

Mapping, synthesis and visualization of Czech dialects

Vit Vozenilek^{a,*}, Martina Ireinova^b, Alena Vondrakova^a, Jakub Konicek^a

^a Palacky University Olomouc, vit.vozenilek@upol.cz, alena.vondrakova@upol.cz, jakub.konicek@upol.cz

^b Czech Language Institute of the Czech Academy of Sciences, ireinova@ujc.cas.cz

* Corresponding author

Abstract: Dialects became a traditional research subject of linguists who collect and analyse data from a linguistic perspective and a spatial perspective. Detailed research on the distribution of dialects and their specifics is an important part of preserving every nation's cultural heritage. With the gradual disappearance of dialects in society, this topic is very relevant. The presented paper briefly describes history of dialects research and the most important outcome in recent years, the Czech Linguistic Atlas (CLA). CLA presents the data only in a very simple form from a cartographic point of view, so a new approach has been applied. A new atlas was created from this data, focusing on a specific area of dialectology – vowels shortening. This new atlas' internal structure corresponds to the cartographic atlas theory as a system and contains both analytical and synthetic maps. Synthetic maps show areas with similar (and unique in the case of regionalisation) combinations of evaluated attributes. The concept of map synthesis is described, and further research questions are designed.

Keywords: language mapping, dialect maps, cartographic synthesis, systematic cartography, linguistic atlas

1. Introduction

Language belongs to the history of the nation, to its cultural heritage, it reflects the conditions in which our ancestors lived, as they thought and perceived the world around them. The traditional territorial dialects of the language are gradually disappearing, dialect differences are fading, dialect words are dying. More systematic records of the gradual extinction of dialects have been emerging in the Czech Territory since the early 19th century. Some researchers even believed that dialects would disappear during the 19th century.

In terms of speech production in the traditional domain of dialectology, researchers have long utilized mapping techniques to isolate the use of a feature or form in geographic space. In doing so, dialectology has brought to light much about the way such features interact with social and geographic barriers and, in wave or gravity models, about the way language change spreads across space and time (Chambers, Trudgill, 1998). Recent work in understanding regional variation has increasingly focused on quantitative and statistical approaches to the analysis and mapping of regional forms.

The advanced social research includes a map and multimedia presentation of selected dialect phenomena and a complex of publicly available tools for the acquisition and managing of dialect data. However, dialect mapping is still hard to understand for non-linguistics and researchers have not explored whether the same techniques that have been used on production data by geolinguists so successfully might be as useful for understanding regional variation in perception.

The aim of the paper is to present the results of the collaborative research of dialectologists and cartographers

on the systematic visualization of the results of the spatial synthesis of the dialects of the Czech national language.

2. Spatial research of dialects

Dialectology is a scientific discipline that studies dialects. Dialect is defined as a traditional territorial unit of the national language determined in a geographical area. Each dialect is formed by relation to another dialect and to the national language (Kloferová, 2017a). In dialectology, the so-called linguistic geography is applied, which deals with the territorial differentiation of linguistic phenomena, i.e., the research of the differences between the studied dialects and their spatial distribution. The result is linguistic maps and linguistic atlases as a comprehensive statement about the dialect situation in the area (Kloferová, 2017b). These are being developed all over the world with a range of mappings from dialects to language families.

2.1 Language mapping

The depiction of the linguistic issues of cartographic products has its history. Many diverse language mapping outputs have emerged, but relatively few deals with dialects (Luebbering, 2013).

Linguistic geography (also areal linguistics or geolinguistics) deals with the spatial differentiation of linguistic phenomena. It analyses and then visualizes language data (Sousa, 2017) in maps using specific geolinguistic methods (Auer, Schmidt, 2010; Chambers, Trudgill, 2014; Boberg, Nerbonne, Watt et al., 2018). Utěšený (1972) discussed the problems of symbols in linguistic cartography.

Cartographic synthesis, typification, and regionalization are also used in linguistics and dialectology (Voženílek et al. 2011; Seabra et al., 2014; Georgakopoulos, 2019).

Mapping of linguistic phenomena builds on fieldwork (Luebbering, 2013). Language is highly variable in space, especially today, when people travel from place to place due to the ability to travel freely, higher education, cheap transport costs, and moving from the countryside to the cities, making it much more difficult for field researchers to work (Upton, Widdowson, 2006). All this significantly reduces the topicality of each language map, when the distribution of languages and linguistic phenomena can change fundamentally at the time of map release.

2.2 Mapping of Czech dialects

Systematic research of the Czech dialects has started after the Second World War. Until this time, partial research has been carried out in the Czech territory, which, however, served as a basis for further work. The study *The Foundations of the Czechoslovak Dialectology* by A. V. Šembera from 1864 is considered to be the founding paper of Czech dialectology. In 1894, V. J. Dušek published a *Dialectical Map of Bohemia*, which already had a geolinguistic character (Balhar et al., 2018).

Systematic geolinguistic research of Czech dialects began in 1947 at the Czech Language Institute. Surveys carried out first by the correspondence method served dialectologists as an initial overview of the initial state of the Czech dialect. Questionnaires were sent to each municipality with a school (Kloferová, 2007). Collected linguistic data were processed and the results came into considering the basic topics for linguistic atlas compilation. Before the field survey itself, the basic issues for the compilation of a linguistic atlas were addressed, in particular the method of field research, the establishment of a network of research locations and the compilation of a special questionnaire that would cover differences in all levels of the language system. The aim was to compile a *Czech linguistic atlas* (Balhar, et al., 2018).

Even after the completion of the *Czech Linguistic Atlas* in 2011, dialectological research is still ongoing in the Czech Republic, mainly trends of traditional territorial dialects and urban and borderland speech (Ireimová, Konečná, 2015, 2016).

For spatial research of dialects, three groups of geolinguistic methods are used in Czech linguistic geography (Kloferová, 2017d):

- methods of linguistic data collection:
 - sound recording of spontaneous speech, incl. text transcription and exploratory methods – they use specialized questionnaires filled in by a dialectologist (so-called explorer) on the basis of an interview with a dialect speaker (so-called informant),
 - sending questionnaires to non-dialectological explorers (e.g. teachers) who return questionnaires to dialectologists after the research,
- cartographic geolinguistic methods – used to display linguistic phenomena on maps with a

network of research locations; research results are recorded in three ways:

- the use text inscriptions (traditional method),
 - the use of special point (geometric) map symbols,
 - the use of areal map symbols for defining areas with the occurrence of a given linguistic phenomenon using hatches or isoglosses (lines delimiting the area of linguistic phenomenon occurrence).
- interpretive geolinguistic methods – based on the expression of a linguistic phenomenon on a map; the basis is the study of the spatial distribution of phenomes, morphological and word-formation prefixes, suffixes, etc., or the search for motivations for naming and creating words; the prerequisite is to take into account linguistic, historical, ethnographic and other contexts.

3. Cartography in linguistics

Over time, several basic methods have been established and used in linguistic atlases (Luebbering, 2013). Currently, with the upcoming of new knowledge, technologies and visual styles, advanced methods of improving and supplementing those established methods are emerging (Luebbering, 2013, Siebler et al., 2012). Using advanced methods, cartographers try to solve three basic problems of creating linguistic maps defined by Siebler et al. (2012):

- individual linguistic phenomena tend to occur in different forms in the same place, which must be expressed cartographically so that the outputs correspond to the basic cartographic rules,
- converting data showing the occurrence of linguistic research objectives from the point elements in which, in the vast majority of cases such data are collected into a flat format; with the advent of GIS and their constant improvement, new possibilities for solving this problem are constantly opening up,
- detachment of sharp outlines between areas with different variants; sharp outline does not capture reality if the regions are strongly interconnected.

The basic visualization procedure of most linguistic maps and atlases came from the origins of linguistics and dialectology of Germany and France (Kretzschmar in Boberg, Nerbonne, Watt et al., 2018). Point map symbols are used as a basic means of expression in linguistics, especially for depicting research locations. Furthermore, the method of areas is applied, which is based on point layers, where the points are interpolated into fields and then visualized as areas. The outlines between the areas are called isoglosses (Kloferová, 2017c).

Dialectometry examines regional patterns in the perception of dialect maps can gain new insights into the perception of language forms of evaluation of geospatial patterns of responses of participants in experiments. It is an attempt to find out whether clusters are in accordance

with regionally defining isoglosses on maps (Kendall, Fridland 2016).

4. Linguistic atlases

Linguistic maps are thematic maps, the content of which forms the language situation in a given area. Kretzschmar (in Boberg, Nerbonne, Watt et al., 2018) defines linguistic atlases as sets of maps that capture the (spatial) distribution of different types of speech and calls them “dialect atlases”. The procedure for creating a linguistic atlas was described by Lee Pederson (in Boberg, Nerbonne, Watt et al., 2018), who compiled a list of seven steps in the process of a linguistic atlas compilation:

- intention to create a linguistic atlas of the required territory,
- establishing a square network over the study area,
- creation of a list of communities or settlements within the network,
- managing the questionnaire regarding the research topic, i.e. if the atlas deals with the lexical or grammatical component of the language,
- selection of informants for each community,
- fieldwork based on interviewing informants,
- recording the answers of informants and further processing.

However, as Kretzschmar (in Boberg, Nerbonne, Watt et al., 2018) argues, this procedure is not applicable in all cases of linguistic atlas compilation, for example for indirect linguistic research.

Georg Wenker studied German dialects (Kretzschmar in Boberg, Nerbonne, Watt et al., 2018) and is considered to be the first widely recognized linguist to attempt creating a work that meets the definition of a linguistic atlas. He focused his research on Westphalia, where he used an indirect exploratory method for his research. He sent questionnaires to teachers them to translate the prescribed sentences from general German into the local dialect.

In Central Europe, more linguistic atlases were compiled, most in Germany. Due to the high linguistic diversity in Germany, a large number of regional studies have been completed, such as the *Sprachatlas von Bayerisch-Schwaben*, the last version of which with 14 volumes was published in 2009 (Universität Augsburg, 2019).

One of the last all-German works is the *dtv-Atlas Deutsche Sprache* (Fig. 1) edited by Werner König from 2015. The atlas contains 155 maps showing the territorial distribution of selected words in Germany and Austria. The atlas focuses on German dialects and modern standard German (König et al., 2019). In the maps, the colour areas show the spatial differentiation of the observed phenomena.

Another German linguistic study is the so-called REDE, which is a long-term project of the German Academy of Sciences and Arts focused on the research of German dialects. Part of this project is, among other things, a relatively large database of linguistic atlases of German

dialects. Another project output is a GIS tool “SprachGIS” (REDE, 2019).



Figure 1. Map of the German *dtv-Atlas Deutsche Sprache* (König et al., 2019).

France is also a pioneering country in linguistic atlas production. French linguists have developed a specific style of language research as a counterweight to the German way. The main difference between the French and German approach is in field explorers. At the first nationwide French *Linguistique de la France* (Gilléron, Edmont, 1904), only one researcher worked directly in the field, by one of the authors, Edmond Edmont. On his bicycle, he mapped the state of linguistic phenomena in 639 places in France and examined almost 2,000 words and phrases (Kretzschmar in Boberg, Nerbonne, Watt et al., 2018). Figure 2 shows the use of label maps to display the mapped phenomenon, i.e. a transcript of a given form of the phenomenon for each research location.

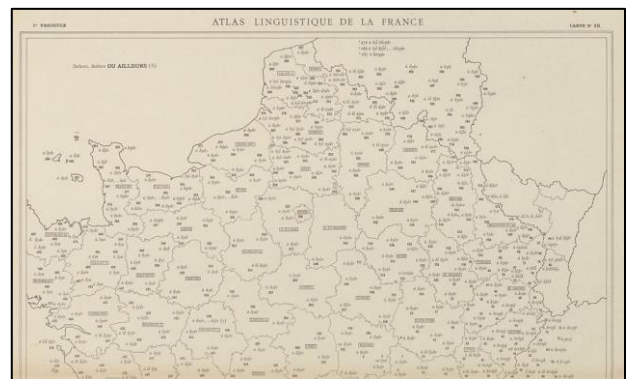


Figure 2. Map of the French *Atlas Linguistique de la France* (source: <http://diglib.uibk.ac.at/ulbtirol/content/zoom/149056>).

The *Atlas of the Slovak Language* consists of four parts (1968–1984). Preparatory work for the atlas began in 1942 (Balhar, 1971). Subsequent research was divided into two stages. The first (1946–1951) used correspondence questionnaires, the second (1964–1970) direct field mapping in 334 municipalities (Hladká, 1996). Maps are two-coloured to distinguish between thematic content and topographic background (Fig. 3).

Another active country in linguistic geography is Poland. The first *Polish regional linguistic Atlas* was issued in

1934 in preparation for the national linguistic atlas of Poland planned since 1939. The nationwide *Mały atlas gwar polskich* in 13 volumes was completed in 1970. Many regional linguistic atlases were being created in Poland. However, they were not coordinated in any way, so the individual results cannot be easily compared (Hladká, 1996).

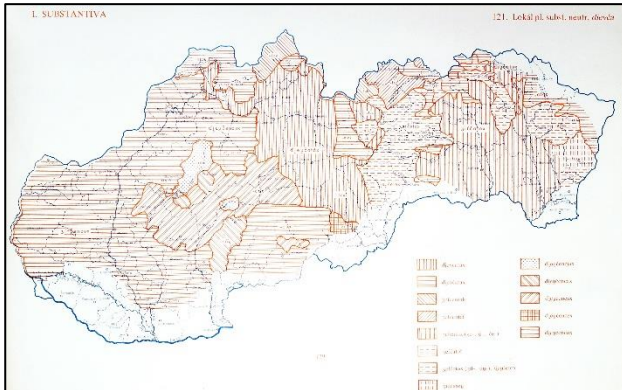


Figure 3. Map from the *Atlas of the Slovak Language II* (Štolc, 1981).

The *Atlas of Bulgarian Dialects* (Fig. 4) (BAN, 1988) consists of eight volumes, which began in the 1950s and were published in 1964–2001. However, the individual parts do not cover the whole territory of Bulgaria, but only designated areas of the country where all linguistic aspects were mapped.

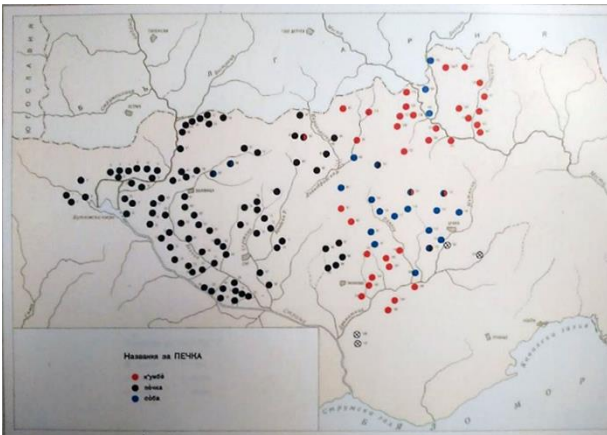


Figure 4. Map from the *Atlas of Bulgarian Dialects* (BAN, 1988)

Other linguistic atlases were released, for example in Slovenia (Škofic et al., 2016), in Macedonia (Vidoeski, 2008), in Croatia (Lončarič, 1997) and in Russia (Avanyesov, Bromley, 1986–1996).

One of the most important works outside continental Europe is *The Linguistic Atlas of England* (Orton et al., 1978). The first intention of the atlas was presented by the authors in 1946 with the work itself published in 1978. Total 313 locations in England were mapped through a field questionnaire (Orton et al., 1978). The atlas contains four types of maps – phonological, lexical, morphological and syntactic. Linguistic phenomena are displayed in the maps by isoglosses and supplemented by point symbols (Orton et al., 1978).

Linguistic atlases are created at national and also multinational levels. Multinational atlases provide searching for relationships between individual languages or language families (Hladká, 1996). The Slavic Linguistic Atlas is the result of the international cooperation of universities and academic institutions of all Slavic nations to map the state of dialect phenomena of Slavic languages. The cooperation began in 1965 and by then 17 phonetic and lexical volumes had been published in several Slavic countries. Another important multinational project produced the *Linguistic Atlas of Europe*. Since 1970, the project runs under the supervision of UNESCO (ILIR, 2019). The aim is to give a complete picture of the linguistic differentiation of traditional dialects of Europe (Jančák, Petr, 1986).

5. Czech Linguistic Atlas

The *Czech Linguistic Atlas* (hereinafter *CLA*) provides the first complete view on the territorial distribution of the Czech language in the Czech Republic. *CLA* divides the territory into smaller and larger dialect or regional units and finds relationships between them.

The *CLA* consists of six volumes, which were published in the years 1992 to 2011. In 2012–2016, the *CLA* was supplemented, corrected and published electronically in PDF version (<https://CJA.ujc.cas.cz/e-CJA/>).

The *CLA* is based on the results of field research conducted in 420 inland rural localities from 1964 to 1972 and in 57 cities from 1973 to 1976. Total 4,364 informants from rural and 1,032 from urban localities took part in the research. The age of the informants ranged from 65 to 75 years, in the cities, they were supplemented by the younger generation at the age of 14 and 15 years. In the border areas (the expulsion of the German-speaking population in 1946), the research was carried out only with the young urban generation, who were already born here and formed their linguistic usage. The research also included data obtained from Czech language enclaves in Poland, Romania and the former Yugoslavia.

The research was conducted on the basis of the Questionnaire for the Research of Czech Dialects. It contains 2649 items and was designed to represent all language forms.

The content of the first three volumes of the *CLA* is the dialect lexical material (the thematic categories: garden, landscape, man or agricultural tools). The fourth volume focuses on dialect morphological phenomena, i.e. differences in grammatical forms of flexible words and is divided according to a part of speech. The fifth volume consists of several parts, the most comprehensive of which deals with phonological phenomena. Unlike other language atlases, the *CLA* also deals with urban speech. An important part of the atlas are the so-called bundles of isoglosses, thanks to which it is possible to determine new knowledge about the relationships between the studied phenomena (Kloferová, 2003). The last volume (so-called Supplements) summarizes the entire research and contains characteristics of the studied municipalities, examples of

dialectal speech in text and audio format or a complete bibliography of Czech dialectology since 1968.

The *CLA* contains almost 1,600 maps and 25,000 language documents. There are four types of dialect maps:

- basic full-page maps of the whole territory,
- half-page maps of the entire territory
- half-page map sections with the Moravian and Silesian situation,
- so-called conducting isoglossic maps, on which two-membered, sharply defined opposites created by the identified equivalents are drawn by isoglosses (Balhar et al., 2018).

Comments supplement all maps the task of which is to provide a brief explanation of the map, to draw attention to the basic mapped phenomena and to incorporate the observed phenomenon into a broader context within Czech and west Slavic languages. For the representation of the mapped phenomena, the area division of the dialect vocabulary was applied, where the individual variants are aggregated from point elements into areas separated by an isogloss. Each area is accompanied by the labelling of the variant. The method is supplemented by point map symbols that express the unsystematic occurrence of the mapped variant.

The uniqueness of the *CLA* lies in several aspects. One of them is the high reliability of the material, which has been collected by field explorers. As one of the few national atlases deals with the dialect situation in Czech emigrant settlements in Europe (Kloferová, 2012).

6. Systematic visualization in spatial synthesis

GIScience has so far rarely dealt with linguistic data, despite its ability to analyse the variability of language in space. Two commonly used paradigms in dialectology for solving dialectal areas in space – the dialect continuum and isogloss – use the dichotomy of fields vs permissions found in GIScience (Voženílek, Vondráková, 2017). These two ways of conceptualizing dialectal boundaries raise the problems faced by quantitative analysis of the definition of language areas, incl. spatial synthesis.

Systematic visualization of spatial synthesis of dialectological data concluded *Atlas of the Dialects the Czech Language* focused on shortening of vowels.

6.1 The first *Atlas of the Dialects of the Czech Language: shortening of vowels*

The aim of the atlas is to show the extension of phonetic variants of selected words in a new cartographic style and to present a synthesis consisting of typification and regionalization. The atlas processed 16 words from the 5th volume of the Czech Linguistic Atlas (Balhar et al. 2016), and 12 selected ones were synthesized. Dialect maps from *CLA* were vectorised and visualized into 16 analytical dialect maps in a new map symbology. The maps of research locations were compiled for each analytical map. The atlas is divided into two sections. The first section contains 16 analytical dialect maps created by cartographic reworking of maps from the *CLA* and supplemented by

other infographic elements. In the first part of the atlas, 16 words are processed, each word is on two pages: on the left side is the main map sheet with the results of dialectological research, on the right side is a map of research locations together with their numerical list. If the so-called regular regional changes are valid for a given word, they are drawn in the map of locations.

The second section deals with the cartographic synthesis, from which two synthetic dialect maps were compiled – a map of types and a map of regions of shortening of vowels. Cartographic and graphic representation of the Czech dialects in the *Atlas of the Dialects of the Czech Language: shortening of vowels* consists of the following elements (see Fig. 5):

A: The main analytical map shows at a map scale of 1:1,400,000 the results of research into phonetic variants of a word using coloured areas outlined by isoglosses. Complementary elements are point map symbols of overlap of phonetic variants in the location. These elements are divided into:

- doublet – when a variant marked with a point symbol and a variant expressed by the colour of the area were recorded,
- underdeveloped – when only the variant expressed by a point map symbol was recorded.

The areas are labelled by phonetic variants. In some maps, there are filled hatches that show phonetic variants in a case other than the 1st case.

B: The map legend consists of three parts:

- extension of phonological variants, in which the 1st case is displayed, expressed by coloured areas and point symbols,
- some maps are supplemented with symbols for another case, which is displayed using a hatch,
- phonological differences expressed using isoglosses, symbols for locations with a documented variant.

The map legend also applies to the maps of the 1st and supplementary fall.

C: The dendrogram displays selected research locations to represent dialect types or subgroups. The locations are sorted according to the dialect areas.

D: The map of dialect areas at a scale of 1:7,000,000 shows the dialect areas of the Czech language and is used to compare the results of mapping the solved word in the 1st case with the definition of dialect areas.

E: The data visualization of the share of the population using phonetic variants is expressed by figures representing 1% of the population. Variants are expressed using colour in the map legend. Only variants in the 1st case are monitored in the diagram. Population data are based on the 2011 Census.

F: Original map from the *Czech Linguistic Atlas* is included for comparison with the new cartographic style of the main map.

Some pages have the map of the additional case at a scale of 1:5,500,000 with the results of the research of the

examined word in a case other than the 1st case expressed by hatching. The map contains symbols from the main map legend.

The map of research localities at a scale of 1:5,500,000 contains map symbols for research locations in which a variant of the solved word was recorded and a map symbol for border areas without continuous dialect research.

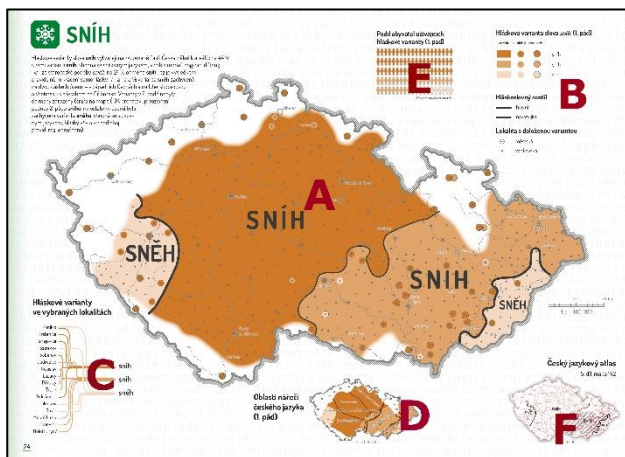


Figure 5. Map layout of the main page in the *Atlas of the Dialects of Czech Language: shortening of vowels* (Ireinová et al. 2020).

6.2 Typology for shortening of vowels

Typification is a kind of spatial synthesis, a scientific method based on the idea that each element in space has its own unique properties. Typification classifies spatial elements into groups according to the similarity of their selected properties. The elements in the groups can be characterized as similar, but not the same. The results of the typification are defined as types and arranged in typology. Besides cartography, typification is used in all scientific disciplines, especially in geography, architecture, healthcare, biology and chemistry (Voženílek, et al., 2011).

Croft (1990) defines a type as “a particular structural feature associated with a particular construction in a particular language”. The process of typification according to the shortening of vowels in Czech dialects took place in six successive steps, which were implemented in the ArcGIS Pro. The first step is to select 16 analytical dialect maps from the *CLA* to define the types. Only 12 analytical dialect maps included for defining types and compiling a typology.

The second step of typification reclassified the maps according to the relationship of phonetic variants to the quantity of vowels. The layers of 12 selected analytical dialect maps were reclassified into two classes – short variants and long variants. For this purpose, a new SHORT attribute was created in each layer, expressing the quantity of the voice. The attribute was created in the numeric data type: +1 for the long variant, -1 for the short variant. According to the new attribute SHORT, individual polygons with the same attribute value were connected. The result was a set of 12 new data layers with the division of the territory into areas with a long variant and a short variant of each word.

In the third step, an overlay analysis of 12 newly created datasets was performed. The output was a data layer composed of intersections of polygons of all 12 input layers. The layer contained 541 polygons with a unique sequence of attributes expressing the representation of long and short variants of selected words.

A large number of small polygons (even in the order of km² units) were generalized in the fourth step. Small polygons do not affect the definition of types. The generalization removed polygons with a minimum size of 150 km² by connecting to the adjacent polygon with the longest boundary. The resulting generalized layer consisted of 54 polygons with 12 attributes specifying long or short phonetic variants of each word included in the typification.

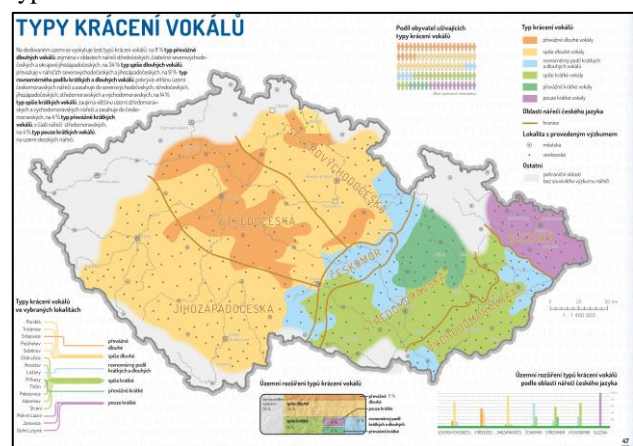


Figure 6. Typology of the Czech dialects as a result of spatial synthesis of dialects data (Ireinová et al. 2020).

In the fifth step, the types of shortening of vowels were defined by quantifying the long and short variants in each polygon (Fig. 6). The sum of the values of the attributes +1 and -1 made it possible to define the type of shortening. In the range from -12 to +10, six intervals were defined as the resulting types. The final, sixth step was to compile a typology of vowel shortening. The default type is the so-called “equal proportion of short and long vowels” (sum of variants -2 to +2). Other types were called “rather short vowels” (sum of -3 to -7) and “rather long vowels” (+3 to +7), followed by “mostly short vowels” (-10 to -8) and “mostly long vowels” (+8 to +10). In addition, the type of “short vowels only” has been defined for short variants, the sum of which is lower than -10.

Cartographic visualization includes representing the types using qualitative colours areas, with the addition of existing boundaries between Czech language dialects areas (chapter 6.1, element D) for possible mutual comparison. The map layout corresponds to the previous chapter's description (Fig. 5), with a few deviations, such as a chart with the representation of types within Czech language dialects groups.

6.3 Regionalization for shortening of vowels

Regionalization is a way of spatial synthesis, a scientific method of defining regions. Regions are unique areas with a characteristic set of selected features. Geographical

names are most often used to name regions. Regions do not repeat, while types do. The result of the regionalization process is a territorially defined region organized into regionalization (Voženilek et al., 2011). Emeneau (1956) defines a linguistic region as an area which includes languages belonging to more than one family but showing traits in common which are found not to belong to the other members of (at least) one of the families.

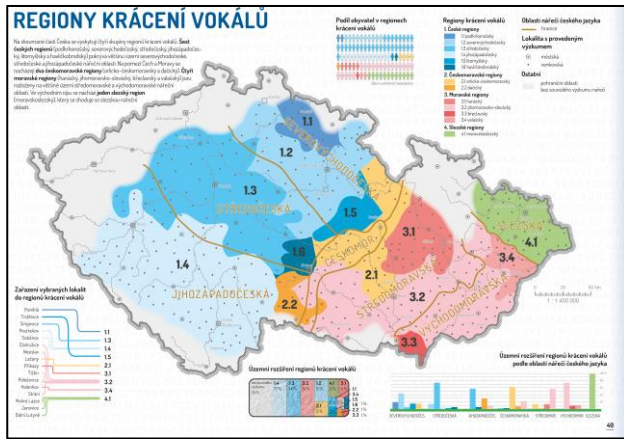


Figure 7. Map of regions of the Czech dialects as a result of spatial synthesis of dialects data (Ireimová et al. 2020).

Regionalization describes the shortening of vowels on 12 analytical dialect maps. Regionalization of shortening of vowels was conducted by cartographic synthesis. Thanks to the uniqueness of the input data, a visual analysis approach was chosen for performing regionalization, the basis of which is the finding of cores in the studied area and the subsequent depiction of the course of regional boundaries. The input layer for the creation of regionalization was a map of types created from 12 analytical dialect maps before generalization. When defining unique areas and their cores, several aspects were taken into account, primarily the occurrence of types in a given area and their spatial interconnectedness, as well as the geographical characteristics of the area.

Regionalization created 13 regions of shortening of vowels in four groups (Fig. 7) according to the geographical distribution into the Czech region (6 types), the Czech-Moravian (2), the Moravian (4) and the Silesian (1). The map corresponds to the map of types.

7. Conclusions

The *Atlas of the Dialects of the Czech Language: shortening of vowels* is to acquaint readers with the issue of shortening of vowels in a new way of working with dialect research data in the Czech Republic and a different form of their visualization. The atlas is unique with its visual style created through analytical dialect maps with a newly compiled character key. It was created for 16 dialect maps from the *Czech Linguistic Atlas* selected on the basis of several criteria, with the main criterion being the shortening of vowels in the Czech dialects.

Dialect maps have become an important material with high information value for the study of mutual relations, especially between the development of language and

society. They can contribute to understanding the social and cultural development of national society.

Atlas with its new knowledge contained in synthetic maps will serve dialectologists not only as an example of alternative visualization of dialectological data to the *CLA*, but also as an analytical document dealing with the shortening of vowels in the dialects of the Czech language. The results of cartographic synthesis are proof of the successful and beneficial approach of dialectological and cartographic research on the one hand and the academic community and the general public on the other. A similar connection between dialectologists and cartographers can be used in solving other dialectological phenomena.

Further research could continue in perceptual dialectology, which studies folk linguistic beliefs (Niedzielski, Preston 1999) and correlates the apparent attitudes and beliefs of speakers with specific places or regions on the map. The authors can expect such research to find out if students' negative attitudes tend to have geographical correlates, especially in areas where stereotyped dialects are thought to be spoken regularly.

Some visual patterns in the maps suggest finer regional patterns of perception than those previously worked on. It should be noted that these patterns are only suggestive. To assess the extent to which these clusters differ significantly from each other, further cartographic research is needed, incl. cartographic synthesis.

8. Acknowledgements

This paper was created on the basis of the project No. DG20P02OVV029 „Czech dialects in an interactive view. Documentation and ways to the accessibility of the disappearing linguistic heritage as an integral part of regional identities“; support provider: Ministry of Culture of the Czech Republic, NAKI II Program. The authors thank the student Marek Pospíšil for his great contribution to data and visualization procedures.

9. References

- Auer, P., Schmidt, J., E. et al. (2010). *Language and Space: An International Handbook of Linguistic Variation*, Vol. 1: Theories and Methods. Wustermark: Walter de Gruyter.
- Avanesov, R. and Bromley, S.V., (eds.) (1989). *Dialektologičeskij atlas russkogo jazyka: Centr evropejskoj časti SSSR*. Moscow: Nauka.
- Balhar, J. (1971). *Atlas slovenského jazyka*, díl I., sv. 1-2. In: SPFFBU. Brno, Vol. 20, Iss. A19.
- Balhar, J. et al. (2016). *Český jazykový atlas 5*. Praha: Academia.
- Balhar, J. et al. (2018). *Český jazykový atlas: O atlasu*. [online]. https://cja.ujc.cas.cz/e-cja/o_slovníku.
- BAS. (1988). *Atlas of Bulgarian Dialects*. Sofia: BAS Academic Press.
- Boberg, Ch., Nerbonne, J., Watt, D. et al. (2018). *The Handbook of Dialectology*. Hoboken: J. Wiley & Sons.
- Chambers, J.K. and Trudgil, P. (2004). *Dialectology*. Cambridge: Cambridge University Press.

- Croft, W. (1990). *Typology and Universals*. Cambridge: Cambridge University Press.
- König, W. et al. (2019). *dtv-Atlas Deutsche Sprache*. München: Verlagsgesellschaft.
- Emeneau, M. (1956). *India as a linguistic area*. Republished in: Anwar S. Dil (ed.), 1980. *Language and Linguistic Area: Essays by Murray B. Emeneau*. Stanford: Stanford University Press.
- Georgakopoulos, T. (2019). *Semantic maps* [online]. https://www.researchgate.net/publication/330399810_Semantic_Maps.
- Gilliéron, J. and Edmont, E. (1904). *Atlas linguistique de la France*. Paris: Honore Champion.
- Hladká, Z. (1996). *Atlasy slovanských jazyků*. In: SPFFBU. Brno, Vol. 45, Iss. C43.
- ILIR (2019). *Atlas Linguarum Europae* [online]. <http://www.lingv.ro/ALE.html/>.
- Ireimová, M. and Konečná, H. (2015). *Nářečí dříve a nyní: Český jazykový atlas*. Praha: Středisko společných činností AV ČR.
- Ireimová, M. et al. (2020). *Atlas nářečí českého jazyka: krácení vokálů*. Olomouc, Univerzita Palackého v Olomouci.
- Jančák, P. and Petr, J. (1986). *Evropský jazykový atlas jako nástroj multilingvální komparistiky*. *Slovo a slovesnost*, Vol. 47, Iss. 3.
- Kendall, T. and Fridland, V. (2016). *Mapping the perception of linguistic form: Dialectometry with perceptual data*. In: Côté, M.-H., Knooihuizen, R. and Nerbonne, J. (eds.). *The future of dialects*. Berlin: Language Science Press.
- Kloferová, S. (2003). *Nad svazky izoglos v Českém jazykovém atlase*. *Naše řeč*, Vol. 86, Iss. 1.
- Kloferová, S. (2007). *Dialektologie*. In: Pleskalová, Jana; Krčmová, M., Večerka, R., Karlík, P. (eds.). *Kapitoly z dějin české jazykovědné bohemistiky*. Praha: Academia.
- Kloferová, S. (2012). *Český jazykový atlas a čeština*. In: Čmejrková, S., Hoffmanová, J., Klímová, J. (eds.). *Čeština v pohledu synchronním a diachronním: Stoleté kořeny Ústavu pro jazyk český*. Praha: Karolinum.
- Kloferová, S. (2017a). *Dialekt*. In: Karlík, P., Nekula, M. and Pleskalová, J. (eds.). *Czech Ency: Nový encyklopedický slovník češtiny*. Brno: Masarykova univerzita.
- Kloferová, S. (2017b). *Jazykový zeměpis*. In: Karlík, P., Nekula, M. and Pleskalová, J. (eds.). *Czech Ency: Nový encyklopedický slovník češtiny*. Brno: Masarykova univerzita.
- Kloferová, S. (2017c). *Izoglosa*. *CzechEncy: Nový encyklopedický slovník češtiny*. Brno: Masarykova univerzita.
- Kloferová, S. (2017d). *Jazykovězeměpisná metoda*. In: Karlík, P., Nekula, M. and Pleskalová, J. (eds.). *Czech Ency: Nový encyklopedický slovník češtiny*. Brno: Masarykova univerzita.
- Lončarič, M. (1997). *Hrvatskijezični atlas*. Zagreb.
- Luebbing, C. (2013). *Displaying the geography of language: the cartography of language maps*. *The Linguistics Journal*, Vol. 1, Iss. 7.
- Niedzielski, N. and Preston, D. (1999). *Folk linguistics*. Berlin: Mouton de Gruyter.
- Orton, H., Sanderson, S. and Widdowson, J. (eds.) (1978). *The Linguistic Atlas of England*. London: Croom Helm.
- REDE. *Startseite*. <https://www.regionalsprache.de/>.
- Seabra, R.D. (2014). *A Brazilian Contribution for Automated Linguistic Cartography*. *American Journal of Linguistics*, Vol. 3, Iss. 2.
- Siebler, P., Weibel, R., Glaser, E., and Bart, G. (2012). *Cartographic Visualization in Support of Dialectology. The 2012 AutoCarto International Symposium on Automated Cartography*.
- Škofic, J. et al. (2011). *Slovenski lingvistični atlas*. Vol. 2. Ljubljana: Založba ZRC.
- Sousa, X. (2017). *From field notebooks to automatic mapping: the „Atlas Lingüística Galego“ database*. *Dialectologia et Geolinguistica*. Berlin: De Gruyter Mouton.
- Štolc, J. et al. (1981). *Atlas slovenského jazyka II*. Bratislava: Veda.
- Universität Augsburg. (2019). *Sprachatlas von Bayerisch-Schwaben*. https://www.philhist.uni-augsburg.de/lehrstuehle/germanistik/sprachwissenschaft/projekte/sbs/sbs_2erh.html
- Upton, C. and Widdowson, J.D.A. (2006). *An Atlas of English Dialects*. Londo: Routledge.
- Utěšený, S. (1974). *K problematice symbolů v lingvistickém kartografování*. *Slovo a slovesnost*, Vol. 35, Iss. 4.
- Vidoeski, B. et al. (2008). *Macedonian Dialect Atlas: Prolegomena*. Skopje: Institute for Macedonian Language.
- Voženílek, V. and Vondráková, A. (2017). *Systematic Cartography in Web Atlas Engineering. Proceedings of the 17th International Multidisciplinary Scientific GeoConference*, Vol. 17, Iss. 23.
- Voženílek, V. et al. (2011). *Metody tematické kartografie: vizualizace prostorových jevů*. Olomouc: Univerzita Palackého v Olomouci.