A User's Guide to This Book

We list here some practical aspects of using this book, with regard to our numbering scheme, our mode of citations, our translations of cited material, our abbreviations and glossing conventions, and details on orthography and transcription.

1 Numbering Within the Text

This work admittedly has a complicated, but nonetheless rational, numbering system for the complex set of sections and subsections within each chapter. To aid the reader, we have provided a comprehensive table of contents in the prefatory material, showing all the section-headings and sub-headings (and sub-sub, and so on), though with numbers only up to three sub-levels; the chapters themselves have the full numbering for all levels of subsections.

Throughout this work when we write "§X.Y . . . ," the first digit is the chapter number, and the remaining digits specify the section and subsection (and sub-subsection, . . .) within that chapter. Thus §1.2.3.3 is to be found in Chapter 1, as sub-subsection 3.3 within subsection 2; the section headings within each chapter lead with the chapter number. Example sentences are numbered consecutively in each chapter (thus (6.1), (6.2), etc. in Chapter 6, (7.1), (7.2), etc. in Chapter 7, and so on), as are tables (thus, Table 6.1, Table 6.2 . . ., Table 7.1 . . .); footnote numbering starts anew in each chapter. Reference to a footnote without a chapter designated signals a note in the same chapter as the reference.

2 Repetitions and Cross-References

We do not expect users of this book to read it cover-to-cover in one sitting. Moreover, some of the material is relevant in different parts of the overall work. Accordingly, we readily acknowledge that there is some repetition here and there; we consider this necessary, especially since readers may turn to a section they feel is particularly germane to their interests without realizing other possible dimensions to the issue at hand. For that reason too, we have attempted to give extensive cross-references internal to the book.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 01:26:45, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139019095

3 Citations

Citations from Ancient Greek texts follow standard Classics abbreviations for authors and works (as listed, e.g., in Liddell, Scott, & Jones, *A Greek–English Lexicon*). Citations from Old Church Slavonic texts follow the common abbreviations as listed in Lunt 1974 (*Old Church Slavonic Grammar*). Reference to works by the same author(s) in the same year is compressed, so that, e.g., 1966abc is used to denote 1966a, 1966b, 1966c.

4 Translations of Quoted Scholarly Material

We provide an English translation of every quotation that is taken from a scholarly work in a language other than English and is cited in the text; the translations in each case are our own, either done individually by one of us or as a joint effort. The presentation of the material to be translated and of the translations is not uniform throughout – sometimes the translation follows, sometimes it is in a footnote – but a translation is always available.

5 Abbreviations and Glossing Conventions

a Language Abbreviations

E	East(ern)
N	North(ern)
S	South(ern)
W	West(ern)
AGrk	Ancient Greek
Alb	Albanian
AMac	Ancient Macedonian
Arbc	Arabic
Arme	Armenian
Aro	Aromanian
Arv	Arvanitika
Aves	Avestan
BCMS	Bosnian-Croatian-Montenegrin-Serbian
Bel	Belarusian
Blg	Bulgarian
BRo	Balkan Romance
BSl	Balkan Slavic
ChSl	Church Slavonic
CoSl	Common Slavic
Cro	Croatian

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 01:26:45, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139019095

Cz	Czech
EMR	Early Modern Romanian
Eng	English
ERT	East Rumelian Turkish
Estn	Estonian
Frn	French
Gag	Gagauz
GBFA	Gorna Belica [Aro Bela di Suprã] Frasheriote Aromanian
Goth	Gothic
Grm	German
Grk	Greek (NB: unmarked sense = Modern Greek)
Heb	Hebrew
Hitt	Hittite
Hung	Hungarian
IE	Indo-European
Ind	Indic
I-R	Istro-Romanian
Itl	Italian
Jud	Judezmo
KPA	Kriva Palanka Arli
Lat	Latin
Lith	Lithuanian
Luv	Luvian
M	Modern (with language names/abbreviations)
Mac	Macedonian
MedGrk	Middle/Medieval Greek
Megl	Meglenoromanian
ModGrk	Modern Greek (where specificity is needed)
Mtn	Montenegrin
NTGrk	New Testament Greek
O	Old (with language names/abbreviations)
OCS	Old Church Slavonic
OEGrk	Ottoman Edirne Greek (cf. Ronzevalle 1911, 1912)
OHG	Old High German
OIr	Old Irish
OT	Old Turkic
Pers	Persian
PIE	Proto-Indo-European
Pol	Polish
Prkt	Prakrit
Rmi	Romani
Rmn	Romanian
Russ	Russian

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 01:26:45, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139019095

SDBR	South Danubian Balkan Romance (= Aromanian + Meglenoromanian)
Skt	Sanskrit
Sln	Slovene
Slo	Slovak
Slv	Slavic
Sp	Spanish
Srb	Serbian
SSl	South Slavic
St	Standard (with language names/abbreviations)
Swed	Swedish
TochA	Tocharian A
TochB	Tocharian B
Trk	Turkish
Trlk	Torlak dialects of BCMS
TsR	Tsărnarekă Meglenoromanian
Ukr	Ukrainian
VLat	Vulgar Latin
Vtn	Venetian
WRT	West Rumelian Turkish
Yid	Yiddish

b OCS Codex Abbreviations

As = Assemanianus (Vajs & Kurz 1929/1955)
Sav = Savvina kniga (Ščepkin 1959)
Supr = Suprasliensis (Zaimov & Kapalo 1982/1983)
Zogr = Zographensis (Jagić 1954)

c Glossing Conventions

In the presentation of data from the various languages covered here, we provide as much grammatical information in the glosses as is needed to make the data intelligible for the particular issue at hand.

We mostly follow the glossing conventions, abbreviations, and symbols in the Leipzig Glossing Rules (www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf), though where needed we have added to or adjusted the Leipzig abbreviations.

. (period)	separating pieces of cumulative morphemes or meanings
- (hyphen)	morpheme boundaries
=	clitic boundaries
′	primary stress (in languages where it is not part of the orthography)
*	reconstructed or hypothetical form (with form in italics)
*	ungrammatical form or string (with form or string in plain font)

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 01:26:45, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139019095

[]	phonetic transcription
/ /	phonemic transcription
< >	graphemes
#	word boundary (thus, word-initial or word-final)
/	used to separate grammatical identification of syncretic forms, e.g., 1/2SG for 1st and 2nd person SG, or to collapse categories for expository purposes, e.g., 1SG/1PL/1DU for first person singular/plural/dual, or for alternatives in forms, glosses, or translations, e.g., *o*/*e* = *o* or *e*
(V)	vowel/zero alternations (i.e., optional vowel)
Ø	zero morpheme (used only when relevant) or marking deletions
+	plus
±	plus or minus (i.e., with or without)
>	develops into, becomes
<	derives from

d Grammatical Abbreviations

Note: Abbreviations that are different or absent from the Leipzig glosses are in **bold**.

ABL	ablative
ACC	accusative
ACT	**active**
ADJ	adjective
ADM	**admirative**
ADV	adverb(ial)
ANIM	**animate**
AOR	**aorist**
ART	article
AUX	auxiliary
C	consonant
CAUS	causative
CMPV	**comparative**
CNFV	**confirmative**
CNJN	**conjugation**
COLL	**colloquial**
COM	comitative
COMP	complementizer
COND	conditional
COP	copula
DAT	dative
DEF	definite
DEM	demonstrative

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 01:26:45, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139019095

DET	determiner
DIM	**diminutive**
DIST	distal
DMS	**dental modal subordinator** (see §4.3.3.1.2, footnote 145)
DO	**direct object**
DOM	**direct/differential object marker** (see §6.1.1.1.2, footnote 34)
DU	dual
DUB	**dubitative**
ED	**ethical dative**
EMP	**emphatic**
ETH	**ethical**
EVD	**evidential**
EXCLM	**exclamation**
F	feminine
FUT	future
GEN	genitive
GPRS	**gnomic present** (a Trk tense, known in Trk as *geniş zaman* 'broad tense')
GRD	**gerund**
GRDM	**gerund marker**
HORT	**hortative**
IDO	**indirect object**
IMPF	**imperfect**
IMPV	**imperative**
IND	indicative
INDF	indefinite
INF	infinitive
INFL	**inflected**
INFM	**infinitive marker**
INS	instrumental
INTR	intransitive
INTRG	**interrogative**
IPFV	**imperfective**
LF	**L-form(for Mac)**
LPT	**L-participle** (for Blg)
LOC	locative
M	masculine
MDP	**mediopassive**
MNEG	**modal negator**
MRKR	**marker**
N	neuter
N	**noun** (in context of discussing syntactic categories)
NACT	nonactive
NCNFV	**nonconfirmative**

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 01:26:45, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139019095

NEG	negation, negative
NOM	nominative
NP	**noun phrase**
OBJ	object
OBL	oblique
OPT	**optative**
OR	**Object Reduplication** (see §7.5.1)
OSV	**Object-Subject-Verb (word order)**
OVS	**Object-Verb-Subject (word order)**
PaGe	**past gerund**
PASS	passive
PC	**particle of concord**
PCOP	**perfective copula** (for Trk *idi*)
PFV	perfective
PL	plural
PLU	**pluperfect**
POSS	possessive
PPCL	**presentational particle**
PPP	**past passive participle**
PREP	preposition
PRF	perfect
PRO	**pronoun**
PROB	**probabilitive**
PROG	progressive
PROH	prohibitive
PRS	present
PST	past
PTCL	**particle**
PTCP	participle
PX	**proximal/proximate**
Q	question particle/marker
RCVD	Received (text, of the Bible; i.e., authoritative version)
REFL	reflexive
REL	relative
REM	**remoteness marker**
SBJV	subjunctive
SG	singular
SOV	**Subject-Object-Verb (word order)**
SP	**singular and/or plural**
STR	**strong**
SUBJ	**subject**
SUP	**supine**
SVO	**Subject-Verb-Object (word order)**
TAM	**tense, aspect, and mood**

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 01:26:45, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139019095

UW	**unwitnessed**
V	**verb** (in the context of discussing syntactic categories)
V	**vowel** (in the context of discussing phonology)
VBL	**verbal**
VBLN	**verbal noun**
VOC	vocative (particle)
VOS	**Verb-Object-Subject (word order)**
VP	**verb phrase**
VSO	**Verb-Subject-Object (word order)**
WH	**question word** (e.g., *what*, *where*, *when*, etc.)
WK	**weak**

e Miscellaneous

See also §5.4 for author abbreviations used in Chapter 5, and the list of bibliographic abbreviations at the beginning of the References.

BCE	before Common Era
BDJ	Brian D. Joseph
CE	Common Era
cent.	century/centuries
ch.	chapter
dial.	dialectal
esp.	especially
fn.	footnote
lit.	literally
NB	nota bene (Latin for 'note well')
p.c.	personal communication
RN	Republic of North (Macedonia)
s.v.	*sub verbo*/*sub voce* (Latin for 'under the word/heading,' referring to dictionary entries)
s.vv.	plural of preceding abbreviation (thus 'under the words/headings')
VAF	Victor A. Friedman
vs.	versus
X	variable over elements in the immediate context (e.g., syntactic categories)

Translations are framed by single quotes ('xxx'), while literal meanings of phrases, often with grammatical information or the like, are preceded by "lit.,".

6 Orthographies and Transcriptions

Orthographies have enormous semiotic power. The choices between Arabic, Armenian, Cyrillic, Greek, Hebrew, and Latin alphabets in the nineteenth-century Balkans, as elsewhere, were (and, to some extent, still are) fraught with

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 01:26:45, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139019095

religious, ethnic, political, and ideological meaning, as were (and sometimes still are) specific variants of a given script. At the same time, the field of linguistics was then, and still is, concerned with phonemic or phonetic representations of sound systems irrespective of nation-states and their languages, be they official or unofficial. The two goals are in obvious conflict in a language like English, where the historical spelling has only a tenuous connection to pronunciation.[1] The situation is similar in French, where /o/ is orthographically *au, aux, eau, eaux, o*, etc. Tibetan is another example of a language with a long historical tradition where there is a disconnect between spelling (and therefore transliteration) and pronunciation (or transcription) for historical reasons, and it is standard practice in Tibetan linguistics to choose between transliteration and a transcription representing current pronunciation, depending on various factors. Here we have opted for a sort of middle way as illustrated by the practice in Sandfeld 1930: non-Latin-alphabet examples will be transliterated, except for Greek (regardless of temporal stage). Greek, which in this respect is like Tibetan, but with a writing system much more closely related to the Latin, is given in Greek orthography unless the example is dialectal, in which case it is transliterated or given in the Greek alphabet, depending on the source and on usual practice among Hellenists. For Greek place names, we follow the recommendations of the United Nations Group of Experts (UNGE 1973), which is available online at <https://unstats.un.org/unsd/geoinfo/ungegn/ungegnSession5.html>. Note that according to their decision, stress is marked on a toponym only when it is not final. Also, if the toponym has a common English form (e.g., Athens, Peloponnese), then that form is used.

In the paragraphs that follow, our decisions concerning Latin-alphabet orthographies and transliterations are elucidated and followed by Table A, which gives the reader a guide to those letters whose value can vary across orthographies. The intention here is not to give complete accounts of all the alphabets and orthographies of the Balkan languages. Such information is readily available online. Rather, we seek here to provide the reader with a key to those elements that are not uniform across the orthographies and transliterations used in this book. Table A is followed by Greek, given in Table B, which gives a guide to pronunciation in Ancient Greek and in Modern Greek.

The orthographies of those Slavic languages traditionally associated with Cyrillic have a tradition of being more or less phonemic or morphophonemic. Cyrillic and Glagolitic, with minor exceptions, were basically phonemic when first created more than a thousand years ago, and the current orthographies of all the Cyrillic-using Slavic languages have undergone reforms such that they have relatively straightforward phonemic or morphophonemic transliterations (some of them legally official in the relevant nation-states, others standard for Slavic linguistic publications, etc.). For Judezmo, one or another version of the Hebrew alphabet (usually Rashi or Soltero) has been used traditionally, but Arabic, Cyrillic,

1 The poem "English is Tough Stuff," widely available on the Internet, is especially illustrative of the challenges of English orthography.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 01:26:45, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139019095

and Latin have also been used, and in modern linguistic works, as well as many popular publications, a Latin orthography is now the norm.

For all languages with internationally recognized standard Latin orthographies, we use that orthography. In the case of Albanian, whose dialectal base for the standard changed from Geg to Tosk during the twentieth century, we use the current standard, unless the point depends on the older standard. For pre-standard orthographies, we use the modern standard equivalents unless preserving graphic peculiarities is crucial to an argument. See also §1.4.8. For Aromanian, we use the orthography of the only country where the language is officially recognized and used in official documents, North Macedonia.[2] For Romani, we also use the official orthography in use in North Macedonia, the only Balkan country where Romani both has constitutional status and is official at the local level.[3] Similarly, for Gagauz, we use the official orthography of the Republic of Moldova. For Meglenoromanian, we follow the orthography of the source, which is based either on Romanian or Aromanian.[4] For Judezmo, we use the orthography of the Autoridad Nasionala del Ladino in Jerusalem, as adapted by Bunis 1999: 15–17. For Cyrillic, we use the current official transliteration for Macedonian and BCMS and the linguistic transliteration in common use for East Slavic, OCS, and Bulgarian. For other non-Latin alphabets, we use standard linguistic transliterations. Some of our sources provide close phonetic transcription or use various non-standard or pre-standard orthographies. For the sake of simplicity, we have normalized these to the corresponding standard orthographies used here unless there was a compelling reason to retain the original transcription. Similarly, we do not indicate stress or tone unless it is relevant to the point in question, or, in the case of Greek, part of the official, everyday orthography.

A special note on Turkish morphophonemic transcription is required here. In citing Turkish morphemes we follow the standard Turcological convention of using capital letters that are involved in morphophonemic alternations of vowel harmony (I = i/ı/u/ü; A = e/a), progressive assimilation of voicing (D=d/t, C=c/ç), and lenition (k > ğ) in intervocalic position. In some instances, where there may have been a shift of vowel height, we use V to mean "the relevant vowel."

Table A gives those orthographic conventions that are not uniform across the various Balkan orthographies (except Greek). All letters for which orthographic conventions are basically unambiguous (e.g., <m>), are not indicated. Phonetic and other minor variations are not included here. Thus, for example <b, d, g> in Judezmo have variants [β, δ, γ] in addition to [b, d, g]; Albanian, Macedonian, and BCMS mellow palatals have slightly different places and manners of articulation, etc. The

2 See Friedman 2001b; Cunia 1999. As with Greek, some specifically dialectal forms are presented as such.

3 For details on Romani orthographies, see Friedman 1991, 1995a; Cortiade et al. 1991; Kenrick 1981; Matras 1999.

4 Meglenoromanian, like Italian, has a contrast between high and low (closed and open) mid-vowels. The contrast is indicated by a hook under the vowel in some of the dialectological literature, whereas publications such as Papatsafa 1997 treat the distinction as in Italian, i.e., unspecified in the orthography.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 01:26:45, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139019095

principle here is that allophonic variation is not indicated but phonemic value is. Thus, for example, the distinction between aspirated and unaspirated voiceless stops is phonemic in Romani (and indicated by <h> after the appropriate voiceless obstruents, as it is in standard orthography), but not in any of the other Balkan languages. Similarly, if an orthographic or transliterated digraph is considered a separate letter for dictionary purposes (for both Latin and non-Latin based orthographies), it is indicated here, but not otherwise. Thus, for example, [ʤ] is indicated as <dž> for Macedonian but not for Bulgarian because the former has a separate Cyrillic letter for this sound while the latter does not. Rhotics and liquids are given when there is a phonemic contrast between different kinds. Although some varieties of Romani oppose /h/ and /x/, for the most part these sounds are in free variation in Romani, and there is no phonemic contrast in the other languages.[5] Our purpose here is to provide a basic guide to what the non-congruent letters – and in the case of transliterations from Cyrillic and Judezmo, their transliterations – stand for phonetically (in broad terms) where the orthography or transliteration system is not uniform across languages.

Table A *Non-uniform orthographic conventions across Balkan orthographies (except Greek)*

IPA	Alb	Rom	Aro	Jud	Trk	Gag	Rmi	BCMS	Mac	OCS	Blg	Rus
æ	–	–	–	–	–	–	–	–	ä[6]	ě	ä[7]	–
ə	ë	ă	ã	–	–	ê	ä	–	ă	–	ă	–
ɨ	–	â=î	î	–	ı	ı	–	–	–	–	–	y
ǐ	–	–	–	–	–	–	–	–	–	ь	–	–
ø	–	–	–	–	ö	ö	–	–	–	–	–	–
ŭ	–	–	–	–	–	–	–	–	–	ъ	–	–
y	y	–	–	–	ü	ü	–	–	–	–	–	–
ð	dh	–	–	–	–	–	–	–	–	–	–	
ʤ	x	–	dz	–	–	–	–	dz[8]	dz	dz	–	–
ʤ̵	gj	–	ghi/e	–	–	đ	gj	ǵ	–	–		

5 Common practice for Bulgarian and Russian is to use < x > where Macedonian, Montenegrin, and Serbian use < h > for Cyrillic < x >; in general, we have used < h > for Cyrillic < x >. The Library of Congress (LC) equivalent is < kh > and we have occasionally left in < x > when, e.g., a title of an item in the bibliography was not in a Slavic language or when a Slavic language with the *x*/*h* contrast was involved. We have kept < kh > for the names of authors who use that transliteration when writing in English and also for the geographical name Makhachkala. For other authors who write in both Western and Cyrillic-using languages, in most instances we have used the spelling that appears in their published work in Western languages (e.g., Makartsev, Trubetzkoy). We have also left < x > in some cited Romani examples. For Romanian documents that were written in Cyrillic, we follow Romanian practice of using modern Romanian orthography. In the case of Romanian < â >, which was replaced by < î > in 1953, brought back in the lexical item *român* 'Romanian' and its derivatives in 1964 and fully resuscitated in 1993, we have kept the spelling used in the original source.

6 Exclusively used for those dialects where it occurs.

7 Exclusively used for those dialects where it occurs.

8 Montenegrin only.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 01:26:45, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139019095

Table A (*cont.*)

IPA	Alb	Rom	Aro	Jud	Trk	Gag	Rmi	BCMS	Mac	OCS	Blg	Rus	
ʤ	xh	gi/e	gi/e	dj	c	c	dž	dž	dž	–	–	–	
j	j	i	i	y	y	y	j	j	j	j	j	j	
k	k	c/ch	c/ch	k	k	k	k	k	k	k	k	k	
l	l	–	–	–	–	–	–	l	lj	–	–	–	
ʎ	–	–	1 j	–	–	–	lj	lj	–	–	–	–	
L	ll	–	l[9]	–	–	–	l[9]	–	l[9]	–	–	–	
ɲ	nj	ni	nj	ny	–	–	nj	nj	nj	–	–	–	
ɾ	r	–	–	r	–	–	r	r	–	–	–	–	
r	rr	–	–	(rr)	–	–	rr[10]	–	–	–	–	–	
ʃ	sh	ş	sh	sh[11]	ş	ş	š	š	š	š	š	š	š
ç	–	–	–	–	–	–	ś[12]	–	–	–	–		
θ	th	–	–	–	–	–	–	–	–	–	–	–	
ʦ	c	ţ	ts	ts	–	ţ	c	c	c	c	c	c	c
ʨ	q	–	chi/e	–	–	–	ć	kj	ć	–	–	–	
ʧ	ç	ci/e	ci/e	ch	ç	ç	č	č	č	č	č	č	č
ʒ	zh	j	j	j	j	ž	ž	ž	ž	ž	ž	ž	
ʝ	–	–	–	–	–	–	–	ź[13]	–	–	–	–	
Ṽ	â . . . [14]	–	–	–	–	–	–	–	–	ę, ǫ	–	–	–
C[h]	–	–	–	–	–	–	Ch	–	–	–	–	–	–
C[j]	–	–	–	–	–	–	–	–	–	–	–	–[15]	C'[16]

Table B gives a guide to pronunciation in Ancient Greek and in Modern Greek; for material from any post-Classical period, the modern pronunciation provides a close approximation, as most of the relevant sound changes were under way by the period of the Hellenistic Koiné (the details of the chronology of particular changes are generally not relevant for our discussion).

9 For Macedonian /l/ is clear before front vowels and /j/, and it is velar otherwise. Aromanian and Romani in RN Macedonia have the same rule.

10 The more archaic retroflex or uvular trill, also represented as <ř>, can also occur. In the standard Romani of RN Macedonia, and the Arli dialects on which it is based, the opposition *r*/*rr* is neutralized and is not spelled.

11 The sequences /s/+/h/ and /ʃ/+/h/ are spelled <s.h> and <sh.h>, respectively.

12 Montenegrin only. 13 Montenegrin only.

14 In Standard Geg, nasalization can be indicated by a circumflex over any vowel.

15 There has been a lively debate about the opposition of palatalization and jotation in Bulgarian phonology, but for our purposes here, <j> suffices for transcriptions. Russian, on the other hand, has a clear distinction such that /j/ can follow a palatalized consonant, which never happens in Bulgarian.

16 Russian <ь> is transliterated <'> and indicates the preceding consonant is not palatalized. Russian <è> indicates a mid-front vowel with the absence of preceding jotation or palatalization.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 01:26:45, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139019095

Table B *Greek alphabet/pronunciation key*

Upper case	Lower case	Ancient pronunciation	Modern pronunciation (IPA)
Α	α	a[17]	a
Β	β	b	v
Γ	γ	g	γ (__a o u) / ʝ (__i e)
Δ	δ	d	ð
Ε	ε	ε	ε
Ζ	ζ	zd	z
Η	η	ē	I
Θ	θ	t^h	θ
Ι	ι	i[18]	i
Κ	κ	k	k
Λ	λ	l	l
Μ	μ	m	m
Ν	ν	n	n
Ξ	ξ	ks	ks
Ο	ο	o	o
Π	π	p	p
Ρ	ρ	r	r
Σ	σ (ς __#)	s	s
Τ	τ	t	t
Υ	υ	y[19]	i
Φ	φ	p^h	f
Χ	χ	k^h	x
Ψ	ψ	ps	ps
Ω	ω	ō	o
ΑΙ	αι	aj	ε
ΕΙ	ει	ej	i
ΟΙ	οι	oj	i
ΑΥ	αυ	aw	av (__+voice), af (__-voice)
ΕΥ	ευ	ew	εv (__+voice), εf (__-voice)
ΟΥ	ου	ū	u
ΜΠ	μπ	mp	(m)b[20]

17 <α> could also indicate [ā], with no special marking for length.
18 <ι> could also indicate [ī], with no special marking for length.
19 <υ> could also indicate a long vowel ([y:]), with no special marking for length.
20 We annotate these groups with the nasal in parentheses because the nasal is pronounced by some speakers, mostly older ones, of the standard language and by speakers of some regional dialects, but generally not by younger speakers (who produce pure oral stops here without any nasalization). Older speakers generally have a nonnasal realization in word-initial position. See Arvaniti & Joseph 2000, 2004 for relevant discussion.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 01:26:45, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139019095

Table B (*cont.*)

Upper case	Lower case	Ancient pronunciation	Modern pronunciation (IPA)
NT	ντ	nd	(n)d
ΓΓ	γγ	ŋg	(ŋ)g
ΓΚ	γκ	ŋk	(ŋ)g
TZ	τζ	dz	dz
Diacritics[21]	᾽	–	–
	῾	h	–
Accents[22]	´	´	´
	`	`	´
	~	~	´

21 These are the so-called "breathing marks," signaling the presence or absence of a word-initial [h] with vowels.

22 The Ancient Greek distinctions of acute, grave, and circumflex pitch accents were neutralized in Post-Classical Greek and yielded a single stress accent, which is all that is marked in Modern Greek now. In the modern orthography, the accent is marked generally only on non-monosyllables, with a few lexical exceptions to differentiate forms that would otherwise be homographs.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 01:26:45, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139019095

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 01:26:45, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139019095