

Multiple Relative Marking in 19th Century West Rumelian Turkish

ANDREW DOMBROWSKI
University of Chicago

Introduction

West Rumelian Turkish (WRT) refers to the dialects of Turkish spoken in the western Balkans. It is now spoken primarily in Macedonia and Kosovo, but was previously spoken more broadly in Bosnia, Greece, Albania, and Serbia. They differ from other dialects of Turkish in that they have been heavily affected by neighboring Indo-European languages like Serbian, Albanian, Aromanian, Romani, and Greek, and have undergone many of the changes characteristic of the Balkan *Sprachbund* (Friedman 2003). Table 1 gives a sense of the magnitude of divergence between WRT and other dialects of Turkish by comparing how various diagnostic syntactic constructions are realized in the two varieties.

Table 1. WRT vs. Standard Turkish syntax

WRT (Friedman 2003:61-65)	Standard Turkish
Lâzım-dır çalış-alım necessary-is work-1PL.OPTV ‘We must work.’	Çalış-ma-mız lâzım work-INF-1PL.POSS necessary ‘We must work.’
Başla-yaca-m çalış-am begin-FUT-1SG work-1SG.OPTV ‘I will begin to work.’	Çalış-ma-ya başla-yacağ-ım work-INF-DAT begin-FUT-1SG ‘I will begin to work.’
Yok-tur biz-im-le gele-sin (is_not)-is us-POSS-INST come-2SG.OPTV ‘You will not come with us.’	Biz-im-le gel-me-yecek-sin us-POSS-INST come-NEG-FUT-2SG ‘You will not come with us.’
Ne zaman git-ti-k sinema-ya what time go-PAST-1PL cinema-DAT ‘When we went to the movies...’	Sinema-ya git-tiğ-imiz-de cinema-DAT go-PAST.PART-1PL.POSS-LOC ‘When we went to the movies....’

Baba-si father-3SG.POSS 'Ali's father'	Ali-nin Ali-GEN	Ali-nin Ali-GEN 'Ali's father'	baba-sı father-3SG.POSS
--	--------------------	--------------------------------------	----------------------------

As the examples in Table 1 demonstrate, WRT thoroughly restructures the Turkic system of subordination and word order to more closely parallel structures found in the Indo-European languages of the Balkans. In fact, languages spoken in close proximity to WRT like Macedonian and Albanian even have morpheme-for-morpheme correspondences to the WRT constructions in Table 1.

Very few written sources reflect this dialect prior to the 20th century, as Turkish-language writing from the Ottoman Balkans tended overwhelmingly to be in the Ottoman Turkish literary language. Since Ottoman Turkish incorporated Persian and Arabic grammatical and lexical items to such an extent that it was markedly distinct from *any* spoken variety of Turkish, texts in Ottoman Turkish do not provide much direct evidence for the evolution of vernacular varieties like WRT.

Therefore, researchers must look beyond Ottoman Turkish documents for sources that provide direct information about the development of WRT. A particularly valuable source is Gjorgji Pulevski's 1875 *Dictionary of Three Languages*, which was published in Belgrade and contains extensive parallel texts in Macedonian, Albanian, and Turkish. The text is mostly formatted as a series of questions and answers covering an encyclopedic range of content including the creation of the world, natural history and geography, and descriptions of the languages and peoples of the world. It is a historically and sociolinguistically notable source in that Pulevski articulates an early conception of distinct Macedonian national identity in three languages (Friedman 2008). Linguistically, the text reflects Pulevski's provenance from Galičnik in northwestern Macedonia; the Macedonian, Albanian, and Turkish texts all reflect corresponding dialect features, although some supradialectal forms occur. Late 19th century Serbian Cyrillic is used for all forms; in this paper, they are transliterated using standard conventions for transliterating Serbian Cyrillic into Serbian Latin orthography. The Turkish text of this document has not yet received any detailed analysis (although Hazai 1963 provides a brief overview of the document with a partial transcription of the first page). It may be noted that Pulevski is not an ethnic Turk. While this is true, his Turkish usage generally reflects phenomena found elsewhere in WRT, a main goal of the text is to reach a trilingual audience in a colloquially accessible idiom, and there is every reason to think that the role of Turkish as a lingua franca in the Ottoman Balkans played a major role in the development of WRT (Friedman 2006:29).

In this paper, I present a pattern of multiply-marked relative clauses in Pulevski's Turkish that has not been attested elsewhere in Turkic, in which relative clauses can be marked with one of six different combinations of (1) overt par-

ticipial morphology, (2) the complementizer *çi*, and the interrogative *angisi* ‘which.’ I argue that this variation is caused by two factors: first, the fusion of the constructions {*çi* + finite verb} and {participle} into a new construction {*çi* + participle} and second, the introduction of relative marking using the interrogative ‘which’ based on models in surrounding Indo-European languages.

1 Multiply-Marked Relative Clauses in Pulevski’s Turkish: Data

A striking feature of Pulevski’s Turkish is that relative clauses can be marked using multiple overt relativizers. Example (1) is a simple example of a strategy for double-marking relative clauses that is found frequently in Pulevski.¹

- (1) Kuş-lar, *çi* ruzjar-a uç-**an**.
 Bird-PL COMP wind-DAT fly-PART
 ‘Birds which fly upon the wind.’ (Pulevski 1875:34)

In (1), the relative clause is marked both by the complementizer *çi*, which is equivalent to standard Turkish *ki* (itself borrowed from Persian), and the participial marker *-an*. This is distinct from standard Turkish, in which the relative clause is marked only with the participle:

- (2) Rüzgâr-a uç-**an** kuş-lar
 wind-DAT fly-PART bird-PL
 ‘Birds which fly upon the wind.’ (modern standard Turkish)

Contrastingly, Indo-European languages spoken in close proximity to WRT form right-branching relative clauses by combining a relativizer derived from an interrogative pronoun with a finite verb. Example (3) illustrates this in Pulevski’s own Macedonian equivalent for (1):

- (3) Piljinjja, koj-i, ljeta-jed po vetor.
 birds which-PL fly-3PL.PRES upon wind
 ‘Birds which fly upon the wind.’ (Pulevski 1875:34 (Macedonian))

In light of examples (2) and (3), constructions like that found in (1) look like a blend of the Macedonian and Turkish constructions, in which the complementizer *çi* is analogized to the Macedonian *koji*, and the basic syntax is that of the Macedonian in (3), but with the addition of the Turkish participle in *-an*. One possible analysis would be to hypothesize that the participle in *-an* has been reanalyzed as

¹ In examples drawn from Pulevski’s text, participial morphology is in **bold**, complementizers are *italicized*, and interrogative forms are underlined.

a finite verb, analogously to Uzbek, where verbal forms in *-gan* can be either participles or finite perfect tense forms depending on the syntactic environment. Likewise, one might hypothesize that *ći* is not really functioning as a relativizer. However, such an analysis is untenable, as both *ći* and participial morphology can be used by themselves to mark relative clauses, as shown in examples (4) and (5):

- (4) mijakčes, *ći* pać mu xabed demek
 Mijak (speech), COMP pure speech means
 ‘Mijak, which means pure speech.’ (Pulevski 1875:1)

- (5) ve herdžins nefeslji suj-un deru n-de bul-un-**an**
 and all_kinds_of creatures water-GEN under-LOC find-PASS-PART
 ‘and all kinds of creatures that are found underwater.’ (Pulevski 1875:8)

Additionally, relative clauses can be marked using the interrogative *angisi* ‘which,’ as shown in (6):

- (6) Ilja adam-n dejlj sade tene-si var ama var ve
 but man-GEN not only body-3SG.POSS exists but exists and

 rux-u daxi angi-si uljumsuz-dir
 soul-3SG.POSS also which-3SG.POSS immortal-COPL
 ‘But man does not only have his body, but also a soul that is immortal.’
 (Pulevski 1875:11)

It is more common, however, for *angisi* to occur with *ći*. This is illustrated by examples (7) and (8) that illustrate relative clauses formed with *ći* + *angisi* + finite verb.

- (7) Ićindži ljisani anil-r sojle-iš-i-ni *ći*
 second language called-3SG.AOR say-NOM-3SG.POSS-ACC COMP

angi-si-*ljlen* bir kavm, muxabet, ed-ejor,
 which-3SG.POSS-INST one people conversation AUX-PROG

 turče, đibi slavjančes, ve arnautčes
 Turkish like Slavic and Albanian
 ‘The second (sense of) language refers to the speech with which a people communicates, like Turkish, Slavic, and Albanian.’ (Pulevski 1875:39)

- (8) Maxalji anl-r, ol, taraf-lar, *ći* angi-lar-da,
 populated called-3SG.AOR this area-PL COMP person
- čok ixsan jaša-jor.
 which-PL-LOC many live-3SG.PROG
 ‘Areas in which many people live are called populated.’ (Pulevski 1875:22)

Structures like those found in (7) and (8) can also occur with participial marking on the main verb instead of finite marking. This results in triply marked relative clauses, as shown in examples (9) and (10):

- (9) Madem-ljer ičun *ći* angi-ljer
 mineral-PL for COMP which-PL
- daa čok iš-e đir-en.
 more much work-DAT enter-PART
 ‘About the mines that are most often worked.’ (Pulevski 1875:25)

- (10) Ol šečer, mev-ljen-ir, kaljem, ust-u-nde,
 this sugar born-PASS-3SG.AOR stalk top-3SG.POSS-LOC
- ći angi-si, misir, saman-i-na, benze-jen.
 COMP which-3SG.POSS corn straw-3SG.POSS-DAT resemble-PART
 ‘Sugar grows on a stalk that resembles a stalk of corn.’ (Pulevski 1875:33)

Table 2 summarizes the variety of relative constructions that occur in Pulevski’s text. The only possibilities that do not occur are (1) null relative marking and (2) *angisi* ‘which’ + participle.

Table 2. Relative Clause Marking Strategies in Pulevski (1875).

Examples	<i>ći</i>	<i>angisi</i> ‘which’	participle
None	-	-	-
(4)	+	-	-
(6)	-	+	-
(5)	-	-	+
(7), (8)	+	+	-
None	-	+	+
(1)	+	-	+
(9), (10)	+	+	+

2 Relative Clauses: The Turkic Context

In order to formulate an analysis to account for the data found in Pulevski's text, it is first necessary to evaluate the data in the context of relativization in modern WRT and elsewhere in Turkic.²

2.1 Relative Clauses in WRT

Multiply marked relative clauses have not been documented in modern WRT, but WRT does have a tendency to thoroughly reorganize the Turkish system of subordinate clause marking. Of particular note is the tendency to use both interrogatives like *ne* 'what,' *kim + ne* 'who' and *nerde + ne* 'where' and the complementizer *ki* to form relative clauses. These patterns are shown in examples (11) – (13). These examples are drawn from Gostivar Turkish, which is spoken in western Macedonia (Tufan 2007: 171-172). However, other dialects of WRT display the same tendencies (Kakuk 1972:246-247, Sureja 1987:107-109, Friedman 2006:39-40).

- (11) O kış-çe (ne / ki) gel-di biz-de şimdi
that girl-DIM what/COMP come-3SG.PAST we-LOC now

yaşa-r Sımbol-da
live-3SG.AOR Istanbul-LOC
'The girl who came to our place lives in Istanbul now.'

- (12) San-a göster-eci-m ev-i
you-DAT show-FUT-1SG house-ACC

nerde ne anne-m otur-ur
where what mother-1SG.POSS live-3SG.AOR
'I will show you the house where my mother lives.'

- (13) O kız kim-in ne fistan-i dır
that girl who-GEN what dress-3SG.POSS is

kırmızı dır biz-im koyşı.
red is we-GEN neighbor
'The girl whose dress is red is our neighbor.'

² I would like to thank Helga Anetshofer for many useful and insightful observations about relativization in the history of Turkic.

With regard to the use of the complementizer *ki* as a relativizer, Tufan (2007:172) states that it can be used in sentences like (11) with a non-restrictive reading.

Some descriptive uncertainty exists about the status of participles in WRT. In a discussion of relativization patterns like those demonstrated in (11) – (13), Matras (2006:54) states that Macedonian WRT has “no alternative structures to express relative constructions.” However, Tufan (2007: 169) does document the existence of headless relative clauses marked with participles in Gostivar Turkish, as shown in (14).

- (14) Dag-lar-da kay-an-lar var çok.
 mountain-PL-LOC ski-PART-PL exist many.
 ‘There are many people who ski in the mountains.’ (Tufan 2007:169)

Nonetheless, participial morphology is not required in headless relatives in Gostivar Turkish, and it is fundamentally clear that participially marked relative clauses in modern WRT are much less productive than in varieties of Turkic that have not undergone such intense influence from Indo-European languages. Pulevski contains some evidence suggesting that participles were being reanalyzed in 19th century WRT, such as example (15) below in which the participle bears inflectional marking that shows that it has been reanalyzed as an adjective. The Macedonian equivalent to (15) also contains an adjective: *nestanoviti* ‘inconstant.’

- (15) Angi jildis-ljer ċi kalk-ma-**jan**, dirlar, di-ver ban-a.
 which star-PL COMP stay-NEG-PART-COPL-PL say-give CONV I-DAT
 ‘Tell me, which stars are inconstant?’ (Pulevski 1875:73)

2.2 Relative Clauses Elsewhere in Turkic

The native Turkic structure of relative clauses is one that makes heavy use of participles, as shown in examples (2) and (14). However, many Turkic languages that have undergone intense influence from other languages display contact-induced innovations in their system of subordinate marking. In this section, I provide context for the three main innovations characteristic of Pulevski’s Turkish: multiple marking of subordinate clauses, the use of the complementizer *ki* (realized as *ċi* in Pulevski), and the use of the interrogative ‘which’ as a relativizer.

2.2.1 Multiply Marked Subordinate Clauses in Turkic

While the precise patterns of subordination found in Pulevski’s Turkish do not have direct parallels in Turkic, doubly marked subordinate clauses are not at all

unprecedented in the history of Turkish. Conditional clauses marked both by *eger* ‘if’ and conditional inflection are common in pre-modern Turkic (see examples in Adamović 1985: 279-300, Kirchner 2005: 309) and occur in modern Turkish (Lewis 1967: 270). Double-marked temporal subordination occurs in Old Anatolian Turkish (Anetshofer 2005: 135-150). Double-marked relatives appear to be rarer, though Kirchner (2006: 168) gives some examples of left-branching double-marked relatives in an interlinear translation of the Qur’an into Old Anatolian Turkish. A parallel exists in the western dialects of the Tungusic language Even, where doubly-marked relative clauses occur under influence from Yakut, but it is unclear whether these constructions are widely used or acceptable (Malchukov 2006: 129). Again, though, it must be stressed that these parallels are broad in nature: triply marked subordinate clauses like in examples (9) and (10) appear to be unique, at least within Turkic.

2.2.2 The Complementizer *ki* in Turkic

The use of *ki* (also found as *kim* ‘who’) has deep roots in Turkic despite being historically triggered by Persian influence. In addition to being used in non-standard varieties of modern Turkish, it occurs in Old Anatolian and Ottoman Turkish (Prokosch 1980: 172-178, Matras 2006: 52, quoting examples from Adamović 1985). Cypriot Turkish has a relativizer *şu* that is structurally parallel to the examples above using *ne/ki* and is used to form right-branching relative clauses (Demir 2002: 108).

The use of *çi* (< *ki*) in Pulevski’s Turkish is strikingly similar to the use of *ki* in other varieties of Turkish in that it is used not only for relative clauses, but also to introduce what Matras (2006:49) terms “realis complements of factual verbs of cognition, utterance, and perception” and optative clauses. Example (16) shows *çi* introducing the complement of ‘see’:

- (16) *alax, nazar ed-ti çi isljax ol-du.*
 God vision AUX-PAST.3SG COMP good be-PAST.3SG
 ‘and God saw that it was good’ (Pulevski 1875:6)

When used to introduce optative clauses, *çi* is most often followed by an infinitive rather than a finite form. This pattern is shown in (17):

- (17) *ve şafła-sin-lar đoj-un čemer-in-den,*
 and shine-OPTV-PL sky-GEN arch-3SG.POSS-ABL 1PL.POSS-DAT

çi toprag-miz-a išig đel-mek surijedt.
 COMP land- light come-INF purpose

‘And let them shine from the sky, so that light may go onto our earth.’
(Pulevski 1875:7)

Constructions like (17) provide an interesting partial parallel to Ottoman constructions in which the complementizers *kim* and *ki* are used to introduce optative clauses that contain a finite verb.

Some examples of *çi* in Pulevski are difficult to characterize. In (18), it is used in a quotative sense, but there are also many instances elsewhere in the text where the verb ‘call’ is not followed by *çi*, as shown in (19).

(18) Ol anl-r-*çi* ta bijet iljim-i
this called-AOR-COMP character science-3SG.POSS
‘This is called morality.’ (Pulevski 1875:13)

(19) Angi taraf anl-r đun dogu-su.
Which side called-AOR day birth-3SG.POSS
‘Which area is called the east?’ (Pulevski 1875:17)

To summarize, the complementizer *çi* is used in Pulevski’s Turkish in a range of contexts – within relative clauses and in other environments – that do not map well onto 20th century WRT usage, but are deeply grounded in the history of Turkish.

2.2.3 ‘Which’ used as a relativizer in Turkic

In 20th century WRT dialects, the interrogative *angisi* ‘which’ is not used to mark relative clauses. Instead, as discussed in section 2.1, other interrogatives like ‘what,’ ‘who,’ and ‘where’ are used to build relative clauses under the influence of corresponding constructions in neighboring languages like Macedonian and Albanian. However, the use of ‘which’ as a relativizer has been well documented in Gagauz (Menz 1999:91-98). Modern Gagauz is spoken predominantly in Moldova and Ukraine, in conditions of intense contact with Russian, but was spoken in eastern Bulgaria until the middle of the 19th century. This connection to the Balkan *Sprachbund* is suggestive, but the innovation of *angisi* ‘which’ as a relative marker in Gagauz is recent. It was first attested in the 1930s, and seems to have become productive under heavy influence from Russian (Menz 1999:99-100).

Nonetheless, Gagauz provides a compelling analogue to Pulevski’s Turkish for two reasons. First, Gagauz has a comparable range of strategies for forming relative clauses, including (1) participles, (2) the complementizer *ki*, (3) a post-posed complementizer *ani*, and (4) *angisi* ‘which’ (Menz 1999:76-98). Unlike

Pulevski's Turkish, though, it does not seem to be the case that these strategies can co-occur within individual Gagauz sentences. Second, the rapid development of *angisi* 'which' in 20th century Gagauz demonstrates that relativization systems can evolve very quickly under conditions of intense language contact, which helps contextualize the striking differences between Pulevski's relative clauses and those found in dialects of WRT documented in the 20th century.

3 Analysis and Conclusions

As shown in section 1, Pulevski's Turkish text displays a pattern of relative marking in which relative clauses can be marked by any one of six combinations of a participle, the complementizer *çi*, and the interrogative *angisi* 'which.' This pattern is novel in Turkic. The behavior of relative clauses in Pulevski's Turkish can be analyzed as the result of two concurrent innovations in 19th century WRT.

The first innovation concerns the status of participles in WRT. While participles may still exist in WRT (Tufan 2007:169 *contra* Matras 2006:54), they are nonetheless much less common and less productive than elsewhere in Turkish. This leads to the hypothesis that Pulevski's language reflects a transitional state, in which participles still existed but were becoming increasingly marginal in the grammatical system of WRT. In this context, it seems that the earlier relative constructions {*ki* + finite verb} and {participle} were fused into a new relative construction {*çi* + participle}. While examples do occur in which the older state of affairs is preserved – like (4) and (5), where *çi* and participles occur by themselves – they are not as frequent as the {*çi* + participle} construction and are therefore easily interpreted as remnants. This explanation accounts for three of the relative constructions attested in Pulevski.

The second innovation is the introduction of *angisi* 'which' as a relativizer due to influence from Macedonian and Albanian, both of which form relative clauses using 'which.' This directly accounts for constructions in which relative clauses are marked with 'which.'

The only remaining step is to hypothesize that in some cases the two types of constructions can be blended. It is not surprising that constructions with the structure {*çi* + 'which'} emerge, since the complementizer *çi* is used in such a wide range of subordinate clauses. Triply marked relative clauses can be then analyzed as a blending of the common relative constructions {*çi* + 'which'} and {*çi* + participle}. The only combination of markers that does not occur is {'which' + participle}, which is not surprising, since relative clauses marked only with participles are rare and archaic in Pulevski's Turkish, and therefore an unlikely target for blending with a more innovative construction.

A close examination of relative clause marking in Pulevski's Turkish reveals

the existence of a pattern of multiple relative marking heretofore unattested in Turkish. These findings also provide a vivid example of the dramatic ways in which subordination systems can evolve under conditions of intense language contact.

References

- Adamović, Milan. 1985. *Konjugationsgeschichte der Türkischen Sprache*. Leiden: E. J. Brill.
- Anetshofer, Helga. 2005. *Temporale Satzverbindungen in altosmanischen Prosa*. Wiesbaden: Harrassowitz.
- Demir, Nurettin. 2002. Kıbrıs ağızları üzerine. In Demir, Nurettin and Fikret Turan, eds., *Scholarly Depth and Accuracy: A Festschrift to Lars Johanson*, 100-110, Ankara.
- Friedman, Victor. 2003. *Turkish in Macedonia and Beyond: Studies in Contact, Typology, and other Phenomena in the Balkans and the Caucasus*. Wiesbaden: Harrassowitz.
- Friedman, Victor. 2006. West Rumelian Turkish in Macedonia and Adjacent Areas. in Boeschoten, H. and L. Johanson, eds., *Turkic Languages in Contact*, 27-45, Wiesbaden: Harrassowitz.
- Friedman, Victor. 2008. Kulturno-jazičnite pojavi kaj Gjorgji Pulevski. *Thirty-Fourth Annual Symposium on Balkan Linguistics and Literature*: 15-20. Skopje: University of Skopje.
- Hazai, Gyorgy. 1963. Rumeli ağızları tarihinin iki kaynağı üzerine. *Türk Dili Araştırmaları Yıllığı Belleten 1963*: 117-120.
- Kakuk, Suzanne. 1972. Le dialecte turc d'Ohrid. *Acta Orientalia Academiae Scientiarum Hungaricae* (26): 227-282.
- Kirchner, Mark. 2005. Adverbial clauses in an Old Ottoman Turkish interlinear version of the Koran. In Csató, Éva Á. et al., eds., *Linguistic Convergence and Areal Diffusion: Case Studies from Iranian, Semitic, and Turkic*, 305-315, London: Routledge Curzon.

- Kirchner, Mark. 2006. Relative clauses in an Old Ottoman Turkish interlinear version of the Koran. in Boeschoten, H. and L. Johanson, eds., *Turkic Languages in Contact*, 166-175, Wiesbaden: Harrassowitz.
- Lewis, Geoffrey L. 1967. *Turkish Grammar*. Oxford: Clarendon Press.
- Malchukov, Andrei L. 2006. Yakut interference in North-Tungusic languages. in Boeschoten, H. and L. Johanson, eds., *Turkic Languages in Contact*, 122-138, Wiesbaden: Harrassowitz.
- Matros, Yaron. 2006. Layers of convergent syntax in Macedonian Turkish. in Boeschoten, H. and L. Johanson, eds., *Turkic Languages in Contact*, 46-62, Wiesbaden: Harrassowitz.
- Menz, Astrid. 1999. *Gagausische Syntax: Eine Studie zum kontaktinduzierten Sprachwandel*. Wiesbaden: Harrassowitz.
- Prokosch, Erich. 1980. *Studien zur Grammatik des Osmanisch-Türkischen*. Freiburg: Klaus Schwarz Verlag.
- Pulevski, Gjorgji. 1875. *Rečnik od tri jezika*. Belgrade: Državna štamparija.
- Sureja, Jusuf. 1987. *Prizrenski turski govor*. Prishtina: Jedinstvo.
- Tufan, Şirin. 2007. Language Convergence in Gostivar Turkish. Ph.D. thesis, University of Manchester.

Andrew Dombrowski
 University of Chicago
 Department of Slavic Languages and Literatures
 405 Foster Hall
 Chicago, IL 60637

adombrow@uchicago.edu