspects of Greek-Turkish language contacts*. Wiesbaden: Harrassowitz.

Schreiber, Laurentia. In prep. The Romeyka corpus.

Schreiber, Laurentia & Ioanna Sitaridou. 2017. Assessing the sociolinguistic vitality of Istanbulite Romeyka: An attitudinal study. *Journal of Multilingual and Multicultural Development* 39. 1–16.

Setatos, Michalis. 1994. Το ποντιακό πα [The Pontic pa]. *Ελληνική Διαλεκτολογία* 4. 39–48.

Sitaridou, Ioanna. 2013. Greek-speaking enclaves in Pontus today: The documentation and revitalization of Romeyka. In Mari Jones & Sarah Ogilvie (eds.), *Keeping languages alive: Documentation, pedagogy, and revitalization*, 98–112. Cambridge: Cambridge University Press.

Sitaridou, Ioanna. 2014a. Modality, antiveridicality, and complementation: The Romeyka infinitive as a negative polarity item. *Lingua* 148. 118–146.

Sitaridou, Ioanna. 2014b. The Romeyka infinitive: Continuity, contact and change in the Hellenic varieties of Pontus. *Diachronica* 31. 23–73.

Sitaridou, Ioanna. 2016. Reframing the phylogeny of Asia Minor Greek: The view from Pontic Greek. *CHS Research Bulletin* 4. 1–17.

Sitaridou, Ioanna. 2021. (in)Vulnerable inflected infinitives as complements to volitionals and modals: Evidence from Galician and Romeyka. In Thórhallur Eythórsson & Jóhannes Gísli Jónsson (eds.), *Syntactic features and the limits of syntactic change*, 161–177. Oxford: Oxford University Press.

Soltic, Jorie. 2013. Late Medieval Greek πάλιν: A discourse marker signalling topic switch. *Greek, Roman & Byzantine Studies* 53. 390–419.

Soltic, Jorie. 2015. *The Late Medieval Greek πολιτικὸς στίχος poetry. Language, metre and discourse*. Ghent: University of Ghent. (Doctoral dissertation).

Stilo, Donald & René Lacroix. 2021. Laz (Arhavi). In Geoffrey Haig, Donald Stilo, Mahîr C. Doğan & Nils N. Schiborr (eds.), *WOWA — Word Order in Western Asia: A spoken-language-based corpus for investigating areal effects in word order variation*. Bamberg: University of Bamberg. multicast.aspra.uni-bamberg.de/resources/wowa/ (27 July, 2023).

Thomason, Sarah G. 2001. *Language contact*. Edinburgh: Edinburgh University Press.

Tombaidis, Dimitrios E. 1992. Η τύχη των μικρασιατικών ιδιωμάτων στον ελληνικό χώρο [The fate of Asia Minor idioms in the Greek area]. *Centre for Asia Minor Studies Bulletin* 9. 241–250.

Tombaidis, Dimitrios E. 1996. Η ποντιακή διάλεκτος [The Pontic dialect]. In Dimitrios Tombaidis (ed.), *Μελετήματα ποντιακής διαλέκτου*, 222–233. Thessaloniki: Kodhikas.

van Emde Boas, Evert, Albert Rijksbaron, Luuk Huitink & Mathieu de Bakker. 2019. *The Cambridge grammar of Classical Greek*. Cambridge: Cambridge University Press.

Vaux, Bert. 2007. Homshetsma: The language of the Armenians of Hamshen. In Hovann Simonian (ed.), *The Hemshin: History, society and identity in the highlands of Northeast Turkey*, 257–278. London/New York: Routledge.