



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2020

The virtue of imperfection. Gjorgji Pulevski's Macedonian–Albanian–Turkish dictionary (1875) as a window into historical multilingualism in the Ottoman Balkans

Sonnenhauser, Barbara

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-187733>

Journal Article

Published Version

Originally published at:

Sonnenhauser, Barbara (2020). The virtue of imperfection. Gjorgji Pulevski's Macedonian–Albanian–Turkish dictionary (1875) as a window into historical multilingualism in the Ottoman Balkans. *Journal of Historical Sociolinguistics*, 6(1):1-29.

Barbara Sonnenhauser*

The virtue of imperfection. Gjorgji Pulevski's Macedonian–Albanian–Turkish dictionary (1875) as a window into historical multilingualism in the Ottoman Balkans

<https://doi.org/10.1515/jhsl-2018-0023>

Abstract: Even though the Balkans constitute one of the most prominent examples of linguistic areas, little is known about the actual processes and mechanisms contributing to the shaping of this area. Most of the assumptions are based on macro-level analyses and describe the linguistic changes observed in terms of generalising tendencies such as increase in analytism or simplification of structures. In order to approach the processes underlying contact-driven change and area formation, however, the focus needs to be shifted to the micro-level, i. e. the individuals and their communicative practices. Among the rare sources allowing to assume this actor-centred perspective is Gjorgji Pulevski's trilingual dictionary of Macedonian, Albanian and Turkish from 1875, which on the orthographic and morphological level allows for insight into a multilingual speaker's perceptions the languages s/he is exposed to and makes use of in her/his every day communicative practice. The present paper discusses the structural parallels between Macedonian and Turkish observed in the dictionary. It illustrates in how far these parallels may contribute to our understanding of the specific kind of individual multilingualism that provided the basis for the morphosyntactic developments observed for the Balkan linguistic area, and may also help to shed light on the more general nature of these developments. It is suggested that these processes evince an increase in morphological transparency, i. e. morphem-to-function mapping, as the most salient and probably most effective outcome of largely imperfect multilingualism.

Keywords: Balkan area, small-scale multilingualism, language contact, language change, Macedonian, Turkish

*Corresponding author: Barbara Sonnenhauser, Slavic Seminar, University of Zurich, Plattenstrasse 43, Zurich 8032, Switzerland, E-mail: barbara.sonnenhauser@uzh.ch

1 Multilingualism in the Ottoman Balkans

Multilingualism in the Balkans during the Ottoman times provides a prime example of ‘small-scale multilingualism’, i. e. “balanced multilingualism practiced in meaningful geographical spaces sustaining dense interaction and exchange at their interior” (Lüpke 2016: 41). Instead of there being one single donor language, the languages involved – among them Balkan Slavic, Balkan Romance, Albanian and Balkan Turkish (see Friedman 2006a for an introductory overview) – each acted as both donor and recipient, with speakers transferring linguistic elements from and into their native languages. By the specific “adult-based long-term, stable, mutual and intense multilingualism” (Lindstedt 2018) interference phenomena are assumed to have become mutually reinforced (Lindstedt 2000). This in turn led to the “conventionalisation of bilingual speech production errors” (Matras 2004: 71) that eventually resulted in the specific structural similarities that are traditionally regarded as characterising the Balkans as a linguistic area. An impressive body of work has been devoted to the linguistic aspects of these developments, such as detailed descriptions of shared morphosyntactic categories and structures and the presumed processes driving these convergences. More recently, it has been pointed out that in order to understand how these similarities emerged and how the linguistic area has been shaped, the primarily linguistic perspective has to be supplemented by insight from social history and patterns of individual-based interaction (e. g. Lindstedt 2016). This calls for an embedding in historical sociolinguistics (as already suggested by, e. g. Stern 2006; Lindstedt 2014) and an investigation of tangible patterns of multilingual communicative praxis in order to uncover the interactional underpinnings of area formation.

It is a commonplace that language contact occurs in the head of a multilingual speaker (e. g. Riehl 2014: 14) and that, therefore, the beginning of language change “should be sought in actual speaker-to-speaker contacts in definite local settings” (Lindstedt 2016: 52). Strictly speaking, thus, insight into the actual driving forces of contact induced change (and maintenance) requires taking an actor-based perspective. However, individual-level investigations of historical contact scenarios are impeded by the difficulty to obtain evidence of speaker-to-speaker interactions and the varieties of individual multilingualism. In order to overcome these empirical restrictions imposed by the lack of quantitatively and qualitatively sufficient data it has been suggested to draw on insight from neighbouring disciplines (e. g. historical linguistics, Joseph 1983) and allegedly similar contact situations (e. g. creolisation, Hinrichs 2004). But this still remains on a macro-level and involves the danger of judging from observable results to underlying processes in a teleological

way. Furthermore, smaller-scale linguistic changes that have not become entrenched in a language community over time are easily left out of consideration, even though they might be revealing as well for a proper understanding of the pathways of language change and maintenance.

One step towards a sociolinguistically based uncovering of the processes shaping the Balkan linguistic area consists in placing historical literacy into its communicative context. On the example of literary sources dating to the late eighteenth/early nineteenth century, Sonnenhauser (2015) illustrates how the particular ways of using lexical and morphological Turkism – as parts of the texts themselves and as parts of explanatory commentaries – provide for insight into the writers' assumptions on the linguistic background of their potential addressees. This allows to infer the specific shape of societal multilingualism as a prerequisite for language contact to occur. The present paper focuses on individual multilingualism, i. e. language contact within the linguistic practices of one specific language user, as a crucial precondition for linguistic change. The analysis is based on the Macedonian and Turkish parts of Gjorgji Pulevski's¹ (1875) trilingual dictionary of Macedonian, Albanian and Turkish. Even though data gained from single individual language users may not tell too much about the larger-scale changes the results of which are observable today, they provide insight into the potential beginnings of these processes that root in short-term accommodation in linguistic practices. Furthermore, such micro-level data allows to uncover also less stable, volatile transfers and smaller-scale interferences that may not necessarily have resulted in longer term entrenched patterns of language change.

In addition, insight into individual language use and perception helps scrutinize structural generalisations related to contact-induced change. As to the Balkans, an area of enduring, intense and multilateral language contact, *prima facie* contradictory views meet with respect to the morphological and/or syntactic coding of grammatical information and the simplification and/or complexification of structures (see Lindstedt 2018 for a discussion). The Balkan development is traditionally assumed to be characterised by analytic tendencies, while on the other hand synthetic innovations and calques are observed as well (e. g. Fiedler 2004; Hinrichs 2004; Sobolev 2004). Furthermore, interpreting analytic tendencies as structural simplification (e. g. Hinrichs 2000) seems at odds with citing the Balkans as a 'residual zone' and as such being characterised by structural complexity (e. g. Nichols 1992).

¹ This is a conventional spelling preferred here to the scientific transliteration of the name as given in the dictionary, which would be *Đorđe Puljevski*.

The paper is structured as follows: Section 2 embeds Pulevski's dictionary into the sociocultural and linguistic context of nineteenth c. Balkan Slavic literacy. The linguistic insight to be gained from the parallel structuring of the dictionary is described in Section 3; Section 4 discusses the evidence imperfect multilingualism may provide of language change and the underlying structural processes. A short conclusion is given in Section 5.

2 A dictionary of three languages

In his grammar of Bulgarian from 1835, Neofit Rilski provides an impression of the general degree of multilingual variation in the Ottoman Balkans in the early nineteenth c., stating that 'if at one place one word is uttered in Turkish or in Greek (or in some other language), at another place the same word can be found to be uttered in perfect Bulgarian' (Grammatika 1835: 5). Among the sources providing access to individual multilingualism are texts being structured as parallel columns of different languages. Such texts were quite popular in the eighteenth/nineteenth c.; Lindstedt (2016: 51) regards them "as an iconic expression of the sociolinguistic situation of the time". One example of this text tradition is the *Rečnik od tri jezika: S. Makedonski, Arbanski i Turski* [Dictionary of three languages: Slavic Macedonian, Albanian and Turkish] authored by the "self-educated mason" (Friedman 1986 [1975]: 285) Gjorgji Pulevski (1875).

2.1 Sociocultural embedding

Pulevski's *Rečnik* (1875) is embedded in the context of a gradually emerging vernacular Balkan Slavic literacy on the one hand, and the complex linguistic and political situation in the Ottoman Balkans towards the end of the nineteenth c. on the other.

Until the sixteenth c., Slavic literacy in the Balkans adhered to Church Slavonic norms, which were quite different from the vernaculars and hence accessible almost exclusively for the well-educated clergy. Starting with the translation of Damaskinos Stouditis' collection of sermons from Greek into Slavic by the mid-sixteenth c., literacy gradually became closer to the vernaculars. This development was accelerated by the aim of disseminating culturally shared contents as one important means to establish and strengthen a common identity under Ottoman rule (see Ilievski 2005; Sonnenhauser and Fuchsbauser 2014).

For instance, in the preface to Pop Punčo's miscellany of 1796 (the *Pop Punčo sbornik*) this is expressed as follows: 'I have written this book in a simple language such that simple people will understand it'. Pop Punčo's 'simple' language is characterised by a considerable amount of morphosyntactic and lexical variation of South Slavic vernaculars, mixed with Church Slavonic insertions (see Šaur 1970 for a dialectological analysis). It also includes a significant number of Turkisms, which illustrates the widespread active familiarity of speakers of Slavic with Turkish. In Petăr Beron's (1824) *Bukvar s različni poučenija* [Primer with various instructions], a textbook for schoolchildren, Turkish elements are explicitly addressed as an indispensable part of the vernacular. Accordingly, Beron gives Turkish equivalents to those Slavic words that he deemed difficult to understand (Beron 1824: 3).

These comprehension oriented, pragmatic attitudes towards variation and multilingualism differ from those found in prescription oriented, normative literacy as exemplified by Neofit Rilski's Bulgarian grammar of 1835. Regarding linguistic variation 'not as wealth but corruption of the language' (Grammatika 1835: 10), it aimed at establishing 'one general grammar for the whole of Bulgaria, which everyone has to follow' (Grammatika 1835: 3). This grammar was intended to be generally binding for an explicitly delimited language (Bulgarian) related to an explicitly delimited geographical and political entity (Bulgaria; see below). It thereby illustrates the close connection of language, nation, territory and – eventually – state, which Friedman (1997) identifies as one of the dominant ideologies underlying the formation of standard languages in the Balkans – an increasingly political question towards the end of the nineteenth c.

The decreasing power and influence of the Ottoman Empire during the nineteenth c. gave rise to movements of national revival, which were accompanied by a tight connection of language and (national) identity. For the Slavic Balkans, one of the main questions was whether there could be one common language for Bulgarians and Macedonians. This idea became obsolete with the establishment of the Bulgarian Exarchate in 1871, the choice of the eastern Balkan Slavic dialects as the basis for the future literary language and the indexing of the western ('Macedonian') ones as 'degenerate' dialects of Bulgarian (Friedman 1986 [1975]: 284). Pulevski's statement that 'here, Bulgarian is not understood' (Pulevski 1875: 3), might thus very well be regarded as a reaction towards this attitude. Remarkably, however, this is not connected to any further pejorative statements, even though Pulevski was the first to formulate in written form the concept of a Macedonian national identity and Macedonian language (see also Friedman 1986 [1975]: 285),

formulated as follows: “Narod se veljid, ljudi koji se od eden rod koji zboruvajed ednakoi zbor [...] Taka i Makedoncive se narod i mestovo njivno je Makedonija” (Pulevski 1875: 49). [A nation we call people that are from one lineage, speak the same language. That way, also the Macedonians are one nation and their place is Macedonia²].

While clearly distinguishing Macedonian as different from closely related Bulgarian, no exclusive connection of nation, language, territory and statehood can be found in the *Rečnik* (see also Friedman 2008: 18). This is quite different from the ideas expressed in the better known manifesto *Za makedonkite raboti* [On Macedonian matters] by Krste Misirkov (1903) (see Friedman 2005: 70–71 for a short comparison of Pulevski’s and Misirkov’s views on language and identity), which formulates decidedly separatist, exclusive attitudes with respect to Serbs and Bulgarians and connects the question of national identity to that of having a national language (Misirkov 1903: 134–135). This national language will then also serve as a means to expel from Macedonian other languages (in particular Bulgarian and Serbian) and, related to this, the political interests of the other Balkan nations (Misirkov 1903: 139–140).

That Pulevski’s concept of a Macedonian language and identity does not prevent him from accepting and valuing the existence of various languages and identities within what he defines as Macedonian territory can also be seen from the fact that he regarded himself as having multiple sociopolitical affiliations: that of a Macedonian and that of a subject to the Ottoman sultan (Pulevski 1875: XX). Friedman (2008: 17) thus regards Pulevski as a “complex and modern personality” that very well understood the complexities of ethnical-national and civilian-national affiliations in the multilingual and multicultural environment of Macedonia. This conception of an inclusive view on the Macedonian people and all peoples living in Macedonia (and, one may add, their languages) presented by Pulevski (Friedman 2008: 17–18) is one that is only gradually gaining ground in contemporary times, with Albanian having been proclaimed a second official language in what is now North Macedonia, and Turkish being among the recognised minority languages.³

² Note the usage of the proximal deictic markers *-ve/-vo* with *Makedoncive* and *mestovo*, which in addition to spatial proximity may express also emotional proximity and possessivity (e. g. Topolinjska 2006; Sonnenhauser 2009); see also Section 3.2.1, ex. (11) and (12).

³ Still, the ideological equation “languages = weakness” is identified by Friedman (2017: 4) as being effective in parts of the public discourse in Macedonia.

2.2 Linguistic relevance

In addition to reflecting linguistic, literary and political developments and attitudes, texts as those cited above also allow for various kind of insight into the linguistic situation at the end of the nineteenth c. In particular, they indicate multilingualism as part of everyday practice on the one hand and as object of linguistic reflection and ideological attitudes on the other. Parallel column texts as exemplified by the *Rečnik* provide yet another type of evidence: the relevance of imperfect multilingualism as reflected in the linguistic perception and practices of an incomplete multilingual speaker as a potential starting point for contact induced language change.

With Macedonian having reached the status of a standard language equipped with all necessary infrastructure only in 1944, the dictionary is printed in Serbian letters. The main body of the text is written in a north-western Macedonian vernacular which displays features of Pulevski's dialect of Galičnik (see Friedman 2005 for a brief dialectological analysis); the Albanian and Turkish translations exhibit dialect features from the same region (Dombrowski 2014: 80).

Highlighting the virtues of knowing further languages besides one's native tongue (Pulevski 1875: I–II), the text aims at reaching “a trilingual audience in a colloquially accessible idiom” (Dombrowski 2015: 80) and thereby allows to infer the linguistic practices of (mostly) incomplete bi- and multilingual speakers, i. e. the ways they used and perceived the languages they were exposed to.

The fact that Pulevski was not a native speaker of Albanian and Turkish has been regarded as problematic from a linguistic point of view, in particular concerning the Turkish part,⁴ since it “contains many mistakes and forms that seem indicative of either intense Slavic influence or incomplete acquisition” (Dombrowski 2015: 81). At the same time, this imperfection makes the text a valuable source concerning “the patterns of multilingualism in the Ottoman

⁴ The *Rečnik* is also one of the scarce documents for the Turkish vernacular found in the western Balkans by the end of the nineteenth century. Dombrowski (2015) provides a more detailed classification of the Turkish idiom in the *Rečnik*, Dombrowski (2014) focuses on of the marking of relative clauses. Both works are among the rare studies of older stages of West Rumelian Turkish. A short description of the most characteristic traits of contemporary Balkan Turkish can be found in Friedman (2006b), Oktaj (2004) investigates the morphosyntax of dialects from the Ohrid–Prespa region. A grammatical description of Gostivar-Turkish, which elaborates also on the role of language contact, is provided by Tufan (2010); Matras (2004) focuses on syntactic convergences resulting from the contact of Turkish with its Indo-European neighbours in the Balkans. Most of the features discussed in these works can be found already in the *Rečnik*.

Balkans that helped lead to the development of the Balkan *Sprachbund*”, as Dombrowski (2015: 82) rightly points out. This is exactly the starting point for the present paper. It is based on the assumption that putative errors may very well be regarded as performance phenomena (see also Zwicky 1979) that reflect the linguistic perception of multilingual individuals not necessarily having been exposed to schooling and prescriptive norms (according to Koneski 1967: 257, Pulevski “was not an educated man, least with grammatical knowledge”). Alleged mistakes may thus reflect the very practical acquisition and handling of non-native languages in everyday communication and are thus indicative of the imperfect bi- and multilingualism that triggered the linguistic changes characteristic of the Balkan languages. In particular adult-based multilingualism and late L2-acquisition seem to have played a crucial role in these processes. This holds for Pulevski’s Turkish, which he, according to Dombrowski (2015: 81) probably acquired “in adolescence or adulthood”. In the absence of normalised orthography, orthographic decisions, too, may provide for linguistic insight into perception and conception of the languages in use.

On the basis of selected examples and with particular focus on the Macedonian and Turkish parts of the text, it will be illustrated in the following how the *Rečnik* provides insight into the processes underlying ‘linguistic balkanisation’.

3 Macedonian – Turkish encounters

The *Rečnik* consists of an encyclopedic part of questions and answers to various topics, and a dictionary part listing wordforms and short combinations thereof. By its language-wise structuring into three columns it can be read as an early parallel corpus. The parallel structure by which the Macedonian text is aligned with its translations into Turkish and Albanian provides three basic types of evidence of how Pulevski perceives their linguistic structures: 1) Perception of L1, 2) Parallels between L1 and L2, 3) L1 influence on L2.

Insight of type 1) relates to lexical retrieval as suggested by the listing of entries in the dictionary part and by orthographic decisions, insight of type 2) concerns the function and interpretation of structures of L1 as suggested by structural and functional parallels in L2 (what Gołąb 1959 calls “isogrammatisms”), type 3) indicates the way unsupervised acquisition may influence and eventually change L2 by L1, as suggested by the deviations from ‘typical’ Turkish structures. These types – which are, of course, not categorially distinguished – will be sketched from the perspective of Macedonian as L1 and Turkish as L2,

but it is reasonable to assume that the respective processes hold for other combinations of languages as well.

3.1 Perception of L1

Contemporary Macedonian is commonly regarded as being highly analytic. Nominal inflection, in particular on substantives, is reduced to a considerable degree, leaving virtually no morphological indication of the grammatical relation a noun bears to its head. Instead, grammatical relations are co-expressed by means of clitic elements attaching to the left of the verb. In (1) *ja* marks *lekcija* ‘lesson’ as P-argument, i. e. the neither more agent- nor more patient-like NP of the three-place predicate ‘give’. By the verbal prefix *mu* in combination with the nominal dependence marker *na* (glossed as DEP), *svojot tatko* ‘her father’ is marked as the most patient-like argument of, i. e. as assuming the G-relation. The DEP-marker *na* is also employed for adnominal modification, see (2). In addition, it has local and directional meanings, see (3a) and (3b).⁵

- (1) Lekcija_i koja edno 3-godišno devojče
 lesson.SG.F which.SG.F one.SG.N 3-year.SG.N girl.SG.N
 mu_{ii}-ja_i-dade na-[svojot tatko]_{ii}
 DAT.M.3SG-ACC.F.3SG-give.AOR.3SG DEP-POSS.REFL.SG.M.DEF father.SG.M
 ‘a lesson one three-year old girl gave to her father’
 (<https://a1on.mk/archives/645867>; accessed Nov. 04, 2018)
- (2) organizacija na-[finaleta na-[fudbalskata Liga
 organisation.SG.F DEP-finals.SG.N.DEF DEP-football.SG.F.DEF league.SG.F
 na-[šampionite]]
 DEP-champion.PL.DEF
 ‘organisation of the finals of the Football Champions League’
 (www.novamakedonija.com.mk/sport/fudbal; accessed Nov. 04 2018)
- (3) a. utre odime na Bistra da skijame
 tomorrow go.IPFV.PRS.1PL to Bistra to ski.PRS.1PL
 (<http://www.makedonski.info>; accessed Feb. 01, 2019)
 ‘tomorrow we go to Bistra to do skiing’

⁵ Note in these latter cases, there is no prefixal marking on the verb. That is, in (2) and (3), *na* is part of different argument coding strategies than in (1). In any case, it functions as a marker of dependency.

- b. Sekoj predmet si stoi na svoeto
 every object.SG.M REFL stand.PRS.IPFV.3SG at POSS.REFL.SG.N.DEF
 mesto
 place.SG.N
 ‘every object is standing at its place’
 (<http://www.makedonski.info>; accessed Nov. 06, 2018)

Regarding the strategies in (1) and (2) to be analytic is based on the notion of ‘word’ as a linguistic entity and shaped to a large degree by present-day orthographic conventions. There are well reasons to analyse elements like *mu* and *ja* as behaving like prefixes attaching to the verbal head (i. e. object agreement markers), involved in the selection of specific verbal argument relations, and *na* as a dependency marker attaching to the left of noun phrases, involved in the selection of adnominal argument relations,⁶ as indicated by hyphenation in the examples, such as their inseparability from their hosts, their fixed position to the left of their hosts or their integration into one phonological entity with their host. The data in the *Rečnik* suggest that this analysis is also more appropriate from the language users’ point of view in that it reflects how speakers perceive the languages they use and are exposed to.

Prefixal elements such as pronominal short forms and the DEP-marker *na* often appear as one orthographic unit with their hosts. This can be taken to indicate that Pulevski regards these combinations as one single unit – which is suggested also by the fact that they constitute the determining factor for alphabetic ordering even if written separately from their hosts: the entry cited in (4) is listed under *T*, that given in (5) under *M*. That is, subject pronouns, as in (4), and object clitics, as in (5), and the hosts they attach to seem to be accessed as one lexical unit. This in turn supports treating the latter as agreement markers on the verb.⁷

- (4) ti reče, taja reče, tije rekoje
 you tell.AOR.3SG she tell.AOR.3SG they TELL.AOR.3PL
 (Pulevski 1875: 151)

- (5) mi reče, ti reče, mi rekoje
 DAT.1SG tell.AOR.3SG DAT.2SG tell.AOR.3SG DAT.1SG tell.AOR.3PL
 (Pulevski 1875: 138)

⁶ Whether local and directional *na* can or should be analysed along the same lines as purely dependency marking *na* is an issue of ongoing debate in Macedonian linguistics, but not relevant for the present purposes.

⁷ That *ti* in (4) is NOM.2SG and *ti* in (5) DAT.2SG is indicated by the Turkish translations.

The perception of pronominal clitics as an integral part of larger entities is also suggested by examples in which their orthographic treatment differs from that of full pronouns. In (6a), the short DAT-pronoun *ti* (as well as *vam*) constitutes an orthographic unit with the nominal host *fala* ‘praise’, whereas the full pronoun *tebe* (as well as *mene* and *nam*) remains separate in the same context, see (7a). As the translations show, this difference is not a merely formal-orthographic one: the incorporated short pronoun is translated by a verbal construction (Mac *falati* – Tr *metix eder-m sana* ‘I praise you’), (6b), the long form by a nominal one (Mac *fala tebe* – Tr *metix sana* ‘praise to you’), (7b).⁸

- (6) a. *fala, falati, falavam*
 praise praise.DAT.2SG praise.DAT.2PL
 (Pulevski 1875: 153)
- b. *metix eder-m, metix eder-m sana,*
 praise do.AOR.PRS-1SG praise do.AOR.PRS-1SG you.DAT.SG
metix eder-m size
 praise do.AOR.PRS-1SG you.DAT.2PL
 (Pulevski 1875: 153)
- (7) a. *fala tebe, fala mene, fala nam*
 praise you.DAT.SG praise me.DAT praise we.DAT
 (Pulevski 1875: 153)
- b. *metix sana, metix bana, metix bize*
 praise you.DAT.SG praise I.DAT praise we.DAT
 (Pulevski 1875: 153)

With the predicative expression *milo mi je* ‘I like (lit: [it] is dear to me)’ in (8a), predicative (*milo*), pronominal clitic (*mi*) and auxiliary (*je*) are written together. The Turkish equivalent, (8b), is a verbal form in AOR.PRS, which points towards *milomije* being perceived as a verbal unit in Macedonian as well.

- (8) a. *Milomije, milotije, milovam*
 dear.DAT.1SG.be.PRS.1SG dear.DAT.2SG.be.PRS.1SG dear.DAT.2PL
je
 be.PRS.1SG
 (Pulevski 1875: 138)

⁸ The glossing AOR in the Turkish examples refers to the tradition in Turkish linguistics of calling the respective forms ‘(gnomic) aorist’ or *geniş zaman* ‘broad tense’. This ‘aorist’ differs in function from the Macedonian aorist as a definite past tense.

- b. *severim,* *seversin,* *severiniz*
 like.PRS.AOR.1SG *like.PRS.AOR.2SG* *like.PRS.AOR.2PL*
 (Pulevski 1875: 138)

Another revealing case is Pulevski's treatment of nominals preceded by *na*, as in (9). The fact that *na čovekod* is listed under *N* suggests that Pulevski interprets *na* as an integral part of the nominal. In the absence of any further context, it is unclear as to whether *na čovekod* is part of a directional/local phrase, an adnominal modification structure or the G-argument coding pattern. The Turkish equivalents in (9) indicate the latter: *adama*, *bukarija* and *kariljere* are marked by the dative morpheme *-a/-e*. This excludes a directional interpretation for *na* (as well as the local interpretation 'on', which is pragmatically not very likely anyway) and the function of dependent marker in adnominal modification.

- (9) a. *na čovekod,* *na ženava,*
 na man.SG.DEF *na woman.SG.DEF.PROX*
 na ženine
 na woman.PL.DEF.DIST
 (Pulevski 1875: 141)
- b. o *adama,* *bukarija,* *kariljere*
 this man.SG.DAT *that.woman.SG.DAT* *woman.PL.DAT*
 (Pulevski 1875: 141)

In sum, the alphabetical arrangement of the lexical entries and the orthographic decisions concerning compound and separate spelling of elements allow for insight into lexical retrieval and thereby into Pulevski's intuitive conception of 'word'. This may very well depart from the linguist's point of view (see Wray 2015 on various non-linguistic, laymen concepts of 'wordhood'), but exactly for this reason provides an entry point for an actor-based perspective on language contact and the potential starting points of contact induced changes.

3.2 Parallels between L1 and L2

The morphosyntactic analogies between unrelated languages in the Balkan area are captured in Jernej Kopitar's well-known dictum of one form and different linguistic material ("eine Sprachform [...], aber mit dreierley Sprachmaterie", Kopitar 1945 [1829]: 253). With the parallel alignment of texts, the data in the *Rečnik* illustrate how multilingual speakers may have contributed to this picture

by drawing analogies between the languages they were exposed to. In addition, the translations provide insight into the functions of linguistic structures by the end of the nineteenth c. that otherwise need to be projected from present day data. These parallels also provide the empirical basis towards assessing of how the perception of non-native languages (here: Turkish) may have exhibited influences on L1 (here: Macedonian).

3.2.1 Nominal determination

Contemporary Standard Macedonian exhibits a three-way system of nominal determination. The respective markers attach to the left-most element in a noun phrase and are based on the roots *-t-*, *-v-*, *-n-*. In addition to determination, they also express a three-way proximity distinction, see (10).

- (10) a. *daj mi ja knjigata*
 give.IMP.2SG 1SG.DAT F.3SG.ACC book.SG.F.DEF
 ‘give me the book’
 (Koneski 1996: 229)
- b. *na ti ja knjigava*
 here 2SG.DAT F.3SG.ACC book.SG.F.DEF.PROX
 ‘here you are, the book [the one close to me]’
 (Koneski 1996: 229)
- c. *daj mi ja knjigana*
 give.IMP.2SG 1SG.DAT F.3SG.ACC book.SG.F.DEF.DIST
 ‘give me that book [the one somewhat further away]’
 (Koneski 1996: 229)

This triple marking is found also in Pulevski’s dialect of Macedonian, see (11a) and (12a). The Turkish equivalents he gives for the Macedonian nouns in the definite form have the demonstratives *bu* (neutral) and *o* (distal), orthographically attached to their host (11b), (12b). In addition, the ACC-suffix *-u/-i* is used, which marks referentially specific direct objects (e. g. Heusinger and Kornfilt 2005). Obviously, thus, Pulevski considers the combination of ‘prefixed demonstrative + suffixed specificity marker’ as an adequate way of rendering the complex semantics of the Macedonian definiteness markers, even though the proximity grades are not congruent (see Topolinjska 2006; Sonnenhauser 2009 on the role of specificity with the Macedonian ‘articles’). Note also the plural marking on the mass nouns *sol* ‘salt’ and *mljeko* ‘milk’. Pluralising mass nouns is quite common in Turkish, where plural-marking primarily serves to induce an

individualising interpretation for a class and hence implies a sortal interpretation (see, e. g. Ersen-Rasch 2001: 25–26).

- (11) a. soljon, soljta, soljevite
 salt.SG.DEF.DIST salt.SG.DEF salt.PL.DEF
 (Pulevski 1875: 150)
- b. butuzu, otuzu, otuzlari
 this.salt.SPEC that.salt.SPEC that.salt.PL.SPEC
 (Pulevski 1875: 150)
- (12) a. mleko, mljekovo, mljeka
 milk.INDEF milk.DEF.PROX milk.INDEF.PL
 (Pulevski 1875: 137)
- b. sut, bu sutu, sutljer
 milk.INDEF this milk.SPEC milk.INDEF.PL
 (Pulevski 1875: 137)

The deictic component involved in the Macedonian articles is also ascribed transposed interpretations, such as the expression of a possessive relation for the proximal marker based on the root *-v-* (see Topolinjska 2006; Sonnenhauser 2014). This is substantiated by (13): the Turkish equivalent to *zemjava* ‘earth. PROX’ is the possessive structure *toprag miz* ‘our earth’.

- (13) a. kakva je zemjava
 what_kind.SG.F be.PRS.3SG earth.SG.F.DEF.PROX
 (Pulevski 1875: 17)
- b. Toprag miz nasldir.
 earth poss.1PL how.EMPH
 (Pulevski 1875: 17)

The orthography applied for *toprag miz* ‘our earth’ could possibly indicate this to be a blend of the Turkish structure with a possessive suffix attaching to the nominal host (*toprağımız* ‘earth.POSS.1PL’) and the Balkan Slavic structure of a possessive clitic postponed to a nominal.⁹

⁹ With different restrictions applying: in Macedonian, this strategy is only possible for nominals denoting kinship relations; in Bulgarian, it is possible for all definite nouns, e. g. Mitkovska (2009).

The most revealing case for the assumption of isogrammatism is (16). Here, the Macedonian past active participle is formed from the IMPF-stem. This constitutes an innovation in Slavic, which otherwise formed this participle only from the AOR-stem. Gołąb (1959) ascribes this to Turkish influence and regards it – together with the retention of the synthetic AOR and IMPF, again parallel to the Turkish patterns – as a decisive step towards the development of (non-)confirmativity to be the main function of the old perfect without the 3rd person auxiliary.

- (16) a. pišel, pišela,
 write.IPFV.IMPF.PTCP.SG.M write.IPFV.IMPF.PTCP.SG.F
 pišele
 write.IPFV.IMPF.PTCP.PL
 (Pulevski 1875: 146)
- b. jazarmiš, jazarmiš, jazarlarmiš
 write.AOR.PTCP.SG write.AOR.PTCP.SG write.AOR.PL.PTCP
 (Pulevski 1875: 146)

Evidence on the possibility of these forms to be used in a non-confirmative sense is provided by examples such as (17), in which the Macedonian participles *osnovalje* and *propadnalje* are rendered with the non-confirmative (indirective) forms *kurulmiš* and (*fet*) *olmišlar* in the Turkish translation. Here, the context provides a non-confirmative embedding, which suggests that Pulevski opted for these forms instead of aorists because of their functional potential. It is thus reasonable to assume that the parallel between the Macedonian and Turkish forms is not merely morphological.

- (17) a. Od kade možeme dase naučime [...] za
 from where can.PRS.1PL DEP.REFL learn.PFV.PRS.1PL about
 carstvana, kako se, osnovalje, ilji
 kindgdom.PL.DEF.DIST how REFL found.PTCP.PL or
 propadnalje [...]
 perish.PTCP.PL
 (Pulevski 1875: 52)
- b. Nereden urele biljiriz ċi [...] padišaxljiklar,
 where.ABL learn.POSSIB can.AOR.1PL that kingdom.PL
 nasl kurulmiš jaot fet olmišlar [...].
 how found.PTCP or conquest become.PTCP.PL
 ‘Where can we learn from those kingdoms, how they *were* (supposedly) *founded*, how they (supposedly) *perished* [Tr. were conquered].’
 (Pulevski 1875: 52)

Isomorphisms in the verbal domain beyond (non-)confirmativity are given in Table 1.

Table 1: (Perceived) structural parallels of verbal forms, Macedonian – Turkish.

Macedonian	Turkish		
da-va-te give-IPFV.PRS-2PL	ver-er-siniz give-AOR-2PL	‘you give’	(gnomic) present
da-l give.PFV-PTCP.SG	ver-miş give.PRS-PTCP.SG	‘must have given’	non-confirmative
da-va-l give-IPFV-PTCP.SG	ver-er-miş give-AOR-PTCP.SG	‘must have given’	non-confirmative
ži-vuv-a live-IPFV.PRS-1SG	yaşa-r-ım live-AOR-1SG	‘I live’	(gnomic) present
ži-vuva-l live-IPFV-PTCP.SG	yaşa-r-miş live.-AOR.-PTCP.SG	‘must have lived’	non-confirmative
ži-vuv-av live-IPFV-IMPF.1SG	yaşa-r-dım live-AOR-PST.1SG	‘I used to lived’	confirmative

3.2.3 Non-assertive structures

A third example of structural parallels that provide insight into the function(s) of Macedonian structures concerns the particle *da*. In contemporary Macedonian, it serves as morphological exponent of the subjunctive, figuring in complex predicate formation on the verbal and clausal level, see (18) and (19).¹⁰ Under the headings of ‘replacement of infinitive by finite subjunctive structures’ and ‘realis vs. irrealis complementizers’ (see Friedman 2006a for a short overview), these structures are commonly regarded as being triggered by contact and areal influence.

- (18) Rabortite na stadionot treba da-započnat
 work.PL.DEF on stadium.SG.M.DEF must SUBJ-begin.PFV.PRS.3SG
 vo april 2019 godina
 in April 2019 year.F.SG
 ‘The construction works on the stadium must begin in April 2019.’
 (www.novamakedonija.com.mk/sport/fudbal; accessed Nov. 05 2018)

¹⁰ Again, the dashes indicate the non-separability of *da* from its verbal host.

- (19) Vodečkite evropski klubovi planiraat
 leading.PL.DEF European.PL club.PL plan.IPFV.PRS.3PL
 da-formiraat Super liga
 SUBJ-form.PFV.PRS.3PL super league.SG.F
 ‘The leading European clubs plan to establish a Super league’
 (www.novamakedonija.com.mk/sport/fudbal; accessed Nov. 05 2018)

The Turkish translations given for the Macedonian entries provide evidence on how *da*-constructions are interpreted and used by the speakers of that time without having to project contemporary analyses onto an older stage.

In (20) and (21), the Turkish equivalents to the Macedonian *da*-constructions exhibit the volitional morpheme *-e-*. This indicates that here, *da* serves the expression of non-assertive meanings typical of subjunctives (present tense in 20, past tense in 21). Moreover, it is orthographically attached to the verbs *dade* ‘give’ and *reče* ‘say,’¹¹ just like Turkish has a morphological exponent of volitionality that attaches to the roots *ver-* ‘give’ and *de-* ‘say’.

- (20) a. dadada, dadaješ, dadaje
da.give.PFV.PRS.1SG da.give.PFV.PRS.2SG da.give.PFV.PRS.2PL
 (Pulevski 1875: 131)
- b. vereim, veresin, veresiniz
give.VOL.1SG give.VOL.2SG give.VOL.2PL
 (Pulevski 1875: 131)
- (21) a. darečev, darečeš, darečete
da.say.AOR.1SG da.say.AOR.2SG da.say.AOR.2PL
 (Pulevski 1875: 131)
- b. dejejdim, dejedi[n], desaidik
say.VOL.PST.1SG say.VOL.PST.2SG say.VOL.PST.1PL
 (Pulevski 1875: 131)

The semantic non-assertiveness of volitionality and the syntactically deranked status of verbs prefixed by *da-* (i. e. the predicate may not be used autonomously, here: the PFV.PRS-form may appear only in combination with markers such as *da*) figures also in the uses of *da* on the clausal level as in (22). The Turkish equivalents have the subordinating conjunction *çî* and the irrealis morpheme *-seydi-*.

¹¹ In this respect (20)–(22) also provide evidence for lexical retrieval, i. e. the perception of *da-* and the auxiliary *sum* ‘be.PRS.1SG’ as parts of complex wordforms.

For both structures a borrowed strategy exists as well, following the Persian model of employing the complementizer *ki* and a finite verb. While this pattern is rare in standard Turkish, in particular for relative clauses (Kornfilt 2000: 47, 60), it prevails in Balkan Turkish (e. g. Matras 2004) and is also attested in the translations to the Macedonian structures in the *Rečnik*. The declarative complement clause introduced by *kako* ‘how; that’ in (25a) is rendered in Turkish not by a non-finite structure but by means of the complementizer *çi*.¹²

- (25) a. govorio sam, kako treba dosta vremena
 say.PTCP.M.SG be.PRS.1SG that is_necessary much time.PL
 (Pulevski 1875: I)
- b. soiljemišim çí ljazim čok zaman
 say.PTCP.1SG that is_necessary much time
 ‘I said that a lot of time is necessary.’
 (Pulevski 1875: I)

The complement clauses in (26) are non-assertive. In Macedonian, this is expressed by means of *da* in combination with the perfective present form *napišem*, i. e. a subjunctive predicate serving as complement to the main predicate *našao sam* (see also Section 3.2.3). The Turkish equivalent has *çi* in combination with the future predicate *jazaim* (which is the phonological spelling of *yazacağım* corresponding in the feature of ‘non-actuality’ to the Macedonian perfective *napišem*).

- (26) a. našao sam za dobro, da ovu
 find.lPTCP.M.SG be.PRS.1SG for good that this.SG.F.ACC
 drugu knjigu ovako napišem
 second.SG.F.ACC book.SG.F.ACC this_way write.PFV.PRS.1SG
 (Pulevski 1875: I)
- b. ben munasib bulmišum çí, buiçindži
 I adequate find.mişPTCP.1SG that this_second
 čitabu, boilje jazaim
 book.SG.ACC this_way write.FUT.1SG
 ‘I found it a good thing to write this second book.’ [lit: that I will/would write]
 (Pulevski 1875: I)

¹² The comma is indeed placed after *çí*; the same can be observed in (26b).

The Balkan Turkish pattern for relative clauses and their structural similarity to Macedonian (with the relative markers *deka* and *što*) is illustrated in (27). Note that the Turkish equivalent exhibits a mixed pattern with the relativiser *ki* and the non-finite form *jaradtiran* (see Dombrowski 2014 for a detailed description of Pulevski's relativisation strategies).

- (27) a. Boga imenuvame navjisoko suštsetvo,
 god.SG.M.OBL name.IPFV.PRS.1PL highest.SG.N being.SG.N
 deka sozdal nas
 that create.PFV.PST ACC.1PL
 (Pulevski 1875: 5)
- b. Jarabi anilrs enjukseg merkumlig,
 god.ACC call.PRS.1PL highest blessed_being
 ċi bizi jaradtiran.
 that ACC.1PL create.PTCP.PRS
 ‘‘God’ we call the highest being that (must have) created us.’
 (Pulevski 1875: 5)

Another instance of mixed strategy can be seen in (28). Here, the first RC ‘the light which governs the day’ makes use of the Turkish strategy of employing a participle (*eden* ‘making’), the second RC has *ċi* plus the participle *eden*.

- (28) a. Teja videlo što upravuvat denot,
 that light what reign.IPFV.PRS.3SG day.SG.M.DEF
 se imenuvat slnce, a tejaa videlo što
 refl call.IPFV.PRS.3SG sun and that light what
 upravuvat noćot se imenuvat,
 reign.IPFV.PRS.3SG night.SG.F.DEF REFL call.IPFV.PRS.3SG
 mesečina i zvezdi.
 moon and stars
 (Pulevski 1875: 7)
- b. Ol išig đunu terpib eden, đuneš anilr,
 that light day.ACC order make.PRS.PTCP sun understand.AOR.3SG
 ve ol išig ċi đedžei tertib eden,
 and that light which night.ACC order make.PRS.PTCP
 ai ve jildisljer, anilr.
 moon and stars understand.AOR.3SG
 ‘That light, which governs the day [Tr. the day ordering light] is called sun,
 and that light, which governs the night [Tr. the light which is night order-
 ing], is called moon and stars’
 (Pulevski 1875: 7)

In particular these mixed strategies disclose the specific repertoire of a multilingual speaker to be characterised less by containing two different and separate systems, but by providing various options, which are not necessarily distinguished language-wise.

4 Evidence from imperfection

The comparative view on the morphosyntactic structures of Macedonian and Turkish in Pulevski's *Rečnik* allows to approximate an actor-based perspective on the linguistic practice under the conditions of small-scale multilingualism. This layman perspective necessarily differs from the more abstract one assumed by linguists aiming at categorizing linguistic entities, capturing structural regularities and describing general trends in diachronic developments. It is also from a genuinely linguistic perspective, often coined by the systems of the contemporary standard languages, that Pulevski's usage of L2 easily appears to be full of mistakes. Approaching his language(s) on a user-centred basis allows to interpret such mistakes as performance phenomena indicating the way multilingual speakers perceive their own and the surrounding non-native languages. The evidence from these performance phenomena, in particular the structural and functional parallels between Macedonian and Turkish, can then be interpreted, with all due caution, as insight into how languages meet in the mind – and practice – of one individual speaker. They support the assumption of one integrative repertoire of possibilities instead of different systems.

In addition, these data allow for insight into the general processes underlying the changes that have contributed to the specific linguistic characteristics of the Balkan languages. Assuming that “the effects of the Balkan contact situation were neither clearly simplifying nor complexifying”, Lindstedt (2018) concludes that “explicit analytic marking is the property common to most grammatical Balkanisms”. However, also an increase in synthetism is observed for the languages in question (see Sobolev 2004). At first sight this contradiction may be resolved by differentiating developments that apply in the nominal from those applying in the verbal domain (e. g. Gvozdanović 2009). The more crucial problem, however, consists in the underlying descriptive-linguistic distinction between wordforms and parts of wordforms. Some of the alleged analytic strategies may as well be described as being morphological in nature. To give only a two examples, Topolińska (1986) analyses ‘analytical’ case marking (see Section 3.1) as morphological marking of syntactic relations on the head, Fiedler (2004) regards the strategy of using particles, e. g. in the coding of grammatical

relations (see Section 3.1) or the expression of the subjunctive (see Section 3.2.3) as morphologisation ‘to the left’. Concerning the above-mentioned question of complexification and/or simplification as underlying trends, situations of intensive language contact in the absence of standardised norms and widespread schooling are said to trigger morphological simplification (see the parallels Hinrichs 2004 draws to creolisation processes). At the same time, however, a tendency towards complexity would conform to the Balkans as a residual zone (in the sense of Nichols 1992), given that such zones are likely to preserve morphological distinctions (which is indeed the case for the synthetic tenses that are retained in Balkan Slavic but lost in the other Slavonic languages).

Insight of the type illustrated in this paper suggests that the observed changes emerge as effects of an increase in morphological transparency, understood as “higher degree of correspondence between semantic or grammatical category and its expression” (Trudgill 2011: 21). An increase in transparency under conditions of intensive language contact without schooling is plausible from the point of view of the individual language users, since, as Rooryck and Vanden Wyngaerd (2015: 213) point out, the “transparency of the mapping between syntactic features and their morphological exponents promotes the learnability of a morphological system”. Regarding language processing, Kusters (2003: 52–57) reports that deviations from the transparency principle are more serious for L2 than for L1 learners, in particular concerning morphology-related deviations such as allomorphy, fusion and fission. That is, an increase in transparency can be related to the amount of L2 speakers. This assumption is supported by what Nichols (2016) discusses for the relation between language spread and linguistic complexity. She distinguishes structural complexity, i. e. inventory size (the number of elements in an inventory, e. g. of phonemes) and opacity, i. e. non-transparency or non-biuniqueness between meaning or category and form (e. g. allomorphy) as two aspects of complexity (Nichols 2016: 118), regarding them as two sociolinguistic effects of language spreads:

[S]ince the spreading language absorbs a sizable population of adult L2 speakers, it undergoes simplification especially of the opacity-reducing (i. e. transparency-increasing) type. [...] Therefore the highland languages (in the eastern Caucasus) and the peripheral languages (around the eastern Eurasian steppe) should exhibit measurably greater opacity than the lowland and/or centrally located spreading languages. In addition, the buildup of inventory size that seems to occur naturally over time in L1 transmission is reduced or halted in spread episodes; thus the spreading languages should be simpler in this regard than the highland and peripheral ones (Nichols 2016: 119–120).

What Nichols (2016) illustrates for the languages in the Caucasus and the eastern Eurasian steppe, i. e. the effects of a language absorbing a considerable amount

of adult L2 speakers, may – on a smaller scale – as well be related to the prestige of a language. More prestigious varieties can be assumed to attract the more L2 speakers than varieties ranking low on this scale. As to the Ottoman Balkans, Lindstedt (2000: 242–43) and Friedman (2006a [1993]: 670) elaborate an illustration of the relative prestige ascribed to the different languages (Friedman 1997: 33–35 offers an overview of how these hierarchies changed in the course of the 20th c. over the territory of what in the 1990s constituted the Republic of Macedonia). While Turkish was not very liked among the Christian population during the Ottoman Empire (e. g. Lindstedt 2000: 239), as the language of the state and administration it ranked highest (Friedman 2006a: 670). At least for practical purposes, it thus attracted a significant number of (incomplete) L2 speakers and can hence be expected to have undergone an increase in transparency under contact conditions. On the other hand, languages figuring low, such as Romani, should be expected to retain opacity and complexity (with opacity related to sociolinguistic isolation and lack of L2 learners, see Nichols 2016: 133). Balkan Slavic (alongside Balkan Romance and Albanian) occupied a position in the middle of that hierarchy, which suggests a high degree of multi-directional multilingualism (Lindstedt 2000: 243; Friedman 2006a: 670). Assuming that language prestige is coupled also to the social prestige of its speakers, regular encounters of speakers of Balkan Slavic with speakers of Turkish and hence mutual exposure to the structure of the other language is very likely. In addition, with Turkish being a highly agglutinating language, this can be expected to trigger an increase in morpheme-to-function correspondence for Macedonian. Both trends are confirmed by Gjorgji Pulevski's linguistic practice as reflected in the *Rečnik*.¹³

The basic evidence provided by the data in the *Rečnik* thus concerns the transparent coding of grammatical distinctions, i. e. the mapping between semantic or grammatical features and their morphosyntactic expression. Transparency manifests itself on two different linguistic levels for the two languages examined

¹³ These prestige relations do not contradict the characterisation of the Ottoman Balkans as an example of small-scale multilingualism as introduced in Section 1. Given that different 'layers' of prestige can be identified, with various languages – displaying quite diverse internal variation in terms of High and Low varieties in the case of Turkish, Greek and Slavic (see Lindstedt 2016 for a sketch) – located at each of them it is hardly possible to speak of one single, dominant language serving as a source language. Even though knowledge of Turkish facilitated economical and administrative exchanges (and was necessary for a career in the Ottoman administration and military), there were no intentions to systematically implement it as an obligatory lingua franca. Neither was there a lingua franca among the Christian population (e. g. Lindstedt 2000: 239).

here: in Macedonian, it is the transparent restructuring on the morphosyntactic level, in Turkish, the transparent restructuring on the clausal level.

5 To conclude

With its almost sentence-by-sentence alignment in the questions-and-answers section and a wordform-based alignment in the dictionary part, Djordje Pulevski's *Rečnik* allows for insight into the perception and usage of languages by a multilingual speaker on the Balkans towards the end of the nineteenth century. Non-native command of at least one L2 can safely be assumed to have been the default situation for the Balkans during the Ottoman times. Communicative practice was thus characterised by efforts to make sense of the other languages against the foil of the native one. Against this background, alleged mistakes allow for insight into the perception and usage of native and non-native languages in multilingual settings.

The examples discussed in this paper are, of course, only selected instances of indicators concerning the linguistic practices of multilingual speakers shaping the languages by their communicative interactions. This first analysis may nonetheless provide a step towards that kind of actor-based investigation that is necessary in order to understand the specifics of contact-induced language change in socio-culturally shaped areas of balanced multilingualism and the ways these specifics may contribute to potential areal developments.

A further step will consist in complementing the linguistic data accessible of individual speakers and their communicative spaces by historical and sociological data. Mapping these data onto their geographical environment by methods of geo-information sciences will allow for the dynamisation of interactions and communicative networks. This requires the collaboration of historians (accessing diverse sources), sociologists (modelling interactional processes) and linguists (modelling processes of language change) and will eventually allow to sketch, on a strictly empirical basis, the mechanisms of areal development applying in the Balkans and other putative areas.

Acknowledgments: The research for this paper was supported by the Swiss National Science Foundation, grant number 100015_176378/1. This support is gratefully acknowledged. I would also like to thank the two anonymous reviewers for their critical and utterly helpful remarks.

References

- Beron, Petăr. 1824. *Bukvarь Sь Različny Poučenę* [Primer with various instructions]. Brašov: Anton'ov Ioannovich.
- Dombrowski, Andrew. 2014. Multiple relative marking in nineteenth century west Rumelian Turkish. *Annual Meeting of the Berkeley Linguistics Society* 38. 79–91.
- Dombrowski, Andrew. 2015. Gjorgji Pulevski's Turkish and Ottoman multilingualism: Syntactic perspectives. *Balkanistica* 28. 79–106.
- Ersen-Rasch, Margarete I. 2001. *Türkische Grammatik für Anfänger und Fortgeschrittene*. Ismaning: Hueber.
- Fiedler, Wilfried. 2004. Der Südosteuropäische Typus des grammatischen Analytismus. Die balkanische Partikelkonstruktion im Verbalsystem. In Uwe Hinrichs (ed.), *Die Europäischen Sprachen Auf Dem Wege Zum Analytischen Sprachtyp*, 362–398. Berlin: Harrassowitz.
- Fielder, Grace E. 1998. Discourse function of past tenses in pre-modern Balkan Slavic prose. In Robert A. Maguire & Alan Timberlake (eds.), *American contributions to the twelfth international congress of Slavists. Cracow, August–September 1998. Literature, linguistics, poetics*, 344–361. Bloomington: Slavica.
- Friedman, Victor A. 1986. Evidentiality in the Balkans: Bulgarian, Macedonian, and Albanian. In Johanna Nichols & Wallace Chafe (eds.), *Evidentiality: The linguistic coding of epistemology*, 168–187. Norwood: Ablex.
- Friedman, Victor A. 1986 [1975]. Macedonian language and nationalism during the nineteenth and early twentieth centuries. *Macedonian Review* 16(3). 280–292.
- Friedman, Victor A. 1997. One grammar, three lexicons: Ideological overtones and underpinnings in the Balkan sprachbund. In Kora Singer, Randall Eggert Egger & Gregory Anderson (eds.), *CLS 33: Papers from the panels on linguistic ideologies in contact; universal grammar; parameters and typology; the perception of speech and other acoustic signals*, 23–44. Chicago: Chicago Linguistic Society.
- Friedman, Victor A. 2004. The typology of Balkan evidentiality and areal linguistics. In Olga Tomić (ed.), *Balkan syntax and semantics*, 101–134. Amsterdam, Philadelphia: John Benjamins.
- Friedman, Victor A. 2005. Za dijalektološkite raboti: Misirkov, Pulevski I Teodorov-Balan [On dialectological matters: Misirkov, Pulevski, and Teodorov-Balan]. In Blaže Ristevski (ed.), *Deloto Na Misirkov. vol. 2*, 69–74. Skopje: MANU.
- Friedman, Victor A. 2006a [1993]. Balkans as a linguistic area. In Keith Brown (ed.), *Encyclopedia of language & linguistics*, 2nd edn., 657–672 Oxford: Elsevier.
- Friedman, Victor A. 2006b. West Rumelian Turkish in Macedonia and adjacent areas. In Hendrik E. Boeschoten & Lars Johanson (eds.), *Turkic languages in contact*, 27–45. Wiesbaden: Harrassowitz.
- Friedman, Victor A. 2008. Kulturno-jazičnite pojavi kaj Gjorgji Pulevski [Cultural-linguistic phenomena in the work of Gjorgji Pulevski]. In XXXIV Naučna Konferencija na XI Međunaroden Seminar za Makedonski Jazik, Literatura i Kultura. *Lingvistika* [Thirty-fourth annual symposium on Balkan linguistics and literature], 15–20. Skopje: University of Skopje.
- Friedman, Victor A. 2017. Language ideology and language change in the Balkans. A view from the early twenty-first century. *Die Welt der Slaven* 62(1). 1–21.

- Gołąb, Zbigniew. 1959. Some Arumanian–Macedonian isogrammatisms and the social background of their development. *Word* 15(3). 415–435.
- Gołąb, Zbigniew. 1997. The ethnic background and internal linguistic mechanism of the so-called Balkanization of Macedonian. *Balkanistica* 10. 13–19.
- Grammatika. 1835. *Neofit Rilski. Bolgarska Grammatika* [Neofit Rilski. Grammar of Bulgarian]. Kraguevac: Kněžhesko Serbska Tipografija.
- Gvozdanović, Jadranka. 2009. Synthetismus und Analytismus im Slavischen. In Sebastian Kempgen (ed.), *Die Slavischen Sprachen: Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung*, 813–828. Berlin, New York: de Gruyter.
- Heusinger, Klaus von & Jaklin Kornfilt. 2005. The case of the direct object in Turkish: Semantics, syntax and morphology. *Turkic Languages* 9. 3–44.
- Hinrichs, Uwe. 2000. Prolegomena zu einer Theorie des Analytismus II. Anhand der Sprachen in Ost- und Südosteuropa. In Uwe Hinrichs & Uwe Büttner (eds.), *Die Südosteuropa-wissenschaften im neuen Jahrhundert*, 107–128. Wiesbaden: Harrassowitz.
- Hinrichs, Uwe. 2004. Südosteuropa-linguistik und Kreolisierung. *Zeitschrift für Balkanologie* 40(1). 17–32.
- Ilievski, Petar H. 2005. *Tradicija i inovacii vo Makedonskite crkovnoslovenski kniževni spomenici od Turskiot period* [Tradition and innovations in the Macedonian Church Slavonic literary monuments from the Turkish period]. Skopje: Makedonska Akademija na Naukite i Umetnostite.
- Joseph, Brian D. 1983. Language use in the Balkans. The contributions of historical linguistics. *Anthropological Linguistics* 25(3). 275–287.
- Koneski, Blaže. 1967. *Za Makedonskiot literaturni jazik* [On the Macedonian literary language]. Skopje: Kultura.
- Koneski, Blaže. 1996. *Gramatika na Makedonskiot jazik* [Grammar of Macedonian]. Skopje: Detska Radost.
- Kopitar, Jernej. 1945 [1829]. Albanische, Walachische und Bulgarische Sprache. In Rajko Nahtigal (ed.), *Jerneja Kopitarja Spisov. II. Del*, 227–273. Ljubljana: Akademija znanosti i umetnosti.
- Kornfilt, Jaklin. 2000. *Turkish*. London, New York: Routledge.
- Kusters, Christiaan W. 2003. *Linguistic complexity: The influence of social change on verbal inflection*. Utrecht: LOT.
- Lindstedt, Jouko. 2000. Linguistic Balkanization: Contact-induced change by mutual reinforcement. *Studies in Slavic and General Linguistics* 28. 231–246.
- Lindstedt, Jouko. 2014. Balkan Slavic and Balkan romance: From congruence to convergence. In Juliane Besters-Dilger, Cynthia Dermarkar, Stefan Pfänder & Achim Rabus (eds.), *Congruence in contact-induced language change: Language families, typological resemblance, and perceived similarity*, 168–183. Berlin, New York: de Gruyter.
- Lindstedt, Jouko. 2016. Multilingualism in the central Balkans in the late Ottoman times. In Maxim Makartsev & Max Wahlström (eds.), *In search of the center and periphery. Linguistic attitudes, minorities and landscapes in the central Balkans*, 51–67. Helsinki: Department of Modern Languages, University of Helsinki.
- Lindstedt, Jouko. 2018. Diachronic regularities explaining the tendency towards explicit analytic marking in Balkan syntax. In Brian D. Joseph & Iliyana Krapova (eds.), *Balkan syntax and (universal) principles of grammar*, 70–84. Berlin, Boston: De Gruyter Mouton.
- Lüpke, Friederike. 2016. Uncovering small-scale multilingualism. *Critical Multilingualism Studies* 2(4). 35–74.

- Matras, Yaron. 2004. Layers of convergent syntax in Macedonian Turkish. *Mediterranean Language Review* 15. 63–86.
- Misirkov, Krste. 1903. *Za Makedonckite raboti* [On Macedonian matters]. Sofija: Liberalnij Klubъ.
- Mitkovska, Liljana. 2009. Expressing prototype possessive relations at the level of the noun phrase in Macedonian and Bulgarian. In Mila Dimitrova-Vulchanova & Olga Mišeska Tomić (eds.), *Investigations in the Bulgarian and Macedonian nominal expression*, 121–147. Trondheim: Tapir akademisk forlag.
- Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago, Ill.: University of Chicago Press.
- Nichols, Johanna. 2016. Complex edges, transparent frontiers: Grammatical complexity and language spreads. In Raffaella Baechler & Guido Seiler (eds.), *Complexity, isolation, and variation*, 117–137. Berlin, Boston: de Gruyter.
- Oktaj, Ahmed. 2004. *Morfosintaksa na turskite govori od ohridsko-prespanskiot region* [Morphosyntax of the Turkish dialects from the Ohrid-Prespa Region]. Skopje: University of Skopje PhD Dissertation.
- Pulevski, Gjorgji. 1875. *Rečnik od tri jezika. S. Makedonski, Arbanski i Turski. Kniga II* [Dictionary of three languages. Slavic Macedonian, Albanian and Turkish. Book II]. Beograd: Državna Štamparcja.
- Riehl, Claudia Maria. 2014. *Sprachkontaktforschung: Eine Einführung. 3*, überarbeitete Auflage. Narr Studienbücher. Tübingen: Narr Verlag.
- Rooryck, Johan & Guido Vanden Wyngaerd. 2015. Morphological transparency and the delay of principle B effect. *Lingua* 155. 121–139.
- Sobolev, Andrej N. 2004. Analytische Tendenzen in den balkanslavischen Dialekten vor dem allgemeinbalkanischen Hintergrund. In Uwe Hinrichs (ed.), *Die Europäischen Sprachen auf dem Wege zum analytischen Sprachtyp*, 243–262. Berlin: Harrassowitz.
- Sonnenhauser, Barbara. 2009. The Macedonian tripartite article: A discourse-oriented account. *Makedonski Jazik* 60. 123–136.
- Sonnenhauser, Barbara. 2014. Constructing perspectivity in Balkan Slavic: Auxiliary variation and the tripartite article. *Balkanistica* 27. 31–66.
- Sonnenhauser, Barbara. 2015. Borrowing in context: A pragmatic perspective on Turkisms in pre-standardised Balkan Slavic. In Emmerich Keliš, Jürgen Fuchsbauer & Stefan N. Newerkla (eds.), *Lehnwörter Im Slawischen: Empirische und crosslinguistische Perspektiven*, 211–236. Wien: Peter Lang.
- Sonnenhauser, Barbara & Jürgen Fuchsbauer. 2014. Corpus-based analysis of changing norms: Tracing the life of Petka Tărnovska from middle Bulgarian Church Slavonic to Balkan Slavic. In Viktor A. Baranov, V. Željzkova & A. M Lavrent'ev (eds.), *Pismenoto nasledstvo i informacionite tehnologii* [Textual heritage and information technologies], 14–16. Sofija, Iževsk: Bulgarian Academy of Sciences.
- Stern, Dieter. 2006. Balkansprachen und kreolsprachen: Versuch einer kontaktypologischen Grenzziehung. *Zeitschrift Für Balkanologie* 42(1–2). 206–225.
- Šaur, Vladimír. 1970. *Pop Punčov Sbornik kak istočnik istoriko-dialektologičeskix issledovanij* [The Pop Punčov Sbornik as a source of historical-dialectological investigations]. Praha: ČSAV.
- Topolinjska, Zuzana. 2006. Trojnjiot člen – da ili ne? [Triple article – yes or no?]. *Južnoslovenski Filolog* 62. 7–15.

- Topolińska, Zuzanna. 1986. Grammatical functions of noun phrases in Balkan Slavic languages and the so-called category of case. In Richard D. Brecht & James S. Levine (eds.), *Case in Slavic*, 280–295. Indiana: Slavica Publishers.
- Trudgill, Peter. 2011. *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford: Oxford University Press.
- Tufan, Şirin. 2010. *Language convergence in Gostivar Turkish – Macedonia: grammar and language contact phenomena*. Saarbrücken: Lambert Academic Publishing.
- Wray, Alison. 2015. Why are we so sure we know what a word is? In John R. Taylor (ed.), *The Oxford handbook of the word*, 725–750. Oxford: Oxford University Press.
- Zwicky, Arnold. 1979. Classical malapropisms. *Language Sciences* 1(2). 339–348.

Online sources

- Digitalen rečnik na makedonskiot jazik [Digital dictionary of Macedonian]. ДИГИТАЛЕН РЕЧНИК НА МАКЕДОНСКИОТ ЈАЗИК. <http://www.makedonski.info> (accessed November 06 2018; February 01 2019).
2016. Лекција која едно 3-годишно девојче му ја даде на својот татко. <https://a1on.mk> (accessed 4 November 2018).
2018. Минхен или Санкт Петербург домаќини на финалето во ЛШ 2021 // Minhen ili Sankt Peterburg domak'in na finaleto vo LŠ 2021 [Munich or Saint Petersburg hosts of the Champions League finales 2021]. <http://www.novamakedonija.com.mk/sport/fudbal> (accessed November 04–05 2018).