DIGITAL CARTOGRAPHY IN THE SERVICE OF PRESERVATION OF CULTURAL LINGUISTIC HERITAGE: IMPLEMENTING THE ELECTRONIC DIALECTAL ATLAS OF CAPPADOCIAN GREEK

Dimitra Melissaropoulou[1], Stavros Bompolas[2] and Charalampos Tsimpouris[3]

[1] Department of Italian Language and Literature,
Aristotle University of Thessaloniki, 541 24 Thessaloniki, Greece
[2] Department of Philology, University of Patras, 265 04 Patras, Greece
[3] Department of Electrical and Computer Engineering, University of Patras, 265 04 Patras, Greece

ABSTRACT

This paper reports on the compilation of the electronic dialectal atlas of Cappadocian Greek, a state-of-the-art major reference work of the *DiCaDLand* research project (http://cappadocian.upatras.gr/en), funded by the Hellenic Foundation of Research and Innovation, which aims at the Digitization of the Cappadocian Dialectal Landscape by visualizing in different maps all different aspects of linguistic variation. Users can access the maps in several ways. The maps consist of two layers, a geographic and a linguistic one, which are user-composable. Different geographic basemaps are integrated, allowing the map reader to correlate linguistic phenomena to geographical space. The linguistic maps are grouped thematically, while the map stock is searchable using several indices (phonetics/phonology, morphology, syntax, etc.) to detect maps featuring specific phenomena. The legends of each map are projected in separate windows facilitating both the demonstration and the analysis of a map, providing an exhaustive range of feature values, and accompanied by examples of authentic use and notes for further reading. Moreover, audiovisual data are incorporated, where feasible. This endeavour is the first attempt, for the Greek standards, to implement a fully digital approach to dialectological research, by applying cutting-edge dialectological methods and informatics tools for the documentation and study of dialectal variation. This output of intangible heritage is meant to constitute, on the one hand, an invaluable cultural artifact securing the sustainable preservation of Asia Minor Greek linguistic heritage, while, on the other, a very useful digital tool for the study of language variation and change.

1. INTRODUCTION

Admittedly, geographically determined linguistic variation experiences an on-going decline both on a cross-linguistic and on a local −dialectal− level, threatening with extinction an enormous number of linguistic systems (Moseley, 2010). Unfortunately, Greek dialectal variation experiences the same serious threats under the influence of mobility, migration and the role of mass media (Kontosopoulos, 1994: 2). From this viewpoint, documentation and analysis of the existing variation is essential for both language sciences and the history and preservation of linguistic culture.

This paper reports on the on-going compilation of the electronic dialectal atlas of Cappadocian Greek, a state-of-the-art major reference work of the *DiCaDLand* research project, which aims at the Digitization of the Cappadocian Dialectal Landscape. Upon completion, the outcome will be the complete linguistic atlas of Cappadocian, which will serve as an empirical database documenting in detail the profile of Cappadocian Greek for a variety of geographical locations, constituting thus both a cultural artifact, securing the preservation of Asia Minor Greek language and a very useful digital tool in the service of language experts for the study of language variation and change.

The choice of this specific dialectal variety are of various reasons, both cultural and linguistic, and are as follows: a) Cappadocian is an endangered dialect of prime importance for the cultural history of (Asia Minor) Greece, b) it is a dialect until recently considered entirely extinct, for which a wealth of newly-discovered material exists, hitherto unexploited by linguistic (dialectological, sociolinguistic, historical) research, c) it is a dialect presenting great linguistic interest due to its long period of isolation from the Mainland Greece and strong contact with Turkish, and d) it is a Modern Greek dialect for which no standard works of reference exist.

In the framework of Digital Humanities, the product of such linguistic research is considered as an output of intangible heritage and cultural artifact which calls for the most appropriate virtual environment in order to secure its dissemination to the public. In the same perspective, the role of linguistic maps as a substantial resource of information for the promotion of impalpable heritage is also undeniable, since it embodies a wide variety of resources in terms of functional and organizational content, offering thus many potentials for the representation of linguistic and cultural heritage (Girnth, 2010; Lameli et al., 2010).

In the light of the above, it should be stressed that this atlas constitutes the first attempt, for the Greek standards, to implement a fully digital approach to dialectological research, by applying cutting-edge dialectological methods and informatics tools for the documentation and study of dialectal variation, with special emphasis on Asia Minor Greek. Moreover, this atlas is meant to serve as an infrastructure for the follow-up mapping of the rest of Modern Greek dialectal varieties, and ultimately to the compilation of a complete linguistic atlas of the Modern Greek Dialects, which is a desideratum both for academic research and the preservation of linguistic heritage.

2. THE DIALECT

The term *Cappadocian* is often used in the linguistic literature as a cover geographical −rather than a linguistic− term capturing other Asia Minor Greek varieties spoken in the wider area, that is in Pharasiot and Silliot (Andriotis, 1948: 10; Anastasiadis, 1975: 163, 1976: 9; see also Manolessou, 2019: 29). For the purposes of this study, we adopt a narrow definition of the term referring separately to all three dialectal varieties. Specifically, Cappadocian Greek used to be spoken in central Turkey, notably in villages within Kayseri, Nevşehir, Aksaray, Niǧde as well as Konya provinces. The relative positions of all three dialectal varieties can be seen on **Map 1**. Following the population exchange imposed by the Lausanne treaty (1923), Cappadocian refugees were relocated in various parts of (mainly Northern) Greece. Hence, no native speakers of Cappadocian are reported to live in Turkey anymore. The dialect was until recently considered extinct, after the end of the life-span of the 1st generation speakers relocated in the Greek mainland. However, it has been discovered to be retained by 2nd and 3rd generation speakers in several villages in Thessaly, Macedonia and Thrace (Janse, 2009).

Cappadocian Greek is a case study par excellence in Modern Greek Dialectology due to its early separation from the rest of the Greek-speaking world (11th century), and its subsequent evolution under circumstances of intense contact with Turkish and the local Anatolian dialects spoken in the area. Hence, the dialect is often used in the literature as a prototypical example of "heavy borrowing", referring to "overwhelming long-term cultural pressure" (Thomason and Kaufman, 1988).

Following "UNESCO Atlas of the World's Languages in Danger" (Moseley, 2010), Cappadocian Greek is critically endangered calling for the most appropriate way of its documentation and reservation. In this line, linguistic maps represent a highly relevant, if not the most appropriate, means of visualizing and documenting linguistic phenomena according to facets of their areal differentiation (Kehrein et al., 2010; Girnth, 2010: 98).

Map 1. Map of Cappadocia, drawn by Dawkins (1916: 725)

3. COMPUTERIZATION OF LANGUAGE MAPPING

The central function of linguistic maps lies in the documentation of the geographic distribution of linguistic forms (phonetic, morphologic, syntactic, lexical) for a specific language area. A collection of such maps constitutes a linguistic (or dialectal) atlas (Girnth, 2010: 98-99; Kretzschmar, 2018: 57; for the history of linguistic atlases, see also Lameli, 2010; Kretzschmar, 2018: 60-63; as regards national and/or language-specific traditions, see Lameli et al., 2010: 158-374).[1] Due to their bi-dimensionality, linguistic maps are highly appropriate for the representation of the areal distribution of linguistic phenomena. Forming part of a linguistic atlas, language maps generate a solid depiction of language variation (Girnth, 2010: 98). Hence, it is not a coincidence that innumerable, and not just recent, linguistic projects have been dedicated to the implementation of linguistic atlases. Grzega (2009) lists 181 projects: 23 for English, 28 for French, 35 for German and 95 for thirty-three other languages.

Traditional mapping techniques have become, the last three decades or so, computer-dependent, or even more recently, web-based. As a result, a great many novelties have expanded the features to be found on linguistic maps and their compilation has been certainly reformed and insanely improved qualitatively, while many atlases are available only online nowadays (cf. Lameli, 2010: 585-587; Lameli et al., 2010: xvii-xx, 375-505). The first projects had used computer methods in language mapping for the preparation and development of printed atlases (e.g., the Atlas Linguarum Europae [ALE], Alinei et al., 1983; the Kleiner Deutscher Sprachatlas [KDSA], Veith et al., 1984-1999; the Computer Developed Linguistic Atlas of England [CLAE], Viereck and Ramisch, 1991-1997). Currently, every atlas is realized virtually with the use of computers, yet the computer-specific approaches for language documentation and their goals vary. Some undertakings develop and use tailor-made computer systems to digitize (historical) data as well as cross-connect them to other information sources with enhanced functions (e.g., the Linguistic Atlas of the Middle and South Atlantic States [LAMSAS], Kretzschmar and Schneider, 1996; the Digital Wenker Atlas [DiWA], Schmidt and Herrgen,

---

[1] It is necessary to distinguish the map of languages (also known as languages map) from the linguistic map. The former illustrates information on the distribution and number of the languages in a given region.

2001-2009; the Bavarian Database of Dialects [BAY-DAT], Zimmermann et al., 2019; the Regional-spsrache.de project [REDE], Schmidt et al., 2008−). Other projects view a digital version of atlases as a valuable supplement to an edited book (e.g., the World Atlas of Language Structures [WALS], Dryer and Haspelmath, 2013; the Atlas of Pidgin and Creole Language Structures [APiCS], Michaelis et al., 2013). In the same line, other projects first publish their databases on the internet as a means of linguistic analysis and, subsequently, produce (printed) atlases of representative maps (e.g., the (Dynamic) Syntactic Atlas of the Dutch dialects [(Dyna)SAND], Barbiers et al., 2006, 2007).

As regards the Greek dialectal landscape, unfortunately little progress has been made in the field the last few decades. Triantaphyllides's (2002 [1938]: 66) observation on the absence of a complete linguistic atlas of Greek and its dialects is still standing. Crucially, till recently the only exception to this scientific lacuna was the printed dialectal atlas of Crete by Kontosopoulos (1988). The last decade some progress was made with the first online dialectal atlas emphasizing specific dialectal phenomena of the island of Lesvos (Ralli, 2010-2015, 2019; Alexelli, 2021) as well as the printed atlas of the Dodecanesian dialects (Minas, 2020).

Undoubtedly, the compilation of the complete atlas of Greek is a huge endeavor entailing many more human, technical and financial resources, let alone time efforts. The DiCadland project aims to contribute towards this direction by implementing the first dialectal atlas of Cappadocian Greek, making good use of the recent advances of digital cartography, aspiring that it could serve as a solid background for the expansion of the attempt in Modern Greek Dialects as a whole.

4. DEVELOPING THE ELECTRONIC DIALECTAL ATLAS OF CAPPADOCIAN DIALECTS

4.1. *The dialectal dataset*

Dialectal data for the implementation of the atlas are drawn from all available written sources, of both primary (folktales, songs, narrations, riddles, etc.) and secondary nature (grammatical descriptions, glossaries, dictionaries) the majority of which were published at the end of the 19th century and the first half of the 20th century. Although Cappadocian dialects used to be spoken in a wider geographical area, available written data is confined to 20 different communities (cf., e.g., Dawkins, 1916; Bağrıaçık, 2018: 14-17; Janse, 2019: 69-70, forthcoming). A great amount of these sources was collected and digitized within the frame of another project, named AMiGre, hosted

by the Laboratory of Modern Greek Dialects and implemented under the supervision of Em. Prof. Angela Ralli (Ralli, 2015: 43-98). AMiGre's list of primary and secondary sources was enriched with extra written sources and academic articles emphasizing specific phenomena or domains of linguistic structure. Unpublished data have also been used as sources depending on availability.

Dataset consists also of oral recordings most of which −with the exception of those available online by Gallica of the Bibliothèque Nationale de France, dating back to 1930s− constitute recent recordings from 2nd or 3rd generation descendants of Asia Minor −both Cappadocian and Pharasiot− refugees which were collected the last 12 years or so. This oral collection consists of transcribed and translated data from AMiGre project as well as of data kindly offered for the purposes of the project by Petros Karatsareas and Metin Bağrıaçık, which were in turn subject to transcription and translation in order to form a homogeneous set.

The amount of raw data of each of the two categories which was either already transcribed/translated or subject to transcription/translation within the frame of the DiCadLanD counts 116,449 words while the total of oral one circa 37 hours.

Given the above, our data cover only the horizontal −geographic− axis and, therefore, our atlas consists of mono-dimensional linguistic maps focusing almost exclusively on the diatopic dimension (Girnth, 2010: 105-106 and references therein; Lameli, 2010: 583-584 and Thun, 2010 for dimensionality in language mapping).

4.2. *The places*

4.2.1. The place names

One major issue every such endeavor has to struggle with is to secure the geographically correct position of the communities on the online linguistic maps. In the same line, consistency and accuracy are essential in referring to a place to prevent confusion in the recreation of the dialectal landscape of Cappadocia. For this purpose, we conducted toponymic research, to establish an exhaustive list with the various geographical names (Greek, Turkish and mixed) used for each place. During the toponymic research, a number of issues should be treated, as follows:

- the place names that should form part of the atlas (only Greek-speaking ones or Turkish-speaking based on common religion);
- the form that should be preferred and the reasoning behind it given that various alternates (in terms of orthography, of morpho-phonology, of (Greek vs. Turkish) origin, learned vs. -learned) are traced in the relevant literature

(for relevant discussion and categorization, see Tsagogeorga et al., 2009) and

- their accurate visualization on the map and the (non-) need for additional information of geographical, dialect or religious nature (prefecture or province, bishopric or dialect membership).

This stage of the project has benefited greatly by the seminal work made by the Center of Asia Minor Studies and the work by Logotheti-Merlier (1948, 1977), recognizing 81 Cappadocian communities, divided in different provinces. Among them, 31 were Greek-speaking, while 50 Turkish-speaking. Additionally, the seminal work by the Historical Lexicon by the Academy of Athens contributed significantly to the finalization of the list with the place names and the preferable name among the existing variants.

Regarding the first issue, both Greek- and Turkish-speaking communities were incorporated in the list of place names in order to visualize the distribution of different types of communities in the area of Cappadocia based on their language. The result can be seen on **Map 2**.

As for the second issue, given the fact that the existing relevant literature is scanty, the decision was made to follow those versions of the place names that are used most often in the Asia Minor Greek literature by the most well-established and expert in the field research centers, such as the Center for Asia Minor Greek Studies and the Historical Lexicon of Modern Greek Dialects. This methodological choice exhibits specific benefits since a) the outputs of this research project will be in line with the well-established relevant literature and b) it saves the project the need for research on the place names from scratch, which would have to be more restricted due to time limitations.
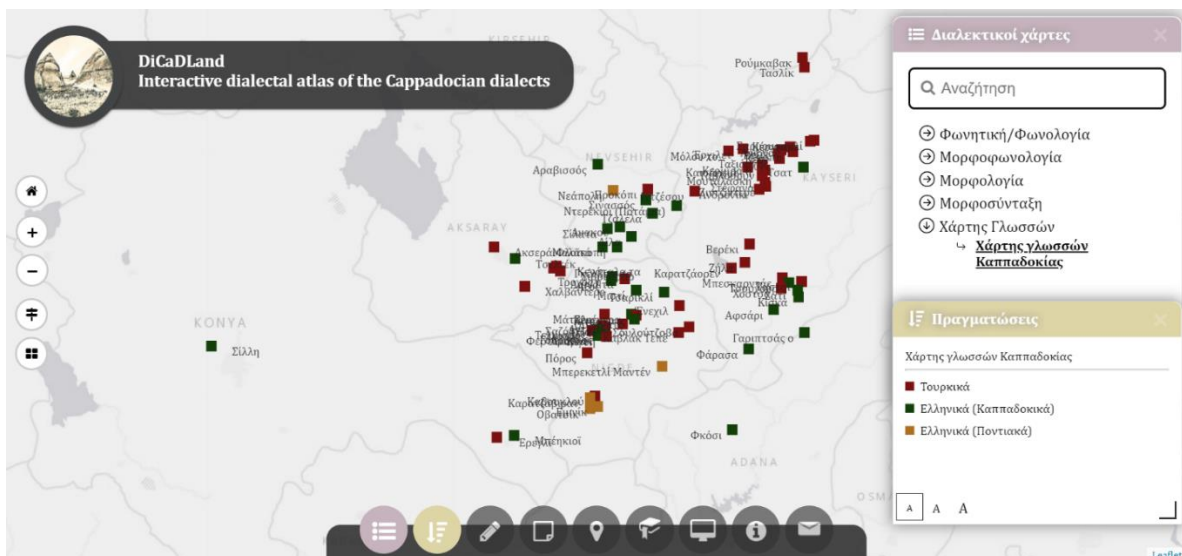
Map 2. Map of the linguistic communities of Cappadocia plateau

Lastly, descendants of Asia Minor Greek Cappadocian refugees were asked to comment on the acceptability of the different versions of the place names. This choice is deliberate and aims to conform to the dialectophons intuitions, aspiring that the compilation of the dialectal atlas will prove accessible and useful not only to the academia but also to native speakers of the dialect and their descendants as well (for a similar discussion see, among others, McDavid, 1958; Kaups, 1966; UN Conferences on the Standardization of Geographical Names, 1967-2017).

4.2.2. The locations

In order to visualize the actual location of the Cappadocian communities onto the map, we needed to acquire reliable geo-references for each one of them. To do so, we needed to associate coordinates with every place name, by aggregating geographical information from all available sources. In this direction, Index Anatolicus database (https://nisanyanmap.com/) proved a precious tool since it lists a highly reliable variety of information of historical and geographical nature for all different place names of Turkey. With the help of this tool (see for **Figure 1** an example), we managed to trace the coordinates of all place names (using DMS coordination system), which served as the appropriate geographical background for the visualization of the distribution of linguistic phenomena.

4.3. The maps

Following Rabanus (2018: 350), a dialectal map in its narrow sense consists of two layers. The basic layer (basemap) is generated from a geographical map of the area under investigation, providing the minimum

information to locate the linguistic data in the geographic space (see 4.3.1). The second −most crucial- layer contains the linguistic data, represented by graphic elements that are more prominent comparing to the topographical background (see 4.3.2).

4.3.1. The basemaps

Since linguistic maps place linguistic data into space, the first step in the compilation of a linguistic map (or atlas) is the specification of the survey area (the basic layer). Based on the size of the research area, our dialectal maps compile a regional (small-area) atlas (Girnth, 2010: 102).

Usually, the topographic layer is graphically coded using neuter colors (Rabanus, 2018: 350). In this line, we integrated ESRI's Light Gray Canvas basemap accompanied by Administrative Boundaries and Places layer, both available freely from ArcGIS Online. This type of basemap contains less detail and a narrow range of colors, which increases the space and palette available for operational overlays, putting the linguistic information in the center. It only shows basic geographic features like administrative boundaries, allowing map reader to correlate features, patterns, place names, etc. to features and names they are familiar with. Nevertheless, a satellite topographic layer background is also available for those who are interested to associate the historic locations of the survey places to the today's geographic and urban context of the area. Furthermore, a fully non-transparent layer has been added as an option for those who are interested solely in the depiction of linguistic data distribution without any geographic contexts and correlations. Finally, two types of non-linguistic layers have been integrated for orientation; one with the place names provided by the toponymic research (see 4.2) and the administrative boundaries and today's place names provided by ESRI.

Graphical computer displays allow superimposing all the above-mentioned layers of geographical information in order to compose user-selectable overlays, which contain some particular information for the basic layer (Kretzschmar, 2013; Kretzschmar, 2018: 67). **Map 3** is actually composed of several geographic layers: the non-linguistic layer, the basemap, the locations of communities and the symbols plotted at the locations (see [a], [b], and [c] on **Map 3**, respectively).

4.3.2. The linguistic information

As for the main layer which contains the linguistic information, we incorporated the qualitative mapping methodology of *point-related maps* (see Girnth, 2010: 108-110; Kretzschmar, 2018: 63-69; Rabanus, 2018: 350-353 for qualitative mapping methods). Generally, in point-related maps, symbols (triangles, circles, or other geometrical signs) are drawn ideally at the points where the exact geographical location of the data is, using coordinates. For our project, we adopt the most frequent subtype of point-related maps, i.e. the *qualitative point-symbol maps* (see the above-mentioned references for examples of atlases adopting this method). Particularly, within our project, the linguistic data are grouped and graphically coded by squares of different colors, which signify the same quantity but different quality. Nevertheless, the reader will notice a definite number of colors which recur throughout the atlas.

Qualitative point-symbol mapping is of increased abstraction as regards the way both the linguistic information and, to some degree, the geographic information is represented. It is particularly preferable when the location network is of medium density and when a purely documentational function ought to be complemented by the function of visual economy (Girnth, 2010: 109). Hence, this mapping method is very objective in that the individual linguistic data points are located in their exact areal distribution, leaving aside any interpretation of the data (Veith, 2006: 521).

Finally, point-symbol mapping can take advantage of the advancements in the field of digital language mapping. More specifically, every point on the map that contains information about its location in the coordination system enables integrating and comparing different types of data (Girnth, 2010: 107). Consequently, not only square-symbols but also texts and sound/image data −when available− are georeferenced for each point illustrated as a pop-up window (see [d] on **Map 3**), forming audiovisual linguistic maps (for exemplary projects of this type, see Atlas linguistique audiovisuel du Valais romand / Vivaio Acustico delle Lingue e dei Dialetti d'Italia, Müller et al., 2001; A Sound Atlas of Irish English, Hickey, 2004; Sprechender Sprachatlas von Bayern, König and Renn, 2007; among many others).
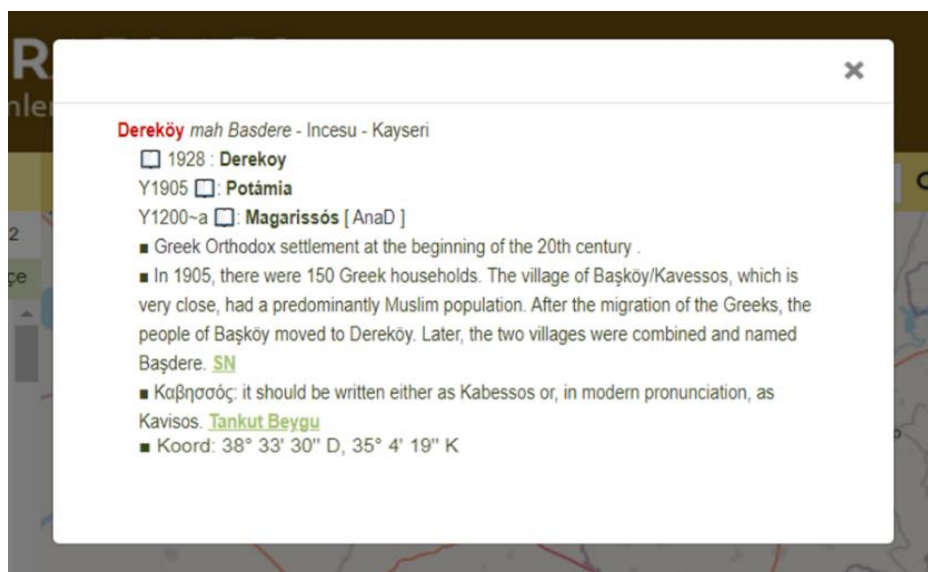
Figure 1. Geo-information taken from Index Anatolicus for Dereköy (adapted to English)
[available at https://nisanyanmap.com/?y=potamia&lv=&t=&cry=TR&ua=5, last accessed 2022/02/04]

4.3.2.1. The features

What dialectal maps visualize is the spatial distribution of linguistic features or, more generally, feature-based areal structures (Rabanus, 2018: 348). Until now, we have mapped ca. 400 linguistic features. Each of the 400 maps shows the distribution of a particular linguistic feature.

While the atlas was initially projected to visualize ca. 15-20 different dialectal phenomena, for all four basic levels of linguistic analysis (namely, phonetics-phonology, morphology, syntax, and vocabulary), a lot of efforts have been put forward in order to provide a complete visualization of the intra-Cappadocian variation based on the following criteria:

- availability of resources depicting intra-dialectal variation;
- degree of intra-dialectal variation;
- degree of divergence with respect to inter-dialectal variation.

To this end, a provisional list of linguistic phenomena/features and their basic values was compiled, obtained primarily from the available general published descriptions of Cappadocian varieties. This methodological choice naturally encompasses the risk of delimiting the choice of phenomena to the range covered by the existing major Cappadocian works. Nevertheless, the risk was eliminated by following a two-staged *bottom-up qualitative (example mining) methodology*, (Szmrecsanyi and Anderwald, 2018: 307 and references therein):

1. firstly, by investigating all different descriptions emphasizing different linguistic communities of Cappadocian, and

2. secondly, by investigating all existing primary sources (oral and written) for both the realization of major linguistic categories as well as for tracing other features and/or values than those already discussed in the existing grammatical descriptions.

The above-mentioned methodology secured that important features and/or values will not remain hidden and all existing variation will be revealed, to the highest possible extent.

The features are indicated in a small window on the map itself (which can be hidden), and in expanded form in a larger window (the features window is called "Dialectal maps", see [e] on **Map 3**). The maps are grouped into six sections corresponding to the basic levels of linguistic analysis: phonetics/phonology, morpho-phonology, morphology, morpho-syntax, syntax and lexicon. Within each of these sections, the features included provide a broad coverage of the most prominent variation. At the top of this window, a search function is integrated, allowing the users to search within the features list.

Dialectal maps require a reference system for the variables (e.g., features) whose variants are depicted on the maps. In the framework of DiCaDLand project, we have adopted *a grammatical/systematic reference system*, according to which the points of reference are grammatical categories (Girnth, 2010: 116; Rabanus, 2018: 348-349). This is an important methodological decision in the sense that it can serve not only for the purposes of the compilation of the atlas of one dialect but also as a solid background (in terms of technical infrastructure and research methodology) for the compilation of the Atlas of Modern Greek Dialects as a whole. Furthermore, a systematic effort was made

to achieve a balance among different grammatical descriptions of the phenomena in order to make the maps accessible not only to the academia but also to everyone interested in Cappadocian dialectal variation.

4.3.2.2. The feature values

Each feature is associated with a set of feature values, which form the basis for distinguishing intra-dialectal variation. The simplest maps display two different values, the majority distinguish between two, three and five feature values, and only a few remaining distinguish up to twelve values. Although many feature values could usually be subdivided into many more subtypes, the values have been limited to twelve at most, since an enormous increase of different colors

(and/or shapes) within the same area would make very unfriendly and hard to read (see WALS and APiCS projects for similar methodological decisions; cf. also Rabanus, 2018: 350).

The feature values are indicated in a window on the map itself (the value window is called "Realizations", see [f] on **Map 3**). For each feature value, this window shows the color and the shape of the point symbol associated with a specific value. Depending on the linguistic level, values can be of phonetic/phonological, morphological, syntactic and lexical nature. We distinguished phonemic representations (embedded to / /) to phonetic ones (embedded to [ ]) −where necessary− in order to be clear to the map reader if values are only phonologically different. This was also necessary to reduce the number of feature values, grouping them under a common head.
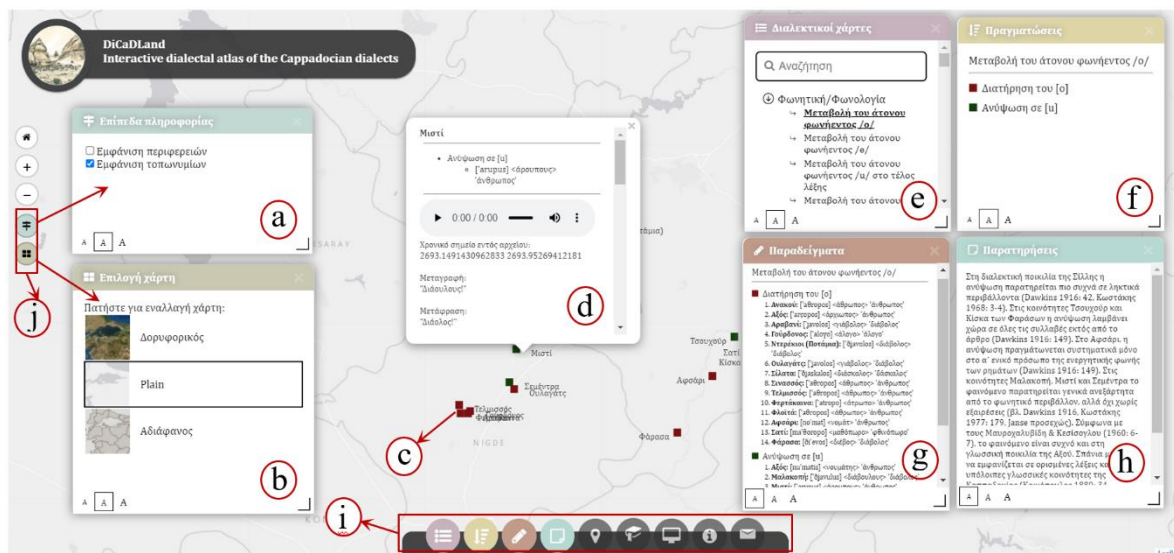
Map 3. Superimposing the different layers of geographical and linguistic information of DiCaDLand's Atlas

One prerequisite on the feature values of each feature is that they must be exhaustive. More particularly, for each feature, every community under consideration must be assigned one feature value; there are no "blank" cases for varieties that do not have a feature value. To meet this requirement, we included feature values that account for various cases involving non-applicability or indeterminacy (e.g., lack of relevant data – sources, no applicability, etc).

4.3.2.3. The examples of authentic use

Special efforts were made so as each feature value to be accompanied by an authentic dialectal example (drawn either by oral or by written textual sources), which would make the feature values easier for interpretation to the map user. The examples of the feature values are indicated in a small window on the map itself (which can be hidden), and in expanded form in a larger window (this window is called "Examples" see [g] on **Map 3**). For each feature value, this window

shows the colour and shape of the associated square, and a list of the dialects characterized by the feature value in question. Next to each variety the example follows transcribed in *International Phonetic Alphabet* (IPA) and as an orthographic form followed by its translation in Standard Modern Greek. For the phonetic and orthographic transcription of the examples coming from the Cappadocian dialects, we are following the list of symbols introduced and applied for the compilation of the Historical Dictionary of Modern Greek and its Dialects by the Academy of Athens (Manolessou et al., 2012). For the communities no reliable dialectal data was available, the examples cell has been left blank, implying that we have information on the specific realization of the feature but no authentic example to accompany it.

4.3.2.4. The notes for further reading

Finally, a "Notes for Further Reading" section is used −if necessary− to explain the feature in question,

define the feature values, provide more examples, discuss the geographical and historical distribution of the values, their frequency, language contact issues, its connection to other maps of the atlas, suggest specialized works for further reading and, generally, situate it in a wider context. Again, this section is indicated in a small window on the map itself (which can be hidden), and in expanded form in a larger window (see [h] on **Map 3**).

4.4. Technical details

This online atlas has been developed with React library (current 16.13.1). React (also known as React.js or ReactJS) is an open-source, front end, JavaScript library for building user interfaces or UI components. It is maintained by Facebook and a community of individual developers and companies. React can be used as a base for the development of single-page or mobile applications. For the purposes of our project, this framework helped proceed with the development of the online map quickly without losing in functionality or extendibility, while also supporting mobile *User Interfaces* (UIs) in the future. Another notable feature is the use of a virtual *Document Object Model* (DOM). React creates an in-memory data-structure cache, computes the resulting differences, and then updates the browser's displayed DOM efficiently. This process is called reconciliation and allows the programmer to write code as if the entire page is rendered on each change, while the React libraries only render subcomponents that actually change. This selective rendering provides a major performance boost which saves the effort of recalculating the CSS style, layout for the page and rendering for the entire page.

One of the main components is the Leaflet mapping library which is included with the use of react-leaflet package (2.7.0). Leaflet is the leading open-source JavaScript library for mobile-friendly interactive maps. Weighing just about 39 KB of JS, it has all the mapping features most developers ever need. Leaflet is designed with simplicity, performance and usability in mind. It works efficiently across all major desktop and mobile platforms, while it can be extended with lots of plugins.

UI has been developed to maximize the user's monitor size. In order to make this happen, while being able to visualize a lot of information, windows can be enabled or disabled at user's will with the use of the menu (see [i] and [j] on **Map 3**). Also, another python script has been developed to automatically create the images of the square symbols that are associated to the feature values within the legends of the maps (see [c], [f] and [g] on **Map 3**).

Dialectal atlas' data is initially stored in Excel sheets and exported manually in semi colon (;) separated format. Subsequently, and with the help of various scripts written in python, all data is transformed in a JSON structure, which is a programmatically friendly way to handle data that is transferred through the internet. The database's structure of entities, while managed from spreadsheets, is depicted in **Figure 2**.

5. CLOSING REMARKS

To sum up, it should be stressed again that this atlas constitutes the first attempt, for the Greek standards, to implement a fully digital approach to dialectological research, by applying cutting-edge dialectological methods and informatics tools for the documentation and study of dialectal variation, with special emphasis on Asia Minor Greek.

It is anticipated that the publication of such a project will not only contribute to new interpretations of the relationship between space and language by offering access to a large quantity of data (Auer and Schmidt, 2010), but could also bring together local and international stakeholders, such as linguists, speech communities and the public, in order to raise awareness and provide opportunities for cultural preservation, all through interactive, user-friendly means in appealing visual and sensory (multimodal) formats. Ultimately, they will contribute to the *European Cultural Heritage exploitation and dissemination*, delivering and enhancing learning experiences for different targets of users (see the Text of the Convention for the Safeguarding of the Intangible Cultural Heritage, UNESCO, 2003; see also Pantano and Tavernise, 2009). Last, but not least, we aspire that the project will serve as an infrastructure for the follow-up mapping of the rest of the Asia Minor Greek dialects as well as other Modern Greek dialects, and ultimately to the compilation of a complete linguistic atlas of the Modern Greek Dialects.
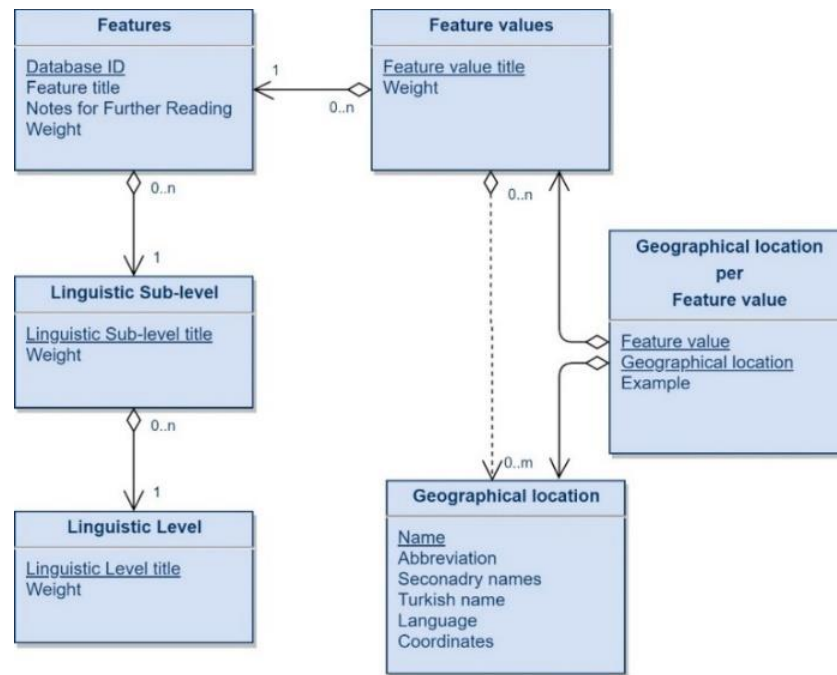
Figure 2. The atlas' database structure

ACKNOWLEDGEMENTS

REFERENCES

Alexelli, V. (2021) Mapping the language variation of Lesvos [in Greek]. Ph.D. Thesis, University of Patras.

Alinei, M., Allières, J., Avanesov, R. I., Itkonen, T., Viereck, W. and Weijnen, A. A. (eds) (1983) *Atlas linguarum Europae (ALE)*. Assen/Maastricht/Rome: van Gorcum/Istituto Poligrafico e Zecca dello Stato.

Anastasiadis, V. K. (1975) History and language of Cappadocia and the idiom of Pharasa [in Greek]. *Mikrasiatika Chronika* 16: 150-184.

Anastasiadis, V. K. (1976) *The syntax in Pharasiot idiom of Cappadocia in comparison to the other idioms of Asia Minor, as well as to the Ancient, Medieval and Modern Greek Language* [in Greek]. Thessaloniki.

Andriotis, N. P. (1948) *The linguistic idiom of Pharasa* [in Greek]. Athens: Ikaros.

Auer, P. and Schmidt, J. E. (eds) (2010) *Language and Space: An International Handbook of Linguistic Variation. Volume 1: Theories and Methods*. Berlin: De Gruyter Mouton.

Bağrıaçık, M. (2018) Pharasiot Greek: word order and clause structure. Ph.D. thesis, Ghent University.

Barbiers, S. C. J., Bennis, H. J., De Vogelaer, G., Devos, M. and van der Ham, M. H. (2006) *Dynamische Syntactische Atlas van de Nederlandse Dialecten (DynaSAND)*. Amsterdam: Meertens Instituut [http://www.meertens.knaw.nl/sand/, last accessed 2022/02/04].

Barbiers, S., Cornips, L. and Kunst, J. P. (2007) The Syntactic Atlas of the Dutch Dialects (SAND): A Corpus of Elicited Speech and Text as an Online Dynamic Atlas. In: J. C. Beal, K. P. Corrigan and H. L. Moisl (eds), *Creating and Digitizing Language Corpora*. London: Palgrave Macmillan, 54-90.

Dawkins, R. M. (1916) *Modern Greek in Asia Minor: a study of the dialects of Sílli, Cappadocia and Phárasa with grammar, texts, translations and glossary*. Cambridge: Cambridge University Press.

Dryer, M. S. and Haspelmath, M. (eds) (2013) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. [http://wals.info, last accessed 2022/02/04].

Girnth, H. (2010) Mapping language data. In: A. Lameli, R. Kehrein and S. Rabanus (eds), *Language and Space: An International Handbook of Linguistic Variation. Volume 2: Language Mapping. Part I. Part II: Maps*. Berlin/New York: De Gruyter Mouton, 98-121.

Grzega, J. (2009) *Sources for Onomasiological Studies*. [https://www1.ku.de/SLF/EngluVglSW/OnOn-4.pdf, last accessed 2022/02/04].

Hickey, R. (2004) *A Sound Atlas of Irish English*. Berlin/NewYork: Mouton de Gruyter.

Janse, M. (2009) Greek-Turkish language contact in Asia Minor. *Etudes Helleniques – Hellenic Studies* 17(1): 37-54.

Janse, M. (2019) Agglutinative Noun Inflection in Cappadocian. In: A. Ralli (ed), *The Morphology of Asia Minor Greek*. Leiden: Brill, 66-115.

Janse, M. (forthcoming) The Cappadocian dialect [in Greek]. In: C. Tzitzilis (ed), *The Greek language and its dialects* [in Greek]. Thessaloniki: Institute for Modern Greek Studies (Manolis Triantaphyllides Foundation).

Kaups, M. (1966) Finnish Place Names in Minnesota: A Study in Cultural Transfer. *Geographical Review* 56(3): 377-397.

Kehrein, R., Lameli, A. and Rabanus, S. (2010) Introduction. In: A. Lameli, R. Kehrein and S. Rabanus (eds), *Language and Space: An International Handbook of Linguistic Variation. Volume 2: Language Mapping. Part I. Part II: Maps*. Berlin/New York: De Gruyter Mouton, xi-xxii.

König, W. and Renn, M. (2007) *Kleiner Sprachatlas von Bayerisch-Schwaben*. Augsburg: Wißner.

Kontosopoulos, N. (1988) *Linguistic atlas of Crete* [in Greek]. Heraklion: Crete University Press.

Kontosopoulos, N. (1994) *Dialects and idioms of Modern Greek* [in Greek]. Athens: Grigoris.

Kretzschmar, W. (2013) GIS for language and literary study. In: R. Siemens and K. Price (eds), *Literary Studies in a Digital Age: An Evolving Anthology*. New York: MLA.

Kretzschmar, W. A. Jr. and Schneider, E. W. (1996) *Introduction to Quantitative Analysis of Linguistic Survey Data: An Atlas by the Numbers*. London/New Delhi: Sage.

Kretzschmar, W. A., Jr. (2018) Linguistic atlases. In: C. Boberg, J. Nerbonne and D. Watt (eds), *The Handbook of Dialectology*. Hoboken: John Wiley and Sons, Inc., 57-72.

Lameli, A. (2010) Linguistic atlases – traditional and modern. In: P. Auer and J. E. Schmidt (eds), *Language and Space: An International Handbook of Linguistic Variation. Volume 1: Theories and Methods*. Berlin/New York: De Gruyter Mouton, 567-592.

Lameli, A., Kehrein, R. and Rabanus, S. (eds) (2010) *Language and Space: An International Handbook of Linguistic Variation. Volume 2: Language Mapping. Part I. Part II: Maps*. Berlin/New York: De Gruyter Mouton.

Logotheti-Merlier, M. (1977) The Greek communities of modern Cappadocia [in Greek]. *Deltio Kentrou Mikrasiatikon Spoudon* 1: 29-74.

Manolessou, I. (2019) The historical background of the Asia Minor dialects. In: A. Ralli (ed), *The Morphology of Asia Minor Greek: Selected Topics*. Leiden/Boston: Brill, 20-65.

Manolessou, I., Beis, S. and Bassea-Bezantakou, S. (2012) The phonetic transcription of the Modern Greek dialects [in Greek]. *Lexikografikon Deltion* 26: 161-222.

McDavid, R. I. (1958) Linguistic Geographic and Toponymic Research. *Names* 6(2): 65-73.

Merlier, M. (1948) *The Archive of Asia Minor Folklore* [in Greek]. Athens: French Institute.

Michaelis, S. M., Maurer, P., Haspelmath, M. and Huber, M. (eds) (2013) *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. [http://apics-online.info, last accessed 2022/02/04].

Minas, K. (2020) *Linguistic Atlas of Dodecanese* [in Greek]. Thessaloniki: Institute for Modern Greek Studies (Manolis Triantaphyllides Foundation).

Moseley, C. (ed) (2010) *Atlas of the World's Languages in Dange*r. Paris: UNESCO Publishing [http://www.unesco.org/culture/en/endangeredlanguages/atlas, last accessed 2022/02/04].

Müller, M. L., Köhler, C. and Kattenbusch, D. (2001) VIVALDI – ein sprechender Sprachatlas im Internet als Beispiel für die automatisierte, computergestützte Sprachatlasgenerierung und -präsentation. *Dialectologia et Geolinguistica* 9: 55-68.

Pantano, E. and Tavernise, A. (2009) Learning Cultural Heritage Through Information and Communication Technologies. *International Journal of Information Communication Technologies and Human Development* 1(3): 68-87.

Rabanus, S. (2018) Dialect Maps. In: C. Boberg, J. Nerbonne and D. Watt (eds), *The Handbook of Dialectology*. Hoboken: John Wiley and Sons, Inc, 348-367.

Ralli, A. (2010-2015) *The Electronic Dialectal Atlas of Lesbos (EDAL)* [http://lesvos.lmgd.philology.upatras.gr/, last accessed 2022/02/04].

Ralli, A. (2019) Linguistic mapping: the electronic dialectal atlas of Lesbos [in Greek]. In: G. Karla, I. Manolessou and N. Pantelidis (eds), *Words. Volume dedicated to Christina Bassea-Bezantakou* [in Greek]. Athens: Kardamitsa, 435-456.

Ralli, A. (ed) (2015) *THALIS Project "Pontus, Cappadocia, Aivali: In search of Asia Minor Greek"* [in Greek]. Patras: Laboratory of Modern Greek Dialects, University of Patras.

Schmidt, J. E. and Herrgen, J. (eds) (2001-2009) *Digitaler Wenker-Atlas (DiWA)*. Executed by A. Lameli, A. Lenz, J. Nickel, R. Kehrein, K.-H. Müller and S. Rabanus. Marburg: Forschungszentrum Deutscher Sprachatlas.

Schmidt, J. E., Herrgen, J. and Kehrein, R. (eds) (2008−) *Regionalsprache.de (REDE). Forschungsplattform zu den modernen Regionalsprachen des Deutschen*. Executed by D. Bock, B. Ganswindt, H. Girnth, S. Kasper, R. Kehrein, A. Lameli, S. Messner, C. Purschke and A. Wolańska. Marburg: Forschungszentrum Deutscher Sprachatlas [https://www.regionalsprache.de/en/home.aspx, last accessed 2022/02/04].

Szmrecsanyi, B. and Anderwald, L. (2018) Corpus-Based Approaches to Dialect Study. In: C. Boberg, J. Nerbonne and D. Watt (eds), *The Handbook of Dialectology*. Hoboken: John Wiley and Sons, Inc., 300-313.

Thomason, S. G. and Kaufman, T. (1988) *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.

Thun, H. (2010) Pluridimensional Cartography. In: A. Lameli, R. Kehrein and R. Rabanus (eds), *Language and Space: An International Handbook of Linguistic Variation. Volume 2: Language Mapping. Part I. Part II: Maps*. Berlin/New York: De Gruyter Mouton, 506-523.

Triantaphyllides, M. (2002 [1938]) *Modern Greek Grammar. Volume 1. Historical Introduction* [in Greek]. Thessaloniki: Aristoteleio Panepistimio Thessalonikis – Institute for Modern Greek Studies (Manolis Triantaphyllides Foundation).

Tsagogeorga, D., Labropoulou, P. and Papageorgiou, H. (2009) From past to present: The electronic thesaurus of place names of Asia Minor [in Greek]. In: K. Valeontis (ed), *Hellenic Language and Terminology, 7th Conference Papers*. Athens: Technical Chamber of Greece, 297-305.

UNESCO (2003) Text of the Convention for the Safeguarding of the Intangible Cultural Heritage. [https://ich.unesco.org/en/convention, last accessed 2022/02/04].

United Nations Conferences on the Standardization of Geographical Names (1967-2017) United Nations. Department of Economic and Social Affairs.

Veith, W. H. (2006) Dialect atlases. In: K. Brown (ed), *Encyclopedia of Language and Linguistics*. Amsterdam/New York/Oxford/Shannon/Singapore/Tokyo: Elsevier, 517-528.

Veith, W. H., Putschke, W. and Hummel, L. (1984-1999) *Kleiner Deutscher Sprachatlas (KDSA). 4 volumes*. Tübingen: Niemeyer.

Viereck, W. and Ramisch, H. (1991-1997) *The Computer Developed Linguistic Atlas of England (CLAE). 2 volumes*. Tübingen: Niemeyer.

Zimmermann, R., Raaf, M., König, W., Eichinger, L. M., Eroms, H.-W., Wolf, N. R., Munske, Horst H. and Hinderling, R. (2019) *Database of Bavarian Dialects (BayDat)*. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University [https://baydat.badw.de, last accessed 2022/02/04].