DATABASE ESTABLISHING AND LINGUISTIC RESEARCH FOR MICROTOPONYMS IN TWO REGIONS IN NORTH MACEDONIA

Xhemaludin IDRIZI
Retired professor at the South-East European University
xhemaludin@yahoo.com

Bashkim IDRIZI
University of Prishtina “Hasan Prishtina”, Geodesy Department
bashkim.idrizi@yahoo.com

Abstract

Microtoponyms are part of geo names in a certain area that represents how the local population call that place. The acquisition, classification, multi-scientific analyses, publishing, archiving and distribution of microtoponyms represents the core task for GI experts, in order to set the preconditions for their involvement in various thematic Spatial Data Infrastructures. In North Macedonia many institutions have included geo names in official spatial databases. The Cultural heritage protection office (CHPO) is state responsible institution for establishing of official database and standardizing of all geo names in North Macedonia. Beside CHPO, the Agency for Real Estate Cadastre (AREC) is developer of many spatial databases such as cadastral, topographic etc, as well developer of National Spatial Data Infrastructure where geo names are included as separate dataset based on INSPIRE specification. In this research, the development of database for microtoponyms based on official state classification and linguistic analyses have been performed. Database was developed in open source platform, by using open layers as basemap, and based on the INSPIRE specifications for geo names, while the linguistic analyses have been performed base on the pronunciation of toponyms by the local population. Regions of Karshiaka and Derven nearby Skopje, are the case study areas.

Keywords: Microtoponyms; geo names; spatial database, Karshiaka, Derven, Skopje, GI

1. INTRODUCTION

The United Nations Group of Experts on Geographical Names (UNEGGN) defines geographical names as a name applied to a feature on Earth, used consistently in language to refer to a particular place, feature or area having a recognizable identity on the surface of the Earth, which include: populated places, civil divisions, natural features, constructed features and unbounded places or areas that have specific local meaning (UNEGGN, 2006).

Geographical names are an important part of the intangible cultural heritage. First of all, it should be underlined that most geographical names were coined at the very spot where they are used – or were used, if they have gone into oblivion, thus, they are memories of the place as well of the people who gave the names. Secondly, geographical names normally contain linguistic elements which are semantically related to the scene of naming, that means that they tell something about the place and the circumstances at the time of the coining of the name. And thirdly, the linguistic material contained in the geographical names is a valuable source for the study of language history (UNSTAT, 2018; Paranina & Paranin, 2017).
Geographic place names as expressions designating a place constitute a significant part of the vocabulary of any language, thus they represent a central element in all languages, since it is exclusively these elements that are appropriate for the denotation of different types of places and objects (Baba, 2014). A geographical name may also be referred to as a topographical name or toponym (UNGEGN, 2006). Geographic place names may form a toponym even on themselves (Baba, 2014).

There are two basic ways to conduct toponymic research — one concentrating on the etymology, meaning, and origin of toponyms, and one focusing on the toponyms of a region and examining patterns of these names. Usually, this distinction is not explicitly recognized (Tent, 2015). Geographical names are widely used in every-day communication for referring to various natural and man-made objects in the real world. Correct usage of geographical names is a principal aspect of everyday communication, consequently the status (official, historical…) linguistic properties (language, spelling, eventual transliteration, etc.) are a prime interest of many users, including press agencies, map publishers, spatial analysts, authorities, etc. (INSPIRE, 2014).

A geographical name serves as a means to identify a location. Gazetteers and gazetteer services associate the names with corresponding features – or locations – by means of coordinates, feature types and/or other necessary information (INSPIRE, 2016). Establishing spatial data set for geo names by including all microregional specific types of geo names followed by linguistic characteristics, in order to incorporate it within the global standardized platforms, is the challenge that occurs researchers on toponymy. All information on toponyms published in a past period in a printed material, by including outcomes from multidisciplinary toponymy analyses, should be coded and converted to the digital version based on Geographical Information System (GIS) and Spatial Data Infrastructure (SDI) platforms, to be reachable via geodatabases and web services.

The analysis of toponyms is highly complicated because they often form an exception to common rules of word formation and semantics, which usually remain in use for a very long period of time so that, between the time of their creation and their first attestation, the morphological patterns and the semantics of words involved may have changed considerably. Therefore, toponyms can reflect a more archaic state of language than the comparative materials which function as the basis for evaluation. They even may belong to foreign languages or remote dialects which have left hardly any traces in written records except for isolated words or a substratum to be found in onomastics. Furthermore, names of places and peoples may have been altered in order to explain them or to provide them with a new meaning by means of re-analysis or folk-etymology (Gundacker, 2017).

Microtoponyms are the names of small features such as fields, pastures, fences, stones, rocks, marshes, bogs, ditches, houses, sections of roads, localities etc. (Nash, 2013). These tend to be used only locally, by a limited group of people, or by special interest groups, such as bushwalkers, fishers etc. Microtoponyms are not often officially recognized or gazetted, and do not generally appear on any published maps. They comprise features that are part of a larger geographical entity, such as the name for the bank on a lake, or for a feature on such a bank, or a waterhole in a river, or the name of a feature on the side of a mountain. A microtoponym is in essence the name of a feature that is itself part of a larger named entity. Because they are usually unofficial placenames, microtoponyms evade the rules of official naming and standardization that normally apply to official toponyms (Tent, 2014).

In this paper we concern ourselves with microtoponyms collected by Xhemaludin Idrizi in a period 1999 to 2019 year, as toponyms given to population places, fields, mountains, forest and so on in the regions of Karshiaka and Derven nearby Skopje. Microtoponyms are particularly interesting for geographers, since they refer to natural features, and thus may allow us to understand ways in which landscapes are partitioned into meaningful elements, while also
providing clues as to the history of, however, microtoponyms, like all toponyms can act as referents without conveying any meaning (Coates, 2006). Over time language can evolve, spelling may be normalized, events can be forgotten, and landscapes and their usage can change making the link between a name and its origins opaque. In linguistics, the field of onomastics explicitly seeks to etymologically disentangle the original meanings of toponyms, typically by exploring historical sources to find and explain the first documented usage of a toponym as referent to a place (Villette 2018). Linguistic analyses for both regions in this paper have been performed by first author Xhemaludin Idrizi, while the database was developed by second author Bashkim Idrizi.

2. CLASSIFICATION OF GEO NAMES IN NORTH MACEDONIA

Geo names in North Macedonia are part of national legislation, as very important element of spiritual cultural heritage, that is under the responsibility of the Cultural Heritage Protection Office (CHPO) under the Ministry of Culture. Law on protection of cultural heritage (Law nr.24, 2004), followed by the regulation on national register of cultural heritage (Regulation nr.25, 2005) and the regulation on national classification of cultural heritage (Regulation nr.37, 2006) are the main national formal documents that protect and standardize toponyms in North Macedonia. On article 26 of the Law (nr.20, 2004), toponyms are defined as names of lakes, rivers, source, and other water objects (hydronyms), cities, villages and other settlements (oiconims), natural or administrative regions (horonyms), roads (dromonyms), agricultural spatial objects (agronyms), mountains and other natural objects related to forest (dendronyms), and other source, local and official names that are subject of toponymy of Republic of North Macedonia (Law nr.20, 2004). According to article 2 and article 32 item 2 point 7 of the Law (nr.20, 2004), toponyms are protected from unauthorized renaming, treated as cultural heritage in danger.

In order to provide formal legal protection, based on the regulation on national registry of cultural heritage (Regulation nr.25, 2005), CHPO is responsible for establishing and managing with the registry, compiled in main book with four chapters and three additional registers. According to article 2 of regulation (nr.25, 2005), spiritual cultural heritage is content of third chapter of main book, so cultural heritage in danger is part of third additional registry. Registration list for spiritual cultural heritages contain (Regulation nr.25, 2005): classification, category, link to other protected good, specific features, spatial position, characteristic entry and author of the entry.

According to the regulation on national classification of cultural heritage (Regulation nr.37, 2006), spiritual cultural heritage contains three types, in which toponyms (geographical names) are listed and classified within third type. A bit different from article 26 of the Law (nr.20, 2004), in point 3.3. (pages 113-115) of the national classification (Regulation nr.37, 2006), toponymes are categorized in eight classes (microtoponyms, oiconims, oronyms, hydronyms, horonyms, memorial toponyms, hodonyms and chrematonyms) that contain 36 subclasses.

3. DATABASES OF GEO NAMES IN NORTH MACEDONIA

Currently in North Macedonia exist many spatial databases which contain geo names, collected in different (standardized and non-standardized) ways and periods. Agency for real estate cadaster (AREC) is governmental institution which has one of the oldest databases with geo names collected within the framework of cadastral survey for scales 1:500-5.000, as well topographic database inherited by the Military Geographic Institute of Yugoslavia in Belgrade within the framework of topographic survey for producing topographic maps in scale 1:25.000-
200,000. Based on the Law for real estate cadaster (Law nr.55, 2013), all data are stored in the geodetic cadastral information system, which is under the full responsibility of AREC.

During the cadastral survey, based on the laws for land cadaster and real estate cadaster, one of main collected data for each cadastral parcel is “place name”, which in fact is geo name. On the latest Law (nr.55, 2013), based on article 66, it is defined as obligatory descriptive data to be collected during field cadastral survey.

Even though the place names represent spatial defined area, which means that certain number of cadastral parcels in same area belongs to same place name, cadastral database of AREC doesn’t contain standardized list of place names defined as domains, in order to select the place name for each parcel instead manual typing of place names for each parcel separately. This results with non-standardized place names in cadastral database of AREC, as well differences in letters of place names even between neighbor parcels which belongs to same area - geo name, recognizable in the property certificate of parcels. In table 1 an example of difference between place names (MLAKA and MLAKLA) between two neighbor cadastral parcels in cadastral municipality Gostivar-1 because of manual typing system in AREC is given. Same problem is recognizable in cadastral database in all country territory.

**Table 1.** Example of nonreal difference between place names of two neighbor cadastral parcels

<table>
<thead>
<tr>
<th>Property Certificate</th>
<th>Number/Part</th>
<th>Cadastral Culture</th>
<th>Area m²</th>
<th>Address/Place</th>
<th>Property right</th>
</tr>
</thead>
<tbody>
<tr>
<td>4271</td>
<td>338/1</td>
<td>н Нива Плодните земјишта</td>
<td>2069</td>
<td>МЛАКЛА</td>
<td>-</td>
</tr>
<tr>
<td>91216</td>
<td>342/5</td>
<td>н Нива Плодните земјишта</td>
<td>904</td>
<td>МЛАКА</td>
<td>СОПСТВ ЕНОСТ</td>
</tr>
</tbody>
</table>

Source: https://osspr.katastar.gov.mk/OSSP/, May 2019, cadastral municipality Gostivar-1

In the database of topographic maps, geo names are listed both, as separate feature class, as well as annotation of spatial objects such as rivers, buildings, roads etc (Regulation nr.159, 2013), which have been used as basic data for National Spatial Data Infrastructure (NSDI) geoportal (http://nipp.katastar.gov.mk/geoportal 2019), classified as: place name, hill, peak, town under 10000 population, town 10000 to 25000 population, town over 25000 population, ridge, valley, canyon, mountain, mountain path, mountain range, village under 1000 population, and village over 1000 population. Technical part of publishing is according to INSPIRE data specification on Geographical names – technical guidelines (Tent, 2015). In figure 1 geo names in North Macedonian NSDI geoportal in scale 1:25000 in the region of Shar Mountain is shown, where duplicate geo names such as “Попова Шапка” and “Лисец” (red circles) can be recognized, without any difference on text style, which doesn’t give any information on any thematic difference between them!
Beside spatial databases of AREC, a huge database of geo names owns the Institute of Macedonian language “Krste Misirkov” in Skopje (http://imj.ukim.edu.mk/en.aspx), as public scientific institution under the University “Ss. Cyril and Methodius” – Skopje, which submit collected and analyzed geo names to CHPO to be included and formalized in the national register.

4. RESEARCH ON MICROTOPONYMS IN KARSHIKA AND DERVEN REGIONS

In a case of the situation described above on geo names databases and existing thematic spatial databases for geo names in North Macedonia, as scientific contribute two regions nearby the capital city of the North Macedonia have been the scope of our research which contain two main objectives: establishing spatial database, and linguistic analyses of microtoponyms.

Establishing thematic spatial database for geo names in order to store, standardize, protect, and distribute, is very important issue for responsible institution, as well for us as researchers and stakeholders. Aimed to contribute in developing official spatial database for geo names in North Macedonia, we performed research analyses for two regions of Skopje, by establishing database structure and operational informal spatial database, by having into consideration national Law conditions, INSPIRE specification on geo names, and specific microregional findings within research analyses, to be explained bellow.

Linguistic analysis of all collected and stored microtoponyms is of a special interest for classifying and core usage. Because of various factors that affects to appearing of microtoponyms, linguistic research outcomes aim to reach to the relation of microtoponyms with local language, history, autochthony, natural phenomena, relief forms, man-made objects etc.

4.1 Case study area
Regions of our chase study are located from southern to north-western part of the capital city of North Macedonia Skopje. Both regions, Derven and Karshiaka are neighbor to each other, and belongs to the rural and mountain areas of the Skopje valley.

The Karshiaka region (figure 2) lies in the southwestern part of Skopje valley, which represents a geographical area 764km², with 62 villages. As the administrative areas, all villages belong to five municipalities (Sopishte, Studenicani, Zelenikova, Kisela Voda and Karposh), three of them belongs to the city of Skopje and others are part of Skopje rural municipalities.

The Derven region (figure 2) lies in the northwestern part of the Skopje valley, which represents a geographical area about 295km², with 24 villages. Shown in figure 2, as part of research analyses, beside Derven region villages, 6 villages in eastern part of river Vardar are included in our research also. From 30 researched village areas, 10 belongs to the Tetova geo region, so 20 belongs to Skopje region (Idrizi, 2019). This means that research area of Derven region have been extended with 6 villages, while 10 out of 24 villages of Derven region belongs to Tetova geo region and 14 of 24 villages belongs to Skopje valley.

### 4.2 Data sources

In order to collect field data with specific details concerning to microtoponyms in both study areas, firstly the questionnaire has been compiled with questions on:

- Placename in original pronunciation at the language of local population.
- Placename in second and third language when population is multiethnic.
- Location (spatial position) of placename and spatial identification.
- Link/relation of placename with its potential origin (relief form, spatial phenomena, man-made object, folklore, legend, etc.), based on the knowledge of local population.
- Specific feature of placename.
- Name of village in which is located the placename.
- Name of municipality in which belongs village.
- Date and place of interview.
- Sign of interviewer.
The interview was realized with 71 interviewed people, 41 from Karshaiaka (Idrizi, 2017) and 30 from Derven (Idrizi, 2019) regions. Age over 50 year, the period of livelihood in that village over 30 year after 15 year age, and knowledge for the history, tradition and culture of the village, were main criteria for election of peoples to be interviewed.

During the process of interviewing, 1562 microtoponyms and 30 oiconims for Derven region (Idrizi, 2019), and 2529 microtoponyms and 62 oiconims for Karshiaka region (Idrizi, 2017), in total 4091 microtoponyms and 92 oiconims have been collected. Obtained information on microtoponyms and oiconims from interviewed people, have been stored in our developed database as platform for performing research analyses.

Open Layers used in GIS software as basemap were utilized for spatial identification of microtoponyms, as well storing in database.

4.3 Database structure

Database for storing of all collected data, in a well-organized database structure in order to create precognition for qualitative and quantitative categorization, as well performing spatial analyzes of microtoponyms with other spatial datasets, is established based on the INSPIRE data specification on geographical names shown in figure 3 (INSPIRE, 2014), and extended to specific data categorization based on national classification in North Macedonia (Regulation nr.37, 2006), as well specific findings from performed research (Idrizi, 2017; Idrizi, 2019). Database have been developed by using FOSS software for GIS, such as QGIS, Post GIS, and Geo Server.

The spatial database with 8 (eight) core datasets and domain codes of toponyms are shown in figure 4. Main objective of this database structure is to contribute on developing national database for geo names in North Macedonia, therefore in the list of datasets and domain codes are covered more types of geo names compared with the findings in both study areas. In this line, INSPIRE platform (INSPIRE, 2014) was used as base for our database, in order to reach to fully compatible database with national rules on NSDI and standards that are internationally defined by the European Commission.
Microtoponyms (331)
- Cultivated land (33111)
- Border area (33121)
- Field (33131)
- Pasture (33141)
- Garden (33151)
- Vineyard (33161)
- Forest (33171)
- Grove (33181)
- Restricted area (33191)
- Forest land (331101)
- Other (3311)

Horonyms (335)
- Populated area (33511)
- Unpopulated area (33512)
- Administrative unit (33521)
- Social phenomenon (33811)
- Institution (33821)
- Item and product (33831)
- Road (33711)
- Street (33721)
- Square (33731)
- Quay (33741)
- Highway (33751)
- Other (3376)

Chrematonyms (338)
- Forest land
- Other

Oiconins (332)
- City (33211)
- Village (33221)
- Part of settlement
- Other
- Lake (33411)
- Marsh (33421)
- River (33431)
- Spring (33441)
- Creek (33451)
- Toponym of monument to a person
- Toponym of monument to a story

Oronyms (333)
- Mountain (33311)
- Hill (33321)

Hodonyms (337)
- Road
- Street
- Square
- Quay
- Highway
- Other

Memorial toponyms (336)

Figure 4. Eight datasets and domain codes of established spatial database for geo names

4.4 Linguistic analyzes of microtoponyms

Microtoponyms are at the intersection of spatial language, culture, and cognition, which provide a way to refer to socially inhabited space by naming the places referred to, rather than the objects or people that occur at the places. Presumably, places referred to by toponyms are places that play a marked role in the life of the language community. Microtoponyms of a language community embody a knowledge structure that figures prominently in the spatial conceptualization of the community's environment, as well as the way reference to places is distinguished from reference to objects, animals, or people at places is an important piece in the puzzle of the 'natural language metaphysics' that underlies spatial reference and conceptualization in the language under study (Bohnemeyer, 2001).

The analyses of microtoponyms (oronyms, phytonyms, hydronyms, zoonims, and names of other fields) of both regions have been performed in lexical and semantic terms. The etymology of microtoponyms have been analyzed, i.e. microtoponyms were categorized in Albanian, Slavic, Turkish, Romanian, Greek etc. origins, while lexemes based on their origin, which represents the source for formation of the microtoponym construction in Karshiaka and Derven regions.

As a second element, the study of microtoponyms from the structural-semantic aspect were taken, as well as the morphological and syntactic elements that are used in the construction of microtoponyms in both regions have been analyzed. According to these criteria are specified microtoponyms that are first words (impersonal names, personal names, simple adjectives, verbs and adverbs), and microtoponyms that are derived from prefixes, name and surname suffixes, as well combined cases with prefix and suffix in same time. There are also a huge number of microtoponyms constructed by word conglomeration and phrases with two or more parts.

The third aspect is related to the basic linguistic elements of Karshiaka and Derven regions. The phonetic, morphological, and lexical features of the denominations that are characteristic of the folk in both regions have been analyzed. Labels, respectively microtoponyms can not contain as many lingual features as the speech content. Therefore, research of this issue has
been performed to the extent that microtoponyms of these regions concerned is enabled. In the field of phonetics have been analyzed emphasis, vocalism and consonantizmus, while in the field of morphology the word construction and grammatical names categories were analyzed. The inclusion of joints, adjectives, numbers, verbs and adverbs in the construction of microtoponyms in both regions have also been reviewed.

In the microtoponymic system of the Karshiaka and Derven regions, we have examined the lexical layers sourced by the Albanian, Slavic, Turkish, and other Balkan languages.

In the end of performed research linguistic analyses in this study, as main conclusion is that microtoponyms in both regions are sprang from Albanian, Slavic, Turkish, and very few other Balkan languages.

4.5 Outputs

As main research outputs we got:

- operational thematic GIS spatial database for geo names based on INSPIRE platform for geographical names, national classification of cultural heritage, and microregional data of study area;
- extendable operational thematic GIS spatial database ready to be used as basement for developing national database for geo names in North Macedonia;
- 4091 microtoponyms defined with their location and attribute table;
- 92 oiconims defined with the area of population centers and attribute table;
- linguistic analyses for each microtoponym and oiconim separately;
- finding the origin and relation of microtoponyms and oiconims with natural/social phenomena nearby their surrounding;
- categorization and classification of microtoponyms based on results from performed research and including them in database;
- categorization and classification of oiconims based on results from performed research and including them in database;
- relation of analyzed microtoponyms and oiconims with languages that are currently in use by local population, as well languages that have been used in past centuries in study area, and
- two maps for each study area.

5. DISCUSSION

Research outputs can be classified in technical (operational GIS spatial database), formal (combined integration of national and continental standards, with findings from microregional data), linguistic (performed linguistic research analyses on microtoponyms and oiconims up to their relation with language and history/etymology), cartographic (compiled maps for both study regions), and publishing (two published books, per one for each study area). Above mentioned outputs represent independent and original results from performed research by both authors of this paper.

Compared with given information in above chapters for the current situation in North Macedonian institutions and legislation, it is so clear that presented research methodology and outputs in this paper are original input of authors in the field of spiritual cultural heritage by using latest GIS technology for collecting, storing, analyzing, archiving, distributing, using, and standardizing geo names. The GIS spatial database is open, extendable and fully interoperable, which can contribute on improvement of current standardization and technical systems in North Macedonia. Contribution to further development of national institutions in a
field of toponymy, by including latest technology as SDI, GIS and remote sensing, was one of our main objectives before powering the idea for research presented in this paper.

6. CONCLUSIONS

Microtoponyms are the names of small features such as fields, pastures, fences, stones, rocks, marshes, bogs, ditches, houses, sections of roads, localities etc. (Nash, 2013). Microtoponyms are not often officially recognised or gazetted, and do not generally appear on any published maps. A microtoponym is in essence the name of a feature that is itself part of a larger named entity (Tent, 2014). There are two basic ways to conduct toponymic research — one concentrating on the etymology, meaning, and origin of toponyms, and one focusing on the toponyms of a region and examining patterns of these names (Tent, 2015).

Currently in North Macedonia exist many spatial databases which contain geo names, collected in different (standardized and non-standardized) ways and periods by many institutions and independent researchers. As main formal actors in this field are Cultural heritage protection office (CHPO) and Agency for Real Estate Cadastre (AREC), as well research Institute of Macedonian language “Krste Misirkov” in Skopje. Other institutions and independent researchers perform analyses for microtoponyms in different regions of North Macedonia, however those are not considering as relevant data for including in national register of cultural heritage by responsible institutions, i.e. they are left as informal data published in books and/or papers.

Chase study areas of performed research are Derven and Karshiaka regions, two neighbor regions which belongs to the rural and mountain areas of the Skopje valley. 4091 microtoponyms and 92 oiconims collected from both regions are stored in developed thematic GIS spatial database, as well as analyzed from a linguistic point of view, which resulted with categorized and classified dictionary and spatial database of microtoponyms and oiconims. From performed linguistic complex analyses for each microtoponym and oiconim, as main conclusion for both regions is that microtoponyms are sprang from Albanian, Slavic, Turkish, and very few other Balkan languages.

Developed thematic GIS spatial database in our research can be used as template or basic technical platform for improvement of current databases in North Macedonia toward establishing central spatial database for all geo names in national level by including all level of geo names in all scales from local up to national level.

Developed database is also open for all interested institutions, organizations, researchers and others, in order to share our findings and established technical platform.

REFERENCES


Gundacker R. (2017). The significance of foreign toponyms and ethnonyms in old kingdom text sources. The University of Chicago. USA.


