

# Chapter 9

## Balkan Turkic as a model for understanding contact-induced change in Turkish

Cem Keskin

University of Potsdam, Leibniz-Centre General Linguistics (ZAS)

How good a model is the Standard Average European–Turkic contact in the Balkans for understanding contact-induced syntactic changes in Turkish in the West? Patterns of head directionality change and innovations in clause combining replicated across several contact situations suggest that the answer to this question is a positive superlative. However, there is no single homogeneous response that Turkic varieties give under language contact but a complex pattern. Also the causal factors behind this pattern are fairly general in nature. So, a more reasonable conclusion that suggests itself is that Balkan Turkic constitutes a good model for understanding contact-induced syntactic change in Turkish in the West insofar as it can help reveal general contact dynamics, not as a Turkic contact variety per se. And on that front the in-depth investigation of Balkan Turkic promises to deliver useful observations.

### 1 Introduction

As is well-known, a significant new era of intensive Standard Average European<sup>1</sup> (SAE)–Turkic contacts can be said to have begun with the arrival of the so-called guest workers in northwest Europe in the late 1950s and early 1960s (Küçükcan & Güngör 2009: 83–84). This new contact constellation naturally stimulated a large

---

<sup>1</sup>The term *Standard Average European* (originally from Whorf 1944) refers in recent studies to a proposed sprachbund that spans Romance, Germanic, Balto-Slavic, the Balkan languages, etc. (see e.g., Haspelmath 1998, 2001, van der Auwera 2011).



body of linguistic research on the changes that were beginning to be observed in Turkish in the region.

One lacuna in this literature has been the lack of references to other earlier SAE–Turkic contact situations as potential models of contact-induced change in Turkic that could help us understand emerging changes in Turkish in the new contact constellation in northwest Europe. This would have been a reasonable approach, given that Turkic varieties have been in contact with SAE languages for several centuries, mostly in the Balkan sprachbund, and have undergone extensive well-documented changes (see Johanson 2021: Ch. 10 for an overview and Johanson 1992, 2013 for a detailed account).

This is the gap that I attempt to fill in this chapter. More precisely, I try to answer the question of how good a model other SAE–Turkic contact situations (more specifically SAE–Turkic contact in the Balkans) can be when trying to understand contact-induced syntactic changes in Turkish in the West. Let me reveal the answer from the beginning: Balkan Turkic (BT) constitutes a good model for understanding contact-induced syntactic change in Turkish in the West insofar as it can help reveal general contact dynamics, not necessarily by virtue of being a Turkic contact variety.<sup>2</sup> In the process of detailing this answer, I will be presenting three sets of findings: (i) those on contact-induced changes in BT,<sup>3</sup> (ii) those on contact-induced changes in heritage Turkish as spoken in Germany and the US and the Turkish of Kurmanji–Turkish bilinguals, and (iii) monolingual Turkey’s Turkish.<sup>4</sup>

The chapter is made up of two main sections, which will follow the next section on data sources, method, and statistics: Section 3 is on contact-induced head directionality changes, and Section 4 details contact-induced changes in clause combining with a focus on subordination. Both sections begin with data from BT, later turning to data from the other contact varieties and showing the connections between the former and the latter.

---

<sup>2</sup>I prefer to use the more general term *Balkan Turkic*, rather than the *Balkan Turkish* of most works in Turkological literature (e.g. Johanson 2021). The reason for this is the following: Even though the Southwest (or Oghuz) Turkic varieties spoken in the Balkans, which the present paper focuses on, are very closely related to Turkish varieties in the east, it is not entirely clear to what degree they have diverged and so whether at least some should be considered separate languages or can still be seen as dialects of Turkish.

<sup>3</sup>Note, however, that I will not be concerned with the so-called Balkanisms here (see e.g. Joseph 2020: 540–541), but only with the more general SAE features seen in BT.

<sup>4</sup>It should be emphasized that the latter two sets of findings are the outcome of extensive collaboration with Kateryna Iefremenko, Christoph Schroeder, and to a lesser extent Jaklin Kornfilt.

## 2 Data sources, method, and statistics

The data used in the analysis of BT come from the Balkan Turkic Corpus (Keskin et al. in preparation). The texts in this corpus, totaling around 80,000 words, were culled from the following sources:

1. Dialect texts
  - a) West Rumelian Turkish
    - i. Kosovar Turkish: personal accounts in Sulçevsi (2019)
    - ii. Macedonian Turkish: folk tales in Destanov (2016) and Kakuk (1972), folk tales and personal accounts in Katona (1969)
    - iii. Western Bulgarian Turkish: folk tales and accounts of traditions in Kakuk (1961a,b)
  - b) North Rumelian Turkic
    - i. Gagauz: Folk tales selected from numerous sources and published by Özkan (2007)
    - ii. Dobruja Turkish: miscellaneous texts in Haliloğlu (2017)
  - c) East Rumelian Turkish: folk tales and accounts of traditions in Hazai (1960) and Kakuk (1958)<sup>5</sup>
2. Historical texts that show Early Balkan Turkic features (the so-called “transcription texts”)
  - a) 14th century: Schiltberger’s *Our Father* published by Helmholdt (1966)
  - b) 15th century: Yusof and Jakob Papas’ letters published by Brendemoen (1980), Pietro Bruto and Hadriano Fino’s Bible verses in Weil (1953)

---

<sup>5</sup>These two sources provide only two small samples from central and southern Bulgaria, selected as representatives of East Rumelian as a whole. Data from the Turkish of Western Thrace (see e.g. Petrou 2021) and elsewhere were not included in the East Rumelian part of the corpus for two reasons. First, the East Rumelian group does not show the syntactic changes that were of interest to the research on which this chapter is based, so a small control sample was judged to be sufficient. Second, some of these texts (particularly the more recent ones in Petrou 2021) were not available during the text collection phase of the corpus. More East Rumelian data are planned to be added in a later version of the corpus.

- c) 16th century: Filippo Argenti's phrases in Adamović (2001), Bartholomaeus Georgievits' dialogue, Our Father, the Apostles' Creed, etc. in Heffening (1942), Marco Antonio Begliarmati's dialogue in Teza (1892), the anonymous phrases in Adamović (1976), Guillaume Postel's phrases in Drimba (1966), Reinhold Lubenau's phrases in Adamović (1977)
- d) 17th century: Pietro Ferraguto's dialogue in Bombaci (1940), Giovan Battista Montalbano's sayings in Gallotta (1986), sample text in Du Ryer (1630), the anonymous dialogue in Blau (1868), Miklós Illésházy's dialogue in Németh (1970), dialogue and Our Father in Herbinus (1675)

These texts were first coded sentence by sentence using a sentence annotation interface for the features below, and the features were stored in a database:

1. Directionality: bare object–verb versus verb–bare object, oblique–verb versus verb–oblique, etc. (a total of 21 pairs of opposing features)
2. Clause type: main, argument, relative, adverbial
3. Finiteness: finite, non-finite
4. Metadata: century, author, genre, provenance

The texts were then analyzed based on directionality, clause type, finiteness, century, and provenance, using a query module that uses the stored grammatical properties.

In addition to Balkan Turkic data, four other sets of data were collected for the studies reported on here following the Language Situations method by Wiese (2020): (i) Turkish with German contact in Germany ( $n = 65$ ), (ii) Turkish with English contact in the US ( $n = 61$ ), (iii) with Kurmanji contact in Turkey ( $n = 30$ ), and (iv) monolingual Turkish ( $n = 66$ ). The Language Situations method brings controlled elicitation together with spontaneous speech and is suited to comparing various contact constellations. It allows for the collection of quasi-naturalistic data across different communicative situations, including formal and informal settings and written and spoken communication. In accordance with this method, participants were shown the video of a car accident and asked to describe it in four different imagined scenarios: (i) WhatsApp voice message to a friend, (ii) WhatsApp text message to a friend, (iii) voicemail to the police, and (iv) written witness report to the police. Each bilingual participant had two

sessions, one in their heritage language and one in the majority language of the society in which they lived.

A few remarks are in order about these data sources. First, the historical sources in the Balkan Turkic Corpus comprise almost all the transcription texts available to date. As the volume of these texts is fairly low, the only selection criterion applied (in addition to whether or not the text showed BT features) was the intelligibility of the author's orthography. Also, some of these sources are religious texts (e.g. translations of the Bible, etc.), and one might wonder whether these texts and their translations are inclined towards contact forms. Indeed, the religious texts in the corpus do have a stronger tendency towards contact forms than the secular sources, but the skewing effect of this can be considered negligible, as religious texts constitute only about 10% of the sources in the corpus in terms of word count. (Genre in general seems not to have a significant effect, but word count per genre is too low to reach a definitive conclusion.) Second, one might wonder whether it is legitimate to compare the material in the Balkan Turkic Corpus with the data collected using the "Language Situations" method. To clarify, the non-Gagauz material in the corpus consists of transcripts of spontaneous oral productions of informants, while the Gagauz texts seem to have been minimally edited for readability and preserve the original style of the narrators as much as possible. So, even though the two sets of data are not products of identical methodologies, this is as close as one can presently get in such a comparative study, and that seems to be an appropriate approximation.

As for statistical tests and basic mathematical operations, three separate tools were used for these: (i) Lancaster Stats Tools online that runs R code (Brezina 2018), (ii) R (version 4.3.0) (R Core Team 2021), and (iii) MS Excel (version 2202).

### 3 Contact-induced head directionality changes

#### 3.1 From head-final to head-initial

A number of syntactic changes associated with a possible shift from head-final to head-initial syntax that appear to be taking place in heritage Turkish in contact with German and English have been subject to research. Some of this research focuses on the use of the postverbal position in the canonically verb-final Turkish (see e.g. Iefremenko 2021c, Schroeder 2018, Schroeder & Iefremenko 2019b,c, Schroeder et al. 2024, see also Bunk et al. 2025 [this volume] and Iefremenko 2019, 2021a,b) and innovative clause combining strategies. The latter I will defer to Section 4 and focus here on directionality in the verbal domain, supplementing

my observations with additional data from noun phrases, adpositional phrases, and clauses.

If a shift from head-final to head-initial syntax is indeed the path that heritage Turkish is treading, it would not be the first time in the Turkic family, Macedonian Turkish (see e.g. Matras & Tufan 2007), Gagauz (see e.g. Menz 1999), Karaim (see e.g. Csátó 2000), and Urum (see e.g. Böhm 2015, Skopeteas 2015) being its most well-known members which already show features of the shift, to varying degrees, as can be gleaned from the literature (see e.g. Balci 2010, Doerfer 1959: 271, Friedman 1982: 33–35, Friedman 2006: 40–41, Gülsevin 2017: 110–111, Günşen 2010: 471–475, İğci 2010: 129, Jable 2010: 148–49, Johanson 2021: 790, 936, Katona 1969: 165, Kirli 2001, Matras 2009: 251, Matras & Tufan 2007, Menz 1999: 40, Menz 2014: 61–62, Özkan 1996: 209–210; see also Petrou 2019: 226 and Petrou 2021: 342 for some initial observations on the shift to head-initial syntax in the Turkish of Western Thrace in contact with Greek).

The current state of heritage Turkish, then, appears to reveal the initial stages of a syntactic shift in the same direction as the head-initial varieties of Turkic, and is probably comparable to (i) Turkic contact varieties in their earlier periods and (ii) Turkic contact varieties that are more on the head-final pole of a range that extends to head-initiality. Thus, a systematic comparative approach promises to help us map out possible trajectories of syntactic change in Turkic, provide insights into the changes taking place in heritage Turkish at present, and perhaps even predict the changes that might occur in the future.

Following this rationale, I will be exploiting, in what follows, two avenues for comparison between BT and heritage Turkish: (i) a dialectological angle that investigates the available range of modern BT varieties with varying preferences for head-final versus head-initial orders, and (ii) a diachronic angle with an eye on older forms of BT on the way to head-initiality.

However, as an anonymous reviewer rightly points out, BT and heritage Turkish in Germany or the US are and have been in very different sociolinguistic settings, meaning that it would be unreasonable to expect wholesale convergences between them. I will mention one sociolinguistic consideration. The contact varieties in the West maintain contacts with Turkey's Turkish, and their speakers still regard the latter as the "proper" variety – even if this proper variety cannot fully assert its influence. Even with these reservations in mind (e.g. the linguistic orientation of later generations, the degree of access to the written standard), this situation is very different than that of BT, for which there has never been a strong Turkic standard towards which the speakers could orient themselves, even during the Ottoman era (Johanson 1989). These differences will no doubt

have their consequences with respect to the patterns that the newer contact varieties will show in the longer term, as well as the timelines of these patterns. Yet, I believe there is still some merit in comparing these contact varieties simply by the fact of their all being Turkic varieties in contact with SAE, and, as we will see during the discussion of the findings, the results seem to justify this approach.

After this introductory discussion, I now move on to detailed descriptions of contact-induced head directionality changes in the relevant contact varieties where I will present detailed data on the constituent orders of 13 head and dependent pairs. As will soon become clear, the data show substantial amounts of variation across syntactic structures and varieties, and reveal spectrums of head directionality (Sections 3.2.1 and 3.2.2), partly brought about by an interplay of progressive and conservative syntactic domains (Section 3.2.3).<sup>6</sup> In addition to this variation, another aspect of these syntactic patterns is what we could call *variability*, which can be briefly described as follows: Progressive and conservative domains are not equally so. As dialects and syntactic pairs shift from head-final to head-initial order, they diverge from each other in the degree to which they do that, presenting a heterogeneous pattern of changes (Section 3.2.4). This multi-faceted pattern can be attributed to the interaction of several factors, such as degree of contact and the differential responses of syntactic structures (Section 3.3).

## 3.2 Change, variation, and variability

### 3.2.1 Two spectrums of directionality in Balkan Turkic

Then, let us begin by addressing the following question: What kind of word order changes does BT show due to SAE contact? As pointed out in the introductory discussion, the answer to this question will set a standard of comparison for the Turkish contact varieties in the West. But before getting into the details of the answer, let me provide in (1) an illustrative example of the kinds of changes that I will be discussing:

- (1) a. Ben ak ekmek al-dī-m.  
1SG white bread buy-PST-1SG  
'I bought some white bread.' (Hazai 1960: 187)
- b. Çocuk al-mış tuz.  
child take-EVID.3SG salt  
'The child took some salt.' (Özkan 2007: 137)

---

<sup>6</sup>The terms *progressive*, *advanced*, *conservative*, etc. contain no value judgments and are intended only to describe the degree of and the responses to the shift to head-initial order.

Example (1a) is from the dialect group referred to as *East Rumelian Turkish*, which is spoken in the southeastern Balkans and known for its conservative syntactic features (see e.g. Johanson 2021: 49). The example instantiates canonical Turkic word order in the verbal domain with the bare direct object *ak ekme* ‘white bread’ preceding the verb *al* ‘buy’. It exemplifies, in other words, the dependent–head constituent order typical of Turkic. Example (1b) is from Gagauz, spoken mostly in Moldova and known for its pervasive contact-induced syntactic features (see e.g. Johanson 2021: 50–51). It contains a head-initial VP, in which the direct object *tuz* follows the verb *al* ‘take’. That is to say, it exemplifies the head–dependent constituent order commonly seen in Turkic varieties going through contact-induced word order shifts.

Comparable shifts (as in Gagauz) or lack thereof (as in East Rumelian) form an overall pattern summarized in Table 1, which shows the distribution of thirteen head–dependent orders in main clauses across BT (see Keskin 2023a for another, similar account of word order changes in BT).

The head–dependent combinations are indicated in the top row.<sup>7</sup> They are sorted on the total percentages of head-initial orders across 25 BT dialect locales (given in the bottommost row) in increasing order from left to right. Each cell indicates the percentage of head-initial order per given head+dependent combination in a given dialect locale. Dialect locales are indicated in the leftmost column, sorted on the total percentages of all head-initial orders observed in those locales (given in the rightmost column) in increasing order from top to bottom.

Table 1 immediately makes the intricacies involved in answering the question above clear. Two spectrums of directionality can be discerned in the pattern it displays, contingent upon (i) type of head and dependent, and (ii) dialect locale. The interaction of these two spectrums, thus, brings about a rather complex pattern of change.<sup>8</sup>

Let us begin with the spectrum of head–dependent pairs. First, we see that there is a gradual transition from very low to very high incidences of head–dependent orders as we move rightward – a transition from head-final to head-initial order, in other words. Clausal dependents (i.e. relative clauses, finite complement and adverbial clauses, and non-finite complement and adverbial clauses)

<sup>7</sup>NAdj=noun–adjective, NGen=noun–genitive, PN=adposition–noun, VAdv=verb–adverb, VSubj=verb–subject, VPP=verb–adpositional phrase, VBar=verb–bare object, VNFin=verb–non-finite dependent clause, VAcc=verb–accusative-marked object, VObl=verb–oblique object, VFin=verb–finite dependent clause, CCl=connector–dependent clause, NRC=noun–relative clause, HD=head–dependent order

<sup>8</sup>Color coding was used in the table for ease of interpretation. Gray cells show cases for which I only have a few examples.



Table 1: Head-dependent orders in Balkan Turkic

	HEAD-FINAL							⇒	HEAD-INITIAL							
	NAdj	NGen	PN	VAdv	VSubj	VPP	VBar	VNFin	VAcc	VObl	VFin	CCI	NRC	HD		
Mamushë	0%	3%	7%	0%	17%	27%	9%	80%	38%	24%	43%	56%	83%	16%		
Kazanlāk	0%	1%	0%	3%	13%	25%	16%	17%	28%	34%	41%	32%	0%	18%		
Gjilani	3%	0%	0%	3%	3%	11%	26%	0%	63%	18%	56%	94%	100%	18%		
Kārdžali	0%	2%	0%	12%	23%	25%	15%	9%	48%	44%	15%	32%	0%	20%		
Resen	0%	4%	0%	13%	5%	18%	8%	50%	22%	29%	76%	69%	67%	23%		
Razgrad	0%	6%	0%	11%	16%	24%	14%	25%	44%	45%	50%	58%	33%	24%		
Prizreni	3%	8%	6%	6%	19%	44%	17%	0%	38%	47%	50%	94%	100%	26%		
Ohrid	0%	0%	0%	11%	9%	13%	31%	40%	18%	25%	68%	85%	50%	26%		
Jeni Maale	0%	9%	0%	26%	24%	36%	39%	38%	62%	46%	41%	63%	0%	28%		
Ali Koč	0%	0%	0%	11%	21%	67%	14%	36%	46%	63%	29%	65%	0%	29%		
Janevë	3%	0%	0%	6%	41%	67%	41%	100%	50%	68%	55%	75%	100%	31%		
Prishtinë	3%	3%	5%	10%	11%	38%	58%	0%	59%	57%	71%	75%	100%	32%		
Konče	0%	11%	0%	19%	24%	22%	37%	33%	68%	53%	56%	69%	50%	34%		
Dobržane	0%	13%	9%	24%	15%	35%	54%	0%	65%	55%	77%	62%	100%	36%		
Pejë	3%	4%	25%	29%	2%	50%	84%	100%	56%	65%	27%	81%	50%	36%		
Chişinău	0%	3%	6%	12%	17%	38%	38%	59%	63%	57%	74%	88%	100%	36%		
Vushtrri	2%	14%	8%	9%	33%	50%	67%	50%	45%	72%	73%	91%	100%	38%		
Montana	1%	4%	0%	23%	32%	14%	35%	17%	69%	70%	77%	82%	100%	39%		
Kyustendil	2%	14%	0%	5%	22%	57%	41%	43%	70%	76%	73%	78%	25%	40%		
Mitrovica	13%	7%	20%	18%	18%	70%	55%	0%	64%	58%	72%	75%	100%	41%		
Bučim	2%	30%	0%	39%	25%	20%	34%	60%	65%	63%	68%	84%	33%	43%		
Odesa	0%	0%	0%	18%	24%	33%	50%	50%	79%	81%	77%	90%	75%	45%		
Bessarabia	0%	5%	0%	25%	19%	67%	74%	53%	73%	74%	85%	73%	33%	45%		
Ossetia	0%	6%	33%	27%	31%	33%	57%	69%	74%	88%	55%	76%	100%	47%		
Tomai	0%	22%	30%	20%	42%	33%	67%	63%	79%	70%	80%	100%	100%	51%		
HD	1%	5%	5%	14%	18%	35%	35%	46%	53%	54%	59%	74%	74%			

appear to be the most progressive constituents as a category. (Non-finite clauses, however, are markedly less progressive.) By contrast, adpositional phrases and noun phrases with non-clausal dependents are the most conservative group. Finally, verbs with non-clausal dependents (i.e. obliques, bare objects, etc.) can be said to be distributed over the middle range.

As for the spectrum of dialects, we see that some dialects, such as Kārdžali and Kazanlāk (members of the dialect group East Rumelian Turkish), are conservative. Some others, such as Tomai and Chişinău (varieties of Gagauz), on the other hand, are progressive.

### 3.2.2 Spectrums of directionality in other contact varieties

Let us now turn to the word order changes in the contact varieties in the West. To repeat, these are Turkish with German contact in Germany (abbreviated as *DEbi*) and with English contact in the US (*USbi*) (which are considered heritage languages). I will also include Turkish with Kurmanji contact in Turkey<sup>9,10</sup> (*TUbi*, spoken by Kurmanji–Turkish bilinguals, not a heritage language) and Early Balkan Turkic (EBT, 14th–17th c.)<sup>11</sup> in the analysis for reference, and will refer to *DEbi*, *USbi*, and *TUbi* as the *new contact varieties* (NCV). The findings are summarized in Table 2 (see also Iefremenko et al. 2023 for another similar analysis).

To begin, it is interesting to note that, broadly speaking, the same pattern of word order changes replicates itself across four different contact situations (Table 2a). As in BT, noun and adposition phrases are the most conservative domains

---

<sup>9</sup>Some grammatical features of Kurmanji that are relevant for this study are as follows (extracted from the relevant sections in Haig & Öpengin 2018 and McCarus 2009). In the nominal domain, possessors, adjectives, and relative clauses follow, while demonstratives and numerals precede the head noun. In the verbal domain, obliques and complement clauses follow the verb, while direct objects precede it. Adverbs can precede or follow the verb, and adverbial clauses can precede the main clause or be postverbal. Conjunctions that introduce clauses are clause-initial. The language makes use of prepositions and circumpositions, as well as locational nouns that precede their dependents.

<sup>10</sup>Turkish and Kurmanji have been in contact since at least the 11th century, but the extent of present-day contact is unprecedented, determined by the language policies of the Turkish state in the post-1923 period in conjunction with the rise of mass media (see e.g. Bulut 2006: 96, 116, Dorleijn 2006: 74, Haig 2017: 398–399, Haig & Öpengin 2018: 222–225, Johanson 2006: 8, Varol 2014: 59, 65). To cite specific but limited figures which point in this direction, the percentage of Kurmanji–Turkish bilinguals in Turkey seems to have increased from around 26% to around 40% just in the 1935–1965 period, the only period from which we have relevant data, collected during the censuses (census data available in Dündar 2000: 171–216). For this reason, I assume that the phenomena that we observe in the Turkish contact variety spoken by Kurmanji–Turkish bilinguals are by and large relatively new developments.

<sup>11</sup>An anonymous reviewer questions the power of transcription texts to represent EBT. The question of whether transcription texts can be considered reliable sources for the historical linguistics of Turkish is a valid one. In the field of Turkology during the 1960s and early 1970s, there was a debate on the extent to which the Turkish language observed in transcription texts (particularly Georgievits 1544 published by Heffening 1942 and Illésházy 1668 published by Németh 1970) could accurately represent a specific Turkish variety (see e.g. Hazai 1990: 64–67, Stein 2016: 161–162). Németh (1968, 1970) argued that these texts were representative of the Balkan dialects of Turkish. On the other hand, Kissling (1968) claimed that the texts contained linguistic mixtures unique to themselves and reflected an imperfectly learned Turkish influenced by the Balkan region. Ultimately, the debate was resolved in favor of the former position, and as a result, a body of scholarly work emerged that utilizes transcription texts as reliable sources for conducting historical linguistic studies of Turkish (see e.g. Csató et al. 2016).

Table 2: Head-dependent orders in NCV and EBT

(a) New contact varieties

	HEAD-FINAL							⇨	HEAD-INITIAL							
	NGen	NAdj	PN	NRC	VAcc	VSubj	VBar	VNFin	VObl	VAdv	VPP	CCI	VFin	HD		
TUbi	0%	0%	0%	0%	4%	5%	4%	5%	8%	12%	21%	17%	45%	8%		
USbi	0%	0%	0%	0%	3%	2%	7%	9%	18%	11%	19%	23%	68%	9%		
DEbi	0%	0%	0%	3%	4%	10%	8%	12%	13%	18%	19%	30%	42%	12%		
HD	0%	0%	0%	1%	4%	6%	6%	8%	13%	14%	20%	21%	48%			

(b) Early Balkan Turkic

	PN	NAdj	VSubj	NGen	VAdv	VBar	VPP	VAcc	VObl	VNFin	NRC	VFin	CCI	HD
EBT	4%	5%	5%	6%	12%	14%	22%	32%	38%	46%	50%	70%	86%	30%

in NCV. Relative clauses, however, present a major difference between BT and NCV. BT tends to use SAE-type finite postpositive relative clauses, the high incidence of noun–relative clause (NRC) order in Table 1 being a consequence of that. As these have not yet clearly emerged in NCV, almost all relative clauses in use are non-finite prepositive Turkish relatives. Next, again as with BT, finite clausal dependents are progressive, with the difference between finite and non-finite clauses being replicated here. Finally, non-clausal dependents of verbs can again be said to be distributed over the middle range.

The pattern just described obtains not only across four contact situations but across different historical periods as well. Consider the distribution of head-dependent orders in EBT given in Table 2b. As before, noun phrases with non-clausal dependents and adpositional phrases are maximally conservative. Note that relative clauses in EBT behave more like those in modern BT, as SAE-type relative clauses had already emerged in BT by that period. (The relatively less progressive position they occupy in the EBT spectrum may be symptomatic of the earlier stage of this development.) Also, clausal dependents are the most progressive elements and non-clausal dependents of the verb occupy the middle of the spectrum. Note, finally, that the asymmetry between finite and non-finite clauses can also be observed in EBT.

Turning now to the dialect spectrum: There are some differences between the three new contact varieties, but they do not differ much from one another with respect to their overall preferences for head-initial order, and they are all more conservative than the most conservative modern BT variety. Finally, in terms of its overall preference for head-initial order, EBT looks like a mid-range modern BT variety, such as the Ali Koč and Janevë dialects.

To sum up, two spectrums of directionality can be discerned in the new contact varieties as well, with much less marked differences along both dimensions when compared to BT (likely due to differences in the duration of contact).

After these general observations, I now move on to a more refined comparison, based on a cluster analysis of these findings.

### 3.2.3 Progressive versus conservative domains

Figure 1 shows three tree plots detailing how the thirteen head-dependent pairs in the three groups of contact varieties (i.e. NCV, EBT, and BT) may be analyzed into six cluster groups in terms of how progressive or conservative they are *within their respective dialect groups* (hierarchical agglomerative cluster analysis, Manhattan distance, average linkage).

To make it easier to compare them, I will use Table 3, which assigns a numerical value to each of these six clusters, cluster one being the most conservative and cluster six the most progressive. In addition, the table gives the means and mean absolute deviations (MAD) of these cluster values per head-dependent pair.

Table 3: Head-dependent clusters

	BT	EBT	NCV	Mean	MAD
CCI	6	6	5	5.67	0.44
VFin	5	5	6	5.33	0.44
NRc	6	4	1	3.67	1.78
VObl	5	2	3	3.33	1.11
VNFin	4	3	2	3.00	0.67
VAcc	5	2	2	3.00	1.33
VPP	2	1	4	2.33	1.11
VAdv	1	1	3	1.67	0.89
VBar	2	1	2	1.67	0.44
VSubj	1	1	2	1.33	0.44
NGen	1	1	1	1.00	0.00
PN	1	1	1	1.00	0.00
NAdj	1	1	1	1.00	0.00

Echoing the previous observations, noun phrases with non-clausal dependents and adpositional phrases are consistently the most conservative group (i.e. are strongly head-final). Verb-subject, verb-bare object, and verb-adverb pairs also tend to be conservatively head-final, but less consistently so than the previous

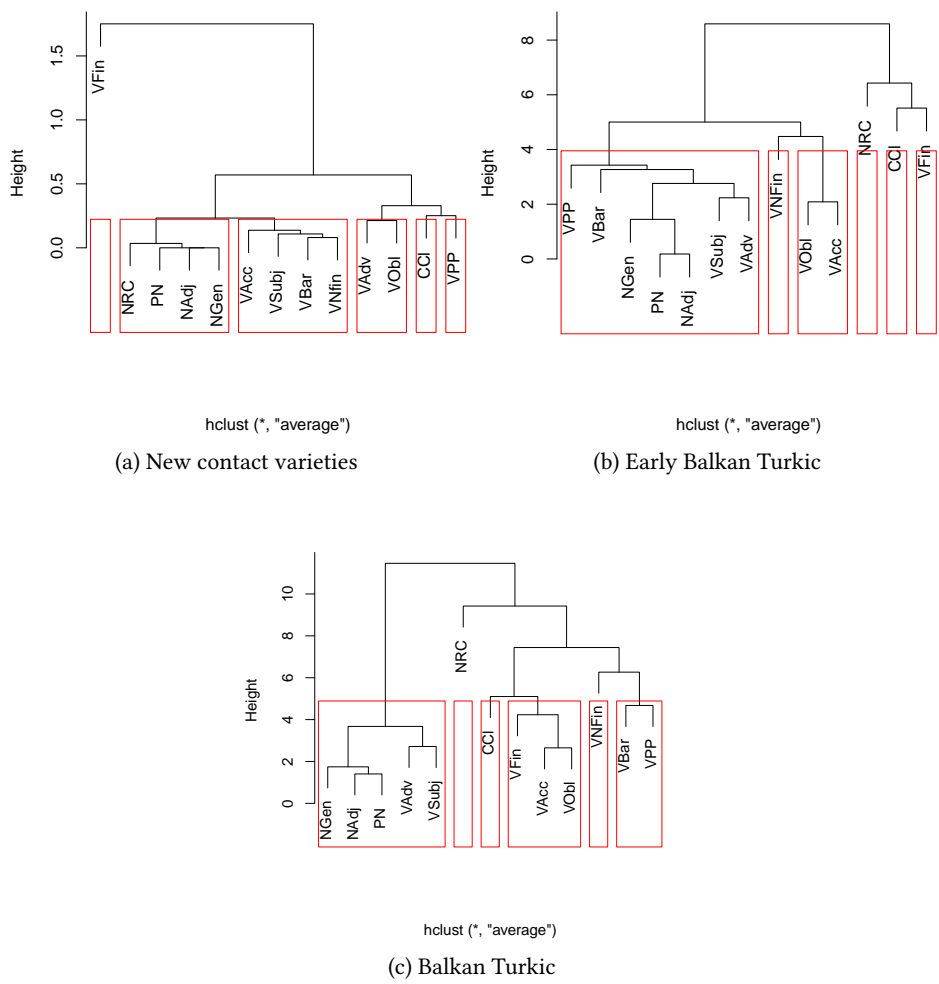


Figure 1: Cluster analysis of Turkic contact varieties

group, as shown by the higher MAD values. Skipping, then, to the top of the table, we see that finite dependent clauses and the clause-initial connectors associated with them are fairly consistently the most progressive group (see the MAD values). The difference between finite and non-finite clauses is clear here as well. In this portion of the table, the lack of SAE-type finite postpositive relative clauses in NCV, mentioned earlier, creates a noteworthy but explainable inconsistency. Finally, verbs with non-clausal dependents are distributed over the middle range (cf. mean  $\approx 3$ ), with the most variation observed in this zone.

### 3.2.4 Change and variability

The discussion that presently follows attempts to answer two questions raised by the data in Table 1:

1. How much do the syntactic domains of a given variety diverge from one another as that variety shifts to head-initial order?
2. How much does a given domain diverge across the varieties in a group (e.g. modern BT) as that domain shifts to head-initial order?

In other words, the following two sections dive into the variation we see along (i) the horizontal dimension (variability within varieties), and (ii) the vertical dimension (variability across varieties) of Table 1. The main generalization that captures the patterns in the data is that there is a strong positive correlation between change and variability along both dimensions.

#### 3.2.4.1 Variability within varieties

The answer to the first question can be visually summarized as in Figure 2.

The scatter plot in Figure 2a shows the link between the incidence of head-initial order in each BT variety as a percentage value (horizontal axis) and MAD across the thirteen syntactic pairs in that variety (vertical axis) (cf. Table 1). The best fitting regression model (i.e. logarithmic regression, adjusted  $R^2 = 0.251$ ), has also been plotted on the graph.

To spell out the generalization suggested by these data: The more a variety shifts to head-initial order, the more the syntactic domains of that variety diverge from one another in head directionality. (I consider some potential internal and external causal factors in Section 3.3.) This finding is supported by a Pearson's correlation test, which showed that there is a strong positive correlation between the incidence of head-initial order and MAD ( $r = 0.517$ ,  $p = 0.008$ , 95% CI: [0.153, 0.757]).

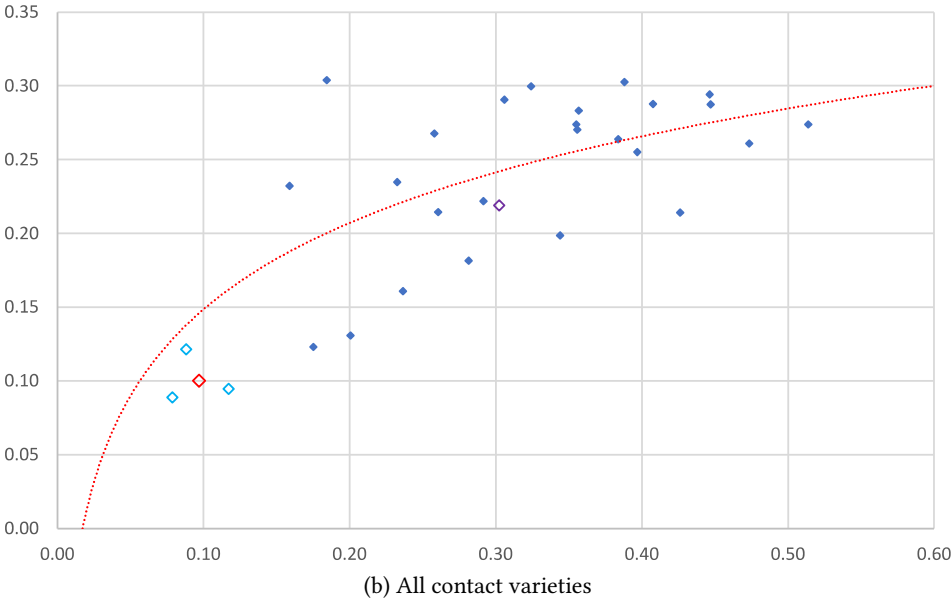
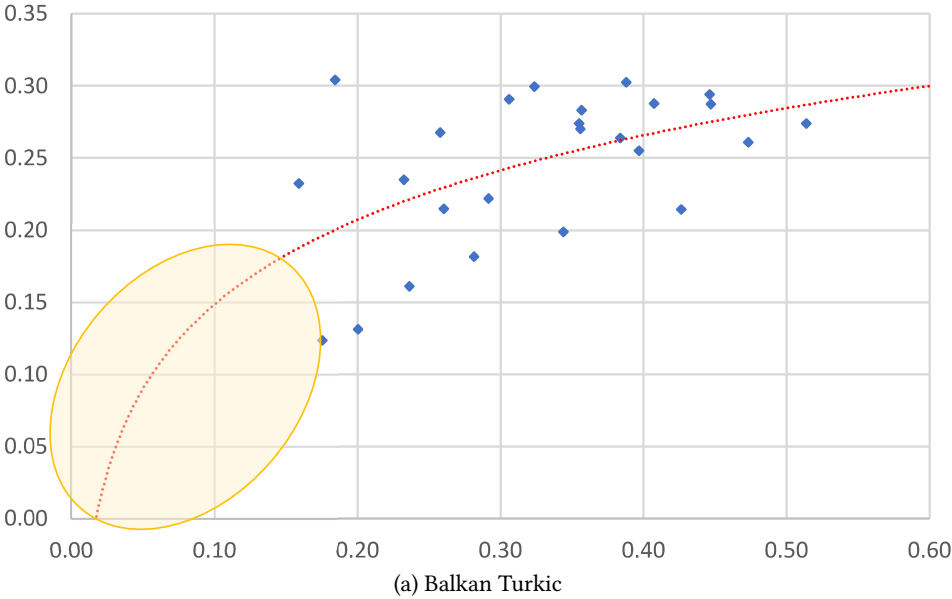


Figure 2: Head-initial order and mean absolute deviation per variety

Based on this generalization, the logarithmic regression model predicts that varieties that are less advanced than modern BT should occur somewhere in the yellow zone in the plot, i.e. that they should be less divergent internally. Indeed, as shown in Figure 2b the three new contact varieties (cyan-colored markers) and monolingual Turkish (red-colored marker) are located close to the trend line. In other words, they show only a limited amount of head-initial word order and are internally less heterogeneous in this regard. Note that the pattern in EBT (purple-colored marker at (0.3, 0.22)), which can be compared to a moderately advanced present-day BT variety, is also consistent with this model.

#### 3.2.4.2 Variability across varieties

Let us now turn to the second question above: How much does a given head-dependent pair (e.g. noun–adjective) diverge across varieties as that pair shifts to head-initial order? The answer to this question is given graphically in Figure 3. Similar to Figure 2, this scatter plot indicates the incidence of head-initial order on the horizontal axis and MAD on the vertical. The relationship represented, however, is between the amount of change observed in a given syntactic pair and the MAD of the amount of change in that syntactic pair across all contact varieties.

The plot makes it clear that the more a head-dependent pair shifts to head-initial order, the more it will diverge across varieties in head directionality. As before, Pearson's correlation shows that this is a very strong positive correlation ( $r = 0.829$ ,  $p < 0.001$ , 95% CI: [0.511, 0.947]).

### 3.3 An explanatory account of change, variation, and variability

Let me now address two questions that arise in connection with the observations in the preceding section:

1. What do these findings mean for the future of NCV?
2. What are the driving factors behind the patterns in Figures 1–3 and Table 3?

I will start with the first question: These findings predict that the more DEbi, USbi, and TUbI adopt patterns of head-initial word order in contact with German, English and Kurmanji, respectively, the more they will diverge to some extent (i) from one another (cf. Figure 3) as syntactic domains undergoing change diverge across varieties, and (ii) internally (cf. Figure 2).



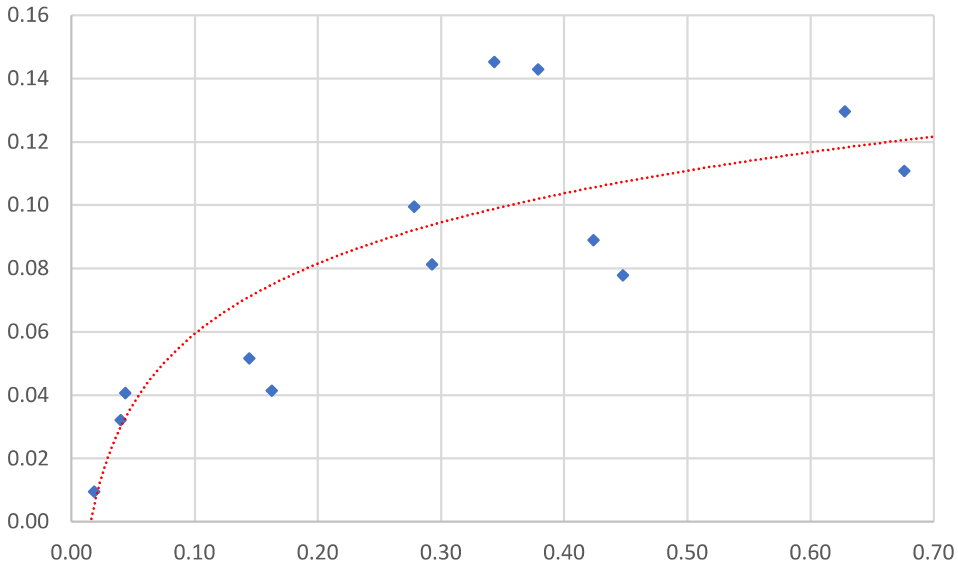


Figure 3: Head-initial order and mean absolute deviation per domain

In an attempt to answer the second question, I will offer some initial thoughts on four explanatory factors common to divergence both across and within varieties, touched upon in the preceding discussion.

### 3.3.1 Degree of exposure to contact

This factor has two facets. The first is time depth or the diachronic angle. BT varieties have been in a multitude of contact situations across the Balkans for a far longer period of time than NCV, and even the most conservative BT variety is more advanced in the shift to head-initial order than the most progressive NCV, as shown in Figure 2b. As the duration of contact between NCV and the majority languages increases, this can be expected to produce stronger effects. As an illustration, consider again the six clusters in Table 3. In that discussion, it was noted that the uppermost and the lowermost ranges of head-dependent pairs are homogeneous across dialect groups, while the middle range (i.e. from NRC down to VPP) contains the highest amount of variation. Despite this heterogeneity, the middle range in BT could be expected to be more advanced than that of NCV due to duration of contact. The means of cluster values in the middle range seem to corroborate this: 4.4 for BT versus 2.4 for NCV. Also, it would not be surprising if NCV and EBT were similar to each other in this regard (as both have had less exposure to contact than modern BT), which they indeed are: Mean of cluster

values in the middle range of both varieties = 2.4. It should be noted, however, that the less consistently conservative VAdv, VBar, and VSubj are surprising from this perspective (and VPP can perhaps also be included among these), as they are more advanced in NCV despite the shorter contact.

The second facet of the degree of exposure involves a synchronic perspective. Keskin (2023a) shows that the incidence of head-initial order in BT strongly and positively correlates with distance from the Turkish border (Pearson's correlation:  $r = 0.438$ ,  $p = 0.029$ , 95% CI: [0.052, 0.71]). Keskin proposes that this is likely due to the decreasing size of Turkish-speaking communities in the Balkans as one travels away from Turkey. This idea is supported by a strong negative correlation between the percentage of Turkish speakers in a municipality and the frequency of head-initial order in that municipality (Pearson's correlation:  $r = -0.53$ ,  $p = 0.04$ , 95% CI: [-0.82, -0.03]). In other words, speakers tend to use more head-initial order as speech communities shrink.<sup>12</sup> This is probably due to increased exposure to the majority language, as it becomes increasingly unlikely for speakers not to be exposed to the majority language as the group that speaks Turkish becomes smaller. This seems to predict that USbi should shift to head-initial order more than DEbi, as it is spoken by a loosely connected community made up of small, scattered micro-communities, while DEbi is spoken by a tightly connected community made up of larger micro-communities (see e.g. Iefremenko, Schroeder, et al. 2021, Özsoy et al. 2022).

A possible counterbalancing factor from both a diachronic and a synchronic perspective is contact with standard Turkish. As pointed out in Section 3.1, there is a strong connection between NCV and standard Turkish thanks to the internet, media, and regular travels, while such contact with standard Turkish would have been unlikely for most of the period that BT varieties were in contact with SAE. An anonymous review suggests that contact with standard Turkish (however weak it may have been) may also have had an influence on the patterns of change in BT, with communities further west possibly having less contact with standard Turkish.

### 3.3.2 Differential responses of different domains

When the syntactic system undergoes change, different elements or components within the system respond in different ways. This is clearly seen in the differences between head-dependent pairs within one variety.

---

<sup>12</sup>The concept of speech community size should be understood in this context as the size of what we could call a *micro-community*, i.e. for instance, the number of speakers in regular contact with one another in a village and not as the whole number of speakers in a given country.

Regardless of the duration of contact, noun and adposition phrases have remained by and large unchanged across the board, and will likely continue to do so, while finite clausal dependents are inclined towards head-initial order and will possibly become increasingly more head-initial. In this connection, the intriguing divergence between non-finite and finite clauses is another case of differential response replicated across periods/groups of varieties.

Further, obliques and accusative objects seem to be slower to react (and change) than finite clauses, as can be judged from their lower cluster values in EBT and NCV in contrast to their higher values in BT. In other words, it may be that finite clauses are likely the first constituents to change (and rather quickly at that), followed by the categories of the middle range.

Finally, two dependent types with surprising behavior are PP dependents of verbs and adverbs: They react differently to change and fall into different clusters, and are far more postverbal in NCV than in (E)BT.

These cross-categorical differences, particularly the differences between the most progressive and the most conservative categories, will lead to an increase in heterogeneity within the system as their responses diverge further, pulling the syntactic system apart, so to speak.

### 3.3.3 Emergence of new elements

As previously mentioned, an important difference between (E)BT and NCV is the use of SAE-type relative clauses in the former. Prolonged contact may bring about the emergence of these relative clauses in NCV, and they will have different characteristics or behaviors compared to Turkic-type RCs with which they will coexist for a period of time, thereby increasing heterogeneity within one variety.

### 3.3.4 Universal constraints

Despite all the factors that contribute to heterogeneity, there is a limit to this tendency. Due to the logarithmic trends that we have seen, heterogeneity will increase less sharply after a certain point, perhaps eventually leveling off. This can perhaps be attributed to universal constraints of grammar (in a theory-neutral sense of the term). What this implies for NCV is that as they evolve, their syntactic domains are likely to converge towards these universal constraints, after some period of divergence. Over time, the available syntactic options may become more constrained, leading to a slower increase in heterogeneity. For instance, we saw in Section 3.2.3 that subordinators (CCI) and finite clauses (VFin) have similar behaviors in that both favor head-initial orders. This can potentially

be due to a “head-parameter” which pushes these two elements towards convergence (i.e. head-initial order).

## 4 Contact-induced changes in clause combining

As pointed out in Section 3.1, a second set of contact-induced syntactic changes that appear to be taking place in heritage Turkish in the West, connected to a possible shift from head-final to head-initial syntax, involves innovative clause combining strategies, which is what I presently turn to. I will be focusing on a phenomenon which is revealing itself as we take into account what may at first glance appear as noise, namely a class of hybrid, blended, or mixed clauses that seem to emerge as part of contact-induced innovations in the subordination system.

I begin the discussion by describing the subordination models used in Turkic languages (Section 4.1) and the changes seen in clause combining in Turkish in the West (Section 4.2). This will set the stage for the treatment of the phenomenon of interest, which begins with a presentation of hybrid subordinate clauses in Balkan Turkic (Section 4.3). Subsequently, I show that similar hybrid subordinate clauses are also being generated in Western contact varieties (Section 4.4).<sup>13</sup>

### 4.1 Subordination models in Turkic

Turkic languages in contact with languages of Indo-European (IE) stock incorporate IE-type subordination strategies into their repertoires and often shift from the Turkic subordination template to the IE pattern (see e.g. Johanson 2021: 55.2.6, 55.3.8, 903–904, 913–916, 923–924). Balkan Turkic and Karaim with SAE contact, and Khalaj, Uzbek, and Azeri with Persian contact can be cited as examples.

The Turkic and the IE subordination models that these Turkic languages make use of side by side are diametrically opposed to each other. A typical subordinate clause (SC) conforming to the Turkic pattern has four main features (see e.g. Csátó & Johanson 1998: 223–224, 229–233, Johanson 1998: 48, 57–66, 2021: 854–931): (i) Its predicate is a non-finite form; (ii) It is positioned before the head noun or the matrix verb, i.e. it is prepositive; (iii) It is subordinated by means of a subordinative element suffixed to its predicate; (iv) if it does involve a free subordinative element, that element is clause-final. Two illustrative examples are given in (2):

---

<sup>13</sup>Note that the material that follows is a summary of various parts of Iefremenko et al. (2022), Keskin (2023b), and Keskin et al. (2024).

- (2) a. [Gel-diğ-in-i]                      bil-iyor-um.  
           come-NIND-2SG.POSS-ACC know-PROG-1SG  
           ‘I know that you have come.’  
       b. sen gel-dik-ten              sonra  
           2SG come-NIND-ABL after  
           ‘after you have come’

The bracketed argument clause in (2a) contains a non-finite predicate, signaled by the nominal indicative suffix which also acts as a subordinative element. The clause is positioned before the verb *biliyorum* ‘I know’ and is marked in accusative case, functioning as the direct object of the verb *bil* ‘know’. The adverbial clause in (2b) is again a nominal indicative, but this time including the postposition *sonra* ‘after’ that functions as an additional subordinative element.

The IE template, by contrast, makes SCs with an opposing set of characteristics possible (see e.g. Göksel & Kerslake 2005: 409–411, 457–460, 463–465, Johanson 1998: 65–66, 2021: 867–868, 894–899, Kerslake 2007, Kornfilt 1997: 3, 46, 60, 321–323, 439–440, 443): (i) Their predicates are finite verb forms; (ii) They are positioned after the head noun or the matrix verb, i.e. they are postpositive; (iii) They are linked to the superordinate clause by means of free subordinative elements; (iv) These subordinative elements are clause-initial. This strategy is exemplified in (3).

- (3) Bil-iyor-um      [ki      gel-di-n].  
       know-PROG-1SG CONN come-PST-2SG  
       ‘I know that you have come.’

The argument clause here can be contrasted with the one in (2a). It has a finite predicate marked in past tense and is positioned after the matrix verb *biliyorum*. It is introduced by the free clause-initial connector *ki* of Persian origin.

#### 4.2 Clause combining in Turkish in the West

Turkish in Western Europe and the US may have started to undergo changes in subordination that are comparable to the changes seen in other Turkic varieties with IE contact, as suggested by the observations of studies on heritage Turkish in the West (see e.g. Bayram 2013, Iefremenko, Özsoy, et al. 2021, Iefremenko, Schroeder, et al. 2021, Karakoç 2007, Onar Valk 2015, Özsoy et al. 2022, Schroeder 2016, Schroeder & Iefremenko 2019a, Treffers-Daller et al. 2006, Turan et al. 2020, see also Schroeder et al. 2025 [this volume]). It has begun to use

non-finite clauses less frequently when compared to Turkey's Turkish. It prefers finite clauses instead, with the consequence that it now makes use of paratactic clause combining in place of subordination. The latter strategy was exemplified in (2a), and an illustrative example of the former is given in (4):

- (4) Gel-di-n            bil-iyor-um.  
      come-PST-2SG know-PROG-1SG  
      'I know you have come.'

Example (4), in contrast to (2a), involves two syntactically independent finite clauses that have simply been strung together sequentially without any differences in their morphological form or syntactic function.

IE-type finite subordination can probably not be said to have developed yet in the Western contact varieties of Turkish. However, concomitantly with the move from non-finite to finite clauses, one observes changes in the functions and positions of paratactic connectors, the emergence of new connectors, and a higher frequency of clause combining connectors overall. These developments suggest the beginnings of finite subordination, and they might culminate in Turkic subordination becoming marginal and IE-type subordination dominating in the Turkish varieties in the West, as with BT.

What is critical at this juncture is an observation on BT subordination: The shift to IE-type dependent clauses in BT seems to have happened through the creation of several kinds of hybrid dependent clauses that persist in present-day BT varieties as marginal types of dependent clauses. These will be detailed in the next two sections, before I return to Turkish in the West.

### 4.3 Hybrid subordinate clauses in Balkan Turkic

To elaborate on the preceding remark, the shift from Turkic to IE-type subordination in the case of BT, as I have argued in Keskin (2023b), involves a process that combines, in an apparently random fashion, the properties of Turkic and IE-type subordination outlined in Section 4.1, creating a range of SCs that show mixed properties (*X-clauses*). *X-clauses* seem restricted to West and North Rumelian, most probably because it is these two groups that make extensive use of IE-type subordination. I have not been able to identify any hybrid SCs in the conservative ERT varieties.

Consider now the examples in (5–10) that illustrate the six subtypes of *X-clauses* which occur at above-average frequency within this class of SCs and account for a vast majority of their occurrences:

- (5) [Ani sırala-dı-m] to urba-lar-ı giy-ē-sin.  
CONN tell-PST-1SG DIST clothes-PL-ACC wear-AOR-2SG  
'You'll put on the clothes that I told you about.' (Murtaza 2016: 81)
- (6) Çuval doqū-du-lā onlā-dan [ani zāre-ler-i quy-mā].  
sack weave.AOR-PST-3PL 3PL-ABL CONN grain-PL-ACC put-NSUB.DAT  
'They weaved sacks from them to put the grains in.' (Haliloğlu 2017: 214)
- (7) Kimsā de-yār-miş [ani ki o nicā gün duuması].  
some say-PROG-EVID.3SG CONN CONN 3SG like sun rise  
'Some were saying that she is like the sunrise.' (Çimpoeş 1988: 5 via  
Özkan 2007: 157)
- (8) Sevin-ēr-im [ani mizin ol-du deye].  
rejoice-AOR-1SG CONN muezzin become-PST.3SG CONN  
'I am happy that he became a muezzin.' (Murtaza 2016: 112)
- (9) O sene [bu cade ne yap-ıl-dı] çalış-i-dı-k biz  
DIST year PROX road CONN make-PASS-PST.3SG work-PROG-PST-1PL 1PL  
or-da.  
there  
'The year that this road was built, we were working there.' (Sulçevsi 2019: 255)
- (10) Sū-müş o [ani bizim yaşa-dī-mız] yer-i.  
plough-EVID.3SG 3SG CONN 1PL.GEN live-NSREL-1PL.POSS place-ACC  
'He ploughed the place where we lived.' (Karaşinik 2011: 230)

The SC in (5) is almost like an IE-type SC with its finite predicate and clause-initial connector (viz. *ani*), but it is prepositive like a Turkic SC. The SC in (6) is partly like an IE-type SC, since it is postpositive and introduced by an initial subordinator, however it is non-finite like a Turkic SC, as indicated by the nominalized subjunctive marker. The SCs in (7), (8), and (9) are finite and postpositive, again partly conforming to the IE template, but the subordinators that introduce them are atypical. The first has two clause-initial subordinators (viz. *ani ki*); the second has two subordinators, one clause-initial and the other clause-final (viz. *ani...deye*), which I refer to as a *circumclausal subordinator*; and the last has a clause-internal subordinator (viz. *ne*). Finally, the SC in example (10) is non-finite and prepositive following the Turkic template, but it is introduced by a clause-initial subordinator like an IE-type SC.

#### 4.4 Hybrid subordinate clauses in Turkish in the West

So to repeat, the shift from Turkic to IE-type subordination in BT seems to involve a process that combines the properties of Turkic and IE-type subordination, creating a range of X-clauses. If this is true, X-clauses could also be identifiable – at the very least in part – in other Turkic varieties with SAE contact, most relevantly in DEbi and USbi, given the present context.<sup>14</sup> That is the possibility that we investigate in Keskin et al. (2024), and I now move on to the discussion of the examples of X-clauses identified in these Turkish contact varieties as part of that study. Using the description of the subordination models in Turkic in Section 4.1 as a point of reference, I present and discuss the data in terms of (i) clause position and (ii) subordinator type and position.

##### 4.4.1 Clause position

Innovations in clause combining in Turkish as spoken in Germany and US are perhaps most clearly observed in clause position. These are seen significantly more in bilingual data than in monolingual data and involve canonically preverbal SCs occurring in the postverbal position. An example is given in (11):<sup>15</sup>

- (11) İnan-mı-ycak-sın     [demin n=ol-duğ-un-a].  
believe-NEG-FUT-2SG just     what=happen-NIND-3SG-DAT  
'You won't believe what just happened.' [DEbi33MT\_isT]

Apart from its position relative to the matrix verb, the SC in this example is entirely ordinary. Now, Turkic SCs can and do occur postverbally under the right conditions, given the flexible constituent order of Turkic languages. However, this has a typical frequency signature: Only about a quarter of SCs are postverbal in standard spoken Turkish (22% in our data). In bilingual data, by contrast, we observe a significantly higher frequency of postpositive SCs (40%, with DEbi = 42%

<sup>14</sup>As Keskin (2023b) observes, components of the X-clause phenomenon are actually attested in several unrelated contact situations, such as Indo-Aryan languages with Dravidian contact (Bayer 2001, Dhongde & Wali 2009: 214, Hock 2021: 541–542, Pandharipande 1997: 2, 6, 70) and Laz with Turkish contact (Demirok & Öztürk 2022). Preliminary research suggests that the phenomenon is also seen in Romeyka with Turkish contact (see e.g. Keskin et al. 2023, Schreiber 2019, 2022).

<sup>15</sup>The codes next to the examples in Section 4.4 provide the following information: Country: DE=Germany, TU=Turkey, US=US; Bi-/monolingual speaker: bi vs. mo; Speaker number and age group: 1–50=adults; from 51 onward=adolescents; Gender: M vs. F (there were no speakers who identified as non-binary); Family language: T=Turkish; Communicative situation: f=formal vs. i=informal; s=spoken vs. w=written; Language of production: T=Turkish.



and USbi = 38%), as shown by a generalized linear mixed model. The model was built with the individual speakers as random effects and bilingualism, clause type, and age group as fixed effects. It showed a significant effect of bilingualism and clause type ( $p = 0.017$  and  $p < 0.001$ , respectively): Bilinguals produce significantly more postpositive SCs than monolinguals, and finite SCs are significantly more postpositive than non-finite SCs.

These observations point towards a blending of the IE feature of postpositiveness with the Turkic features of non-finiteness, etc.

#### 4.4.2 Subordinator type and position

I begin the discussion of subordinators and their features with the issue of the possible emergence of free clause-initial connectors that introduce finite SCs (cf. the IE model of subordination). I then present potential clause-initial connectors that introduce non-finite clauses – X-clauses per excellence. I finally connect these two sets of data to irregularities in subordinator position, which present further examples of X-clauses.

The potentially emerging connectors that occupy the central stage in these findings are *hani*, *yani*, and *işte*, all derived from homophonous discourse markers. An important feature of most of the examples which contain these elements is an ambiguity between subordinate and coordinate readings, which can be taken as a sign of the beginning stages of the development of finite subordination in DEbi and USbi.<sup>16</sup>

##### 4.4.2.1 Subordinator type 1: Clause-initial connectors introducing finite subordinate clauses

Consider, first, the example in (12):

- (12) Araba [hani bun-u gör-dü] dur-du.  
 car CONN this-ACC see-PST.3SG stop-PST.3SG  
 ‘The car that saw this stopped.’ [DEbi06FT\_isT]

<sup>16</sup>I should point out an important difference between this section and the preceding ones. Here, I will not be presenting any statistical analyses and I confine myself to a presentation of several examples which seem to involve clause combining connectors. The reason for this is that the uses of the above-mentioned elements in the data are open to interpretation and hence their analysis is marred by a high degree of subjectivity. A more systematic and objective investigation of these data is left to future research.

In this example we have, in square brackets, what appears to be a typical IE-type relative clause introduced by *hani*: finite, postpositive, with a free clause-initial subordinator.

The next example in (13) exemplifies the use of *yani* as what again seems to be a subordinator:

- (13) Anla-dı-n [yani bi çift var-dı top-lu].  
 understand-PST-2SG CONN a couple exist-PST.3SG ball-ATTR  
 ‘You understood that/well there was a couple with a ball.’  
 [DEbi18MT\_isT]

Here, the bracketed clause is a possible complement clause but the interpretation of the example is ambiguous: The bracketed segment can also be interpreted as the member of a coordinate structure together with the preceding segment.

Finally, consider the example in (14), demonstrating the use of *işte* as connector:

- (14) Karşı taraf-ta da bi kadın var [işte alışveriş yap-mış  
 opposite side-LOC FOC a woman exist CONN shopping do-PRF.3SG  
 san-ır-ım].  
 think-AOR-1SG  
 ‘There is a woman on the opposite side who/and she had been shopping I think.’ [DEbi08FT\_isT]

As with the preceding two examples, the bracketed clause here can be interpreted as an SC, an extraposed relative detached from its head noun. An alternative interpretation is as a second conjunct.

#### 4.4.2.2 Subordinator type 2: Clause-initial connectors introducing non-finite subordinate clauses

In addition to the cases in (12–14) which strongly resemble IE-type SCs, another pattern that we observe in DEbi and USbi data is the frequent occurrence of *yani* and *işte* immediately preceding various non-finite SCs, as exemplified in (15–17):

- (15) Top=la [işte gid-er-ken] oynuyo-du.  
 ball=with CONN go-AOR-CONV play.PROG-PST.3SG  
 ‘S/he was playing with the ball as he went.’ [USbi74FT\_fsT]
- (16) [Yani top-u düş-ür-en] adam kadın-a yardım et-ti.  
 CONN ball-ACC fall-CAUS-SREL man woman-DAT help-PST.3SG  
 ‘The man who dropped the ball helped the woman.’ [DEbi60FT\_isT]

- (17) Bi adam [yani dokuz yüz on bir-i ara-ma-ya] çalış-tı.  
 a man CONN 911-ACC call-NSUB-DAT try-PST.3SG  
 ‘A man tried to call 911.’ [USbi52FT\_isT]

The bracketed segment in (15) contains a preverbal converb (viz. *giderken*). Immediately to the left of the converb is *işte* which could be interpreted as introducing the converbial clause. Next, in example (16), we have a comparable structure which involves *yani* on its left periphery and a prepositive participial relative clause typical for Turkic languages. Finally, (17) contains an infinitival clause preceded by *yani*, producing the same structure as before.

These observations suggest that the IE feature of free clause-initial subordinators is being blended with the Turkic features of non-finiteness and prepositiveness in DEbi and USbi.

#### 4.4.2.3 Subordinator position

I now move on to non-canonical uses of subordinator position that give rise to further examples of X-clauses. Recall from Section 4.3 that BT developed, in addition to canonical free clause-final subordinators (e.g. *diye*) in its repertoire, free subordinators that occupy non-clause-final positions, using them as part of its X-clause inventory, such as double clause-initial (7), circumclausal (8), clause-initial (5–6 and 10), and clause-internal (9) subordinators.

When we turn to Turkish in Germany and the US in the light of this possibly inexhaustive list, we readily observe the occurrence of circumclausal subordinators. This phenomenon involves both well-established clause-initial and final connectors, such as *çünkü* ‘because’, *için* ‘for’, *ki* ‘which, that’, *diye* ‘that’, but also the potential emergent connectors presented above, i.e. *hani*, *yani*, and *işte*. This appears to be a largely bilingual phenomenon: We have identified three relevant examples in monolingual data (all with the emergent connectors) as opposed to 30 in bilinguals (with both new and old connectors).

Consider the example in (18) with well-established connectors:

- (18) Gerek-iyor-muş [çünkü bunlar-a şahit ol-duğ-um için].  
 be.necessary-PROG-EVID because these-DAT witness-NIND-1SG for  
 ‘It was necessary because I witnessed these.’ [DEbi05FT\_isT]

The adverbial clause in this example has a non-finite predicate, as attested by the nominalized indicative suffix on it – a hallmark of Turkic SCs. The clause contains not only the free clause-final subordinator *için* (Turkic-type) but also the

clause-initial subordinator *çünkü* (IE-type), forming a circumclausal free subordinator. Finally, the clause is positioned after the main verb like an IE-type clause. This blending of Turkic and IE features make this adverbial clause an X-clause.

Next is an example in (19) with an emergent connector:

- (19) Öndeki araba dur-du      [hani köpeğ-e çarp-mıyım diye].  
in.front car   stop-PST.3SG CONN dog-DAT hit-NEG.OPT.1SG CONN  
'The car in front stopped so as not to hit the dog.' [USbi52FT\_isT]

The SC in this example follows a pattern close to that in example (18): postpositive with a circumclausal free subordinator. It is, however, finite, which means that it is a twin of the X-clause in (8).

## 5 Conclusion

Let me begin the concluding section by remembering the question with which I set out: How good a model is the SAE–Turkic contact in the Balkans for understanding contact-induced syntactic changes in Turkish in the West?

As an answer to this question, I presented patterns of change (i.e. the two spectrums of directionality, resistant versus progressive domains, and variability within and across varieties) replicated across contact situations past and present. In the light of these observations, one may be inclined to answer this question with a positive superlative.

Still, two things need to be underlined. First, what can be observed across periods and contact situations is not a single homogeneous response to language contact but a complex pattern with a broad range into which all the contact varieties investigated fit rather well. Second, this pattern may be brought about by explanatory factors which are hardly specific to any given language or family of languages: degree of exposure to contact, differential responses of different syntactic domains, emergence of new elements, and universal constraints.

The same comments apply, *mutatis mutandis*, to the contact effects observed in the clause combining system as well. The emergence of X-clauses in both Balkan Turkic and the Western contact varieties is a striking convergence. However, the reader will recall that they are attested in other contact situations too (e.g. Romeyka–Turkish contact).

So, the longer and more reasonable answer to the research question above seems to be that Balkan Turkic constitutes a good model for understanding contact-induced syntactic change in Turkish in the West insofar as it can help reveal general contact dynamics, not as a Turkic contact variety per se. And on that front the investigation of Balkan Turkic can be said to have been fruitful.

## Abbreviations

1	first person	LOC	locative
2	second person	NEG	negation
3	third person	NIND	nominalized indicative
ABL	ablative	NSREL	nonsubject relative
ACC	accusative	NSUB	nominalized subjunctive
AOR	aorist	OPT	optative
ATTR	attributive	PASS	passive
CAUS	causative	PL	plural
CONN	connector	POSS	possessive
DAT	dative	PRF	perfective
DIST	distal	PROG	progressive
EVID	evidential	PROX	proximate
FOC	focus particle	PST	past
FUT	future	SG	singular
GEN	genitive	SREL	subject relative

## Acknowledgments

The academic research that constitutes the foundation of this article was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) grant, as part of the projects *Head directionality change in Turkic in contact situations: A diachronic comparison between heritage Turkish and Balkan Turkic* (grant number: 394841858) and *Clause combining in Balkan Turkic: Pathways and stages of contact-induced grammaticalization* (grant number: 313607803) within the Research Unit *Emerging Grammars in Language Contact Situations: A Comparative Approach* (FOR 2537).

First and foremost, I thank Christoph Schroeder for his unwavering encouragement and support, as well as his comments, suggestions, exhortations, etc. I am grateful to Heike Wiese for allowing me to join the RUEG network with my supplemental projects and to Shanley E. M. Allen to link my research with the project that she co-directed with Heike and Christoph. I also thank the two anonymous external reviewers, the internal reviewers Onur Özsoy and Judith Purkarthofer, and the following colleagues (in alphabetical order): Eystein Dahl, Kateryna Iefremenko, Jaklin Kornfilt, Nils Picksak, Sergey Say, Ilja Seržant, Anna Shadrova, and Nadine Zürn. Finally, I thank Lea Coy for help with pre-publication formatting. The usual disclaimers apply.

## References

- Adamović, Milan. 1976. Vocabulario Nuovo mit seinem Türkischen Teil. *Rocznik Orientalistyczny* 38. 45–50.
- Adamović, Milan. 1977. *Das osmanisch-türkische Sprachgut bei R. Lubenau* (Beiträge zur Kenntnis Südosteuropas und des Nahen Orients 25). München: R. Trofenik.
- Adamović, Milan. 2001. *Das Türkische des 16. Jahrhunderts: Nach den Aufzeichnungen des Florentiners Filippo Argenti (1533)*. Göttingen: Pontus Verlag.
- Balci, Mustafa. 2010. Kosova ve Makedonya Türk ağızlarında devrik cümle meselesi. *Avrasya Etüdləri* 38(2). 43–60.
- Bayer, Josef. 2001. Two grammars in one: Sentential complements and complementizers in Bengali and other South Asian languages. In Rajendra Singh (ed.), *The yearbook of South Asian languages and linguistics*, 11–36. New Delhi: Sage.
- Bayram, Fatih. 2013. *Acquisition of Turkish by heritage speakers: A processability approach*. Newcastle: Newcastle University. (Doctoral dissertation).
- Blau, Otto. 1868. *Bosnisch-türkische Sprachdenkmäler*. Leipzig: Deutsche Morgenländische Gesellschaft.
- Böhm, Stefanie. 2015. Differential object marking in standard Turkish and Caucasian Urum. *Language Typology and Universals* 68(4). 421–438.
- Bombaci, Alessio. 1940. Padre Pietro Ferraguto e la sua Grammatica turca (1611). *Annali del R. Istituto Superiore Orientale di Napoli*. 205–236.
- Brendemoen, Bernt. 1980. Labiyal ünlü uyumunun gelişmesi üzerine bazı notlar. *Türkiyat Mecmuası* 19. 223–240.
- Brezina, Vaclav. 2018. *Statistics in corpus linguistics: A practical guide*. Cambridge: Cambridge University Press. <http://corpora.lancs.ac.uk/stats/index.php>.
- Bulut, Christiane. 2006. Turkish elements in spoken Kurdish. In Hendrik Boeschoten & Lars Johanson (eds.), *Turkic languages in contact*, 95–121. Wiesbaden: Harrassowitz Verlag.
- Bunk, Oliver, Shanley E. M. Allen, Sabine Zerbian, Tatiana Pashkova, Yulia Zuban & Erica Conti. 2025. Information packaging and word order dynamics in language contact. In Shanley E. M. Allen, Mareike Keller, Artemis Alexiadou & Heike Wiese (eds.), *Linguistic dynamics in heritage speakers: Insights from the RUEG group*, 379–412. Berlin: Language Science Press. DOI: 10.5281/zenodo.15775181.
- Çimpoș, Lübov S. 1988. *Literatura okumakları 9–10 klasslar*. Chișinău.
- Csató, Éva Á. 2000. Syntactic code-copying in Karaim. In Östen Dahl & Maria Koptjevskaja-Tamm (eds.), *The circum-baltic languages: Their typology and contacts*, 265–277. Amsterdam: John Benjamins.

- Csató, Éva Á. & Lars Johanson. 1998. Turkish. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages*, 203–235. London: Routledge.
- Csató, Éva Á., Astrid Menz & Fikret Turan (eds.). 2016. *Spoken Ottoman in mediator texts*. Wiesbaden: Harrassowitz Verlag.
- Demirok, Ömer & Balkız Öztürk. 2022. The emergence of clausal nominalizations in Laz. Talk delivered at Formal Approaches to Contact in/with Turkish, UiT, Tromsø.
- Destanov, Ekrem. 2016. *Doğu Makedonya sözlü geleneğinde masallar üzerine bir inceleme*. Istanbul: Yildiz Technical University. (Dissertation).
- Dhongde, Ramesh V. & Kashi Wali. 2009. *Marathi*. 3rd edn. (London Oriental and African Language Library 13). Amsterdam: John Benjamins.
- Doerfer, Gerhard. 1959. Das Gagausische. In Jean Deny, Kaare Grønbech & Scheel, Helmuth, Togan, Zeki Velidi (eds.), *Philologiae turcicae fundamenta I*, 260–271. Wiesbaden: Steiner.
- Dorleijn, Margreet. 2006. Turkish–Kurdish language contact. In Hendrik Boeschoten & Lars Johanson (eds.), *Turkic languages in contact*, 75–94. Wiesbaden: Harrassowitz Verlag.
- Drimba, Vladimir. 1966. « L’instruction des mots de la langue turcesque » de Guillaume Postel (1575). *Türk Dili Araştırmaları Yıllığı-Belleten*. 77–101.
- Du Ryer, André. 1630. *Rudimenta grammatices linguae turcicae*. Paris: Antoniu Vitray.
- Dündar, Fuat. 2000. *Türkiye nüfus sayımlarında azınlıklar*. 2nd edn. Istanbul: Çiviyazıları.
- Friedman, Victor A. 1982. Balkanology and Turkology: West Rumelian Turkish in Yugoslavia as reflected in prescriptive grammar. In Anthony A. Barentsen, Ronald Sprenger & Marinus G. M. Tielemans (eds.), *South Slavic and Balkan linguistics*, vol. 2 (Studies in Slavic and General Linguistics), 1–77. Leiden: Brill.
- Friedman, Victor A. 2006. West Rumelian Turkish in Macedonia and adjacent areas. In Hendrik Boeschoten & Lars Johanson (eds.), *Turkic languages in contact*, 27–45. Wiesbaden: Harrassowitz Verlag.
- Gallotta, Aldo. 1986. Latin harfleri ile yazılmış bir kaç Osmanlı atasözü. *İstanbul Üniversitesi Edebiyat Fakültesi Türk Dili ve Edebiyatı Dergisi* 24. 235–249.
- Georgievits, Bartholomaeus. 1544. De turcarum ritu et caeremoniis.
- Göksel, Aslı & Celia Kerslake. 2005. *Turkish: A comprehensive grammar*. London: Routledge.
- Gülsevin, Gürer. 2017. *Yaşayan Türkiye Türkçesi ağızlarının tarihî dönemlerini belirleyebilmek için bir yöntem denemesi örneği: XVII. yüzyıl Batı Rumeli Türkçesi ağızları*. Ankara: Türk Dil Kurumu Yayınları.

- Günşen, Ahmet. 2010. Rumeli ağızlarının söz dizimi üzerine I: Makedonya ve Kosova Türk ağızları örneği. *Turkish Studies* 5(1). 462–494.
- Haig, Geoffrey. 2017. Western Asia: East Anatolia as a transition zone. In Raymond Hickey (ed.), *The Cambridge handbook of areal linguistics*, 396–423. Cambridge: Cambridge University Press.
- Haig, Geoffrey & Ergin Öpengin. 2018. Kurmanji Kurdish in Turkey: Structure, varieties, and status. In Christiane Bulut (ed.), *Linguistic minorities in Turkey and Turkic-speaking minorities of the peripheries*, 157–230. Wiesbaden: Harrassowitz Verlag.
- Haliloğlu, Neşe. 2017. *Zavet köyü Türk ağızı*. Kırşehir: Ahi Evran University. (Master's thesis).
- Haspelmath, Martin. 1998. How young is Standard Average European? *Language Sciences* 20. 271–287.
- Haspelmath, Martin. 2001. The European linguistic area: Standard Average European. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals*, vol. 2, 1492–1510. Berlin: De Gruyter.
- Hazai, György. 1960. Textes turcs du Rhodope. *Acta Orientalia Academiae Scientiarum Hungaricae* 10(2). 185–229.
- Hazai, György. 1990. Die Denkmäler des Osmanisch-Türkisch-Türkischen in nicht-arabischen Schriften. In György Hazai (ed.), *Handbuch des türkischen Sprachwissenschaft*, 63–73. Budapest: Akadémiai Kiadó.
- Heffening, Willi. 1942. *Die türkischen Transkriptionstexte des Bartholomaeus Georgievits aus den Jahren 1544–1548*. Leipzig: Brockhaus.
- Helmholdt, Wolfgang. 1966. Das türkische Vaterunser in Hans Schiltbergers Reisebuch. *Central Asiatic Journal* 11(2). 141–143.
- Herbinius, Johannes. 1675. *Horae Turcico-Catecheticae: Sive institutio brevis catechetica, cuiusdam Turcae circumcisi*. Danzig: David-Fridericus Rhetius.
- Hock, Hans H. 2021. *Principles of historical linguistics*. 3rd edn. (Trends in Linguistics 34). Berlin: De Gruyter.
- Iefremenko, Kateryna. 2019. *Placement of the locational goal elements in Kurmanji Kurdish in contact with Turkish*. Talk delivered at the 4th International Conference on Kurdish Linguistics, University of Rouen, France.
- Iefremenko, Kateryna. 2021a. *Placement of the locational goal elements in Kurmanji Kurdish in contact with Turkish and German*. Talk delivered at the 9th International Conference on Iranian Linguistics (ICIL 9), Österreichische Akademie der Wissenschaften, Vienna, Austria.



- Iefremenko, Kateryna. 2021b. *Post-predicate position in Kurmanji in contact with Turkish and German*. Talk delivered at the Workshop on Turkic and Languages in Contact with Turkic (TU+), University of Toronto, Canada.
- Iefremenko, Kateryna. 2021c. *Word order in Turkish in contact with German and Kurmanji*. Talk delivered at the RUEG Conference - Dynamics of Language Contact, Humboldt Universität, Berlin, Germany [online].
- Iefremenko, Kateryna, Cem Keskin, Jaklin Kornfilt & Christoph Schroeder. 2022. *Individual variation in heritage Turkish clause combining*. Talk delivered at Formal Approaches to Contact in/with Turkish, Tromsø.
- Iefremenko, Kateryna, Cem Keskin & Christoph Schroeder. 2023. The post-predicate position in Turkish in language contact situations: How resistant to change? Talk delivered at DGfS 2023, Universität zu Köln.
- Iefremenko, Kateryna, Onur Özsoy & Christoph Schroeder. 2021. *Paratactic clause combining in heritage Turkish in Germany and the US*. Talk delivered at the International Conference on Turkish Linguistics, Anadolu University, Eskişehir, Turkey [online].
- Iefremenko, Kateryna, Christoph Schroeder & Jaklin Kornfilt. 2021. Converbs in heritage Turkish: A contrastive approach. *Nordic Journal of Linguistics* 44(2). 130–154. DOI: 10.1017/S0332586521000160.
- İğci, Alpay. 2010. *Viçitırın-Kosova Türk ağzı*. Izmir: Ege University. (Master's thesis).
- Illésházy, Miklós. 1668. *Dictionarium turcico-latinum*.
- Jable, Ergin. 2010. *Kosova Türk ağızları*. Sakarya: Sakarya University. (Dissertation).
- Johanson, Lars. 1989. Substandard und Sprachwandel im Türkischen. In Günter Holtus & Edgar Radtke (eds.), *Sprachlicher Substandard II* (Konzepte der Sprach- und Literaturwissenschaft), 83–112. Tübingen: Niemeyer.
- Johanson, Lars. 1992. *Strukturelle Faktoren in türkischen Sprachkontakten* (Sitzungsberichte der wissenschaftlichen Gesellschaft an der Johann Wolfgang Goethe-Universität Frankfurt am Main). Stuttgart: Steiner.
- Johanson, Lars. 1998. The structure of Turkic. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages*, 31–66. London: Routledge.
- Johanson, Lars. 2006. Turkic language contacts in a typology of code interaction. In Hendrik Boeschoten & Lars Johanson (eds.), *Turkic languages in contact*, 4–26. Wiesbaden: Harrassowitz Verlag.
- Johanson, Lars. 2013. *Structural factors in Turkic language contacts*. London: Routledge.
- Johanson, Lars. 2021. *Turkic* (Cambridge language surveys). Cambridge: Cambridge University Press.

- Joseph, Brian D. 2020. Language contact in the Balkans. In Raymond Hickley (ed.), *The handbook of language contact*, 2nd edn., 537–549. Chichester: John Wiley & Sons.
- Kakuk, Suzanne. 1958. Textes turcs de Kazanlyk II: Zehra Džafer. *Acta Orientalia Academiae Scientiarum Hungaricae* 8(3). 241–311.
- Kakuk, Suzanne. 1961a. Die türkischen Mundart von Küstendil und Michailovgrad. *Acta Linguistica Academiae Scientiarum Hungaricae* 11(3/4). 301–386.
- Kakuk, Suzanne. 1961b. Türkische Volksmärchen aus Küstendil. *Ural-altaische Jahrbücher* 33(1/2). 90–94.
- Kakuk, Suzanne. 1972. Le dialecte turc d’Ohrid en Macédoine. *Acta Orientalia Academiae Scientiarum Hungaricae* 26(2/3). 227–283.
- Karakoç, Birsıl. 2007. Connectivity by means of finite elements in monolingual and bilingual Turkish discourse. In Jochen Rehbein, Christiane Hohenstein & Lukas Pietsch (eds.), *Connectivity in grammar and discourse* (Hamburg Studies on Multilingualism 5), 199–227. Amsterdam: John Benjamins.
- Karaşinik, Bahtışen. 2011. *Silistre (Silistra) ili ve yöresi ağızları*. Edirne: Trakya University. (Master’s thesis).
- Katona, Louis K. 1969. Le dialecte turc de la Macédoine de l’ouest. *Türk Dili Araştırmaları Yıllığı-Belleten* 17. 57–194.
- Kerslake, Celia. 2007. Alternative subordination strategies in Turkish. In Jochen Rehbein, Christiane Hohenstein & Lukas Pietsch (eds.), *Connectivity in grammar and discourse* (Hamburg Studies on Multilingualism 5), 231–258. Amsterdam: John Benjamins.
- Keskin, Cem. 2023a. On the directionality of the Balkan Turkic verb phrase: Variationist and theoretical perspectives. *Languages* 8(1). 1–20.
- Keskin, Cem. 2023b. Transient subordinate clauses in Balkan Turkic in its shift to Standard Average European subordination: Dialectal and historical evidence. *Folia Linguistica* (57). 155–197.
- Keskin, Cem, Kateryna Iefremenko, Jaklin Kornfilt & Christoph Schroeder. 2024. Hybrid clause combining strategies in Turkish language contacts. *Studia Linguistica* (Early View). DOI: 10.1111/stul.12256.
- Keskin, Cem, Laurentia Schreiber, Kalliopi Katsika & Metin Bağrıaçık. 2023. *Contact-induced blended subordination in Romeyka*. Talk delivered at the 3rd International Conference on Language Contact in the Balkans and Asia Minor, Aristotle University of Thessaloniki, Thessaloniki, Greece.
- Keskin, Cem, Taner Sezer, Christoph Schroeder, Lale Diklitaş, Zeynep Güler, Nil Özkul, Süleyman Yüceer, Ayşe Nur Güler & Ebru Özgenç. In preparation. *Balkan Turkic corpus*.

- Kirli, Nadejda. 2001. *N. N. Baboglu'nun 'Legendanın izi' adlı eserinin cümle bilgisi yönünden incelemesi*. Adana: Çukurova University. (Master's thesis).
- Kissling, Hans Joachim. 1968. Bemerkungen zu einigen Transkriptionstexten. *Zeitschrift für Balkanologie* 6. 119–127.
- Kornfilt, Jaklin. 1997. *Turkish*. London: Routledge.
- Küçükcan, Talip & Veyis Güngör (eds.). 2009. *Turks in Europe: Culture, identity and integration*. Amsterdam: Turkevi Research Centre.
- Matras, Yaron. 2009. *Language contact*. 1st edn. (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.
- Matras, Yaron & Şirin Tufan. 2007. Grammatical borrowing in Macedonian Turkish. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 215–227. Berlin: Mouton de Gruyter.
- McCarus, Ernest N. 2009. Kurdish. In Gernot Windfuhr (ed.), *The Iranian languages*, 587–633. London: Routledge.
- Menz, Astrid. 1999. *Gagausische Syntax: Eine Studie zum kontaktinduzierten Sprachwandel* (Turcologica 41). Wiesbaden: Harrassowitz Verlag.
- Menz, Astrid. 2014. Gagauz. *Journal of Endangered Languages* 3(3). 55–69.
- Murtaza, Durgul Metin. 2016. *Razgrad ili Kubrat ilçesi Alevi-Bektaşî köyleri ağzı (inceleme, metinler, sözlük)*. Edirne: Trakya University. (Master's thesis).
- Németh, Gyula. 1968. Die türkische sprache des Bartholomaeus Georgievits. *Acta Linguistica Academiae Scientiarum Hungaricae* 18(3-4). 263–271.
- Németh, Gyula. 1970. *Die türkische Sprache in Ungarn im siebzehnten Jahrhundert* (Bibliotheca Orientalis Hungarica 13). Budapest: Akadémiai Kiadó.
- Onar Valk, Pelin. 2015. *Transformation in Dutch Turkish subordination? Converging evidence of change regarding finiteness and word order in complex clauses*. Utrecht: LOT Publications.
- Özkan, Nevzat. 1996. *Gagavuz Türkçesi grameri* (Türk Dil Kurumu Yayınları 657). Ankara: Atatürk Kültür, Dil ve Tarih Yüksek Kurumu.
- Özkan, Nevzat. 2007. *Gagavuz destanları* (Türk Dil Kurumu Yayınları). Ankara: Atatürk Kültür, Dil ve Tarih Yüksek Kurumu.
- Özsoy, Onur, Kateryna Iefremenko & Christoph Schroeder. 2022. Shifting and expanding clause combining strategies in heritage Turkish varieties. *Languages* 7(3). 242. DOI: 10.3390/languages7030242.
- Pandharipande, Rajeshwari V. 1997. *Marathi* (Routledge descriptive grammars). London: Routledge.
- Petrou, Maria. 2019. Diglossia features and bilingualism in the Turkish-speaking community of Western Thrace. In C. Tzitzilis & G. Papanastasiou (eds.), *Γλωσσικές επαφές στα βαλκάνια και στη μ. ασία*, 2nd edn., vol. 1, 222–232. Thessaloniki: Ινστιτούτο Νεοελληνικών Σπουδών (Ιδρυμα Μ. Τριανταφυλλίδη).

- Petrou, Maria. 2021. *Der türkische Dialekt von Westthrakien* (Turcologica 125). Wiesbaden: Harrassowitz Verlag.
- R Core Team. 2021. *R: A Language and environment for statistical computing*. Version 4.3.0. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Schreiber, Laurentia. 2019. Romeyka. In Geoffrey Haig & Geoffrey Khan (eds.), *The languages and linguistics of Western Asia: An areal perspective*, 892–934. Berlin: De Gruyter Mouton.
- Schreiber, Laurentia. 2022. *A (contact-)grammar of Romeyka*. Bamberg & Gent: Otto-Friedrich-Universität Bamberg & Universiteit Gent. (Dissertation).
- Schroeder, Christoph. 2016. Clause combining in Turkish as a minority language in Germany. In Mine Güven, Didar Akar, Balkız Öztürk & Meltem Keleşir (eds.), *Exploring the Turkish linguistic landscape: Essays in honor of Eser Erguvanlı-Taylan* (Studies in Language Companion Series), 81–102. Amsterdam: John Benjamins. DOI: 10.1075/slcs.175.06sch.
- Schroeder, Christoph. 2018. *The post-verbal position in heritage Turkish in Germany: Focus on register*. Talk delivered at EuroSLA, Colloquium Turkish Heritage Speakers in Germany, Münster.
- Schroeder, Christoph & Kateryna Iefremenko. 2019a. Göçmen Türkçesinde cümle birleştirme: Pilot çalışma. In Kamil İşeri (ed.), *Dilbilimde güncel tartışmalar*, 243–250. Ankara: Dilbilim Derneği Yayınları.
- Schroeder, Christoph & Kateryna Iefremenko. 2019b. *The post-verbal position in Turkish in two contact settings: Turkish–German bilinguals in Germany and Turkish–Kurdish bilinguals in Turkey*. Talk delivered at the Workshop ‘Turkish in Europe’, Universität Hildesheim.
- Schroeder, Christoph & Kateryna Iefremenko. 2019c. *The postverbal position in heritage Turkish in Germany: Focus on register*. Poster presented at the 4th workshop on Turkic and languages in contact with Turkic, New York, USA.
- Schroeder, Christoph, Kateryna Iefremenko, Kalliopi Katsika, Annika Labrenz & Shanley E. M. Allen. 2025. Clause combining in narrative discourse: A contrastive analysis across heritage and majority languages. In Shanley E. M. Allen, Mareike Keller, Artemis Alexiadou & Heike Wiese (eds.), *Linguistic dynamics in heritage speakers: Insights from the RUEG group*, 251–286. Berlin: Language Science Press. DOI: 10.5281/zenodo.15775173.
- Schroeder, Christoph, Kateryna Iefremenko & Mehmet Öncü. 2024. The postverbal position in heritage Turkish: A comparative perspective with a focus on non-clausal elements. In Zeynep Kalkavan-Aydın & Yazgül Şimşek (eds.), *Deutsch-Türkische Zweisprachigkeit mit besonderem Fokus auf Ju-*

- gendliche [Turkish-German bilingualism with a special focus on adolescents] (Mehrsprachigkeit [Multilingualism] 58), 109–132. Münster: Waxmann.
- Skopeteas, Stavros. 2015. From OV to VO: Observables and linguistic structures. Talk delivered at the University of Athens, Philosophical Faculty, 24.04.2015.
- Stein, Heidi. 2016. The dialogue between a Turk and a Christian in the *Grammatica Turchesca* of Pietro Ferraguto (1611). Syntactical features. In Éva Á. Csató, Astrid Menz & Fikret Turan (eds.), *Spoken Ottoman in mediator texts*, 161–171. Wiesbaden: Harrassowitz Verlag.
- Sulçevsi, İsa. 2019. *Kosova Türkçesi ve fil yapıları*. Istanbul: Istanbul University. (Dissertation).
- Teza, Emilio. 1892. Un dialogo turco fatto in Italia nel cinquecento. *Rendiconti della R. Accademia dei Lincei* 1(6). 391–407.
- Treffers-Daller, Jeanine, A. Sumru Özsoy & Roeland van Hout. 2006. Oral language proficiency of Turkish–German bilinguals in Germany and Turkey: An analysis of complex embeddings in Turkish picture descriptions. In Hendrik E. Boeschoten, Lars Johanson, Sevgi Ağcagül & Vildan Milani (eds.), *Turkic languages in contact*, 203–219. Wiesbaden: Harrassowitz.
- Turan, Dilek, Elena Antonova-Ünlü, Çiğdem Sağın-Şimşek & Mehmet Akkuş. 2020. Looking for contact-induced language change: Converbs in heritage Turkish. *International Journal of Bilingualism* 24(5–6). 1035–1048.
- van der Auwera, Johan. 2011. Standard Average European. In Bernd Kortmann & Johan van der Auwera (eds.), *The languages and linguistics of Europe*, vol. 1, 291–306. Berlin: Mouton de Gruyter.
- Varol, Orhan. 2014. *İkidilli konuşuculara ait verilerdeki dilbilimsel görünüm*. Ankara: Ankara Üniversitesi. (Dissertation).
- Weil, Gotthold. 1953. Ein unbekannter türkischer Transkriptionstext aus dem Jahre 1489. *Oriens* 6. 239–265.
- Whorf, Benjamin Lee. 1944. The relation of habitual thought and behavior to language. *ETC: A Review of General Semantics* 1(4). 197–215.
- Wiese, Heike. 2020. Language situations: A method for capturing variation within speakers' repertoires. In Yoshiyuki Asahi (ed.), *Methods in dialectology XVI* (Bamberg Studies in English Linguistics 59), 105–117. Frankfurt am Main: Peter Lang.

