



**HAL**  
open science

## Nashta grammatical sketch with examples linked to online corpus

Evangelia Adamou

► **To cite this version:**

Evangelia Adamou. Nashta grammatical sketch with examples linked to online corpus. Slavische Mikrosprachen im absoluten Sprachkontakt, Harrassowitz, In press, 10.24397/pangloss-0000305 . halshs-02417117

**HAL Id: halshs-02417117**

**<https://shs.hal.science/halshs-02417117v1>**

Submitted on 18 Dec 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Nashta grammatical sketch with examples linked to online corpus

Evangelia Adamou

CNRS

## Background on Nashta

The Nashta variety, literally ‘our (language)’, is part of the group of South Slavic spoken in the Balkans and represented in the area by two official languages, that is, Literary Bulgarian (ISO 639-3: bul) and Literary Macedonian (ISO 639-3: mkd). Unlike these languages, however, the speakers of the Nashta variety have been in contact with Greek at least for the past century during which the shift to Greek took place. In the scale of endangerment elaborated by Krauss (2006), Nashta can be considered as a critically endangered variety (D).

## Recordings and annotation

The texts in Nashta were recorded in the small town of Liti (approx. 2,000 inhabitants), which is located 10 km from the city of Thessaloniki in Greece. The recordings were made in the years 2000 among three fluent speakers, two female and one male, in their 70s at the moment of the recordings, and one younger female semi-speaker. They all use Nashta very rarely, their everyday language being Greek. The recordings took place at home, bringing together acquaintances who were asked to speak Nashta although Greek would have been the language used in these circumstances. Interviews with the researcher completed the corpus.

The texts in this corpus were transcribed in broad phonetics using the International Phonetic Alphabet (IPA). Glosses follow the Leipzig Glossing rules.

## Historical background

Liti, where the Balkan Slavic Nashta variety has been recorded, was known as Aivati or Aivatovo during the Ottoman period; see Map in Figure 1. The *Tabula Imperii Romani* K 3 (1976: 78) attests to the existence of a ‘*demos* of Liti’ in 117 BC, whose inhabitants appear to be of Thracian origin. The presence of Slavic agricultural populations in Liti is noted in a seventh-century Byzantine source, entitled *The miracles of Saint Demetrius*. The relevant excerpt is presented below:

On put alors voir nos concitoyens, [semblables par l’effet de la famine à] des morts et des fugitifs, se rendre avec femmes et enfants aux habitations [sklavènes] des environs de Lite et autres lieux voisins, en ramener du blé et des légumes secs...

[We could then see our fellow citizens, [who because of the famine looked] dead and like fugitives, with their women and children, going to the [Slavic] inhabitants around Lite and other neighbouring localities, to get wheat and fruits...]

(Lemerle 1979: 207. My translation from French).

It describes the rush of the people of Salonica to the countryside following the end of one of the city's sieges.

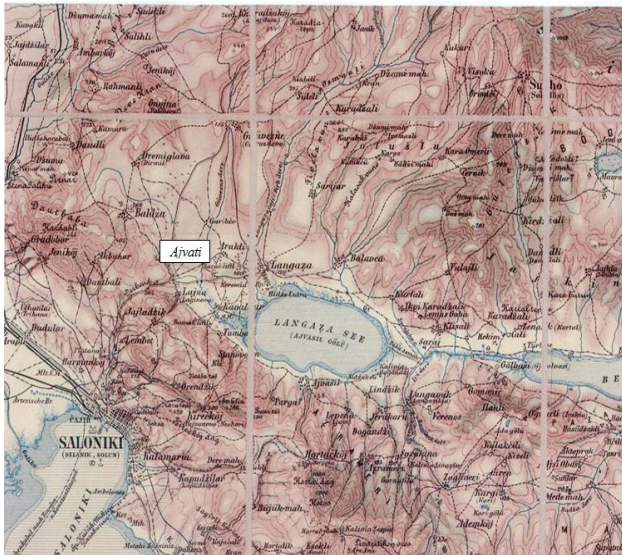


Figure 1. Map from 1903<sup>1</sup> indicating the village of Ajvati (nowadays Liti) nearby the city of Saloniki (nowadays Thessaloniki)

In the beginning of the twentieth century, the Hilmi Pasha census (1904) reports 1,395 so-called Greeks in the village, a term which is more adequately understood as Christian Orthodox. Indeed, during the Ottoman era the people of Liti were part of the Greek *millet*, a grouping based on religious belief. In 1912–1913 Liti was integrated in the Greek state together with part of the geographical region of Macedonia. The new socio-political setting triggered the shift of the Slavic speakers to Greek.

### Consonants, vowels, and stress

Nashta has voiced and voiceless plosives, voiced and voiceless fricatives, nasals, laterals, a palatal approximant and a short alveolar trill. Voiced plosives and fricatives become devoiced word-finally, e.g., ['noʒ-et] 'knife-ART' but [noʃ] 'knife' in sentence 21, Text <https://doi.org/10.24397/pangloss-0000305>. See Table 1 and Table 2. Nashta also has affricates, e.g., [tʃ], [dʒ], [ts]. We note that [s] realizations do not only occur in front of back vowels [i, e], as in contact varieties of Northern Greek, but also in front of [a, æ], as in sentence 3, making its phonological status unclear. Indeed, Nashta has two phonemes, /s/ and /ʃ/, while

<sup>1</sup> Published by Artaria and Co. Landkartenhandlung. Reproduced with the kind authorisation of the Hellenic Literary and Historical Archive (code HLHA Archive K No 06-04. E 1777).

Northern Greek, has one phoneme /s/ realized [sʲ] in front of the vowels /i/, /e/, (see <https://doi.org/10.24397/pangloss-0000317>, sentences 55, 56 *aθa'nas'ios*, 69 *ðimopra'sia*) and [s] in front of vowels [a], [o], [u]. In the Nashta corpus we observe the realizations [s], [sʲ] and [ʃ] but their distribution is not the direct outcome of the combination of the Nashta and the Northern Greek systems. For example, [sʲ] can be realized in front of [a], and can also alter the realization of inherited phonemes as in ‘village’, realized [ʃelo] or [sʲelo] while in other Slavic dialects and languages of the area it is [selo]; see <https://doi.org/10.24397/pangloss-0000317> sentences 84 *'sʲel-to*, 86, 87. Also of interest is the status of syllabic /r/ in Nashta; see ‘church’ in <https://doi.org/10.24397/pangloss-0000317>, sentences 6, 23, 58, 61, 100 *'tsrkvə-ta*.

Table 1. Consonants

		bilabial	labiodental	alveolar	postalveolar	palatal	velar
plosives	unvoiced	p		t			k
	voiced	b		d			g
affricates				ts	tʃ dʒ		
fricative	unvoiced		f	s	ʃ		
	voiced		v	z	ʒ		x
nasal		m		n			
approximant						j	
lateral				l			
trill				r			

Table 2. Consonants in initial, medial, and final position

	INITIAL		MEDIAL		FINAL	
p	pop ‘priest’		'klepem ‘I wink’		sərp ‘sickle’	
b	'brafno ‘flour’		ra'bota ‘work’		-	
t	'tæjko ‘father’		'kotka ‘cat’		i'vurt ‘yoghurt’	
d	'dætsa ‘children’		i'dno ‘one.SG.N’		-	
k	'kamen ‘stone’		'kotka ‘cat’		luk ‘garlic’	
g	'gozba ‘dinner’		'druga ‘other.SG.F’		-	
f	'furna ‘oven’		mə'rafka ‘ant’		krəf ‘blood’	
v	'vəlna ‘wool’		tʃu'viæk ‘man’		-	
s	'setne ‘after’		'meso ‘meat’		nos ‘nose’	
z	zet ‘groom’		vi'zape ‘embroidered’		-	
ʃ	ʃu'kaj ‘scarf’		'niʃto ‘nothing’		noʃ ‘knife’	
ʒ	'zəna ‘woman’		'kazem ‘I say’		-	
x	'xubo ‘good’		sa'xat ‘time/hour’		prax ‘dust’	
m	'majka ‘mother’		'vræme ‘time’		tam ‘there’	
n	nos ‘nose’		'zəna ‘woman’		'kamen ‘stone’	
r	'rətse ‘hands’		'sirne ‘cheese’		'vetʃer ‘evening’	

l	'listo	'leaf	'xabalka	'apple'	'daskal	'teacher'
j	'jagne	'lamb'	'lajno	'dung'	fu'kaj	'scarf'

Nashta has six vowels, see Table 3. [æ] appears in the corpus but its status as a phoneme is problematic.

Table 3. Vowels

	front	central	back
close	i		u
close-mid		ə	
mid	e		o
open-mid		æ	
open		a	

Table 4. Vowels, stressed and unstressed

	STRESSED		UNSTRESSED	
i	'listo	'leaf'	i'dno	'one.SG.N'
e	'meso	'meat'	'sirne	'cheese'
ə	'vəlna	'wool'	mə'rafka	'ant'
æ	'dætsa	'children'		
a	'xabalka	'apple'	'zəna	'woman'
o	'kotka	'cat'	'meso	'meat'
u	'furna	'oven'	fu'kaj	'scarf'

Nashta has stress, that is, stressed syllables are longer than unstressed syllables and have differences in vowel quality. Stress falls mainly, although not exclusively, on the penultimate in the native parts of its vocabulary.

Most vowels present substantial variation in pronunciation, within individual speakers and between speakers. The phonetic variation in the realization of the vowels is captured by a pilot phonetic study of this text based on a digital speech analysis software. Semi-automatic alignment was run with EasyAlign by phonetician Martine Toda.<sup>2</sup> It was then possible to run a semi-automatic script in Praat with the “log file 4” and “triangle vocalique” scripts.<sup>3</sup> Figure 2 illustrates the mean F1 and F2 for the stressed (in grey) and unstressed (in black) vowels based on 455 tokens produced by a female speaker in Text <https://doi.org/10.24397/pangloss-0000305>. It can be seen that the mid vowels /e/ and /o/ are not raised to [i] and [u] as is the case in the contact-varieties of Northern Greek. Figure 3 shows the variability of F1 and F2 values of stressed vowels and Figure 4 of unstressed vowels. It can be seen that unstressed /a/ and /e/ are realized in a very similar area. We also note a significant overlap for /a/ and the non-phonological [æ], (see Text <https://doi.org/10.24397/pangloss->

<sup>2</sup> EasyAlign is developed by Jean-Philippe Goldman, University of Geneva.

<sup>3</sup> Developed by Cédric Gendrot, see <http://gendrot.ilpga.fr/scripts.htm>

0000305, sentences 1 'vræme, 3 'tæjko, 12 'sfætixa, 14 'bæfe, 16 tsæ, among many others), as well as for /e/ and /ə/. In contrast, there is no overlap between /o/ and /u/ as well as between /i/ and /e/.

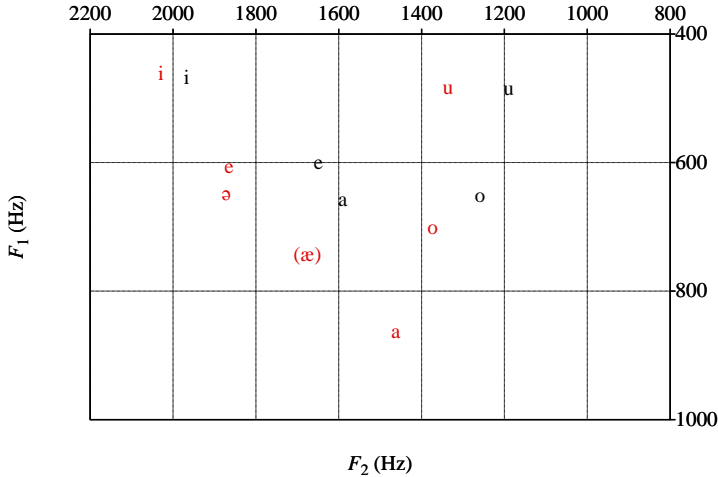


Figure 2. Mean F1 and F2 for the Balkan Slavic Nashta stressed (in grey) and unstressed (in black) vowels based on 455 tokens produced in the spontaneous speech of a female speaker (Adamou 2016)

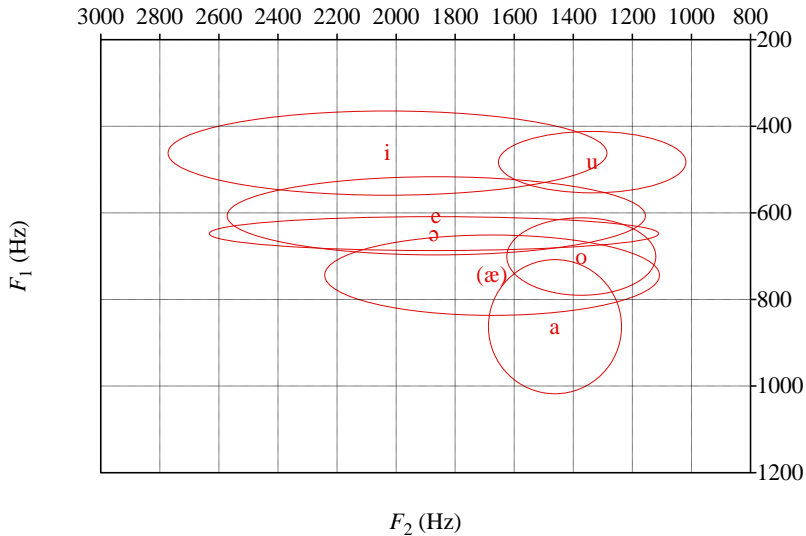


Figure 3. Variability of F1 and F2 values of stressed vowels in Nashta produced in the spontaneous speech of a female speaker (Adamou 2016)

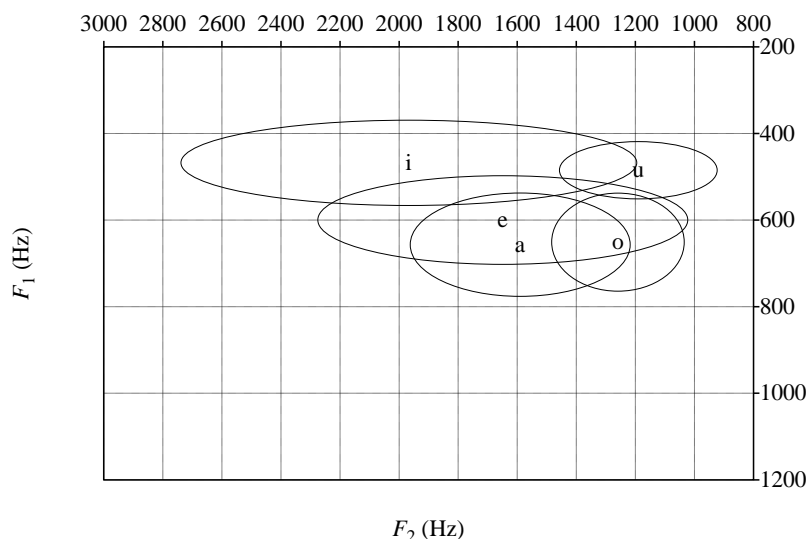


Figure 4. Variability of F1 and F2 values of unstressed vowels in Nashta produced in the spontaneous speech of a female speaker (Adamou 2016)

## Noun phrase

### *Gender*

Nashta has grammatical gender, that is, we find different agreements on the verb, adjective, numeral, and definite article, depending on the gender of the noun (Corbett 2013, WALS 30). Anaphoric pronouns also agree with the gender of their antecedent. Nashta has three grammatical genders: feminine, masculine, and neutral.

### *Number*

Nashta marks plural on the nominals. Interestingly, we note several instances of the old collective *-je*. See an example in Text <https://doi.org/10.24397/pangloss-0000305>, sentences 1, 2, 4, '*komperje* translated as 'ghosts' by our consultants. However, as collective is no longer in use, we gloss it as a mere plural marker. We note in sentence 32 of the same text the use of the old collective with an article in singular form, '*komperje-to*.

### *Case marking*

In Nashta, we can observe the loss of case marking, similar to Bulgarian and Macedonian, but with the exception of several Rhodope and Macedonian varieties that have retained various case markers. In the Nashta corpus we can observe some traces of the vocative; see <https://doi.org/10.24397/pangloss-0000341>, sentences 24, 26, 27 '*babo* 'grandmother', but in nominative '*baba*. We note the use of the same word in Northern Greek dialects, '*vavo*.

### *Demonstratives*

Nashta has a three-way contrast for distance in demonstratives: mid *t*, distal *-n*, and proximal *v*, which is relatively rare in the corpus. Demonstratives in Nashta inflect for gender and number; see Table 5. In Nashta, the demonstrative precedes the noun as commonly found in most languages of Eurasia (Dryer 2013, WALS 88). For examples see Text <https://doi.org/10.24397/pangloss-0000305>

sentence 1 *u'no 'vræme*. The *t* forms in sentences 30, 38, 40, the *n*- forms in sentences 1, 17, 25, 28, and the *v*- forms in sentences 7, 15, 45.

Table 5. Demonstratives

	Singular			Plural
	masculine	neutral	feminine	
<b>proximal</b>	vo	vo	va	-
<b>mid</b>	to	to	ta	tie
<b>distal</b>	on	u'no	'ona	'oni/u'nie

### *Definite article*

Most languages of Europe have definite articles or affixes (Dryer 2013, WALS 37). However, only a few Slavic languages have grammaticalized definite articles: this is the case for South Slavic. The grammaticalization of postposed articles in South Slavic results from both internal and contact-induced factors, as it coincides with a similar development in the Romance languages of the area (Romanian, Aromanian, and Meglenoromanian) and Albanian. Definite articles in South Slavic were grammaticalized from demonstratives, a cross-linguistically common development. Written sources show that the postposed demonstratives attested in Old Church Slavonic documents (between the ninth and eleventh centuries) were grammaticalized into clitic demonstratives and then into clitic articles. Literary Bulgarian has a single definite article but Literary Macedonian and the Rhodope varieties have three definite articles.

Nashta has a single definite article, based on the *-t* form, inflecting for gender and number; see Table 6. The analysis of the corpus shows that 83% of all the nouns in the corpus are determined by a definite article, a high percentage indicating that definite articles are probably fully grammaticalized in Nashta. Greek influence can be seen in the use of the Nashta definite articles with proper nouns, impossible in most Slavic languages of the area; e.g., Text <https://doi.org/10.24397/pangloss-0000321>, sentences 4, 5, 6.

Table 6. Definite article in Nashta

<b>Singular</b>	<b>Masculine</b>	-et
	<b>Feminine</b>	-ta
	<b>Neutral</b>	-to
<b>Plural</b>		-te

### *Indefinite article*

Nashta has an indefinite article inflecting for gender and number; see Table 7. The definite article has grammaticalized from the numeral 'one' (Dryer 2013, WALS 38: demonstrative word same as 'one'). In a quantitative perspective, the analysis of the Nashta corpus reveals that 15% of all the nouns are determined by the numeral *i'din* 'one'. A qualitative analysis of the examples indicates the use of 'one' either as a numeral, which can combine with the definite article *i'din-at* as in sentence 12, <https://doi.org/10.24397/pangloss-0000301>, or with referential nouns, e.g., *i'din a'ndrojino* 'a couple' in sentence 1, <https://doi.org/10.24397/pangloss-0000301>, but not with non-referential and generic nouns. We note 2% of bare plural or singular nouns with an indefinite value. However, we do find occurrences of an indefinite article in plural (see



<https://doi.org/10.24397/pangloss-0000299>, sentence 30; *i'dni 'zene* ‘some women’; <https://doi.org/10.24397/pangloss-0000313>, sentence 13 *i'dni ku'tie* ‘some boxes’).

Table 7. Indefinite article in Nashta

<b>Singular</b>	<b>Masculine</b>	i'din
	<b>Feminine</b>	i'dna
	<b>Neutral</b>	i'dno
<b>Plural</b>		i'dni

### Numerals

Numerals in Nashta precede nouns (NUM-N). We note that distributive numerals are marked by a preceding word, *po/pu*, see sentence 31, <https://doi.org/10.24397/pangloss-0000317> *po sto 'franga*; sentence 15, Text <https://doi.org/10.24397/pangloss-0000305> *'jimafe pu i'din 'kamen*.

### Degree words

In Nashta, the order of a degree word, e.g., '*mnogo* ‘very’, *po* ‘more’ (see sentences 28, 29 for ‘very’ and 4, 9, 16 for ‘more’, <https://doi.org/10.24397/pangloss-0000317>), and the adjective it modifies is DEG-ADJ, where the degree word precedes the adjective. This is another common feature among the languages of Eurasia (Dryer 2013, WALS 91).

### Prepositions

Nashta has prepositions, similar to most languages in Europe (Dryer 2013, WALS 85). See Tables 8 and 9 for the temporal-spatial prepositions. See examples in Text <https://doi.org/10.24397/pangloss-0000321>: *u* ‘locative, punctual’, sentences 2, 18, 27, 33, *na* ‘allative, durative’, *ut* ‘ablative’, sentences 25, 27, 32, 38, *dur* ‘limitative’, sentence 17.

Table 8. Prepositions formed with the locative/punctual *u*

adverb	adposition	adposition	meaning
<i>u'pret</i>	+ <i>u</i>	= <i>u'pret 'u</i>	‘in front’
<i>u'dgore</i>	+ <i>u</i>	= <i>u'dgore 'u</i>	‘on’
<i>'netre</i>	+ <i>u</i>	= <i>'netre 'u</i>	‘in’

Table 9. Prepositions formed with the allative/durative *ut*

adverb	adposition	adposition	meaning
<i>u'dgore</i>	+ <i>ut</i>	= <i>u'dgore 'ut</i>	‘above’
<i>u'dol</i>	+ <i>ut</i>	= <i>u'dol 'ut</i> (spatial)	‘below’
<i>u'pret</i>	+ <i>ut</i>	= <i>u'pret 'ut</i>	‘in front of’
<i>'setna</i>	+ <i>ut</i>	= <i>'setna 'ut</i>	‘after’

### Possessor NP

The order of a possessor noun phrase in relation to the head noun is POSS-N, the possessor preceding the noun it modifies (see text <https://doi.org/10.24397/pangloss-0000305> for examples for kinship, sentence 37

*kate'rin-ta 'nejn-ət peθe'ros*, and ownership, sentence 11 *xadži'joŋeve-te 'kəf(ə)*. The locus of marking in the possessive noun phrases is on the dependent (possessor) and not on the head of the phrase (possessed noun). Possessors in the genitive case agree with the possessed in number and gender.

### *Adjectives*

In Nashta adjectives, expressing a descriptive property such as 'big', 'good', or 'red', precede nouns (AdjN). This order is widespread among the languages of Eurasia, with the exception of Western Europe (Dryer 2013, WALS 87). Adjectives in Nashta inflect for gender and number.

### *Predicative adjectives*

Predicative adjectives in Nashta have nonverbal encoding as typical of languages spoken in the geographical area of Eurasia (Stassen 2013, WALS 118 Predicative adjectives). A verbal copula is typically used although zero copula constructions are allowed (see <https://doi.org/10.24397/pangloss-0000303>, sentence 4 *taj'fa gu'kama* 'it was a) big family' with inversion of the order of the adjective and the noun from AdjN to NAdj). Nominal and locational predications are encoded through the same devices, as frequently encountered in languages of Eurasia (Stassen 2013, WALS 119 Nominal and locational predications: Identical).

## **Linear order**

Nashta has a dominant verb-object order (VO). The order of the object with respect to the verb can be associated to a number of other typological features. Nashta has prepositions and VO order, two features that correlate in the following way: "if a language is OV then it is usually postpositional and if a language is postpositional then it is usually OV" (Dryer 2011, WALS 95). Nashta has an AdjN order, the adjective preceding the noun. According to Dryer (2013, WALS 97), there is no correlation between the order of object and verb and the order of adjective and noun. In intransitive clauses there is no clear dominant order (Dryer 2013, WALS 82, Order of S and V), but in transitive clauses, the order is generally SVO.

## **Pronouns**

### *Alignment of verbal person marking*

Nashta has subject suffixes on the verb (Dryer 2013, WALS 101: Expression of pronominal subjects: Subject affixes on verb). It shows accusative alignment in indexing, where the unique argument of intransitive verbs (S) and the agent-like argument of transitive verbs (A) are indexed on the verb through a single set of suffixes; see Table 10. Notice that third person is zero (Siewierska 2013, WALS 100: Alignment of verbal person marking).

Table 10. Person suffixes

Person	Subject/Agent
	t
1SG	-m
2SG	-ʃ
3SG.M/F/ N	Ø
1PL	-me
2PL	-te
3PL	-t

*Free person pronouns*

Nashta is a pro-drop language, i.e., the free pronouns are not obligatory but appear in marked contexts; see Text <https://doi.org/10.24397/pangloss-0000305> sentences 5 'komperje ja na 'komper ja ci gu u'bija, 7 ja vo 'vetʃer ci 'poja 'vonka, 45 ar'tlik 'ela da ti 'kaze so tʃi'nix ja vo 'vetʃer, 50 ja 'veʎe pak za 'komper 'veʎe gu (ud)və'rvjax. Table 11 shows the forms of the free person pronouns.

Table 11. Free person pronouns

Person	Subject/Agent (NOM)	Indirect object following a preposition (DAT)
1SG	ja	'mene
2SG	ti	'tebe
3SG.M/ N	to	'nego
3SG.F	ta	'neja
1PL	ni	nas
2PL	vi	vas
3PL	'tie	nix

*Person clitics and clitic doubling*

Nashta has two series of person clitics, one for direct objects (glossed ACC) and one for indirect objects (glossed DAT); see Table 12. Both series precede the verb, including clause initially. When the two clitics co-occur, the indirect object clitic always precedes the direct object clitic. We note some variation in the forms of the clitics for first and second person. The person clitics can cross-reference a noun when the noun is determined by a definite article or a possessive pronoun; this is the so-called “clitic doubling”. See examples in Text <https://doi.org/10.24397/pangloss-0000305>, sentences 5, 10, 30, 43, 46, 47. See <https://doi.org/10.24397/pangloss-0000319>, sentence 78 for a case of co-occurrence of the free person pronoun and the clitics; *na nix mu='davaxa* ‘they gave to them’.

Table 12. Person clitics

Person	Direct object (ACC)	Indirect object (DAT)
1SG	(=)me=	mi=
2SG	(=)te=	ti=
3SG.M/N	(=)gu=	mu=
3SG.F	(=)a=	a=
1PL	(=)ne=	ni=
2PL	(=)ve=	vi=
3PL	(=)gi=	mu=

### *Differential object marking*

Based on the Nashta data, Adamou (2006) was the first scholar to suggest an analysis of the Balkan Slavic marking of direct objects through the adposition *na* as a case of differential object marking (Bossong 1985; Lazard 2001). Differential object marking can be seen in the Nashta texts for human objects (see <https://doi.org/10.24397/pangloss-0000309>, sentence 1 *e u 'gledam tsæ na to sə 'bere 'prisiŋe*), or in Text <https://doi.org/10.24397/pangloss-0000305> for human-like objects, as in sentence 46 for the 'ghost' *fo tji'ni gu prova'lix na 'komperət*.

The analysis of the Balkan Slavic data in the light of the differential object marking frame allowed to establish a link between this phenomenon in contemporary varieties and genitive-accusative in Old Church Slavonic (Adamou 2009). This approach also weakened the claim that the Slavic data developed under Romanian and Arumanian or Greek influence (Adamou 2009), a hypothesis that is dominant in the literature (Koneski, Vidoeski & Jašar-Nasteva 1968; Cyxun 1981; Topolinjska 1995; Markovik 2007; Sobolev 2008).

### *Possessive pronouns*

Possession is expressed by a series of free possessive pronouns, that inflect for gender and number, and are generally formed with the definite article, '*mojet* 'my', although indefinite forms are possible, e.g., *moj* 'my', '*negof* 'his'. Possessive pronouns precede nouns. See examples in Text <https://doi.org/10.24397/pangloss-0000305>, sentences 3, 13, 37. The series of indirect object person clitics presented in Table 12, also serves to express possession. In that case the clitics follow the noun; see examples in Text <https://doi.org/10.24397/pangloss-0000305>, sentences 3 and 23 '*majka=mu*, 14 '*kəfce-te=mo*, 16 '*majciji=ne*.

## Verb

### *Tense, mood, aspect markers*

Although the Nashta corpus has few contact words (less than 5%), it shows significant convergence with Greek at the level of the verbal system; see Table 13. This convergence results both from century-long processes within the Balkan Sprachbund and very quick changes within the shifting process to Greek.

Nashta distinguishes between a perfective and an imperfective verbal aspect, the latter being marked through the suffix *-uva*. We note that some verbs are lexically perfective or imperfective. Preverbs, which are well known for Slavic languages, are not very productive in Nashta. We note the use of preverbs such as *na*, *iz*.

The past is generally marked through the suffix *-x*, and in combination with the aspect it is glossed as either “aorist” (AOR) or (IPRF) “imperfect”. For example, in <https://doi.org/10.24397/pangloss-0000311>, sentence 1, the verb ‘to watch’ is inherently imperfective and thus the *-x* suffix is glossed as “imperfect”, *'gledaxme*. However, in sentence 8 of the same text *-x* combines with a verb that is lexically perfective ‘to see’ and thus *-x* is glossed as an “aorist”, *vi'd'axa*. The third person for imperfect has a specific form, *-e*, as can be seen in sentence 1, *'berefe*, and it can combine with the imperfective suffix *-uva*, as in sentence 2, *sa ka't'fuvafe*.

Nashta has a volitive future formed with the particle *ki*, an optative with *da*, and an exhortative with *neka*, all used with finite verbs. These particles can combine with the perfect and the imperfect forms of the verb, but they are mutually exclusive. The rise of a future based on the volitive ‘want’ is among the most well-known Balkan features: a ‘want’-future is found in Greek, Tosk Albanian, Rumanian, Macedonian, Bulgarian, Serbian and Croatian, and Romani. The use of an optative particle follows the loss of the infinitive which characterizes the languages of the Balkans in general.

Nashta — unlike Bulgarian but like Literary Macedonian — has developed a fully grammaticalized ‘have’-perfect (Adamou 2012). It is characterized by an invariable verbal form based on the neuter past participle and ending in *-no/to*. Sentence 14 in <https://doi.org/10.24397/pangloss-0000311> illustrates the use of the ‘have’-past perfect with a transitive verb, *sa 'imafe 'zeto*. Unpublished data also show its use with intransitive verbs such as the verb ‘to die’, e.g., *jima umrjiano* ‘he has died’. Several scholars consider the ‘have’-perfect as a Balkan feature, either under Romance influence or under Greek influence, but these analyses remain controversial.

The grammaticalization of a ‘have’-perfect with the past passive participle in Nashta probably led to the loss of the former Slavic ‘be’-perfect as well as to the even more remarkable loss of all the *-l* verb forms based on the active participle. Greek was probably a catalyst for the loss of the *-l* verbal forms; this loss being a unique feature within the Slavic branch. The gradual loss of the *-l* forms has also been reported for other Balkan Slavic varieties in contact with Greek and is most probably a side-effect of the grammaticalization of the ‘have’-perfect. Indeed, in most Macedonian varieties which have no contact with Greek, the rise of the ‘have’-perfect led to the use of the *-l* forms to express evidentiality. The evidential uses of the *-l* forms are lacking in Nashta, probably because there is no

grammaticalized evidential in Greek. Indeed, under Greek influence, Nashta speakers consistently use the aorist or the narrative present for tales as can be observed in this corpus.

Table 13. Verb morphology in Balkan Slavic Nashta and Greek (Adamou 2016)

	Nashta	Greek
volitive	<i>ki</i>	<i>θa</i>
optative	<i>da</i>	<i>na</i>
exhortative	<i>neka</i>	<i>as</i>
imperfective	<i>-uva-</i> or stress	stem morphology
'have' perfect	<i>imam Vinv(-no)</i>	<i>exo Vinv.</i>

#### *Reflexive, reciprocal, middle*

Nashta has an uninflected reflexive, *sa*, which is also used as a reciprocal, e.g., <https://doi.org/10.24397/pangloss-0000301>, in sentences 1 *sa 'karaxa*, 3, 4, and as a middle, e.g., in sentences 15 *sa stanaxo'resa*, 16 *ci sa i'stemne*, 19 *sa var'na*. It always precedes the verb, including clause initially. Also see <https://doi.org/10.24397/pangloss-0000307> sentences 4 *sa 'klavaxme*, 18 *sa 'imafe* for more examples. The reflexive, reciprocal, and middle always precedes the verb, including clause initially. Table 14 summarizes all the possible uses.

Table 14. Reflexive, reciprocal, and middle

Form	Construction	Use
<i>sa</i>	intransitive	Middle
	intransitive	Naturally reciprocal
	transitive, attributive	Reflexive

## Syntax

### *Relative clauses*

Nashta has externally-headed relative clauses where the relative clauses follow the head nouns (NRel).

### *Subordinators*

Adverbial subordinators, such as 'because', 'when', 'if', are found in the initial position of the subordinating clause.

## Borrowings and codeswitching

For the annotation of this corpus, the morphologically and phonologically integrated, single words were considered as borrowings. When borrowings come from the current-contact language they were tagged and appear in italics. When borrowings come from a past-contact language they are followed in parentheses by the code of the language from which they come, e.g., *tur* for Turkish and *rup* for Arumanian. Some words can be found in the current contact-language, Greek, and in the past contact-language, Turkish. For example, in <https://doi.org/10.24397/pangloss-0000317> in sentences 20, 21, 22, 24, and 26 among others, the word *pa'ri* 'money' comes from Turkish and is most likely a

borrowing that took place during the Ottoman era since it can also be found in all the Balkan languages. In Modern Greek, this borrowing from Turkish is integrated with a plural suffix, as *parades*, or in singular as *paras*, and co-exists with the words *lefta* and *xrimata*, ‘money’. In this case, borrowings were tagged as ‘multiple’ (*mult*) to indicate that although the origin may be a past contact-language, the same word is found in the present contact-language in a similar form.

We considered as codeswitching, the phonologically and morphologically non-integrated or multi-word insertions from the current-contact languages. For example, the Greek insertion in <https://doi.org/10.24397/pangloss-0000317>, sentence 55, ‘*ajos aθa'nasios*, is followed in sentence 58 by its equivalent with native material, *sfi'ti ata'nas*. In sentence 55, the first occurrence of the cluster ‘Saint-Athanasius’, ‘*ajos aθa'nasios*, shows partly phonological adaptation with [j], immediately followed in the repetition by the Greek [j] in sentence 56 ‘*ajos aθa'nasios*. The Greek [θ] is kept in both cases as opposed to the integrated form seen in sentences 58 and 64 as *sfi'ti ata'nas*, with [t]. In sentence 55, the Greek insertion of the cluster ‘Saint-Athanasius’ is not integrated into Slavic as it keeps the Greek nominative case. Also see sentence 2, <https://doi.org/10.24397/pangloss-0000315> for similar variation between the Greek form, with the genitive case, and the Slavic equivalent, *a'jiu aθana'siu sfi'tij ta'nas*. We note that word order is identical in the two languages, i.e., adjective-noun.

The Balkan Slavic Nashta corpus, shows less than 5% Greek words from all word classes. Very few word classes from Greek are concerned, namely nouns (33% of all nouns) and a small number of verbs. Tokens from other word classes may occur but in a peripheral way (Adamou et al. 2016).

#### *Noun borrowings from Greek*

The Nashta corpus shows a variety of nouns in lexical domains which are generally known to be affected by contact. Table 14 illustrates the Greek nouns found in this corpus following an organization by lexical semantic fields.

Table 14. Lexical semantic fields of Greek nouns in the Nashta corpus

Semantic fields	Word
<b>Food and drink</b>	<i>paato</i> ‘plate’
	<i>tindzir</i> ‘pot’
	<i>skurdar</i> ‘garlic mixture’
	<i>liparide</i> ‘kind of fish’
	<i>kulur</i> ‘round loaf’
	<i>lanjide</i> ‘fritters’
<b>Animals</b>	<i>kotopulo</i> ‘chicken’
	<i>kuku</i> ‘cocoon’
<b>Physical world</b>	<i>mura</i> ‘mulberry tree’
	<i>kalampe</i> ‘rose branches’
<b>Basic actions and technology</b>	<i>putir</i> ‘glass’
	<i>paner</i> ‘basket’
<b>House</b>	<i>lamba</i> ‘lamp’
	<i>kandil</i> ‘candle’

	<i>porta</i> ‘door’
	<i>skala</i> ‘ladder’
	<i>kalmotie</i> ‘rose tree mat’
	<i>disko</i> ‘tray’
<b>Time</b>	<i>april</i> ‘April’
	<i>avyusto</i>
	‘August’
<b>Clothing</b>	<i>fustan</i> ‘woman’s dress’
	<i>mandil</i> ‘headband’
	<i>peto</i> ‘lapel’
	<i>sakaki</i> ‘jacket’
	<i>kordjele</i> ‘ribbons’
<b>Religion and beliefs</b>	<i>prika</i> ‘trousseau’
	<i>stefano</i> ‘wedding crown’
<b>Cognition</b>	<i>sxo(ɔ</i> ‘school’
	<i>daskal</i> ‘teacher’
	<i>matima</i> ‘class’
<b>Social and political relations</b>	<i>kratos</i> ‘state’
	<i>ðimoprasia</i> ‘bidding’
	<i>sklovac</i> ‘little slave’
<b>Law</b>	<i>ðikastire</i> ‘court’
<b>Possession</b>	<i>foros</i> ‘tax’
	<i>ispraktor</i> ‘collector’
<b>Modern world culture</b>	<i>aftocinta</i> ‘cars’
	<i>tinikie</i> ‘tin jerry cans’
	<i>fortoti</i> ‘shipper’
	<i>raceta</i> ‘racket’
<b>Quantity</b>	<i>çilde</i> ‘a thousand’
<b>Location, buildings</b>	<i>çendra</i> ‘clubs’
<b>Spatial relations</b>	<i>bala</i> ‘ball’
<b>Kinship</b>	<i>andronjino</i> ‘couple’
<b>Body</b>	<i>ijia</i> ‘health’
<b>Sense and perception</b>	<i>kroto</i> ‘crack’

#### *Verb borrowings*

Greek verbs are integrated in Nashta with the Greek aorist marker *-s*. In Nashta, however, the *-s* aorist marker does not keep the initial aoristic value as it can combine with an imperfective; see sentence 8 in <https://doi.org/10.24397/pangloss-0000299> *ci pla'kosaxme* and compare with the Greek imperfective form *plakoname*. Moreover, as can be seen in <https://doi.org/10.24397/pangloss-0000343#S10>, for the verb ‘to finish’, *biti'sax*, the Greek *-s* is used to accommodate Turkish verbs. It is therefore considered as a loan verb marker (glossed LVM).

#### *Borrowability of conjunctions*

We note that the adversative ‘*ama*’ ‘but’ from Turkish is totally absent from the Nashta corpus unlike in most Slavic varieties of the area (Adamou 2016). Indeed, Nashta uses the Greek adversative *a'la*, as in sentence 25,



<https://doi.org/10.24397/pangloss-0000317> *a'la i 'd'atsa-ta*. We note the use of a different 'ama, from Greek, with a temporal and conditional meaning, as can be seen in sentences 7 'ama 'jide bu'zik, 25 'ama ci sa 'zenixa, 28 'ama ci 'b'axa 'mnogo i'snaf, in <https://doi.org/10.24397/pangloss-0000317>. The replacement of the adversative from the current contact-language indicates its high degree of borrowability as suggested in Matras (2007: 54). This is rendered as a borrowability hierarchy: *but* > *or* > *and*. We note, in sentence 22, <https://doi.org/10.24397/pangloss-0000317> 'is'kaʃ da ti 'kupime i 'is'kaʃ da ti 'daima pa'ri, the use of the Greek *i* 'or', and in sentence 21 the homophonous native *i* 'and', 's'akoj i'din 'berət pa'ri i mu 'davat (note that the Greek coordinator *ke* 'and' is not borrowed).

### *Borrowability of numerals*

In Nashta, numerals below 5 are native and we also note that this is the case for 100; see sentence 34 in Text <https://doi.org/10.24397/pangloss-0000321>, *sto 'franga*. The years are generally in Greek; see sentences 8, *na'pret ut trianda'eksi*, and 31, *to ks'inda o'kto eks'inda e'ja*, both in Text <https://doi.org/10.24397/pangloss-0000321>.

### *Reduplication*

Reduplicated forms in Nashta can be used in order to express intensity; see Text <https://doi.org/10.24397/pangloss-0000305>, sentence 48, e.g., *'rano 'rano* 'very early'.

### *Lexical semantics*

One of the goals of this project was to give access to the original recordings and their annotations. More specifically, the data elicited through visual stimuli allow to unambiguously interpret lexical meanings which may otherwise be misinterpreted by speakers of closely-related languages. For example, in Adamou (2006: 88) the word *dəp/dəb-at* was glossed and translated as 'tree'/'the tree'. In their citation of this example, however, Asenova and Aleksova (2008: 11) based on their knowledge of Bulgarian changed the translation of the word to *chêne* 'oak', as can be seen in the examples below.

Asenova and Aleksova (2008: 11)

На вичеро\_т иле\_нчето хо\_дише, **гу** ли\_жише **на** дъ\_бат, гу чи\_нише  
о\_ште по  
гул'а\_м. (A 88 Ajvatovo, région de Thessalonique)  
« Le soir venait le petit cerf, il le léchait, le **chêne**, ça le rendait encore plus grand. »

На вичеро\_т, но\_шке кава\_ бе\_ше, ути\_де пак, **гу** лиза\_ уно\_ **Ø** дъ\_бат,  
са чини\_  
о\_ште по гул'ам. (A 88 Ajvatovo, région de Thessalonique)  
« Le soir, pendant qu'il faisait nuit, il vint à nouveau, il le lécha, le **chêne**, il devint encore plus grand. »<sup>4</sup>

<sup>4</sup> We also note the change of the original transcription into Bulgarian script, the absence of the original glosses as well as the absence of the citation of the author and change of the village's name according to the Bulgarian dialectological tradition from Liti to Aivatovo.

Text <https://doi.org/10.24397/pangloss-0000309>, based on the video stimuli Pear Story, clearly reveals the mistake of the Bulgarian scholars, as it would be rather amusing that the speaker refers to the ‘pears’ in the video as being collected from an ‘oak tree’; *ji 'bere 'pris'inji-te ji 'klava ut 'fut-ta ji 'klava u 'kofnɔtsə-ta. 'skal-ta tam 'stoi u 'dɔb-at ud'gore*. Indeed, it is common in several Balkan Slavic dialects that the specialized meaning for ‘oak’ has been extended to the more generic meaning ‘tree’.

## References

- Adamou, E. (2006). *Le Nashta. Description d'un parler slave de Grèce en voie de disparition*. Munich: Lincom.
- Adamou, E. (2009). Le marquage différentiel de l'objet en nashta et en pomaque (Grèce). Retour sur l'hypothèse du contact, *Bulletin de la Société de Linguistique de Paris*, 104/1: 383-410.
- Adamou, E. (2012). Le parfait avec ‘avoir’ dans l'aire balkanique : approche multifactorielle et diversifiée d'un balkanisme. Chamoreau, C. & Goury, L. (eds), *Conséquences linguistiques du contact de langues et changements linguistiques. Tendances dans le domaine de la prédication*, 101–119. Paris: CNRS éditions.
- Adamou, E. (2013). Nashta corpus. [doi](#) | [doi](#) | [doi](#) | [doi](#) | [doi](#) | [doi](#) | [doi](#) | [doi](#) | [doi](#) | [doi](#) | [doi](#) | [doi](#) | [doi](#) | [doi](#) | [doi](#) | [doi](#)
- Adamou, E. (2016). *A corpus-driven approach to language contact: Endangered languages in a comparative perspective*. Berlin & New York: Mouton de Gruyter.
- Adamou, E., Breu, W., Scholze, L. & Shen, R. X. (2016). Borrowing and contact intensity: A corpus-driven approach from four Slavic minority languages. *Journal of Language Contact* 9/3: 515-544.
- Asenova, P., & Aleksova, V. (2008). L'aspect balkanique de la nota accusativi personalis. *Zeitschrift für Balkanologie* 44/1: 1-23.
- Bosson, G. (1985). *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Gunter Narr.
- Cyxun, G. A. (1981). *Tipologičeskie problemy balkanoslavjanskogo jazykovogo areala*. Minsk: Nauka i texnika.
- Corbett, G. G. (2013). Number of Genders. In: Dryer, M. S. & Haspelmath, M. (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/30>, Accessed on 2017-07-25.)
- Dryer, M. S. (2013). Definite Articles. In: Dryer, M. S. & Haspelmath, M. (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/37>, Accessed on 2017-07-25.)
- Dryer, M. S. (2013). Indefinite Articles. In: Dryer, M. S. & Haspelmath, M. (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/38>, Accessed on 2017-07-25.)
- Dryer, M. S. (2013). Order of Subject and Verb. In: Dryer, M. S. & Haspelmath, M. (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max

- Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/82>, Accessed on 2017-07-25.)
- Dryer, M. S. (2013). Order of Adposition and Noun Phrase. In: Dryer, M. S. & Haspelmath, M. (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/85>, Accessed on 2017-07-25.)
- Dryer, M. S. (2013). Order of Adjective and Noun. In: Dryer, M. S. & Haspelmath, M. (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/87>, Accessed on 2017-07-25.)
- Dryer, M. S. (2013). Order of Demonstrative and Noun. In: Dryer, M. S. & Haspelmath, M. (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/88>, Accessed on 2017-07-25.)
- Dryer, M. S. (2013). Order of Degree Word and Adjective. In: Dryer, M. S. & Haspelmath, M. (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/91>, Accessed on 2017-07-25.)
- Dryer, M. S. (2013). Relationship between the Order of Object and Verb and the Order of Adposition and Noun Phrase. In: Dryer, M. S. & Haspelmath, M. (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/95>, Accessed on 2017-07-25.)
- Dryer, M. S. (2013). Relationship between the Order of Object and Verb and the Order of Adjective and Noun. In: Dryer, M. S. & Haspelmath, M. (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/97>, Accessed on 2017-07-25.)
- Dryer, M. S. (2013). Expression of Pronominal Subjects. In: Dryer, M. S. & Haspelmath, M. (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/101>, Accessed on 2017-07-25.)
- Koneski, B., Vidoeski, B., & Jašar-Nasteva, O. (1968). Distribution des balkanismes en macédonien. In *Actes du premier congrès international des études balkaniques et sud-est européennes, 1966*, 517-546. Sofia.
- Krauss, M. (2006). Classification and terminology for degrees of languages endangerment. In M. Brenzinger (ed.), *Language Diversity Endangered*, 1-8. Berlin: Mouton de Gruyter.
- Lazard, G. (2001). Le marquage différentiel de l'objet. In Haspelmath M., König E., & Oesterreicher W., *Language Typology and Language Universals, vol. 2*, 873-885. Berlin & New York: Mouton de Gruyter.
- Lemerle, P. (1979 t. 1, 1981 t. 2). *Les plus anciens recueils des miracles de Saint Démétrius*. Paris: CNRS éditions.
- Markovik, M. (2007). *Aromanskiot i makedonskiot govor od ohridsko-strushkiot region*. Skopje: MANY.
- Matras, Y. (2007). The borrowability of structural categories. In Y. Matras, & J. Sakel, *Grammatical borrowing in cross-linguistic survey*, 31–73. Berlin & New York: Mouton de Gruyter.

- Siewierska, A. (2013). Alignment of Verbal Person Marking. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/100>, Accessed on 2017-07-25.)
- Sobolev, A. (2008). On some Aromanian grammatical patterns in the Balkan Slavonic dialects. In *The Romance Balkans, Collection of papers presented at the International Conference The Romance Balkans*, 113-121. Belgrad.
- Topolinjska, Z. (1995). *Makedonskite dijalekti vo Egejska makedonja*. Skopje: Makedonska akademija na naukite i umetnostite.
- Stassen, L. (2013). Predicative Adjectives. In: Dryer, M. S. & Haspelmath, M. (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/118>, Accessed on 2017-07-25.)
- Stassen, L. (2013). Nominal and Locational Predication. In: Dryer, M. S. & Haspelmath, M. (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/119>, Accessed on 2017-07-25.)
- Tabula Imperii Romani K 34. (1976). Slovenska: Union académique internationale.