

SERBIAN ACADEMY OF SCIENCES AND ARTS
INSTITUTE FOR BALKAN STUDIES

LV



2024

BALCANICA

V. G. PAVLOVIĆ, *Ninety Years of the Institute for Balkan Studies*
• L. DIERS, *Reading the Subtext – Site Location and Settlement Systems in Roman Moesia* • O. LIGORIO, *Antrešelj. An Early Romanian Remnant in Serbo-Croatian* • E. ADAMOU & A. N. SOBOLEV, *The Atlas of the Balkan Linguistic Area program*
• P. HRISTOV & T. MANOVA, *Urbanisation, Migration, Depopulation and Virtual Ritual Community – The Village Kurban as a Shared Meal* • D. POPOVIĆ, *The Landscape of the Monastic Endeavour: The Choices of St Sava of Serbia* • M. BACCI, *Latin-Byzantine Artistic Interactions and the Church of Saint Basil in Mržep (Montenegro)* • S. G. Markovich, *Debating Balkan Commonalities: Is There a Common Balkan Culture?* • A. BASCIANI, *Beyond Nationalism? The Inter-war Period and Some Features of the Complex Transformation of Southeastern Europe* • M. VARTEJANU-JOUBERT, *Ritual Objects for the Feast of Sukkot: Theoretical Analysis of the Talmudic Prescriptions and Some of their Ethnographical Achievements in the Balkans*
• G. VALTCHINOVA, *From “Religion” to “Spirituality” in Socialist Bulgaria: Vanga, Nicholas Roerich, and the Mystique of History* • N. LACKENBY, *You are what you don’t eat – Fasting, Ethics, and Ethnography, in Serbia and Beyond* ❧

ANNUAL OF THE INSTITUTE FOR BALKAN STUDIES

UDC 930.85(4-12)

BELGRADE 2024

ISSN 0350-7653
eISSN 2406-0801

Evangelia Adamou*

*Oral Tradition Languages and Cultures Lab
French National Center for Scientific Research
Paris, France*

Andrey N. Sobolev**

*Institute for Linguistic Studies
Russian Academy of Sciences
St. Petersburg, Russia*

The Atlas of the Balkan Linguistic Area program

Abstract: This article presents the Atlas of the Balkan Linguistic Area (ABLA), a French-Russian research program that created an online database of language contact phenomena documented in the languages of the Balkans. This resource will be open access after its launch in 2025, enriching the fields of Balkan and areal linguistics. Specifically, ABLA consists of 93 phonological, morpho-syntactic, semantic, and lexical features. Each feature is matched to a map covering 60 localities across Balkan countries. Each map is accompanied by a chapter co-authored by the project contributors. The paper offers some preliminary results for the feature “Infinitive: Forms”. The online database in Wordpress is hosted by Huma-Num in France. ABLA, to be published by de Gruyter, is not only the first online database for the Balkans, an area shaped by multilingualism in forms that are rapidly disappearing, but also an example for other linguistic areas in the world

Keywords: areal linguistic typology, Balkan linguistic area, linguistic databases, linguistic atlases, dialectology, minoritized languages, methods of linguistic research.

The concept of “linguistic area”

When contact-induced change and borrowing become systemic across a number of languages and persist over time, language clusters emerge containing languages which, although not genetically related, have become more similar. The areas in which such languages are spoken are referred to as “Sprachbund,” “convergence area,” and “linguistic area” (Trubetzkoy 1923; 1928).

Though roughly consistent with a geographic area, a linguistic area is not defined by natural boundaries but by the relations between the human groups

* evangelia.adamou@cnrs.fr

** andrey.n.sobolev@mail.ru; sobolev@staff.uni-marburg.de

present. The most well-known linguistic area is the Balkans, but more recently, several authors have drawn attention to the existence of numerous linguistic areas around the world (Hickey 2017; Adamou and Matras 2021).

This has revived discussions about the adequacy of the term “linguistic area” to capture such versatile phenomena. How many languages are needed to identify a linguistic area? Clearly, it takes more than one language to form a group, but is two enough? How many language families should be involved? How many linguistic features are needed to consider that there is a linguistic area? What kind of linguistic features are needed (structural and lexical)? In the face of these challenges, some authors take a different stance and suggest: “We should abandon the search for a definitive definition of “linguistic area” (Campbell 2006, 21).

Despite the difficulty of convincingly defining linguistic areas, the topic remains very popular in contact linguistics, and increasing numbers of new linguistic areas are being discussed, such as North America, Amazonia, West Africa, Anatolia, Southeast Asia, and Melanesia, among others. Overall, modern linguists prefer to pay attention to the processes of convergence and relate them as much as possible to specific socio-cultural contact processes.

The Balkans as a linguistic area

The Balkans are a geographic area in the southeastern part of Europe, often delimited in the North by the Danube River. However, rather than a geographic region, the Balkans are best understood as a socio-political and cultural area where people share a centuries-long common history, in particular during the Byzantine Empire (4th century to 15th century) and Ottoman Empire (15th to 19th and early 20th centuries). Within this socio-political space, multilingualism was widespread and constant over time. The modern Balkan states have been reshaped multiple times, including recently, after the Yugoslav wars in the early 1990s, and yet despite the current national borders, exchanges and contact in the Balkans are ubiquitous.

Although the Balkans have been central to theoretical discussions regarding linguistic areas since the early 20th century, there is still no consensus over the precise linguistic features and their distribution. On the one hand, Balkan linguistic features (also known as Balkanisms) are not necessarily restricted to the Balkans but can be found in other languages of Europe and the Middle East. So why draw a linguistic boundary at that specific level? In a way, this question parallels the ambiguity in delimiting the Balkans as a geographic or cultural region since it is also part of a larger geographic and cultural area, namely Europe. Indeed, some Balkan states are still members of the European Union while forming smaller regional alliances, and speakers are in mutual contact in the Balkans but also beyond.

On the other hand, Balkan linguistic features are instantiated differently in different language varieties, and multilingualism is distributed differently depending on the region in the Balkans. In that sense, it is more accurate to say that, rather than a homogeneous linguistic area, one could identify a number of small linguistic areas that compose the larger Balkan linguistic area. Wiemer (2004) nicely captures this superposition of areas by the metaphor of the Russian *matryoshka* dolls where a big doll has a smaller one inside it that contains a smaller one and so on.

Despite the complexity of superimposed areas and heterogeneity, many specialists in Balkan languages argue that there is good reason to consider the Balkans as a linguistic area as its inhabitants were traditionally multilingual, had many ties across linguistic communities, as attested in many historical documents, and linguistic changes can often be traced in the written documents of the languages. In addition, a comparison between the languages of the Balkans and languages of the same branch that are spoken outside of the Balkans suggests the presence of unique linguistic developments (Ledgeway 2017; Gardani et al. 2021).

Most languages of the Balkans that are considered to be “Balkan languages” belong to the Indo-European family, albeit to various branches: the South Slavic branch includes Macedonian, Bulgarian, many non-standard Balkan Slavic varieties, and dialects of Serbian, such as Torlak; the Romance branch includes the so-called Balkan Romance languages, such as Aromanian and Meglenoromanian, as well as Romanian, but also Judeo-Spanish or Ladino, a Spanish variety spoken by Jews who arrived in the Ottoman Empire after they were driven out from Spain; the Albanian language and, to some extent, the Greek language, and more significantly some Greek dialects in Northern Greece; as well as some of the Romani dialects belonging to the Indic branch of Indo-European; and possibly also Balkan Armenian. The non-Indo-European languages belong to the Balkan Turkic branch of the Turkic family (Sobolev 2004; Friedman 2021).

In terms of linguistic features, some are well-established and well-researched while others are still in need of more careful consideration. For instance, linguists agree that the rise of a modal future particle based on the volitive ‘want’ resulted from several parallel contact processes. This feature is found in Greek, Albanian, Romanian, Macedonian, Bulgarian, Serbian and Croatian, and Romani (Joseph 1992).

Another well-known Balkan convergence phenomenon concerns sentences like ‘I want to know’. Balkan languages use an optative particle (e.g. Balkan Slavic *da*, Albanian *të*, Romani *te*, Balkan Romance *să, si, s’*, Greek *na*) and a finite verb (i.e. a verbal complement that has tense and subject marking) following the loss or reduction of the infinitive (i.e. a verbal complement that is

nonfinite, that is, has no tense or subject marking), a process described in Joseph (1983). Infinitive reduction is also noted in Balkan Turkish, Armenian, Judeo-Spanish and Corfioto (Judeo-Italian spoken in Corfu).

The use of an enclitic definite article is another convincing result of convergence. Indeed, very few Slavic languages that have grammaticalized definite articles are spoken in the Balkans. It is argued that the grammaticalization of postposed articles in South Slavic languages (e.g., in Macedonian, Bulgarian, Torlak Serbian and other Balkan Slavic varieties) results from both internal and contact-induced factors, as it coincides with a similar development in the Romance languages of the area (e.g. Romanian, Aromanian, and Meglenoromanian) and Albanian (Asenova 2002; Runić 2019).

Regarding the lexicon, Friedman and Joseph (2017) capture the trend by referring to borrowings that are Essentially Rooted in Conversation (ERIC). ERIC borrowings are generally lexical items that are either known to be rarely borrowed, such as pronouns, numerals, kinship terms, and bound morphology or are more frequently borrowed lexical items, such as discourse particles, interjections, and taboo expressions, among others.

Finally, as Friedman (2021) notes, phonology offers a different view of the Balkan linguistic area, with several localized processes of convergence rather than a single process that would apply to all Balkan languages.

The Atlas of the Balkan Linguistic Area (ABLA)

To obtain an up-to-date and detailed picture of what the Balkans as a linguistic area look like, we created an online database that carefully maps the linguistic features of various languages and dialects and related them to the available socio-historical information.

Friedman (2021) points to the need for such an atlas in an overview chapter about the Balkans: “The concept of a multilingual Balkan linguistic atlas, one that could take into account the various dialects of the various languages and oriented toward mapping the actual instantiations and occurrences of various Balkanisms, is, close to a century after it was first mooted [...], a desideratum” (Friedman 2021, 398).

Indeed, while the volumes of the *Minor Dialectological Atlas of Balkan languages (MDABL)* offer a solid basis, it is restricted to 12 locations covering morphosyntactic and lexical-ethnographic features for south Aromanian, central Geg, northern Tosk, northern Greek, and southern Greek, as well as Slavic in southwestern North Macedonia, southwestern Bulgaria, the Bulgarian Rhodopes, northeastern Bulgaria, southeastern Serbia, southern Montenegro, and Dalmatia in Croatia.

The ABLA project aims to fill this gap by bringing together the scientific and technological experts needed to create an online linguistic atlas of the Balkans. Such an atlas is not only a technological achievement aligned with current open science goals, where the datasets will be FAIR (findable, accessible, interoperable and reusable) (Wilkinson et al. 2016), but would also allow for extensive empirical testing of the concept of the linguistic area by examining how linguistic features intertwine in a given geographical area by taking into consideration smaller contact areas within the larger contact area.

The ABLA also paves the way and provides a model for the creation of databases for other linguistic areas around the world. The authors drew their inspiration from the “Atlas of Pidgin and Creole Language Structures Online” (<https://apics-online.info/>), elaborated by Michaelis, Maurer, Haspelmath, and Huber (2013) with the collaboration of numerous language experts. This database has not only made data accessible to the academic community and wider audiences, but has also given rise to numerous publications relying on the quantitative analysis of the data compiled in the database, and the Atlas itself was published by Oxford University Press. The recent database “Languages of hunter-gatherers and their neighbors” is another source of inspiration (<https://huntergatherer.la.utexas.edu/>), elaborated by Bower, Epps, Hill, and McConvell (2020) in collaboration with many linguists.

The Balkans offer a unique example of how languages that were not closely related (in that they belonged to different branches) have structurally converged over the centuries. This convergence was driven by the multilingualism of the people of the Balkans, who, in many cases, lived in the same villages and cities and formed linguistically mixed households. The Atlas of the Balkan Linguistic Area makes available a much-needed database illustrating the multiple, complex convergence processes across a larger number of datapoints than currently available in the literature.

Language and location selection

The ABLA team selected 60 representative linguistic varieties and corresponding locations across all Balkan countries and their periphery where data are available and relevant for the discussion of the Balkans as a linguistic area to be depicted on maps. We were careful to include non-standard varieties in which the effects of long-term language contact can still be documented among the elders and changes observed among the younger generations (Adamou 2021).

The Indo-European language family is represented by 16 language datasets from the Slavic branch, 11 from the Romance branch, 11 from the Albanian branch, 7 from the Greek branch, 10 from the Indic branch, and 2 from the Armenian. For the Turkic family, we include 3 language datasets from the Oghuz

branch. The comprehensive and detailed list of language datasets is available at <https://abla.cnrs.fr/languages>.

ABLA offers a unique opportunity to document Balkan languages and dialects, many of which are endangered. To illustrate the linguistic diversity that we capture, we refer to (Sobolev 2021) where the changing multiethnic and multilingual profiles of peoples in various countries of the Balkans were examined, including L1 or L2 speakers of dialectal Greek (Tsakonian in the Peloponnese, Greece; Himariotika in Albania), Albanian (Dibra dialect in Golloborda; Laberia dialect in Himara, Albania; Ana e Malit dialect in Montenegro; Prespa idiom in North Macedonia), Romanian (Iabalcea variety in Karashevo/Carașova, Romania), Aromanian (Prespa variety in the Republic of North Macedonia), Macedonian (Golloborda dialect in Albania; Prespa dialect in North Macedonia), Serbo-Croatian (Karashevo/Carașova dialect in Romania; Mrkovići dialect in Montenegro). These are all non-standard varieties in which the effects of long-term language contact can still be documented among the elders and changes observed among the younger generations.

Feature selection

Regarding linguistic features, we have chosen both well-established and extensively researched characteristics, as well as those discussed in typological literature, requiring further careful consideration for the Balkans.

We start by presenting borrowings, greetings, and lexical features that play a key role in the Balkans as they are Essentially Rooted in Conversation (ERIC) (Friedman and Joseph 2017).

We examine about 40 features pertaining to the noun phrase and the verb phrase. About 10 features illustrate complex clause phenomena, about 15 simple clauses, and one is specific to word order.

We also dedicate about 15 features to phonology, phonetics, and prosody. Such features are not as commonly discussed even though they offer a different view of the Balkan linguistic area, with several localized processes of convergence rather than a single process that would apply to all Balkan languages (Friedman 2021).

The full list of linguistic features is present at <https://abla.cnrs.fr/features>.

Data collection, database and website creation

The data were collected from available existing resources such as standard language grammars and dictionaries, as well as dialectal grammars, dictionaries, and language corpora covering the 20th and 21st centuries. Additionally, proj-

ect participants gathered naturally spoken, written, or elicited data in the field. The project contributors filled in the online database by selecting the appropriate value(s), providing bibliographical references and examples illustrating the feature.

The database and website, developed on the WordPress platform, are archived for long-term preservation.

Infinitive forms: Preliminary results

The ABLA chapter “Infinitive: Forms” illustrates the project’s results (Sobolev et al. 2025).

Background

Since (Miklosich 1861) and (Sandfeld 1900; 1930), “the loss of the infinitive” has been considered one of the primary balkanisms (Joseph 1983). Even the most advanced and broadest approach today still regards “all the developments associated with the infinitive – the loss of the infinitive, the prevalence of finite subordination, and the occurrence of two types of finite subordination” (Friedman and Joseph [forthcoming], 743–745) as a Balkan contact-induced convergence. Against this background, the remnants and propagation potential of the Balkan infinitive appeared less relevant and have not been addressed in contact linguistics (Matras 2020, under entry *infinitive (reduction of)*).

Nevertheless, the inherited infinitive, along with its non-finite substitutes and equivalents, here collectively referred to as the “infinitive cluster”, play a significant role in the structure of individual Balkan languages, showcasing notable formal and functional similarities, worthy of being presented in ABLA. In this chapter, we are particularly interested in non-finite predicates (Fiedler 2018), primarily the ones like:

- Romanian *se poate face*_{INF} ‘can be done’;
- Romanian *trebuie făcut*_{P_{TCP}} ‘must be done’;
- Albanian Gheg *duhet bërë*_{P_{TCP}} ‘must be done’;
- Aromanian *va lukrari*_{VN} ‘needs working’.

These Balkan predicates, relatively underexplored from a comparative standpoint, are functionally equivalent to the infinitive in a diachronic perspective (from which the forms in the Romanian and Aromanian examples above directly derive), cf. also (1), (2) and (3):

- (1) Latin
debet dicere
debet *dicere*
 should.PRS.3SG say.INF
 ‘should say’
- (2) Old Albanian (Pjetër Budi. *Rituale Romanum*, 1621)
dubete me *ëëane*
dubete *me=* *ëëane*
 should.PRS.3SG with= say.PTCP
 ‘should say’
 (Schumacher, Matzinger 2013, 384–389)
- (3) Old Albanian (Pjetër Budi. *Rituale Romanum*, 1621)
dubete *ëëane*
dubete *ëëane*
 should.PRS.3SG say.PTCP
 ‘should say’
 (Schumacher, Matzinger 2013, 384–389)

In general, two categories of forms that can function as nonfinite complements of a genuine modal verb should be examined (Joseph 1983; Gabinskii 2002; Fiedler 2018):

Analytic: Albanian Gheg *me=shkru* with=write.PTCP ‘to write’; Albanian Tosk *së=shkruari* ART=write.VN.N.ABL.SG; Romanian *a=scrie* INF=write.INF; Romanian *de=scris* of=write.PTCP; Albanian Tosk *për=të=shkru(ar)* for=ART=write.PTCP; Albanian Gheg *për=me=shkru* for=with=write.PTCP;

Synthetic: Greek *γράφειν* write(IPFV)-INF; *γράφω* write(PFV)-VN.N.SG ‘writing’; Serbian and Croatian *pisa-t(i)* write(IPFV)-INF ‘to write’; Serbian and Croatian, Bulgarian *pisa* write(IPFV).INF (the so-called short infinitive) ‘to write’; Macedonian *pisa-nje* write(IPFV)-VN.N.SG; Megleno-Romanian *cânta-ri* sing-INF ‘to sing’; Romanian *scrie* write.INF ‘to write’; *scris* write.PTCP ‘written’; Albanian Gheg and Tosk *shkru(e)* and *shkruar* write.PTCP ‘written’; Turkish *yaz-mak* write-INF ‘to write’.

In scholarly traditions, these forms are usually labeled “infinitive proper”, “short infinitive”, “Gheg infinitive”, “Tosk infinitive”, supine, participle, and verbal noun. The primary research objective here will be to enquire into the possibilities of enhancing the systematicity of our understanding of the structure, functioning, and patterns of variation in Balkan languages and dialects in both synchrony and diachrony and to establish a comprehensive and logically coherent framework for systematically delineating and linguistic-geographically map-

ping the regional and systemic variations of infinitives, their equivalents, and substitutes across Balkan languages.

Nevertheless, there is an abundance of forms of different origins, oscillating widely between verbal (person, tense, transitivity, etc.) and nominal (gender, case, definiteness, etc.) properties, and there is also a broad palette of their distribution patterns that cannot be considered in full in ABLA. Therefore, we limit ourselves here to nonfinite verbal forms, both synthetic and analytic, both inherited and developed in the Balkans, used as complements of modals, preferably dependent on core modals of verbal origin (Hansen 2009, 470-471; Arapi 2010; Dragomirescu 2013) in dependent clauses with identical subjects (*tautoprosopy*).

This is illustrated with an example from Megleno-Romanian (Archange-los) in (4), Romanian (Bucharest) in (5) and Albanian in (6):

(4) *Nu la pot priflari*

<i>nu</i>	<i>la</i>	<i>pot</i>	<i>prifla-ri</i>
NEG	3SG.ACC	can.1SG	find.again-INF

'I can't find him/her/it again.'

(Capidan 1935, 8)

(5) *dorința de a veni*

<i>dorinț-a</i>	<i>de</i>	<i>a</i>	<i>ven-i</i>
desire(F)-DEF	of	INF	come-INF

'the desire to come'

(Pană Dindelegan 2013, 211-215)

(6) *Pse s'desha për të martuar*

<i>pse</i>	<i>s</i>	<i>desha</i>	<i>për</i>	<i>të</i>	<i>martuar</i>
why	NEG	want.AOR.1SG	for	SBJV	marry.PTCP

'Why I didn't want to marry.'

(Andon Zako Çajupi "Burri i dheut" [1908])

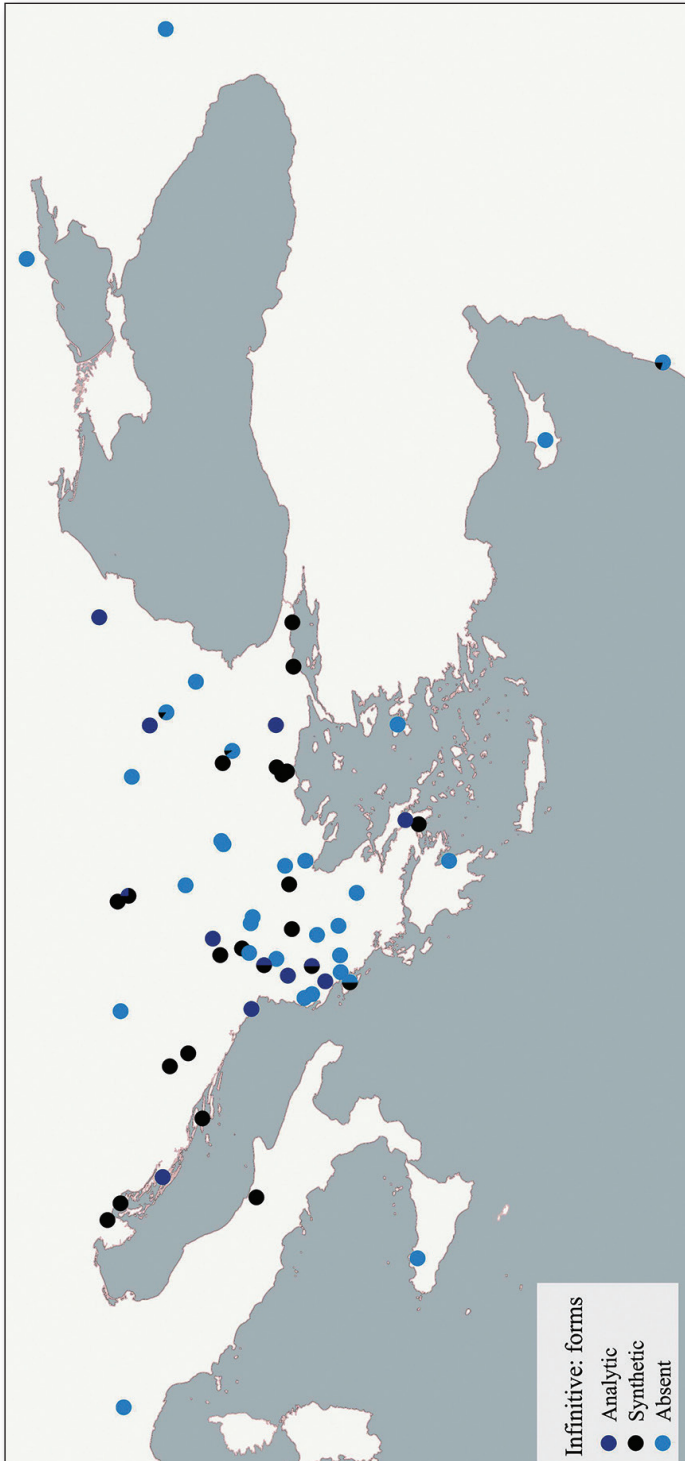


Figure 1. ABLA language datasets coded for the feature Infinitive: Forms

participles as full synthetic equivalents of prototypical infinitives. In contrast, the same criterion forces us to rule out the verbal nouns of Greek, Balkan Romance and Balkan Slavic as presented in example (9) from Trebisht:

- (9) *zm'ijata s'akat t'epajne*
zmij-a-ta *saka-t* *tepa-jne*
 serpent(F)-SG-DEF.NTR will(IPF)-PRS.3SG kill(IPFV)-VN.N.SG
 'One should kill serpents (Lit. The serpent wants killing).'
 (Sobolev and Novik 2013, 187)

Absent. The category under study is absent from Greek (Thessaloniki), Greek of Southern Albania (Dropull), Tsakonian (Tyros) and the varieties Northern Chiotic (Lagkada), Cypriot (Dymes), Azov Greek (Maloianysol'), and Pontic (Yessentuki). No infinitives are found in Albanian (Piana degli Albanesi), Aromanian (Selenica), Bulgarian (Sofia, Starozagorski mineralni bani), Standard Macedonian (Skopje), Macedonian (Dojran)

Balkan Slavic (Nestorio, Trebisht, Vranishte), nor in any varieties of Romani (Fieri, Kaspičan, Knjaževac, Parakalamos, Piacenza, Pitesti, Skopje, Sofades, Sofia, Šid).

A very restricted usage can be found in Bulgarian of Pozharevo, and variation due to contact in the Balkan environment is witnessed in Corfioto (Corfu City, Tel Aviv).

Summary

ABLA clearly demonstrates that the use of high-quality, representative linguistic materials, meticulous observation of linguistic contexts, enumeration of intrasystemic constraints in the distribution of forms and constructions, the quest for isosemantic, isomorphic, and isofunctional linguistic parallels, as well as contactology and historical interpretation of areal connections, can be applied to Balkan non-finite predicates dependent on modal verbs, namely members of the "infinitive cluster" – infinitives proper, diverse analytic forms and constructions, participles, and deverbal nouns.

The purpose of this chapter has been to showcase the abundance of the nonfinite forms that can serve in the Balkan languages as complements to a genuine modal verb. It is not possible to consider the infinitive as lost in these languages, and its "reduction" in the Balkans (Joseph 1983) and beyond (Masica 2001) can in no way be seen as a teleological process.

Another striking observation is that Balkan languages exhibit maximum complexity. They appear to possess a highly extensive set of non-finite forms,

which are not always entirely free but rather lexically bound, used to express dependent actions with modal verbs.

The isomorphism of finite subjunctive forms in Balkan languages is in full contrast with the complete non-isomorphism of a series of non-finite structures featuring rare, unique, and idiosyncratic combinations of grammatical elements, particles, prepositions, and connective articles, which selectively and freely combine with synthetic forms of true infinitives, participles, and deverbative nouns. The grammatical analysis of such constructions is often ambiguous and challenging.

Despite the variability in different Balkan languages and the linguistic specificity of their inventory of structures and rules of variation, the contact-induced nature of certain binary isosemantic and isomorphic inter-Balkan parallels remains a promising field of research.

The scientific impact of the program

To conclude, the scientific impact of the ABLA program lies in 1) the methodological novelty for the field of areal linguistics, 2) the unprecedented quantity of linguistic data from different Balkan linguistic varieties, and 3) fostering international cooperation between linguists in the field of Balkan and areal linguistics.

In addition, by introducing state-of-the-art practices in digital humanities into the established field of Balkan linguistics, we will set a new benchmark for research in the medium and long term, by adopting a quantitative approach to languages and dialects and linguistic features at all levels and by examining interdependencies across linguistic features, space, and social factors that play a role in the formation of linguistic areas.

Beyond the specific interest for the Balkans, ABLA will also provide a general model for documenting and digitally mapping linguistic areas across the world.

Acknowledgements: This work has been carried out within the framework of the “Atlas of the Balkan Linguistic Area” project supported by the French National Research Agency (ANR-21-CE27-0020-ABLA) granted to Evangelia Adamou and the Russian Science Foundation (22-48-09003) granted to Andrey Sobolev.

Bibliography

- Adamou, E. 2021. *The Adaptive Bilingual Mind: Insights from Endangered Languages*. Cambridge: Cambridge University Press.
- Adamou, E. and Y. Matras (eds). 2021. *The Routledge Handbook of Language Contact*. London: Routledge.
- Adamou, E. and A. N. Sobolev (eds). 2025. *Atlas of the Balkan Linguistic Area Online*. Available online at <https://abla.cnrs.fr>, Accessed on 2024-05-03.
- Arapi, I. 2010. *Der Gebrauch von Infinitiv und Konjunktiv im Altalbanischen mit Ausblick auf das Rumänische*. Hamburg: Verlag Dr. Kovač.
- Asenova, P. 2002. *Balkansko ezikoznanie: osnovni problemi na balkanskiya ezikov süyuz* [Balkan Linguistics: Main Issues of the Balkan Linguistic Union]. Sofia: Faber. 1st edition in 1989.
- Bowern, C., Epps, P., Hill, J., and K. Hunley. 2020. *Hunter-gatherer language database: A collection of lexical, grammatical, and other information about languages spoken by hunter-gatherers and their neighbors*. Available online at <https://huntergatherer.la.utexas.edu/lexical>.
- Campbell, L. 2006. "Areal linguistics: a closer scrutiny". In *Linguistic Areas*, eds. Y. Matras, A. McMahon and N. Vincent, 1–31. London: Palgrave.
- Capidan, Th. 1935. *Meglenoromâni. Vol. III. Dicționar meglenoromân*. București: Cultura Națională.
- Dragomirescu, A. 2013. *Particularități sintactice ale limbii române în context romanic. Supinul*. București: Academia Română.
- Fiedler, W. 2018. *Vergleichende Grammatik der Balkansprachen (Morphosyntaktisch-typologischer Vergleich des Albanischen mit den anderen Balkansprachen)*. Prishtina: Kosova Academy of Sciences and Arts.
- Friedman, V. A. 2021. "The Balkans". In *The Routledge Handbook of Language Contact*, eds. E. Adamou and Y. Matras, 385–403. London: Routledge.
- Friedman, V. A. and B. D. Joseph. 2017. "Reassessing sprachbunds: A view from the Balkans." In *Cambridge handbook of areal linguistics*, ed. R. Hickey, 55–87. Cambridge: Cambridge University Press.
- Friedman, V. A. and B. D. Joseph. [forthcoming]. *The Balkan Languages*. Cambridge: Cambridge University Press.
- Gabinskii, M. A. 1967. *Voznikovenie infinitiva kak vtoričnyj balkanskij jazykovoj process (na materiale albanskogo jazyka)* [Emergency of the infinitive as a secondary linguistic process]. Leningrad: Nauka.
- Gabinskii, M. A. 2002. *Posobie po morfologii dakoromanskogo glagola* [Manual on Daco-Romanian verb morphology]. Kishinev: AN Respubliki Moldova.
- Gardani, F., Loporcaro, M. & A. Giudici. 2021. "In and Around the Balkans: Romance Languages and the Making of Layered Languages". *Journal of Language Contact* 14 (1), 1–23.
- Hickey, R. (ed.) 2017. *Cambridge handbook of areal linguistics*. Cambridge: Cambridge University Press.
- Joseph, B. D. 1983. *The synchrony and diachrony of the Balkan infinitive: A study in areal, general and historical linguistics*. Cambridge: Cambridge University Press.
- Joseph, B. D. 1992. "The Balkan languages". In *International encyclopedia of linguistics*, ed. W. Bright, 153–155. Oxford: Oxford University Press.

- Joseph, B. D. 2010. "Language contact in the Balkans". In *The Handbook of Language Contact*, ed. R. Hickey, 618–633. Chichester: Wiley-Blackwell.
- Kopitar, J. 1829. "Albanische, walachische und bulgarische Sprache". *Wiener Jahrbücher der Literatur* 46, 56–106.
- Ledgeway, A. 2017. "The Romanian definite article in a comparative Romance perspective". In *Sintaxa ca mod de a fi: Omagiu Gabrielei Pană Dindelegan, la aniversare*, eds. A. Dragomirescu, A. Nicolae, C. Stan & R. Zafiu, 231–247. Bucharest: Editura Universității din București.
- MDABL – Sobolev, A. N. (ed.) 2003–2018. *Malyi dialektologicheskii atlas balkanskikh iazykov*. Probnii vypusk (2003). Serii leksicheskaja. Tom I. Leksika dukhovnoi kul'tury (2005). Tom II. Chelovek. Sem'ia (2006). Tom iii. Zhivotnovodstvo (2009). Tom IV. (Avtor: M. V. Domosiletskaia). Landshaftnaia leksika (2010). Tom V. (Avtor: M. V. Domosiletskaia). Meteorologija (2012). Tom VI. Polevodstvo. Ogorodnichestvo (2013). Tom VII. (Avtor: M. V. Domosiletskaia). Pčelovodstvo (2018). Serii grammaticheskaja. Tom I. Kategorii imeni sushchestvitel'nogo (2005) [Minor dialect atlas of the Balkan languages. Test launch (2003). Lexical series. Vol. 1. Cultural vocabulary (2005). Vol. 2. The person and human relations (2006). Vol. 3. Animal husbandry (2009). Vol. 4. (M. V. Domosiletskaia, author). Landscape (2010). Vol. 5 (M. V. Domosiletskaia, author). Meteorology (2012). Vol. 6. Field-tending and Gardening (2013). Vol. 7. (M. V. Domosiletskaia, author). Apiculture (2018). Grammatical series. Vol. 1. Noun categories (2005)]. St. Petersburg, München: Nauka, Verlag Otto Sagner.
- Michaelis, S., Maurer, P., Haspelmath, M. and M. Huber (eds.) 2013. *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://apics-online.info> Accessed on 2021-03-17.
- Miklosich, F. 1861. *Die slavischen Elemente im Rumunischen*. Wien: K. K. Hof- und Staatsdruckerei.
- Pană Dindelegan, G. (ed.) 2013. *The Grammar of Romanian*. Oxford: Oxford University Press.
- Petrovici, E. 2005. *Folclorul din Valea Almăjului*. Timișoara: Editura Waldpress.
- Runić, M. 2019. "Indefinite articles and licensing of nominals in two Slavic varieties". In *Italian dialectology at the interfaces*, eds. S. Cruschina, A. Ledgeway & E.-M. Remberger, 295–318. Amsterdam: Benjamins.
- Sandfeld, K. 1930. *Linguistique balkanique: problèmes et résultats*. Paris: Champion. 1st version in Danish [1926].
- Sandfeld-Jensen Kr. 1900. *Rumaenske Studier. I. Infinitiv og udtrykkene derfor i Rumaensk og Balkansprogene*. København: E. Möller.
- Schumacher, St. and J. Matzinger. 2013. *Die Verben des Altalbanischen. Belegwörterbuch, Vorgeschichte und Etymologie*. Unter Mitarbeit von Anna-Maria Adaktylos. Wiesbaden: Harrassowitz.
- Sobolev, A. 2004. "On the areal distribution of syntactical properties in the languages of the Balkans". In *Balkan Syntax and Semantics*, ed. O. Mišeska-Tomić, 59–100. Amsterdam and Philadelphia: John Benjamins.
- Sobolev, A. (ed.) 2021. *Between Separation and Symbiosis: Southeastern European Languages and Cultures in Contact*. Berlin: de Gruyter.
- Sobolev, A. N. and A. A. Novik (eds.). 2013. *Golo Bordo (Gollobordë), Albaniia. Iz materialov balkanskoi ekspeditsii RAN i SPbGU 2008–2010 gg.* [Golo Bordo (Gollobordë)

- Albania. From the materials of the RAS and SPSU expedition of 2008–2010*. Munich: Otto Sagner Verlag, St. Petersburg: Nauka.
- Sobolev, A. N., M. L. Kisilier, M. S. Morozova, I. Padure, V. V. Kozak, A. Y. Rusakov, S. Gadjeva, D. D. Ivanova, Z. Jovanovic, C. Yebra López, E. Bužarovska, E. Ivanova, D. V. Konior, L. Mitkovska, A. Scala, M. A. Suloeva, S. Cirkovic, A. Dragomirescu, M. Forlano, A. Giudici, A. V. Kharlamova, M. Miric, J. Mücke, G. Vardakis, E. Adamou. 2025. "Infinitive: Forms". In *Atlas of the Balkan Linguistic Area Online*, eds. E. Adamou and A. Sobolev. <https://abla.cnrs.fr/feature/infinitive-forms/> Last accessed 2024-10-11.
- Trubetzkoy, N. S. 1923. "Vavilonskaja bashnja i smeshenie jazykov". *Evrazijskij vremenik* 3, 107–124.
- Trubetzkoy, N. S. 1928. "Proposition 16." In *Actes du Premier Congrès International de Linguistes du 10–15 avril, 1928*, 17–18. Leiden: Sijthoff.
- Wiemer, B. 2004. "Population linguistics on a micro-scale. Lessons to be learnt from Baltic and Slavic dialects in contact". In *Dialectology Meets Typology (Dialect Grammar from a Cross-Linguistic Perspective)*, ed. B. Kortmann, 497–526. Berlin: Mouton de Gruyter.
- Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. 2016. "The FAIR Guiding Principles for scientific data management and stewardship". *Scientific Data* 3, 160018.
- Ylli, Xh. and A. N. Sobolev. 2002. *Albanskii toskskii govor sela Leshnia (Krajina Skrapar). Sintaksis. Leksika. Etnolingvistika. Teksty.* [The Albanian-Tosk subdialect of the town of Leshnia (Skrapar Krajina). Syntax. Lexicon. Ethnolinguistics. Texts]. Marburg: Bibliion Verlag.

ISSN 0350-7653



9 770350 765003 >