DiG 18 (2010), 73–100 *DOI 10.1515/DIG.2010.005.*

Quantification and Statistical Analysis of Structural Similarities in Dialectological Area-Class Maps[*]

Jonas Rumpf, Simon Pickl, Stephan Elspaß, Werner König, Volker Schmidt

Abstract

Dialect atlases comprise considerable numbers of linguistic feature maps, i.e. dialect maps representing one linguistic feature each. Large amounts of data like these are often difficult to handle. This article presents a new quantitative method for the automatic analysis of such large corpora of linguistic feature maps. It makes use of geographical similarities between single maps to establish a system of criteria for structural relatedness. Furthermore, it employs statistical techniques to test whether given linguistic relations between the maps coincide significantly with structural relations. To achieve this, each underlying point-symbol map is converted into an area-class map (with all the original information still available). These area-class maps yield additional information regarding their structural composition. Cluster analysis is then employed to obtain groupings of similar maps. Such groupings facilitate the search for language-internal factors that influence the geographical distribution of linguistic variants, as the relevance of any given linguistic parameter for spatial patterns can be tested using statistical methods. Moreover, language-external factors, such as topographical conditions, can be tested in the same way. Thus, this new method allows for a profound and substantiated investigation of the regularities that can be found in the geographical distributions of linguistic variants.

1. Introduction

1.1 Spatial Distribution of Single Linguistic Features

Some of the main aims of dialectology have always been the division of a given geographic space into areas where different dialects are spoken, the detection of dialect boundaries, and the investigation of the strength of these boundaries. For this purpose, a sub-discipline of dialectology, 'dialectometry', has introduced the method of counting or measuring the differences between lects spoken at different locations. (For a concise overview

[*] The project is funded by the DFG (Deutsche Forschungsgemeinschaft). The authors would like to thank Michael Böhm and Matthias Kugler for their help in the implementation of the experiments reported on in this study. See http://www.uni-ulm.de/en/mawi/mawi-stochastik/forschung/projekte/sprachatlas.html

of dialectometric methods, see, for example, Heeringa 2004, 9–24.) The higher the number or degree of differences between two locations is, the higher is the chance that they are placed in two disjoint dialect areas. This implies, however, that each pair of locations sometimes exhibits agreement, sometimes disagreement, even within one dialect area. Otherwise, there would only be completely disjoint, mutually unintelligible lects. Instead, we find that while some of the variants fit neatly into the dialect areas, others show somewhat divergent or even completely different geographical distributions. Consequently, the distributions of single variants must be considered more than mere deviations from one underlying pattern as represented by the dialect areas; in fact, they differ so greatly from one another that they need to be studied, too.

The study of the distributions of single linguistic variables or variants has, so far, not been widely considered. Some scholars, however, have attempted a classification of patterns that appear on the maps recurrently (see, for example, Bach 1969: 39–226 for an overview or Hildebrandt 1983 for a more recent, lexical approach), although to date no quantitative approach has been undertaken.

The methods that we will discuss in the following are the results of the joint research project “New Dialectometry using Methods from Stochastic Image Analysis” of the Chair of German Linguistics at the University of Augsburg and the Institute of Stochastics at Ulm University.[1] Instead of dividing the area under investigation into dialect areas, we look at the spatial distributions of single linguistic variants in order to gain insight into factors that determine these distributions. Since a quantitative approach to this question is desired, pre-processing of the maps and quantification of their structural characteristics is necessary. A suitable method was introduced in Rumpf et al. (2009) and will appear in Pickl and Rumpf (forthcoming), the former providing a more technical, the latter a more conceptual account of the procedure. With this method, it is possible to obtain values that correspond to intuitive concepts such as the “complexity” or “homogeneity” of a map. Thus, from a corpus of maps, a dataset containing these values for each map can be generated.

[1] For a more detailed account of the methods, cf. Rumpf et al. (2009), Pickl and Rumpf (forthcoming). All results and examples are obtained using data from the *Sprachatlas von Bayerisch-Schwaben* (SBS 1996–2009).

1.2 Detecting Structural Similarities

With the methods discussed in this article, it is possible to find clusters of similar maps in a corpus of linguistic feature maps. These groupings, which are determined by congeneric spatial structures, allow for a substantiated interpretation of these structures as results of language-external factors, such as topographical circumstances or (former) political borders. More importantly, they also allow for an analysis that aims at an interpretation of the patterns as results of language-internal factors. It seems natural to arrive at the hypothesis that linguistic similarity entails similar spatial structures. In structuralist phonology, such corresponding spatial developments of systematically dependent phonemes are known as "Reihenschritte": If one sound changes, other sounds that are phonologically related will develop in the same way, which leads to matching patterns on the maps. This might also be the case on other linguistic levels: If similar spatial distributions are found in linguistic features that are at first glance mutually independent, such as, for example, certain lexical items, this suggests that there are reasons for the conformity of the distributions other than purely phonological ones, which could be previously unthought-of typological linguistic similarities between the variables involved.

Different notions of similarity between maps will, of course, lead to different groupings. But when do we consider two maps as similar?

A first, rather concrete, concept accepts similarities between maps only if they are location-dependent, i.e. structures that contribute to the similarity of two maps have to appear in the same region of the investigation area, although they do not have to match exactly. If, for example, a certain structure appears in the upper left corner of one map and a similarly shaped structure in the upper left corner of the second, then this will increase the similarity of the two. If, however, the structure appears in the lower right corner in the second map, this will result in a lower similarity.

A more abstract notion of similarity is not so much determined by the shapes and positions of structures on maps, but by the overall texture of maps. This location-independent approach assesses the similarities between maps by looking at features such as the average size of areas, the total length of boundaries or the mean heterogeneity of the areas on maps.

These two concepts or loose definitions of similarity between maps can have various implementations: all of the methods discussed in this article, however, refer to the location-dependent concept. Location-independent approaches are currently in development.

2. Data

2.1 The *Sprachatlas von Bayerisch-Schwaben*

The analyses discussed in this article have been performed on the basis of data from the *Sprachatlas von Bayerisch-Schwaben* (SBS 1996–2009), which was developed and compiled at the University of Augsburg under the direction of Werner König during the years 1984–2005. It comprises 14 volumes with a total of approximately 2,700 maps. On these maps, the raw data is mapped by point symbols representing different variants recorded at the 272 record locations (which we will sometimes refer to as 'points of measurement') of the investigation area.

To obtain the data for the SBS, up to six informants were interviewed at every record location. In the database, a number of entries are assigned to each of these localities. Each entry corresponds to one distinct answer given by the informant(s). For our investigations, (phonetic) transcriptions of the actual records and codes for the symbols that were used in the mapping are used. The identification of variants in our analyses is based on these symbol codes. In this way, we make use of the pre-classification of the records, thus adhering to the established linguistic principles employed by the cartographers.

So far, we have restricted our investigations to a corpus of 736 maps representing a large portion of the word-geographical maps of the SBS. This set of point-symbol maps was first subjected to preprocessing and preliminary analysis to obtain so-called area-class maps. The concept of this procedure will be described in the following section.

2.2 Preparation: Area-Class Maps and Their Structural Characteristics

In a first step towards the analysis of the maps from the SBS, we employed methods from spatial statistics to transform each of the original point-symbol maps into an area-class map, which decomposes the investigation area into disjoint sub-areas. The different sub-areas (marked by different colour hues in our figures) represent regions of predominance of different variants. By calculating 3 different map characteristics, we then analyze each map with respect to the questions of how complex its structure is (C), how well the division into sub-areas represents the original data ($\bar{L}$), and how homogeneous these sub-areas are ($\bar{B}$). In this section, we will give a brief summary of the statistical methods used in this process. For a detailed description and explanation, see Rumpf et al. (2009) and the references given there.

For the purpose of creating area-class maps, we consider the data on the point-symbol maps to be point patterns in the plane (cf. Figure 1a), so they can be investigated with methods from point process statistics (see, for example, Baddeley et al. 2006, and Illian et al. 2008). The occurrences of each variant are marked by distinct symbols at the 272 record locations. Thus, as the initial step, we separate a point-symbol map into “variant-occurrence maps”, one for every variant occurring (cf. Figure 1b). On these variant-occurrence maps, we mark all locations according to whether the variant under consideration occurs at that location or not. Additionally, we specify an “occurrence weight” $l_x(t)$ of variant x at location t that denotes what fraction of all variants occurring at t is represented by x.

Using multivariate kernel density estimation (see, for example, Silverman 1986, and Scott 1992), we then estimate an intensity field for the point pattern formed by the occurrences (and non-occurrences) of each variant. Thus, the intensity map for variant x is simply a smoothed version of its variant occurrence map (cf. Figures 1b and 1c). Roughly speaking, for all locations in the investigation area, it indicates the likelihood of x occurring there. This estimated intensity at a location t is mainly based on the frequency of actual occurrences of x and t’s proximity to them: the more occurrences of x appear in the proximity of t, the higher is the estimated intensity of x at t.

Once we have obtained a set of intensity maps for all variants occurring on a point-symbol map, we create the area-class map by re-merging the intensity maps in the following way: Each of the 272 record locations is assigned to that variant whose estimated intensity is the highest at that location; we call this the dominant variant at that location. The rest of the investigation area is split into cells through the Voronoi mosaic (also known as Thiessen polygons; see, for example, Okabe et al. 2000) generated by the record locations, and each Voronoi cell is assigned to the same variant as its corresponding location. In this way, we obtain a decomposition of the investigation area into areas of dominance of different variants and, automatically, a set of boundaries between them.

Figure 1d shows the area-class map of variants of “Kartoffelkraut” as an example. Areas with different dominant variants are denoted by different colour hues. The boundaries between areas of dominance of different variants are marked by the orange lines, and the record locations are marked by black points. The brightness of the colour at a location t corresponds to the degree of dominance of the dominant variant at t. A darker colour means that the relative intensity of the dominant variant, i.e. the contribution of the dominant variant to the total sum of estimated intensities, at t is larger. In other words, darker colours denote a higher dominance and less interference of other variants. As a natural consequence of this, colours tend to be brighter in regions near the boundaries separating different variant areas.

As a next step, we suggest calculating certain characteristics that help to describe the geometric features of the area-class maps obtained from the original data. Firstly, we propose calculating the total length of boundaries between the areas on a map as an indicator of the overall complexity of a map; we thus denote it by C. A larger value of C means that there are more boundaries where a change from one dominant variant to another occurs, making the map more complex on a larger scale. Note that C does not just represent a different way of counting the distinct areas, since areas can e.g. be irregularly shaped or disconnected, causing C to increase even though the number of areas is constant. The area-class map for "Kartoffelkraut" in Figure 1d has a complexity of $C = 240.8$ km.

To characterize the area on a map that is assigned to variant x with respect to its fidelity to the original data, we calculate $\bar{l}_x$, the average of all occurrence weights of x at locations in that area. This value shows what fraction of the variants actually occurring in that area are represented by assigning them to the variant x. From a different point of view, it can be called the compactness of that area, since larger values of $\bar{l}_x$ indicate a higher concentration of occurrences of variant x with fewer other variants mixed in. Calculating the average $\bar{L}$ of all $\bar{l}_x$ on a map gives us the overall compactness (or overall fidelity) of the area-class map. The overall compactness of the map in Figure 1d is 0.72, which is slightly above average for the considered corpus.

Finally, we calculate the homogeneity of an area on an area-class map by taking the average of the dominances of variant x at all locations assigned to x. Remember that the dominance of a variant at a location is represented by the brightness of the corresponding colour at that location. Hence, the homogeneity value $\bar{b}_x$ of the area of variant x is higher for areas with less interference from other variants. A value of overall homogeneity of a map can be obtained by simply calculating the average $\bar{B}$ of all $\bar{b}_x$ on a map. The map in Figure 1d has the above-average overall homogeneity of $\bar{B} = 0.71$.

Obviously, the characteristics C, $\bar{L}$, and $\bar{B}$ are not independent. For example, C and $\bar{B}$ are negatively correlated, i.e. maps with large values of C tend to exhibit small values of $\bar{B}$ and vice versa. This is very plausible since maps that switch between dominant variants very often (and thus are very complex) can be expected to have these variants interfere with each other quite strongly (and thus not be very homogeneous). For histograms of the values of these characteristics in the corpus under consideration, more elaborate examples, and a more detailed discussion of the definition and meaning of these characteristics, see Rumpf et al. (2009).

3. Mathematical Methods

In this section, we will explain and discuss the mathematical tools used in our investigations. In doing so, we restrict ourselves to mostly heuristic explanations, sometimes in algorithmic form, since rigorous mathematical deductions would only obstruct the view of the essence of the methods. Firstly, we will define measures of distance between area maps, which will allow us to quantify the similarities between them. In Section 3.2, we will show how these numerical values will be used by certain methods of cluster analysis to obtain groupings (i.e. clusters) of similar area-class maps. To check how similar the maps in a cluster are and how different they are from the maps of other clusters, we will then introduce the notions of uniformity and distinctness by means of certain types of statistical hypothesis tests. The section concludes with an algorithmic summary that shows how the various methods interact to produce the results presented in Section 4.

3.1 Measures of Distance between Area-Class Maps

One of the main goals of our investigations is the quantification of structural differences and similarities between area-class maps, which can then be used as a distance measure for further analysis. Before such quantification can be done, we have to clarify what determines a linguistically relevant concept of structural similarity between area-class maps.

Looking at Figure 1d, it is clear that the viewer's impression of the map's structure is greatly influenced by the position, length, and shape of the boundary lines marked in orange. A map that, for example, has a single straight boundary running from north to south instead of the horizontal boundary positioned roughly in the middle of Figure 1d would be considered vastly different from the map for variants of "Kartoffelkraut" given in Figure 1d. On the other hand, a map exhibiting almost the same boundaries as Figure 1d, only shifted slightly to the north, would not be considered much different from it; the exact degree of difference would presumably depend on the amount of displacement of the boundaries between the two maps.

A closer inspection of the example given in Figure 1d yields that the distribution of brightness levels within the different areas can be another important determinant of the map's structure. Maps that have identical boundaries can nevertheless be quite different in other aspects. If, for example, one map has areas that exhibit strongly varying brightness levels (such as in Figure 1d), and a second, conceived map with the same boundaries has areas that are uniformly bright, then their similarity is only macroscopic. On a smaller scale, much more variation is occuring on a brighter map. Recall that the actual colour hues of the different areas have no meaning other than to distinguish between different variants (cf. Section 2.2). This means that in Figure 1d, red, green, purple, and turquoise can be interchanged arbitrarily without changing the meaning or the structure of the map, as long as the brightness levels remain unchanged.

Based on these observations, we will now propose different ways of calculating numerical values for the distance between two area maps. These values will be necessary for cluster analysis of a corpus of area-class maps, cf. Section 3.2.

3.1.1 The Sector-Method

The first method we propose for the calculation of a distance measure is based on the positions and geographical distances of the nearest boundaries in relation to each location. In order to determine these positions, the investigation area is split into disjoint sectors, which are considered separately. Thus, we will call distances between area-class maps calculated in this way "sector-method distances". The sector-method distances are calculated by going through the following algorithm:

1. Choose two area-class maps whose distance is to be calculated. Perform step 2 separately for each of the maps, then proceed to step 5.
2. For each of the 272 points of measurement on the area-class map, follow steps 3 and 4.
3. Partition the observation area into a fixed, pre-specified number d of sectors delimited by evenly spaced rays emanating from the respective point of measurement. The parameter d can be chosen arbitrarily; the higher the number d is, the more precisely the sectors are defined. However, for values larger than 12, no visible differences can be found in the results. For an illustration, see Figure 2a, where $d = 4$ and the sectors for two points are marked by the white and black lines, respectively.
4. For each of the sectors, calculate the distance to the closest point of measurement in the same sector that belongs to a different variant area. If such a point does not exist (i.e. all locations in the sector belong to the same area), calculate the distance to the farthest point of measurement in that sector. For an illustration, see Figure 2b, where the calculated distances are indicated by red lines.
5. For each area-class map, $272 \cdot d$ values of "boundary distances" have been calculated. We now calculate the absolute differences between all corresponding values of the two maps. The distance between the two

area-class maps is then defined as the sum of all these absolute differences.

It is important to note that this definition of distance between area-class maps implies that two maps that exhibit exactly the same boundaries between areas will have distance zero. The inner structure of any of these areas on the maps – which is, as mentioned above, primarily formed by the distribution of brightness values inside the area – is completely ignored.

3.1.2 Differences in Relative Intensities

While the sector-method distance relies completely on the positioning of boundaries to determine the distance between two area-class maps, we will now introduce a measure that utilizes the changes in relative intensities within and across areas. Recall that the relative intensity of a variant at a location is given by the fraction of the total sum of estimated intensities, and that a location’s brightness is determined by the relative intensity of its dominant variant (cf. Section 2.2).

The “relative-intensity distance” is computed in the following way:

1. Choose two area-class maps whose distance is to be calculated. Perform step 2 separately for each of the maps, then proceed to step 4.
2. For each of the $\frac{271 \cdot 272}{2}$ pairs of points of measurement, perform step 3.
3. If both points of the pair have the same dominant variant, calculate the absolute difference of the relative intensities of this variant (i.e. the brightness values) at the two locations. Otherwise, i.e. if the two points have different dominant variants, proceed in the same way for both these variants and calculate the average of the two differences.
4. For each area-class map, 36,856 values of “relative intensity differences” have been calculated. We now calculate the absolute differences between all corresponding values of the two maps. The distance between the two area-class maps is then defined as the sum of all these absolute differences.

This distance measure does not focus on the boundaries between areas as strongly as the sector-method distance. Still, the distance values here are influenced indirectly by the boundaries on a map, since the colours on a map tend to be brighter (i.e. the dominances smaller) near boundaries between areas (cf. e.g. Figure 1d).

Note that, in theory, it is possible for two area-class maps to have relative-intensity distance zero. However, one cannot deduce the distance to be zero simply from the brightness values of the maps being equal, since this method also considers the relative intensities of some variants that are not the dominant variant (see step 3), which are not displayed on the area-class maps.

3.2 Cluster Analysis

Cluster analysis is a mathematical method for the division of a set of multivariate objects into disjoint subsets (clusters). Its main goal is to obtain a partition in which the objects in each cluster are as similar to one another as possible, while the clusters are as dissimilar from each other as possible. Note that a similarity measure is required, which is usually implemented as a distance function. In linguistics, cluster analysis is commonly used in dialectometrical studies, where it serves to obtain groupings of dialects, which can then be charted as dialect areas on a map (cf., for instance, Goebl 2005). In our investigations, we use cluster analysis to partition the investigated corpus of area-class maps into groups of maps with a similar spatial structure. For a thorough treatment of cluster analysis, see, for example, Jobson (1992), and Arabie et al. (1996). There are many possible variants of cluster analysis, but, for reasons of simplicity and efficiency, we confine ourselves here to the simplest case, i.e. hierarchical methods. We will briefly explain some of these methods in the following; details can, for example, be found in Chapter 10 of Jobson (1992).

The basic idea of hierarchical clustering is very simple: Initially, calculate the distances between all pairs of objects (i.e. in our case, area-class maps), using a specified distance measure (here, one of the distance measures discussed in Section 3.1). Then, consider each object to be a cluster of its own. In each step, form a new, larger cluster by joining together two separate clusters according to specified criteria. This will reduce the number of clusters until a specified termination criterion is satisfied. The result will be a partition of the considered objects, i.e. a set of clusters of area-class maps, with each of the maps of the corpus being assigned to exactly one cluster. The rules that decide which clusters are joined in each step are what distinguishes the various hierarchical methods from each other; some of them will be explained here.

In one approach, a distance between any two clusters is calculated, which then determines which clusters are to be merged. The two clusters with the smallest distance are the ones to be joined together. Hence, a measure for the distance between clusters (not only objects) is required. The methods “complete linkage”, “average linkage”, and “single linkage” are different ways of determining the distance between two clusters. When both clusters consist of a single object each, it is natural to take the previously calculated distance between the two objects as the distance between two clusters. However, the methods differ on how to calculate the distances between clusters that contain more than one object. In complete linkage, the distance between two clusters is the maximum of distances between all pairs of objects consisting of one map from each of the two clusters. In contrast to this, single linkage takes the minimum of all these distances of object pairs, while average linkage calculates the average of this set of distances to obtain the distance between two clusters of objects.

A different approach is taken by "Ward's method". Instead of calculating distances between clusters, it quantifies the variability within clusters in each step. It then selects those two clusters for joining whose union will cause the smallest increase of variability from one step to the next. Note that instead of a distance function, the objects need to be characterized by a number of interval-scaled features (such as, for example, the total length of boundaries on the map), so that an "average object" of a cluster, and with it, a cluster's variance as a measure of its variability, can be calculated.

As mentioned above, a termination criterion is required to determine at what point the hierarchical clustering methods should stop clustering, which can be a pre-specified number of clusters. In our investigations, we terminated the clustering algorithm when the smallest distance between clusters (or, in the case of Ward's method, the increase in variability) exceeded a certain level. This threshold was given by $m + k \cdot s$, where m denotes the average of all distances between two clusters that are joined, s the variability of these distances (measured by their empirical standard deviation), and k is a parameter that was chosen to be $k = 1.7$ as an adaptation of the rule of thumb value 1.96 (cf. also Mojena 1977). The goal of this is to make sure that the distances between maps within a cluster do not beome unreasonably high.

3.3 Statistical Hypothesis Testing

The aim of clustering is to obtain groupings of elements that are similar to each other, whereas the groupings themselves should be clearly distinguishable from each other. Assuming that a partition into clusters is an approximation of a natural division of the elements, it is desirable to test whether the clusters really reflect such natural groupings, and to what extent they do so. For the testing of the quality of a partition with regard to what we call uniformity of clusters and distinctness between clusters, it is advisable to look at characteristics that were not part of the distance calculations, in order to avoid circular conclusions. There are several types of statistical tests that are suitable for the testing of the quality of a partition. For this purpose, the expectations we have of the clusters in a partition have to be rephrased as hypotheses that can be either true or false.

Statistical tests are objective rules to decide whether a certain sample fits well with a given hypothesis about its distribution or whether enough evidence against this hypothesis exists, so that it should be rejected for a specified alternative. In our case, samples will be clusters of area-class maps in a specific partition, and the hypotheses will be the following: 1) "A given cluster is not more uniform within itself than a typical random cluster would be", and 2) "Two given clusters are not well distinguished from each other". Usually, an upper bound $\alpha \in (0,1)$ on the probability of the error of type I, i.e. rejecting the hypothesis although it is true, is specified for such tests; α is called the level of significance of the test. Note that it is not appropriate to speak of a hypothesis as being "accepted" when it is not rejected; all that non-rejection implies is a lack of evidence against the hypothesis. Of the many possible types of tests (see, for example, Lehmann and Romano 2005), we will briefly discuss only those two relevant for our investigations. Later on, we will explain the application of these testing procedures in our context of characterizing partitions of a corpus of area-class maps.

3.3.1 Two-Sample Tests

Generally speaking, so-called two-sample tests check the hypothesis that the distributions underlying two statistical samples are the same. For the purposes of our investigations, the two samples will be given by the values of one of the map characteristics discussed in Section 2.2 in two clusters of area-class maps. One popular example for a non-parametric two-sample test is the so-called Kolmogorov-Smirnov test (see, for example, Gibbons 1985: 127 ff.). Formally, it tests the hypothesis that the distribution functions underlying two given samples are equal against the alternative that they differ in at least one value. Informally speaking, the rejection of its hypothesis by this test simply means that the two clusters are statistically distinct from each other.

The Kolmogorov-Smirnov test is sensitive to all kinds of differences in the distribution functions. In contrast to this, many other tests restrict their attention to one aspect of the data. The median-test, for example, only checks whether the values in the two clusters tend to have equal size on average and disregards their variability completely, while the Ansari-Bradley test does the opposite. For a detailed definition and discussion of these and other two-sample tests, see, for example, Gibbons 1985: 122 ff. and 179 ff. In Section 4, we will only discuss results obtained with the Kolmogorov-Smirnov test to avoid confusion. However, we have performed all investigations using various other tests, and it appears that the results are not very sensitive to the choice of two-sample test.

3.3.2 Monte-Carlo Tests

With the two-sample tests mentioned in the previous section, we compare the distributions of certain characteristics in two clusters of area-class maps, to see if these clusters are well distinguished from each other. On the other hand, we are also interested in the question of whether the clustering procedure has any influence at all on the variability within a cluster, i.e. whether the values of a certain characteristic (for example *C*) within a single cluster are more uniform than they would be within a random set of area-class maps. To be able to answer this question, we need information about the distribution of the characteristic of interest in a typical random cluster. Since a priori, no such information is available, we generate it by simulating large amounts of random clusters. This approach is called Monte-Carlo testing. Monte-Carlo tests were first proposed by Barnard (1964), and generalized by various authors, for example by Besag and Clifford (1989). For a more rigorous introduction to and discussion of Monte-Carlo tests, see, for example, Edgington (1995). We will clarify the procedure with the following example.

Assume the sample to be given by the values of one of the characteristics from Section 2.2, say *C*, in a single cluster of maps *M*. We would now like to check the hypothesis that the variability of the values of *C* in *M* – which has been obtained using an objective, deterministic clustering procedure – is equal to the variability in a cluster created purely at random. To this end, we create "random clusters" by randomly selecting the same number of maps from the whole corpus as contained in *M*. For such a random cluster, we can now calculate the same measure of variability as for *M*, say the empirical standard deviation. If we repeat this procedure *k* times, say $k = 99$, we have a total of 100 standard deviations which we sort in ascending order. In this way, we can detect whether the variability of *M* is lower (or higher) than that of most of the random clusters. To be precise, we reject the hypothesis of equal variabilities for the alternative that the variability in *M* is lower (or higher) than that of a random cluster if the position of the standard deviation of *M* in the ascending sequence of standard deviations is lower than $\alpha \cdot 100$ (or higher than $(1 - \alpha) \cdot 100$), where α is the pre-specified level of significance of the test, see above. In this case of rejection, we consider the investigated cluster of maps to be uniform from a statistical point of view.

3.3.3 Uniformity and Distinctness

In the preceding section, we have provided statistical tools that will enable us to evaluate partitions of a corpus of area-class maps (see Section 2.1).

The purpose of this is to check in which respect and to which degree the clusters of maps created by cluster analysis using different distance measures (see Sections 3.1 and 3.2) satisfy the requirements of being uniform within and distinct compared with one another. Note that we check uniformity and distinctness with respect to one of the characteristics of Section 2.2, while the clusters have been created with regard to one of the distance measures of Section 3.1. This discrepancy is intentional: had the clusters been created using as a measure of distance the same characteristic that is used for checking the cluster results, the conclusions would be meaningless, since the clusters would be almost perfectly uniform and/or distinct by construction.

For the evaluation of the uniformity of a cluster of area-class maps, we use the Monte-Carlo tests of Section 3.3.2. We say that a single cluster M is α-uniform with respect to the characteristic C if the hypothesis that the standard deviation of the values of C in M is equal to that of a random cluster is rejected at level of significance α. The alternative hypothesis can in this case be formulated as "the standard deviation within M is smaller than that of a random cluster". This definition is justified by the fact that the rejection of this hypothesis implies that the variability of C within M is small compared with random clusters, i.e. the values of C in M are much more similar to each other than what could be expected of a random cluster.

To be able to draw conclusions about the distinctness of a pair of clusters, we employ the two-sample tests of Section 3.3.1, in particular the Kolmogorov-Smirnov test. The pair of clusters M_1 and M_2 will be called α-distinct with respect to the characteristic C if the hypothesis that the distribution functions of the values of C in the clusters M_1 and M_2 are equal everywhere is rejected at level of significance α for the alternative that the distribution functions differ at one or more values. It should be obvious that this definition makes sense, since the fact that the values of C in the two clusters are significantly different is exactly what we perceive as distinctness of M_1 and M_2 with respect to C. Obviously, identical definitions for uniformity as well as distinctness can be given with respect to the characteristics $\bar{L}$ and $\bar{B}$.

3.3.4 Characteristic Vector

Having specified the terms *uniformity* and *distinctness* in the previous section, we can now define the so-called characteristic vector at level α with respect to a characteristic of a partition of our corpus of area-class maps. This two-dimensional vector will summarize the information about a partition's uniformity and distinctness in two numbers from the interval [0,1].

Let the first component of the characteristic vector, the uniformity-value of a partition, be given by the relative frequency of clusters in the partition which are α-uniform with respect to the considered characteristic. This means that a partition consisting of 40 clusters, 25 of which are α-uniform while 15 are not, will have a uniformity-value of 0.625.

Accordingly, we define the second component of the characteristic vector, i.e. the distinctness-value of a partition, to be the relative frequency of pairs of clusters in the partition which are α-distinct with respect to the considered characteristic. As an example, consider again a partition consisting of 40 clusters. Then, one can distinguish $\frac{39 \cdot 40}{2} = 780$ distinct pairs of clusters. If 600 of these pairs are α-distinct and the others are not, the distinctness-value of the partition will be approximately 0.769.

In summary, we can say that the characteristic vector defined above gives a concise overview of the degree to which a certain partition of the corpus of area-class maps satisfies the criteria of uniformity and distinctness with respect to a certain characteristic. Large uniformity- and distinctness-values, i.e. values close to 1 in the characteristic vector, indicate a good fulfilment of these criteria, while smaller values (closer to 0) indicate that the partition must be considered less uniform and/or distinct regarding the considered characteristic. In Section 4, we will discuss some results of cluster analysis of the corpus of area-class maps from the SBS (1996–2009) as well as the corresponding characteristic vectors together with an interpretation of the values.

3.4 Algorithmic Approach

To conclude our discussion of the mathematical aspects of the research presented here, we will briefly summarize our approach to the quantification and statistical analysis of structural similarities in a corpus of area-class maps in a simple, algorithmic way. The aim is to illustrate how the various techniques described in Sections 2.2 through 3.3 fit together to obtain the results which we will discuss in Section 4.

For our investigations, we proceed by performing the following steps:

1. Select a corpus of point-symbol maps from the SBS (1996–2009) and prepare the maps for analysis in the way described extensively in Rumpf et al. (2009) and summarized in Section 2.2.
2. Select one of the measures of distance between area-class maps described in Section 3.1 and calculate the distances between all possible pairs of area-class maps in the corpus.

3. Select one of the clustering methods described in Section 3.2, and select the parameter k for the stopping rule defined there.
4. Perform the clustering method selected in step 3 on the corpus, using the distance measure from step 2. Stop clustering when the termination criterion is satisfied.
5. In this way, a partition of the selected corpus is obtained. Each of the clusters in this partition can then be investigated linguistically with respect to possible language-internal or language-external factors which cause the similarities between the maps which the partition contains; cf. Section 1.2. For a statistical evaluation of the partition, proceed to step 6.
6. Select a level of significance α and a map characteristic (C, $\bar{L}$, or $\bar{B}$; see Section 2.2) with respect to which the partition should be evaluated.
7. Obtain the characteristic vector of the partition by calculating the uniformity and the distinctness-value of the partition as described in Section 3.3.5 with respect to the characteristic selected in step 6, and with α as selected in step 6.
8. Use the characteristic vector obtained in step 7 to evaluate the partition’s uniformity and distinctness with respect to the characteristic selected in step 6. The comparison with the characteristic vectors of other partitions (see steps 1–5) or with respect to other map characteristics (see steps 6–7) will help to put the obtained values into perspective.

Following these steps, we have obtained all the results discussed in Section 4, in which we will give the selected measure of distance (see step 2), and the clustering method (see step 3) to identify a partition. All partitions were based on the same corpus of 736 area-class maps (see Section 2.1). When specifying a characteristic vector for a partition, we will state the level of significance α and the selected map characteristic (see step 6).

4. Results

In this section, we will provide some examples of clusters obtained with the distance measures introduced in 3.1, including the respective characteristic vectors as indicators for the clusters’ and partitions’ quality. We will also show how these clusters can be a useful tool for discerning linguistic relationships between similar maps. All results are obtained using complete linkage as the clustering algorithm.

4.1 Example Cluster: Sector-Method Distance

The cluster shown in extracts in Figure 3 is one of 39 in a partition of our corpus of 736 maps. Here, the sector method with 12 sectors has been used to calculate the distances between the maps. As the sector-method takes into account only differences in the positions of boundaries, the brightness values have played no role here. A general pattern underlies this particular grouping, in that a large area in the south-west as well as smaller areas in the north and the east can be discerned, which determine the calculated similarity through the borders by which they are delimited. Note, however, that the variability that is covered by both the distance measure and the cluster method allows for a certain amount of deviation, such as small areas appearing in the south-west corner (e.g. the maps for “Kinn” (‘chin’) and for “Holzstoß” (‘pile of wood’) in Figure 3).

The characteristic vector as introduced in Section 3.3.4 can give us an insight into the quality of the partition to which this cluster belongs. A look at the characteristic vector at level $\alpha = 0.05$ reveals that the impression that this cluster is strongly influenced by boundaries is also true for most of the rest of the partition: With respect to the complexity C (given by the total boundary length), this partition has the very high uniformity value of 0.944 and a distinctness value of 0.710. In other words: More than 94% of the clusters obtained have a significantly low variability with respect to C, and 71% of all possible pairs of clusters are statistically distinct from each other regarding C. This shows that the sector method is a good way to obtain clusterings of maps with similar boundaries. In contrast, the characteristic vector for the homogeneity values ($\bar{B}$) has the much lower values of 0.694 for uniformity and 0.378 for distinctness, which reflects the fact that the brightness values are not regarded in this method. (Note, however, that there is an indirect influence of the homogeneity values on the clustering according to C, as C and $\bar{B}$ are not independent, as explained in Section 2.2.)

In a nutshell, this cluster contains maps that have similar patterns of borders. The visual impression that one gets by looking at Figure 3 supports this view. Obviously, the sector-method distance in combination with complete linkage does yield clusters that are constituted by structural similarity of the maps, as was intended. In Section 1.2, we hypothesised that structural similarity can be a consequence of linguistic similarity. In order to test this hypothesis, both kinds of similarity have to be defined; the definition for structural similarity applied here is the one given by the sector-method distance measure. For a definition of linguistic similarity, various approaches are conceivable, as similarity can occur on different linguistic levels. The present cluster provides a very simple example of what linguistic similarity can be in this context. We find two variables that share all semantic features, except for one (sex): "Witwe" ('widow') and "Witwer" ('widower'). Obviously, the close semantic relationship of the two has caused the spatial distributions of their variants to develop in similar ways, presumably influencing each other in the course of this development due to their systemic relationship. This is an indication that close semantic relatedness is a factor of structural similarity in lexical maps. In this case, it is very likely that this is due to a similarity of the lexemic structure.

While further research is necessary to confirm this, one can look a bit further to see if there is evidence that more distant semantic relatedness is likewise a factor of (dis-)similarity between maps. The division of the lexical maps into thematic groups as given in the SBS provides a rudimentary framework to check if membership in the same thematic group is a factor in the assignment to clusters, assuming that a thematic group is a relevant entity for semantic relatedness (for example, belonging to the same semantic field). If this is the case, then there should be significant accumulations of maps from certain thematic groups in a cluster. In the present cluster, maps from 16 thematic groups are represented. A one-sided binomial test has been performed to find out if a specific thematic group is significantly overrepresented in this cluster, i.e. if the relative frequency of occurrences from any group is significantly higher in the cluster than in the complete corpus. The results show that the occurrences of none of the thematic groups in this cluster are significant at a level of $\alpha = 0.05$, which contradicts the hypothesis that affiliation to one of the groups is a factor in the development of structural similarity between maps as defined by the sector-method distance.

A look into different clusters reveals that in many other cases, the fraction of a thematic group is significant, as, for example, the occurrences of the group "Kinderspiele" ('children's games') in Cluster 22, which have $p = 0.0287$. At a total of 2 occurrences, however, it is questionable whether this statistical significance bears any actual relevance.

It must be noted that this test, using thematic groups as a rough approximation of semantic relationship, can also be applied to various other concepts of linguistic relatedness, for example, frequency or word class. Which of these concepts, if any, has a significant influence on the development of structural similarity will be the subject of future investigations.

4.2 Example Cluster: Relative-Intensity Distance

Figure 4 (only web-version) shows a cluster that has been generated applying the relative-intensity distance in combination with complete linkage clustering. In contrast to the sector-method, here the boundaries are not taken into consideration, at least not directly. Instead, the relative intensity values, as displayed in the brightness values of the cells, are used to calculate the structural distance between maps. Indirectly, however, the boundaries do play a role, as the colours tend to be brighter along the boundaries. This is why also here, we can find a basic structure that underlies all (or most) of the maps in this cluster, and to which the calculated similarity can at least partly be contributed. Roughly speaking, a boundary that extends from the north west to the south east can be found in virtually all of these maps. Other boundaries, belonging to smaller areas, and all other deviations from this basic structure, can either be attributed to the distance measure, which does not exclusively rely on boundaries, or to the amount of variation that is a consequence of the clustering process, and which can, to a certain extent, be considered as 'noise'.

The characteristic vector of this partition (at level $\alpha = 0.05$) reflects the fact that here, brightness values (representing intensities) and not boundaries between areas have been the focus of the distance measure: With respect to the closely related characteristics $\bar{L}$ and $\bar{B}$, the partition has the high uniformity values of 0.913 and 0.826, respectively, while for the complexity C, the uniformity is only 0.739. The distinctness values vary around 0.5 for all three characteristics. This means that while the relative-intensity distance considers all three characteristics, it puts the least emphasis on complexity C, as it is not taken into account directly. As this distance measure does not yield very good results for distinctness, it can be considered a compromise between all three characteristics at the expense of distinctness.

Again, we find two maps in this cluster which have an obvious semantic relationship: "herumkriechen (von kleinen Kindern)" ('to crawl [of babies]') and "kriechen (vom Wurm)" ('to crawl [of a worm]'). The similarity between these two maps can be attributed to the fact that in many regions, the same variant is used for both variables. A significance test of the thematic groups, as in Section 4.1, yields the following results:

thematic group	*occurrences in cluster*	*cluster size*	*occurrences in corpus*	*corpus size*	p
Getreide	5	22	53	736	0.0181
Pflanzen / Obst / Gemüse	2	22	32	736	0.2478
Düngung	1	22	11	736	0.2820
Kleidung	1	22	12	736	0.3035
Wohnung / Einrichtung	1	22	16	736	0.3834
Geflügelhaltung / Imkerei	1	22	17	736	0.4020
Kinderspiele	1	22	19	736	0.4375
menschliche Gemeinschaft	2	22	50	736	0.4462
das Bauernhaus	1	22	23	736	0.5027
Transport	2	22	58	736	0.5264
Zeiteinteilung / Grüßen	1	22	30	736	0.5997

Rindvieh / Milchverarbeitung	1	22	43	736	0.7340
Heuernte	1	22	54	736	0.7940
Wald / Holz	1	22	61	736	0.8509
Ernährung / Kochen / Backen	1	22	98	736	0.9569

Again, most of the thematic groups are not overrepresented significantly, with the exception of the group “Getreide” (‘crop’). The *p*-value of 0.0181 is well below $\alpha = 0.05$, suggesting that the semantic relationship between these maps actually has something to do with their similar structuring. Interestingly, four of these five maps that are concerned with crop-related matters that can be summarised as “waste or side products of crop production” (the fifth map is “Roggen” [‘rye’]). Thus, it is conceivable that the variables from this semantic field have developed in a similar way because they are used similarly in everyday life, e.g. in similar situations, by people with similar occupational backgrounds and with similar frequencies. Taking into consideration the low *p*-value, this might be evidence that also somewhat looser semantic relatedness has a positive effect on structural similarities as determined by the relative-intensity distance.

In most other cases, however, p is well above α, indicating that in these cases, the similarity and the thematic grouping have nothing to do with each other. The fact that all thematic groups apart from “Getreide” are represented by only 1 or 2 instances supports this conclusion.

To summarize the results of Sections 4.1 and 4.2, we can state that close semantic relationship of variables does go along with similar spatial distributions of their variants, especially when there are morpholexical implications for the variants. This can be observed when only the boundaries are regarded, as well as when the inner-areal variation is taken into consideration. For a similar conclusion regarding loose thematic groupings, there is little evidence in the data. This kind of thematic grouping is obviously a negligible factor in most cases. Of course, the testing of hypotheses regarding the given thematic groups is only a preliminary, tentative experiment. More substantial testing will use other, linguistically more relevant categories, such as, for instance, frequency. Thus, further research will show if there is any influence of such linguistic factors on the spatial behaviour of variables.

5. Conclusions and Outlook

As we have shown, the computational analysis of the structural characteristics of linguistic feature maps is a rewarding application of statistical methods on corpora of geographical language data. This new approach facilitates the search for language-external as well as internal factors for the development of variables in space. As far as language-internal factors are concerned, the examples of Section 4 have shown that a close semantic relationship is an influential factor in the emergence of similar spatial structures, while this is not, or only to an insignificant degree, the case with loose thematic relations. Factors on other linguistic levels are yet to be tested. While an exemplification of the detection of language-external factors has not been part of this article, the testing of predefined spatial structures (such as, for example, given in the runs of rivers or mountain ridges) against structures found in clusters of linguistic feature maps seems promising.

With this methodology, it is – as far as we can see – for the first time possible to perform a quantitative analysis of spatial structures enclosed in linguistic feature maps. All studies up to this point have to be considered qualitative approaches. This new quantitative method makes it possible to pre-process geographical data to a point where hypotheses about factors for spatial developments can be tested without the need to examine all maps in the corpus one by one. The quantification of the data is a necessary abstraction for an objective assessment of underlying regularities. Thus, the (qualitative) linguistic interpretation of the results is moved to the end of the process, making the outcome less susceptible to the influence of the expectations of the dialectologist.

Since this article is only an introduction to this new method, it is clear that further methods and implementations can be developed as extensions of this dialectometrical toolbox. Some further distance measures that have not been introduced here have already been developed, most of them based on the location-dependent concept of structural similarity (cf. Section 1.2); the development of a further method, which will employ a location-independent approach, is also in progress. Furthermore, the testing of the clusters can be adapted to different interests, for example by using other characteristics than C, $\bar{L}$ or $\bar{B}$ to evaluate clusters, or by using different measures for distinctness and uniformity than the distribution function and the standard deviation. Of course, the application of these methods is not restricted to maps that have been generated as described in Section 2.2. As long as a certain data format is given and the maps are area-class maps with intensity values for the locations, nothing prevents the same or similar tests as presented in this article from being performed. Currently, for example, we are developing a method that uses linguistic similarities instead of geographical distances to generate area-class maps from the raw data. For different methods of calculating linguistic similarities, see, for example, Goebl 1984: 74–85 or Heeringa 2004: 9–24.

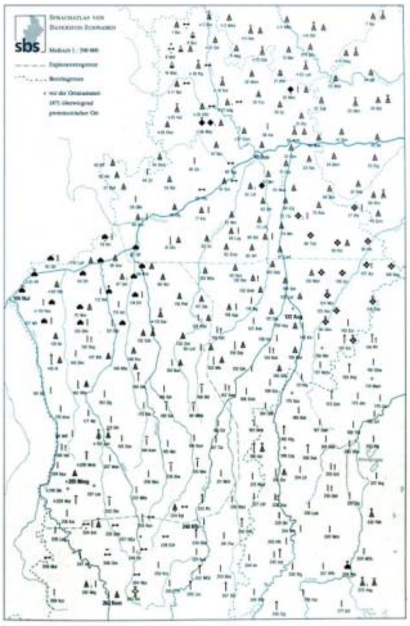


Figure 1a: Example: original point-symbol map 80 "Kartoffelkraut" from the SBS, vol. 8.

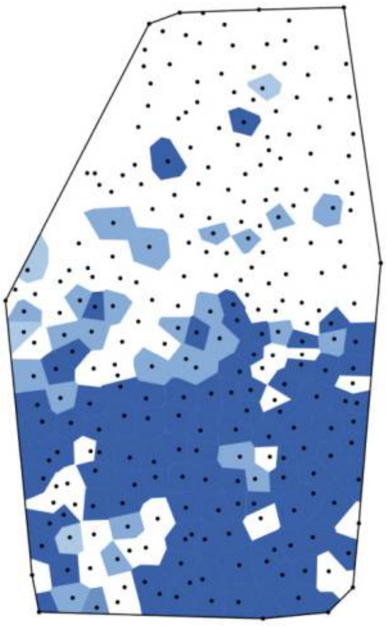


Figure 1b: Example: variant occurrence map of variant *Staude* in map 80 "Kartoffelkraut" (SBS, vol. 8). In Figure 1a, this variant is represented by the vertical line. The corresponding area in Figure 1d is marked purple.

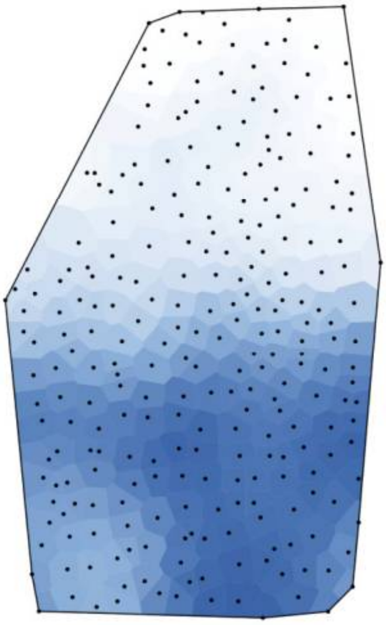


Figure 1c: Example: estimated intensity field of variant *Staude* in map 80 "Kartoffelkraut" (SBS, vol. 8).

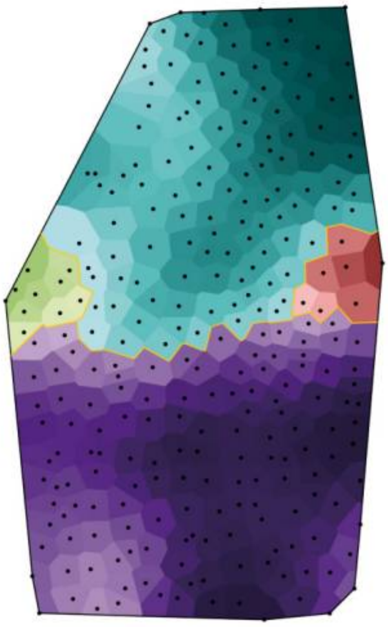


Figure 1d: Example: area-class map of the data underlying map 80 "Kartoffelkraut" (SBS, vol. 8).

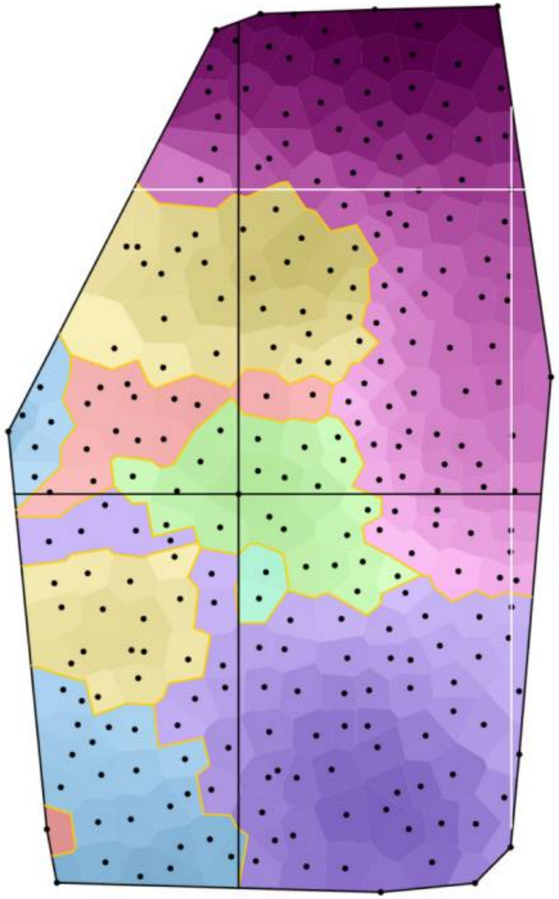


Figure 2a: Example for the partition of the observation area into 4 sectors. The sectors emanating from two points are marked by the black and white lines, respectively.

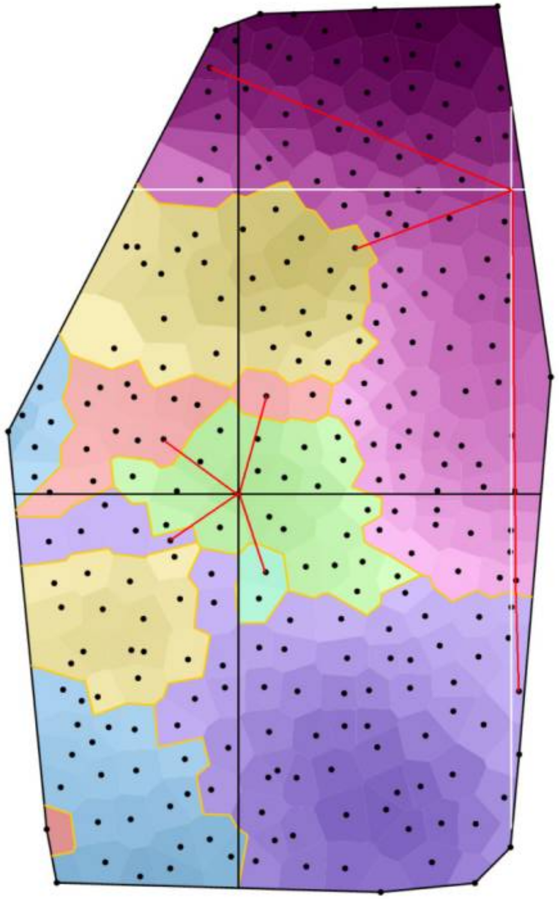
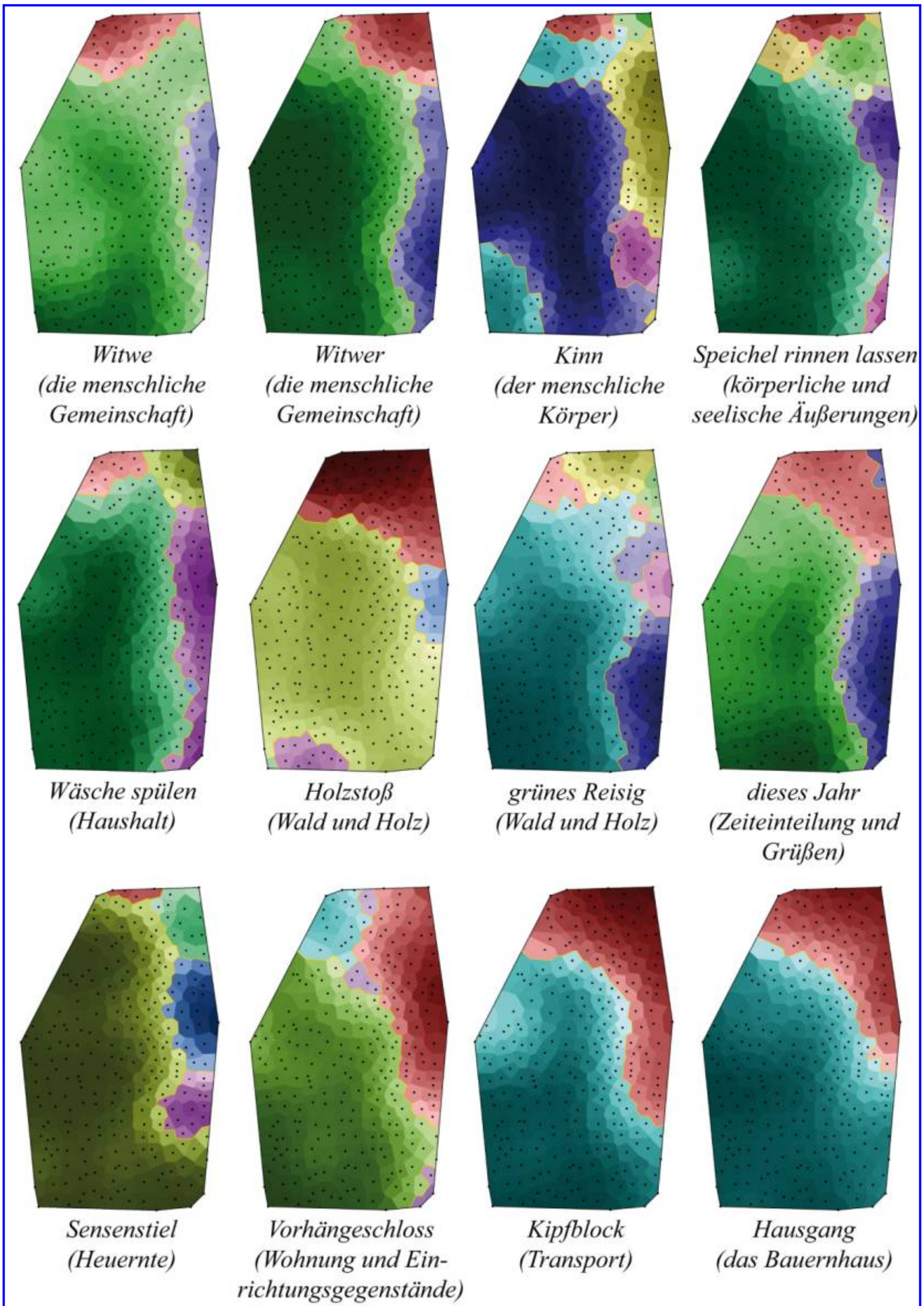


Figure 2b: Example for the calculation of the distances to the nearest point of measurement located in the respective sector but in an area of dominance of a different variant. Extending Figure 2a, the calculated distances are indicated by red lines.



Witwe (die menschliche Gemeinschaft)

Witwer (die menschliche Gemeinschaft)

Kinn (der menschliche Körper)

Speichel rinnen lassen (körperliche und seelische Äußerungen)

Wäsche spülen (Haushalt)

Holzstoß (Wald und Holz)

grünes Reisig (Wald und Holz)

dieses Jahr (Zeiteinteilung und Grüßen)

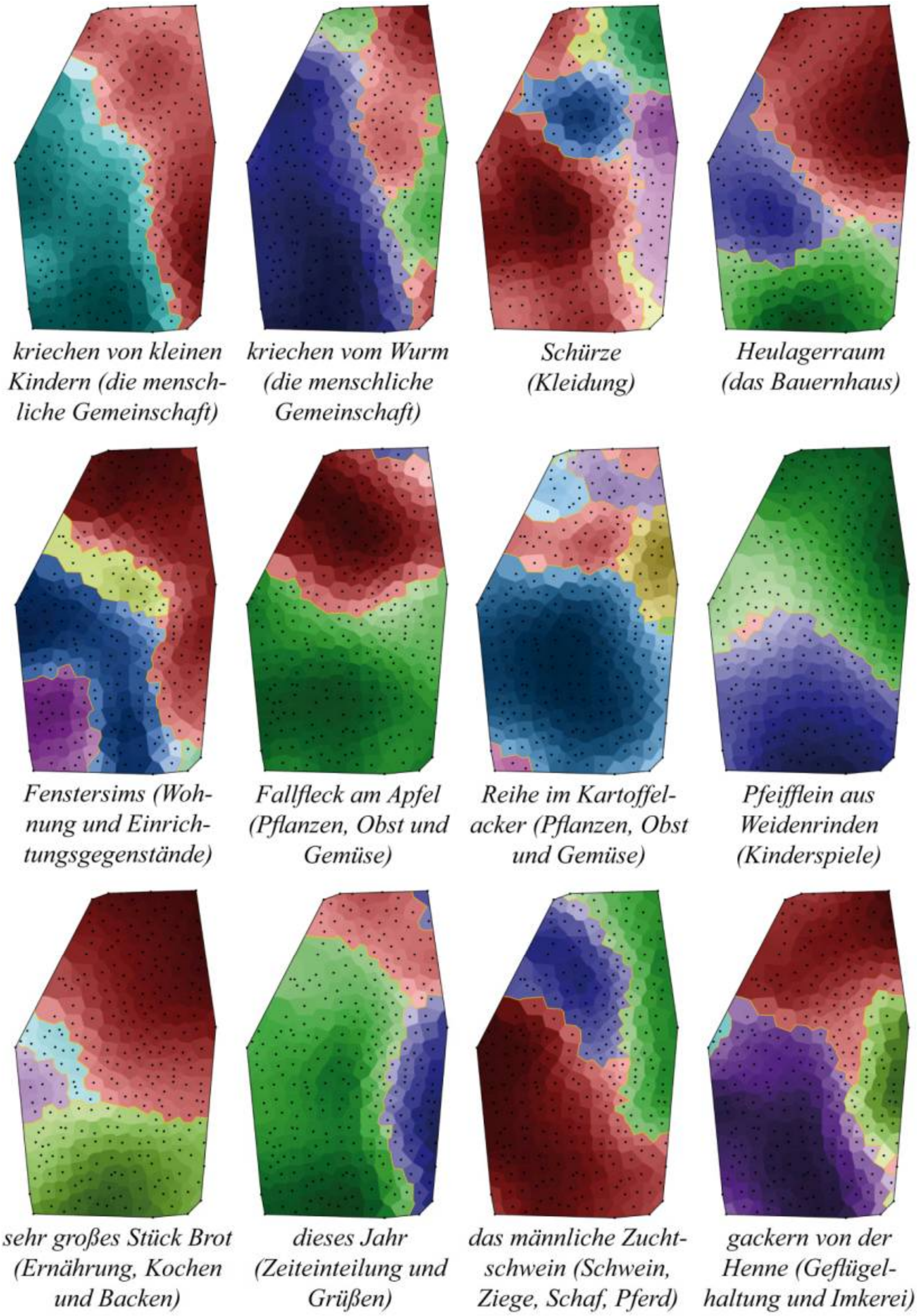
Sensenstiel (Heuernte)

Vorhängeschloss (Wohnung und Einrichtungsgegenstände)

Kipfblock (Transport)

Hausgang (das Bauernhaus)

Figure 3: Extract of Cluster 23 from a partition of 736 lexical maps, which has been obtained using the sector-method distance with 12 sectors and complete linkage as the clustering algorithm.



kriechen von kleinen Kindern (die menschliche Gemeinschaft)

kriechen vom Wurm (die menschliche Gemeinschaft)

Schürze (Kleidung)

Heulagerraum (das Bauernhaus)

Fenstersims (Wohnung und Einrichtungsgegenstände)

Fallfleck am Apfel (Pflanzen, Obst und Gemüse)

Reihe im Kartoffelacker (Pflanzen, Obst und Gemüse)

Pfeifflein aus Weidenrinden (Kinderspiele)

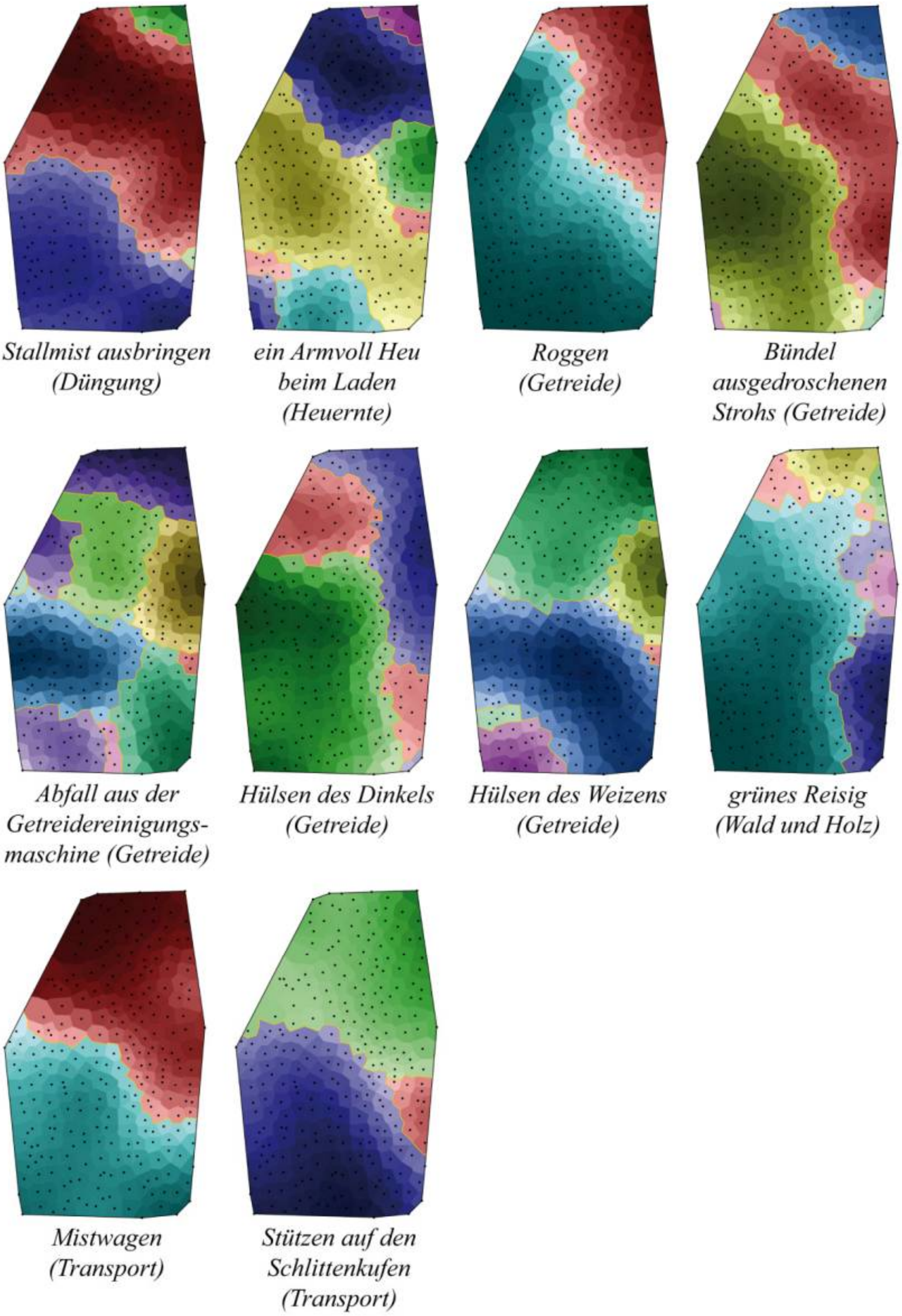
sehr großes Stück Brot (Ernährung, Kochen und Backen)

dieses Jahr (Zeiteinteilung und Grüßen)

das männliche Zuchtschwein (Schwein, Ziege, Schaf, Pferd)

gackern von der Henne (Geflügelhaltung und Imkerei)

Figure 4, Part I: Cluster 21 from a partition of 736 lexical maps, which has been obtained using the relative-intensity distance and complete linkage as the clustering algorithm.



Stallmist ausbringen (Düngung)

ein Armvoll Heu beim Laden (Heuernte)

Roggen (Getreide)

Bündel ausgedroschenen Strohs (Getreide)

Abfall aus der Getreidereinigungs-maschine (Getreide)

Hülsen des Dinkels (Getreide)

Hülsen des Weizens (Getreide)

grünes Reisig (Wald und Holz)

Mistwagen (Transport)

Stützen auf den Schlittenkufen (Transport)

Figure 4, Part II: Cluster 21 from a partition of 736 lexical maps, which has been obtained using the relative-intensity distance and complete linkage as the clustering algorithm.

References

Arabie, Ph., J. H. Lawrence, G. De Soete (eds). 1996. *Clustering and Classification*. Singapore: World Scientific.

Bach, A. 1969. *Deutsche Mundartforschung. Ihre Wege, Ergebnisse und Aufgaben* (Germanische Bibliothek. Dritte Reihe: Untersuchungen und Einzeldarstellungen). Heidelberg: Winter.

Baddeley, A. et al. (eds). 2006. *Case Studies in Spatial Point Process Modeling* (Lecture Notes in Statistics 185). New York: Springer.

Barnard, G. A. 1963. Discussion of paper by M.S. Bartlett. J. R. Statist. Soc. B 25, 294.

Besag, J. and P. Clifford. 1989. Generalized Monte Carlo significance tests. *Biometrika* 76 (4): 633–642.

Edgington, E. S. 1995. *Randomization Tests*. 3rd ed. New York: Marcel Dekker.

Gibbons, J. D. 1985. *Nonparametric Statistical Inference*. 2nd ed. New York: Marcel Dekker.

Goebl, H. 1984. *Dialektometrische Studien. Anhand italoromanischer, rätoromanischer und galloromanischer Sprachmaterialien aus AIS und AIF.* Vol. 1. Tübingen: Max Niemeyer.

–––. 2005. Dialektometrie. In: R. Köhler, G. Altmann and R. G. Piotrowski (eds), *Quantitative Linguistics. An International Handbook* (Handbooks of Linguistics and Communication Science 27). Berlin/New York: de Gruyter, 498–531.

Heeringa, W. 2004. *Measuring Dialect Pronunciation Differences using Levenshtein Distance*. Groningen: Univ. Diss.

Hildebrandt, R. 1983. Typologie der arealen lexikalischen Gliederung deutscher Dialekte aufgrund des Deutschen Wortatlasses. In: W. Besch et al. (eds). 1983. *Dialektologie. Ein Handbuch zur deutschen und allgemeinen Dialektforschung* (Handbücher zur Sprach- und Kommunikationswissenschaft 1). Vol. 2. Berlin/New York: de Gruyter, 1331 –1367.

Illian, J. et al. 2008. *Statistical Analysis and Modelling of Spatial Point Pattern*s. Chichester: Wiley.

Jobson, J. D. 1992. *Applied Multivariate Data Analysis.* Vol. 2: *Categorical and Multivariate Methods*. New York: Springer.

Lehman, E. L. and J. P. Romano. 2005. *Testing Statistical Hypotheses*. 3rd ed. New York: Springer.

Mojena, R. 1977. Hierarchical grouping methods and stopping rules: an evaluation. *The Computer Journal* 20: 359–363.

Okabe, A. et al. 2000. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams.* 2nd ed. Chichester: Wiley.

Pickl, S. and J. Rumpf (forthcoming): *Dialectometric concepts of space: towards a variant-based dialectometry.* Proceedings of the FRIAS conference “Dialectological and folk dialectological concepts of space – current methods and perspectives in sociolinguistic research on dialect change”, Freiburg, 27–29 November 2008.

Rumpf, J. et al. 2009. Structural analysis of dialect maps using methods from spatial statistics. *Zeitschrift für Dialektologie und Linguistik* 76/3: 280–308.

SBS: König, W. (ed.) 1996–2009. *Sprachatlas von Bayerisch-Schwaben* (Bayerischer Sprachatlas. Regionalteil 1). 14 vols. Heidelberg: Winter.

Scott, D. W. 1992. *Multivariate Density Estimation: Theory, Practice, Visualization*. New York: Wiley.

Silverman, B. W. 1986. *Density Estimation for Statistics and Data Analysis.* New York: Chapman & Hall.

Jonas Rumpf and Volker Schmidt • Institute of Stochastics
Universität Ulm • 89069 Ulm • GERMANY

jonas.rumpf@uni-ulm.de
volker.schmidt@uni-ulm.de

Simon Pickl, Stephan Elspaß and Werner König • Universität Augsburg
Deutsche Sprachwissenschaft • Universitätsstraße 10 • 86159 Augsburg
GERMANY

simon.pickl@phil.uni-augsburg.de
stephan.elspass@phil.uni-augsburg.de
werner.koenig@phil.uni-augsburg.de