Postprint. Appeared in: *Literary and Linguistic Computing* 28(1): 63–81.
https://doi.org/10.1093/llc/fqs050

Lexical Meaning and Spatial Distribution
Evidence from Geostatistical Dialectometry

Simon Pickl

1. Motivation

In the German dialectological literature of the last century, there are several recurring ideas concerning the relation between the meaning of a lexical variable (i.e. a semantic concept)[1] and the way in which its variants (i.e. a concept's designations) are distributed in space. They are mainly based on observations of regularities in the corpora of maps the authors in question had at hand, and can be seen as attempts to systematize the vast amount of seemingly unstructured lexical distributions. However, the accounts that they give are somewhat fragmentary, as they are mostly based on individual observations that lack the support of hard quantitative evidence, and the explanations provided are, on the whole, not very methodical.[2] The main themes that can be distinguished in the relevant literature are the following:

- Concepts belonging to the fields of *agriculture*, *crafts* or *household* tend to have a richly varied terminology with variants that are clearly distributed in space.

  The reason that is given for this is mainly cultural: Purportedly, in these domains cultural practices are spatially limited and the readiness to adopt cultural influences varies from region to region, which readily leads to the formation of distinct variant areas.

  Cf. Wenzel (1930, p. 13); Bach (1950, p. 174).

- Concepts that belong to the field of *child language* tend to have a broad range of variants that are dispersed irregularly in space.

  It is assumed that children have a strong tendency to alter words playfully due to their play instinct and their untamed imagination (not to be confused with the idea of inadvertent imperfect learning or altered replication[3]).

  Cf. Wenzel (1930, pp. 59, 66); Bach (1950, pp. 173–174); Hildebrandt (1983, p. 1364); Lötscher (2005, p. 304).

- Concepts that implicate a certain *emotional involvement*, are *expressive*, or convey notions of *affect*, have a tendency to develop small, heterogeneous areas.

  It is argued that expressiveness and emotionality foster creative innovations and spontaneous modifications of existing forms, which leads to the beginnings of spatial diversity,[4] also impeding the emergence of uniform areas. Examples typically given are obscene expressions, terms of superstition, and – strangely – some plant and animal names, which again is ascribed to the imaginative power of children.

  Cf. Wenzel (1930, p. 59); Bach (1950, pp. 170–174); Hildebrandt (1983, p. 1364); Lötscher (2005, p. 310).

- More *general* and *prototypical* concepts tend to have larger, more homogeneous areas than specific or marginal ones.

  One of the reasons given for this is that it is easier to distinguish between prototypes than between different entities of similar makeup, simply because the latter tend to share more common characteristics. For instance, it is rather easy to tell apples and pears apart, but it is not so easy to distinguish between different apple varieties. Thus, more specific terms have a disposition to be mixed up easily because the concepts they refer to are confused easily, which results in heterogeneous geographical distributions. Prototypes and general terms, on the other hand, represent rather clear-cut and distinct concepts, which allows them to develop regular spatial distributions.

  Cf. Bach (1950, p. 170–172); Hildebrandt (1983, pp. 1333, 1364); Lötscher (2005, pp. 304–305, 310; 2010).

There are several other alleged patterns of associations between meaning and spatial distribution brought up by different authors, but there are noticeable differences between them and they are sometimes even contradictory. However, most authors seem to be in general agreement about the four different patterns named above. These patterns illustrate the way in which the issue of spatially relevant lexical properties has been dealt with so far theoretically. If the claims made by Wenzel, Bach, Hildebrandt and Lötscher are justified, it should be possible to substantiate them by way of statistical hypothesis testing. It should also be possible to test whether there are any connections between meaning and spatial distribution that go beyond those postulated so far.

To this end, it is necessary to quantify the spatial patterns that are expected to be of relevance. Lee and Kretzschmar (1993) have proposed measuring the 'spatial clustering' of variants. Then, 'Combinations of words may be put together either by experiments or by prior understanding of relations among words. For example, it may be possible to observe significant spatial clustering of whole groups of variants from semantic fields such as the vocabulary of farming, or of the household.'[5] Speelman and Geeraerts (2008)[6] have taken a slightly different approach: they measure the 'heterogeneity' of lexical spatial distributions by taking into account the number of variants of a concept, the spatial expanse of their occurrence and their dispersion across space. By means of linear regression analysis, they can show that, for instance, the 'negative affect' of a concept on the whole increases the heterogeneity of its distribution, which partly corroborates the third of the assertions quoted above. It appears, however, that there are not only one, but two characteristics that could play a role in assessing spatial distributions: the size of the variants' areas of occurrence, and the degree of interference from other variants within these areas. For this reason a more detailed analysis of geographical variation might prove worthwhile.

This article attempts to shed new light on the question of whether and to what extent there is a relation between a lexical variable's meaning and the spatial distribution of its variants. By looking at the variation of a lexical variable, not at the meanings that a single word can have, this question is tackled from an onomasiological perspective. While this study is explicitly focusing on lexical variation, this is not a typical area of interest in dialectology (esp. of German), primarily because variation on the lexical level is generally perceived to be less systematic than on other linguistic levels (e.g. phonetic variation).[7] One of the objectives of this article is to investigate whether there is a systematic dimension to lexical variation in space, and whether there is a relation between the meaning of a concept and the way in which its designations are distributed in space. To this end, indices are calculated that describe the size of the variants' areas and their internal uniformity (Sections 3 and 4). It is then tested whether significant values of these indices occur in particular subject areas (Section 5 and 6). Subsequently, the findings are used to set up an explanatory framework that allows them to be accounted for in terms of language variation and change in space (Section 7). It should be emphasized that the focus of this study is mainly on methodological advancement, as it aims to provide the necessary tools for testing and interpreting relationships between linguistic properties and types of spatial distributions.

2. Data

The analyses discussed in this article have been performed using data from the *Sprachatlas von Bayerisch-Schwaben* (SBS), which was developed and compiled at the University of Augsburg under the direction of Werner König during the years 1984–2005. Comprising 14 volumes with a total of approximately 2,700 maps, it documents the dialectal variation in the administrative region of Bavarian Swabia (southern Germany) and adjoining areas. For the elicitation of the data, between one and six informants were interviewed at 272 record locations. Intended to record the most original state of the dialects still accessible, the typical informant is born between 1900 and 1930 and has a farming or crafting background (men and women were interviewed indifferently).[8] The onomasiological lexical variation was elicited using a catalogue of questions, each aiming at a specific concept or lexical variable. Typically, each question regarding a lexical item in the questionnaire corresponds to a lexical map in the SBS.[9]

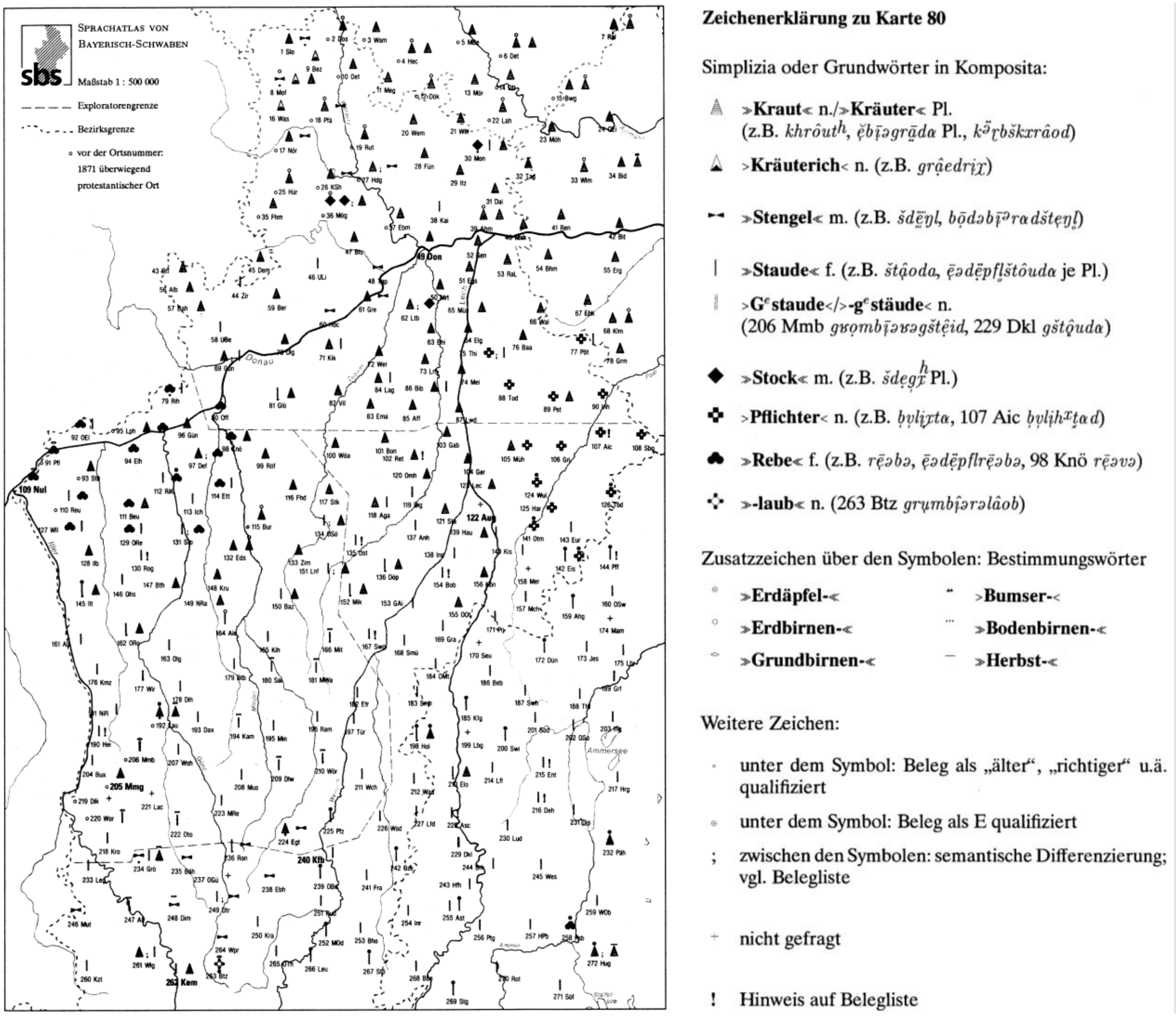

Fig. 1: Original point-symbol map "Kartoffelkraut" ('potato haulm') of the SBS (vol. VIII, p. 294)

Fig. 2: Legend for the original point-symbol map "Kartoffelkraut" ('potato haulm') of the SBS (vol. VIII, p. 295)

Figure 1 is a reproduction of an original point-symbol map from the SBS, the map for the lexical variable 'potato haulm' ("Kartoffelkraut"). Fig. 2: Legend for the original point-symbol map "Kartoffelkraut" ('potato haulm') of the SBS (vol. VIII, p. 295) is the legend for this map, which gives explanations for the symbols used and some additional information. Please note that the symbols only indicate whether a variant has been recorded at a location or not; information about their rank or relative frequency is not given.

For the present study, a sub-corpus of 735 lexical variables has been selected, which covers the majority of the onomasiological lexical maps of the SBS. For the categorization of the raw data with respect to lexical criteria, the original classification performed by the authors of the SBS was used, who arranged the individual records in variants with different degrees of abstraction (e.g. by denoting them with similar symbols), of which the highest degree was used for the present purpose: Thus, lexically related records (e.g. forms that consist of the same lexemes but have a different morphological configuration, such as *Kraut* and *Kräuterich* in the example above) were assigned to the same variant; purely phonetic differences were ignored completely.[10] As the methods that are introduced in Sections 3 and 4 require nominal-scale lexical data, taking into account gradual relatedness was not expedient. For further processing and analysis, the data they were converted into an SQL database format. In the database, one or more entries per variable are assigned to each of the 272 localities. Each entry corresponds to one distinct response given by the informant(s). Thus, whether or not a variant has been recorded at a location is documented, but no further information, e.g. about the variants' relative frequencies, is available. Accordingly, the working data in the database consist of information about the occurrence or non-occurrence of variants at 272 individual locations.

The locations being geo-referenced, the individual records can be related to each other in terms of geographical proximity, which is essential in order to make statements about the geographical distribution of the recorded variants and to discern meaningful distributional patterns. These goals are achieved by applying methods from Geostatistical Dialectometry.

3. Geostatistical Dialectometry

Dialectometry has acknowledged the fact that the information obtained in dialect surveys is sometimes unreliable, in the sense that individual records are susceptible to all sorts of sources of errors, to the extent that the data about individual variant distributions can be so sketchy that they are unsuitable for analysis. Especially if one assumes that a considerable amount of local variation is to be expected at a given point in space, a small number of records per location are insufficient material for the assessment of its lectal variation. The approach that traditional dialectometry takes to cope with this problem is to aggregate many variants' geographical distributions in order 'to compensate for the noisiness of individual distributions',[11] which is also a means of finding general spatial patterns in the data. By taking many distributions into account at the same time, the weak points of their individual representations in the data are balanced out. However, this way of proceeding renders it impossible to make any statements about the relations between the distributions of single variables or their alignment with extra-linguistic constraints,[12] as the aggregation reduces the many dimensions of diatopic variation to only one that accounts for the overall dialectal variation but not for the inter-feature variability within the collection of variables.

The recent advances in the field of Geostatistical Dialectometry[13] that have been achieved in a collaboration of the Department for German Linguistics at the University of Augsburg and the Institute for Stochastics at Ulm University[14] are the result of an effort to eliminate the blind spot of dialectometry concerning individual variants' distributions.[15] Geostatistical Dialectometry tries to take the individual geographical configurations of linguistic variables and their variants into account.[16] It also acknowledges the need to compensate for the unavoidable inaccuracy inherent in raw geolinguistic data, which fails to reproduce the subtleties of geographical and local variation, but its approach aims to preserve the features' individuality and the whole of the relevant variation. Treating the records as mere statistical samples, it tries to assess a record's reliability by taking into account the variants recorded in the area surrounding the record in question. Thus, the variance in the data is stabilized not by utilizing evidence from other variables, but evidence from other records of the same variable in the record's neighbourhood. Stray outliers are regarded as less reliable sources of information than records that are surrounded and thus "backed" by records of the same variant. This is done by applying intensity estimation (which is closely related to kernel density estimation) to the geo-referenced data points.[17] This method utilizes the degree of geographical proximity to estimate the mutual relevance that the records at two locations have for one another in order to calculate an occurrence probability (or estimated relative frequency) for each variant at every location. This way, a relative-frequency *graded area-class map*[18] can be generated out of non-frequency data recorded at isolated points in the plane, providing the basis for further analyses. (Instead of using geographical distance, linguistic distance measures can also be employed,[19] which prevents dialect islands and sharp dialect boundaries from being "smoothed over". In this article, however, geographical distances are used, as the aim of the procedure is the quantification of very general distribution patterns and not the detailed reproduction of variant distributions.)

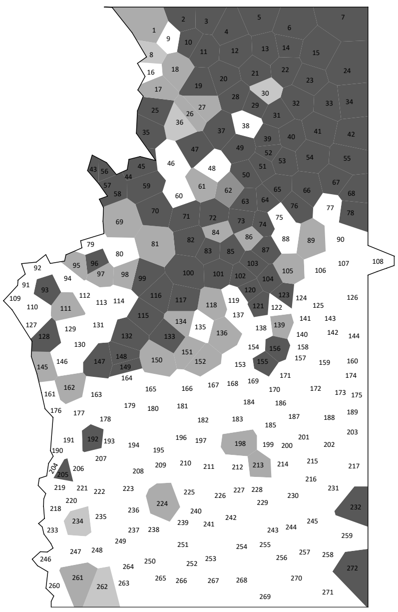

Fig. 3: Occurrence map of the variant *Kraut* of the variable 'potato haulm' from the SBS

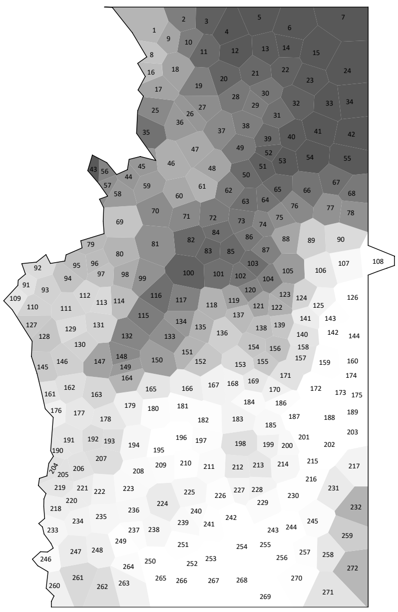

Fig. 4: Intensity map of the variant *Kraut* of the variable 'potato haulm' from the SBS

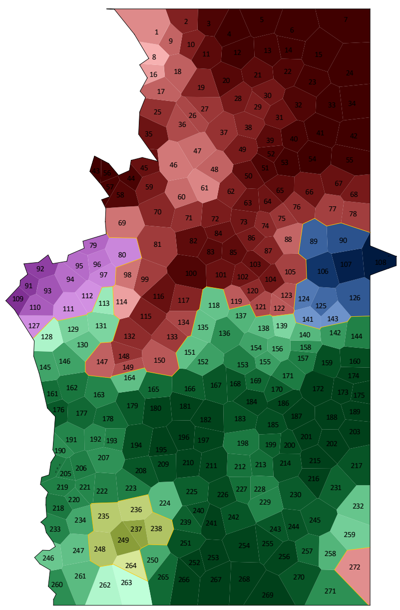

Fig. 5: Graded area-class map of the variable 'potato haulm' (combination of the intensity fields of all its variants)

Figure 3 shows the occurrences of the variant *Kraut* of the variable 'potato haulm' in the data of the SBS (cf. the triangle symbols in Fig. 1).[20] The differences in the shades of grey represent the relative proportion of total *Kraut* records at a location: The darkest locations have exclusively *Kraut*, in the lighter cells *Kraut* has been recorded alongside other variants (in differing proportions), and in the white cells *Kraut* has not been recorded at all. These proportions, ranging between zero and one, are a very rough estimate of the expected relative frequency of the variant.

In Fig. 4, the intensity field, consisting of the estimated occurrence probabilities of *Kraut*, is depicted, which is a much more sophisticated approximation of its relative frequencies, as it integrates the information about the presence of *Kraut* in the nearby locations. It is thus a geographically informed estimation of local variant frequencies on the basis of records that do not give any information about frequency if considered individually.

Figure 5 combines the intensity fields of *Kraut* (red) and the remaining six variants of 'potato haulm' (*Staude*, *Stock*, *Rebe*, *Pflichter*, *Stengel*, *Laub*), in such way that each cell is assigned to the variant that has the highest intensity at that point. In this example, red stands for *Kraut*, i.e. *Kraut* is the most probable (or the "dominant") variant at the locations that are coloured red. As can be seen, only five of all seven variants appear as dominant somewhere on the map (*Kraut* (red), *Staude* (dark green), *Rebe* (violet), *Pflichter* (blue), *Stengel* (light green)), the other two (*Stock*, *Laub*) are too dispersed or too sparse to show up at all (cf. Fig. a), even though they are latently present: they are globally "recessive". They do, however, cause a weakening of the dominant variants, which appears as a lightening of the shades of colour. The combination of the intensity fields of all seven variants in one area-class map fails to reproduce the information about the intensity of locally recessive variants; this can only be inferred through the lowered intensity values of the dominant variants. The orange lines in Fig. 5 are the boundaries of the variants' dominance areas. They are not isoglosses in the classical sense, in that they divide clear-cut, uniform variant areas, but rather they mark the centre of a gradual transition zone, which is probably much closer to linguistic reality:

> Closer investigation would probably reveal that most speakers over the whole area know both words, and in some areas the two words are interchangeable, perhaps with a preference for one over the other [...]. The isogloss, therefore, does not mark a sharp switch from one word to the other, but the center of a transitional area where one comes to be somewhat favored over the other.[21]

Graded area-class maps like Fig. 5 are, in a sense, geographically "smoothed" versions of the original data; in this sense they can be criticized for not reflecting the original data faithfully, but that is not their purpose. Instead, they provide an abstraction from the original data by emphasizing general, overall patterns of the spatial variation of a single variable that are captured neither by the original data nor by aggregative dialectometrical techniques. These patterns and the intensity data underlying them, i.e. the estimated frequencies of the variants, are the basis for the analyses in this study and a whole range of further methods of analysis that are currently being developed in Augsburg and Ulm.[22] One of their applications concerns the association of certain types of spatial distributions with properties of the linguistic features, in the present case, with lexical meaning. The intensity data that are obtained by the procedure described in this section enable us to measure certain characteristics of spatial distributions that are the basis for statistical tests.

The fact that intensity estimation yields estimated relative frequencies raises the question of how accurate these estimations are. Two kinds of validating the estimations are conceivable. The first possibility would be to revisit one or a few of the locations and to do a more exhaustive survey that allows making empirical statements about relative frequencies, which could then be compared to the estimated frequencies. In our case, however, this option is ruled out because the original elicitation dates too far back to obtain a comparable informant sample today. The second possibility, which is currently being implemented, is to apply a bootstrapping method that validates the predications of the intensity estimation by calculating intensities for locations whose data is temporarily ignored. Then the result can be compared to the original data of that location. Applied successively to all locations in a map, the average agreement between estimation and original data gives an idea of how reliably the intensity estimation performs. This procedure can also be used to validate different parameters of the method.

4. Distributional Characteristics

As mentioned in Section 1, there are two characteristics of spatial distributions that are of interest in the current investigation. One is the overall spatial expanse of variants, i.e. the size of the areas they cover; the other is the uniformity of these areas, i.e. the presence or absence of interference from other variants within them. The intensity data obtained using the method described in the previous section allows us to calculate indices that express exactly these characteristics in a simple way. I call the former the *complexity*, the latter the *homogeneity* of the distributions of a variable's variants in space.

1. **Complexity**

   With the term *complexity* (C) of a variable's distribution I refer to the degree to which space is split up into variant areas, or its onomasiological spatial "fracturedness". This is not exactly the same as the average size of the variant areas, because the areas can themselves be fractured, e.g. if they have very jagged contours. C is very easy to derive from the intensity data, which becomes evident if one looks at an area-class map like that in Fig. 5. The orange lines that delimit the variants' areas of dominance are a crucial indicator for the fracturedness of the map. The longer they are, the more fractured the map is. Thus, it seems natural to measure the total length of boundary lines on a map in order to quantify the degree to which it is divided into smaller areas. This gives an idea of the large-scale amount of geographical variation in a distribution, ignoring, however, the amount of small-scale "fluctuation" within the areas, which is captured by the so-called homogeneity of a distribution (see next paragraph). In order to make the complexity comparable across different areas of investigation, the total boundary length is divided by the surface of the study area, which is in our case 11,315 km^2. The result has the unit km^{-1}, and thus indicates the number of boundaries you would cross on average by walking 1 km within the study area. The map in Fig. 5 has a C-value of 0.037 km^{-1}, the theoretical range being between 0 km^{-1} and – for our area – a theoretical 0.764 km^{-1}, but values above 0.10 km^{-1} are quite rare.

2. **Homogeneity**

   The *homogeneity* (B) of a distribution describes the amount of variation that occurs within the variants' areas of dominance. For example, for one variable the areas might describe the distribution pretty well because they are compact and very uniform in themselves, whereas for another variable the general picture given by the division of the map into areas might be disturbed by records of locally recessive variants. The interference from other variants is given, on a local level, by the intensity of the recessive variants. The local homogeneity, on the other hand, is quantified by the inverse, viz. the intensity of the dominant variant at the location in question. The overall homogeneity of a distribution, therefore, is calculated as the mean intensity of the respective dominant variant. It is visualized as the overall "lightness" of a map. The map in Fig. 5, for example, has a B-value of 0.79, the theoretical range is between 0 and 1.

To give an idea of the visual impression of these two indices on the maps, the following figure provides a few examples that cover the range of variability found in the data pretty well.

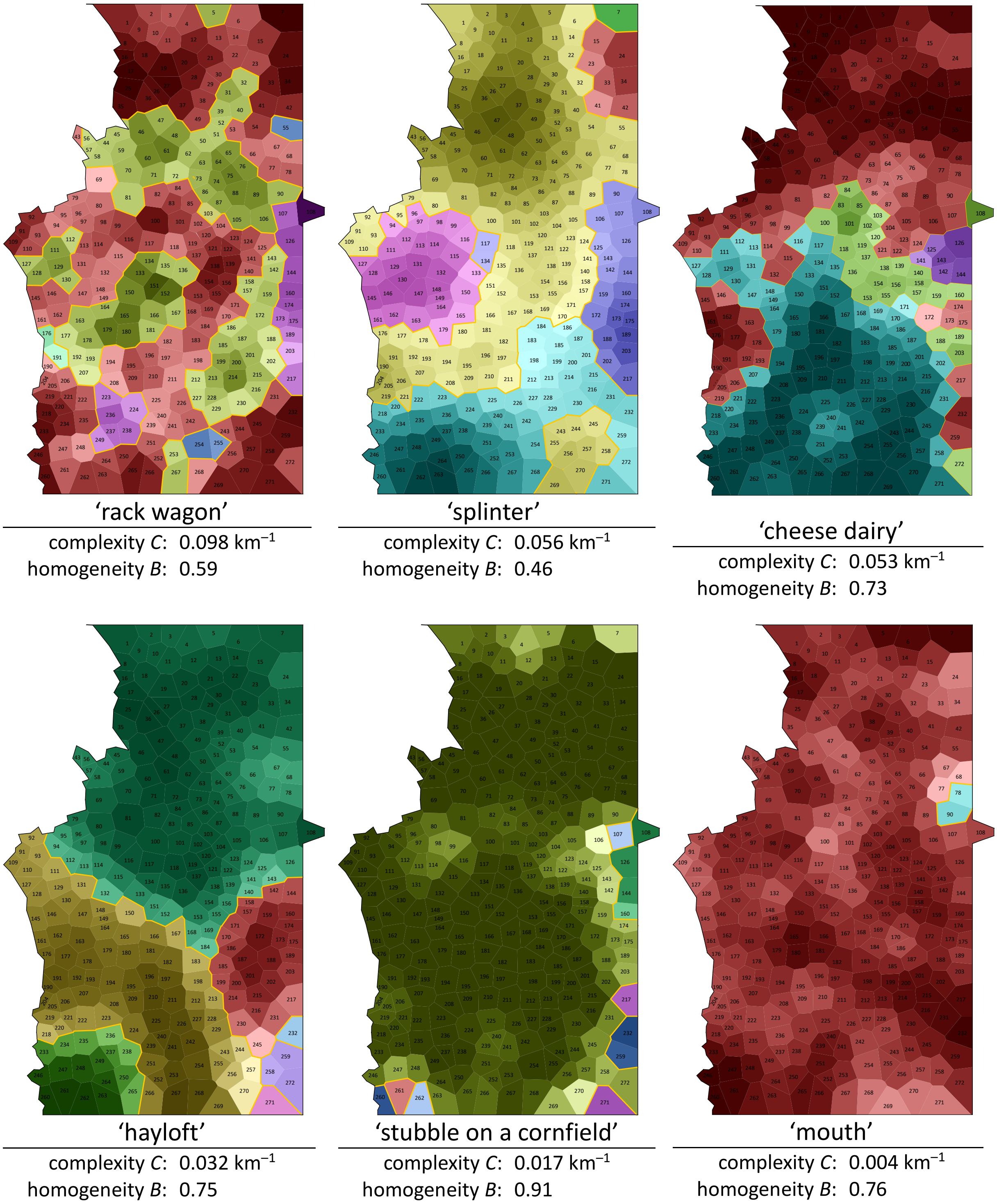

Fig. 6: Six examples from the SBS for maps with different values for complexity and homogeneity. For each map, different colours indicate different variants; matching colours across maps do not indicate identical variant.

As can be seen, complexity and homogeneity give a pretty good indication of the fracturedness and the intra-areal variation of these lexical distributions. As for complexity, the map for ‘rack wagon’ also shows why it is important to look not at area sizes, but at boundary lengths instead: even though there are not many dominant variants, which means that the space they occupy individually is rather large, the areas are broken up and jumbled to a great extent. Concerning homogeneity, the relationship between the overall lightness of a map (visualizing the interference from recessive variants) and its homogeneity becomes clear if one compares, for instance, the maps for 'splinter' and for 'stubble on a cornfield'.

Unsurprisingly, the two indices are not unrelated. With a correlation coefficient of $\varrho(C,B) = -0.65$, higher values of complexity tend to have lower values of homogeneity, and vice versa. There are two reasons for this. Firstly, the two indices measure geographical variation at two ends of a spectrum from small-scale (intra-areal) variation (= homogeneity) to large-scale (inter-areal) variation (= complexity). As there is no natural division between small-scale and large-scale variation, both blend into each other and therefore influence each other. As isoglosses are the dividing lines between transition zones, a higher complexity leads to an increase in transition zones, which influences the overall homogeneity negatively. Secondly, the latter effect is augmented by the process of intensity estimation, as it smooths seemingly sharp distinctions along the isoglosses of a distribution. This artefact of the method is not undesirable, however, because such sharp transitions are mostly due to data limitations.[23]

With C and B, two major characteristics of spatial distributions are quantified. Applied to all 735 lexical items in the corpus, we get 735 pairs of these indices. Are their values distributed randomly across the corpus, or are there any regularities as to which concepts tend to have higher or lower values? Are there any linguistic properties of variables that can be identified as constraints of their spatial distributions? Motivated by the statements reported in Section 1, the next section will use statistical methods of hypothesis testing to determine whether any particular subject areas yields significantly high or low values for complexity or homogeneity, and whether the findings corroborate the cited authors' assertions.

5. Statistical Testing

In the previous section, we defined the indices of C and B, which will be the dependent variables for the testing. We have to choose those aspects of lexical items for independent variables in whose effect on C and B we are interested. For the present study, the variables in the corpus were divided into subject areas which represent different domains of everyday life (cf. Table 1). This turned out to be relatively easy and straightforward to do, as the printed version of the SBS already features such a categorisation. These categories or "topics" are rather loosely defined but reflect the domains in which a variable would typically appear in everyday life quite well and seem to be a suitable expedient. Although this categorisation has clear deficits (e.g. one of the groups is called *miscellaneous*), it is difficult to come up with cogent criteria for a categorisation that allows establishing hypotheses about effects on spatial variation. It may prove rewarding to use different and more sophisticated categorisations to see how they perform in comparison to the results of this study.[24]

As subject area membership is a categorical or nominal-scale variable, the use of parametric tests is ruled out. Instead, non-parametric tests have to be used, which often require no assumptions regarding the distribution of the data. One common approach is the use of so-called Monte Carlo methods. The basis for Monte Carlo methods is repeated random sampling in order to determine the range that can be assumed to be random. If the real sample to be tested has an extreme value compared to the simulated random samples, this indicates that it is not a random sample following the same rules, but that other constraints are exerting an effect on it.

For our application, random sampling means that a high number (e.g. 999) of random groups of lexical variables is selected from the corpus. Each of these groups must contain the same number of variables as the group to be tested (i.e. a group with variables of one of the subject areas). The value to be tested is the mean value of C or B within a group. More precisely, it is tested whether the mean value of C or B of the maps in a real sample (= one of the subject areas) is equal to the means found in random samples, or whether it is significantly higher or lower than in the real sample. These means are calculated for each random group and for the test group, and then ranked in ascending order. If the position of the test group in this sequence is very high or very low, it is concluded that it has a significantly high or low mean value of C or B. In order to decide this, a significance level α has to be specified; 5 % is a typical threshold. If, for instance, the mean value of C in the test group (a subject area to be tested) is within the lowest 5 % of 999 random groups, it is deemed to be significantly low. In statistical terms, the p-value is the proportion of groups (including the test group) that have a value lower or equal to the test group. If the p-value is greater or equal to 95 %, the group has a significantly high value of C or B.

6. Results and Interpretation

The following table (Table 1) shows the p-values for mean C and mean B of all subject areas in the corpus.

Table 1 p-values obtained using Monte Carlo simulation (number of random sampling: 999). Numbers in brackets are the sizes of the groups.

	complexity C	homogeneity B
the human body (34)	27,5%	84,2%
physical & emotional expressions (46)	41,4%	65,6%
community (44)	**2,6%**	89,8%
clothing (9)	22,6%	26,1%
the farmhouse (22)	92,3%	**3,0%**
habitation & furnishing (16)	9,7%	67,5%
weather phenomena (14)	**96,6%**	**0,1%**
wild animals (7)	44,3%	17,8%
plants, fruits, vegetables (31)	33,6%	58,9%
must production (3)	80,6%	29,0%
flowers (4)	32,0%	65,1%
children's games (16)	**98,0%**	**4,1%**
household (37)	**95,1%**	27,2%
nutrition, cooking & baking (45)	83,8%	23,9%
farmers & agricultural workers (4)	**99,5%**	**0,6%**
time & greeting (24)	21,1%	**98,0%**
adverbs (8)	54,2%	21,0%
cattle & dairy (62)	77,6%	17,9%
pig, goat, sheep, horse (38)	43,6%	80,6%
poultry & bee keeping (13)	**2,6%**	93,8%
pets (10)	91,6%	56,0%
terrain (11)	34,6%	41,9%
soil & agriculture (26)	17,0%	56,1%
fertilisation (11)	**4,3%**	**95,1%**
hay harvest (45)	5,8%	79,7%
crop (48)	87,8%	36,3%
wood & timber (42)	8,1%	**99,0%**
fences (9)	63,2%	52,4%
transport (41)	74,2%	39,5%
baskets, vessels and carrying frames (12)	86,0%	5,5%
miscellaneous (3)	46,9%	39,2%

significantly low | not significant | significantly high

As is to be expected, most of the subject areas show unremarkable p-values for both indices. Especially in the case of the somewhat awkward groups *adverbs* and *miscellaneous*, this is little surprising.

Strictly speaking, Table 1 contains the results of a total of 124 single statistical tests: 31 subject areas were tested at both ends of the scale with regard to two indices (31·2·2). The risk of obtaining a type I error (i.e. an accidental significance)[25] is therefore a lot higher than in the case of one single test. Assuming that the single tests for one index and at one end of the scale are completely independent, a theoretical number of 1.55 type I errors would be expected to occur when performing 31 tests ($31 \cdot \alpha = 1.55$) for one index and at one end of the scale. In the results of the tests as shown in Table 1, there are between three and four groups with significant p-values for each index and at each end of the scale, which is remarkably more than the theoretical 1.55 significances that would be expected randomly as type I errors. Because they are in fact not independent but interrelated to a great extent (cf. Section 4), the actual expected number of type I errors is even lower. This is a first indication that there is a relation between the subject areas and the C- and B-values that goes beyond random associations. Nonetheless, this means that up to half of the significances shown in Table 1 are presumably false positives. Since it is impossible to determine which ones are "true" and which ones are "false" positives, it should be attempted to give a plausible interpretation for all of them. The statistical significances that the tests yielded, though they are to be taken with a grain of salt, are an incentive to look closer.

Community has a significantly low p-value for C and a high, though not significant value for B. It can be concluded that *community* contains maps that have rather large, smooth, uniform areas. In this case, the explanation is quite straightforward. In this group there are many variables that concern social relationships (e.g. kinship relations), or other socially relevant concepts. Naturally, variables of this kind will tend to be a regular part of everyday communication, particularly in chitchat-type conversations and especially across village borders, and thus they have a special disposition for interpersonal accommodation. This fosters spatial diffusion, differences are levelled out quickly, and this, ultimately, leads to large, uniform areas.

The significantly low p-value for B in the group *the farmhouse* is accompanied by a high, though not significant value for C, which is at least partly due to the correlation of the two indices. The high complexity probably has to do with the domestic and place-bound nature of the topic, which does not encourage spatial diffusion, but rather impedes it. The low homogeneity is to be seen in the context of a lack of conceptual and terminological distinctness that seems to occur with the items in this group. It is stated quite often in the corresponding commentaries in the SBS that the informants 'did not always know how to distinguish things, and that the expressions' semantic scope could not be delimited precisely'[26]. This can be due partly to material differences in the way houses are built, partly to the fact that the different parts of the house are not always distinct from each other, which can lead to differences in the division of the building into conceptual parts.

Both indices for *weather phenomena* have significant values: the maps are very complex and very heterogeneous. This may seem surprising at first, because weather is a regular part of everyday conversations and has a central function in small talk, which would rather suggest a stronger tendency for diffusion and low complexity. If one looks closer, however, it again appears that many of the concepts in this topic are not so clear-cut that a distinct representation in the lexicon is probable. Instead, some of the concepts rather form part of a continuum, such as the variables 'drizzle', 'heavy rain', 'snowstorm', 'soft hail', and 'hail'. Given the vague semantic scope of these variables, it is not surprising that the records show a high degree of terminological variability or 'referential variance'[27], which inevitably leads to maps of reduced homogeneity. What is more, one of the alleged mechanisms described in Section 1 might play a role here. Weather is a field that often gives rise to emotional reactions. Emotionally charged concepts are assumed prone to re-motivation and re-innovation. Their emotive function becomes manifest in formal expressivity, which is conditioned by affect and spontaneity. This can lead to pronounced onomasiological and geographical diversity, first on a local and regional level (as expressed by B), and as a result of spatial diffusion also on a larger scale (as expressed by C). This interpretation can also explain why the variables for 'to drizzle' and 'to rain heavily' have much higher C-values than the much less affective wind directions.

The domain of *household*, quite naturally, comprises concepts that concern stationary activities. These activities are usually performed at home, in the house, and often alone. This promotes geographically diverse forms, as the requirements for a wider spread of fewer variants, such as mobility of the speakers (at least in the relevant situation) or socially relevant character of the concepts, are not met. Likewise, the group *children's games* also does not involve mobility of the speakers: the concepts are socially relevant, but the social interaction is confined to a small spatial perimeter, as children usually play predominantly in the neighbourhood, and they learn how to play tag, for example, in the yard or in the street. This, again, fosters the emergence of variants with a small spatial expanse, which finds expression in high C-values. It is noteworthy that B also reaches a significant value, for which its negative correlation with C cannot be the only explanation, because in the case of *household*, the p-value for B is far from significant irrespective of a similarly high p-value for C. Perhaps here a mechanism described by Wenzel (1930), Hildebrandt (1983) and Lötscher (2005), who describe children as especially active and emotional innovators, actually comes into play (cf. Section 1). They explain uneven distributions of child-related variables as a result of children's play instinct and imagination, which lead to spontaneous, emotionally motivated forms.

In the case of the (rather small) group *farmers & agricultural workers*, we find significant results for both C and B. The variables that contribute to the especially high C- and low B-values concern the name for a farmer with respect to the size of his holding, e.g. 'smallholder' or 'large-scale farmer'. It seems that we have to do with 'referential variance' once again, since the categories that are applied are not clear-cut, but rather fuzzy segments of a continuum. The borders between them are unclear, which leads to differences in the way the continuum is divided into concepts. The instability of concept brings about instability of expression, resulting in a multitude of re-motivations and nonce words,[28] which can reduce the overall homogeneity. Another possible explanation for the extreme values is again emotionality. The size of a farmer's holding, which is closely linked to his economic situation and social status, can often lead to derogative, scornful or envious expressions (e.g. the pejorative terms *Grattler* or *Pfröpfler*). As described in the context of *weather phenomena*, the emotional potential of concepts is assumed to entail high onomasiological diversity.

The group *time & greeting* has a very high, significant p-value for homogeneity (B), which means that the distributions in this group have very uniform, interference-free variant areas. The explanation for this is obvious: Salutations are linguistic phrases whose pragmatic function is almost exclusively phatic, while having a pronounced social relevance. They are vital elements of many conversations, especially with unknown or unfamiliar interlocutors. These are ideal preconditions for rapid spatial spreading and stable levelling, which accounts for the high $\bar{B}$-values of the concerned maps. Time, on the other hand, is not primarily a social subject area, but it does play an important role in interpersonal contacts: Whenever experiences are narrated or activities have to be coordinated, the conventionalized division of time provides a common frame of reference. This is even more important if the contact between two persons is of an irregular nature, because this renders an efficient arrangement crucial. Thus, highly conventionalized time terminology is essential for arrangements to be successful, which also contributes to high homogeneity in this group.

For the low complexity in the group *poultry & bee keeping* there is more than one possible explanation. Firstly, this group contains many variables that tend to have onomatopoetically motivated variants (e.g. the variables 'cock', 'to cackle', 'chick', 'to cluck'), which restrains the arbitrariness and thus also the diversity of innovations, which in effect reduces their number. Fewer variants share the same space, which results in larger areas on average. Secondly, the fact that this group belongs to the wider field of agriculture might have some relevance. One might think that there is little tendency for spatial diffusion in agricultural settings, as they are often associated with immobility, which would in theory lead to high complexity. On the other hand, the agricultural topics contain a lot of "outdated" material: agriculture played an important role in rural life in the past, but lost much of its significance during the last century. Many of the tools and techniques are no longer in use today, as agriculture became less common as an occupation and machines took over the roles of humans. Consequently, these obsolete linguistic variants had a lot of opportunity for long-term diffusion and levelling, which was largely unimpaired by the development of new innovations. Such a "late stage" of geolinguistic development, accordingly, is characterized by large, consolidated areas. As for *poultry & bee keeping*, this is only significant for C, whereas for *fertilisation*, which also belongs to agriculture, both C and B have significant p-values. In the case of *hay harvest*, the p-values display the same tendency, but do not reach a significant level. For the high homogeneity of *wood & timber*, it seems plausible to apply the same explanation. It is interesting that the low complexity and high homogeneity in the fields of agriculture and forestry contradicts the statements made by Wenzel and Bach (cf. Section 1). They argue that variables that refer to agricultural items or activities have a multitude of variants of limited geographical spread.[29] Furthermore, Wiesinger (2005, p. 1109), following Kranzmayer (1956, pp. 3–4), states that terms of agriculture and related subjects have a bias against long-distance mediation, which should result in small areas, i.e. high complexity. Such a connection cannot be substantiated with the lexical maps of the SBS; on the contrary: all of the agriculture-related topics that exhibit significant values point in the opposite direction.

7. Summary and Explanatory Framework

The interpretations of the results in the previous section can be argued about if considered individually, but the general picture seems quite plausible as the mechanisms described above can be condensed into relatively few basic tendencies. For instance, the concepts with a socially relevant or even phatic component and those that involve mobility of the speakers can be subsumed under the common heading "concepts with a disposition to spatial spread and levelling"; they are attributable to a principle that takes effect at the level of diffusion. Other principles, by contrast, work at other levels, for instance the "emotionality" of a concept, which influences its propensity for innovation.[30] Based on such levels of influence, the mechanisms can be categorised as follows (see also Table 2).

1. **Innovation affinity**
   What I call the innovation affinity of a variable is its tendency to develop new variants regularly. If a variable has a low innovation affinity, only rarely do new forms emerge. This can be the case, for instance, when the denoted concept is obsolete, e.g. an outdated tool or an activity that is no longer performed; the linguistic variable itself is dying out. Existing conventions remain undisturbed and stable; this benefits homogeneity, but impedes the emergence (or maintenance) of complexity. A high innovation affinity is frequently found in the case of emotional concepts, which often show spontaneous and expressive neologisms. Many innovations emerge locally, interfere with existing homogeneity, and can diffuse into the environment, causing an increase in complexity.[31]

2. **Arbitrariness**

   Some variables allow only a limited range of possible variants: their arbitrariness is restricted. This is especially true in the case of onomatopoetically motivated expressions. This mechanism, which influences the process of innovation, leads to a reduced diversity of variants in relation to variables with unrestricted arbitrariness. While the innovation affinity can be the same as with other variables, there is an increased probability that two or more innovators come up with the same innovation. Consequently, the homogeneity of these variables is on average higher than that of otherwise comparable variables, because there are fewer competing variants locally. At the same time, the complexity is lower, because local innovations lead to identical variants, whose areas merge in the course of diffusion.

3. **Diffusion affinity**

   The diffusion affinity describes the disposition of a variable's variants to spread in space and – by way of selection out of competing variants – to be levelled to regionally uniform areas. Possible factors for a high diffusion affinity are social relevance and a phatic component of meaning, as well as speaker mobility. These lead to a quick diffusion in space, accelerating the process of spatial levelling and resulting in large and uniform variant areas, which entails low complexity and high homogeneity. The opposite is the case if a concept has no social relevance or if, say, an activity is place-bound and the typical speaker rather immobile. In that case, spatial diffusion is impeded, the areas remain small and for that reason alone are heterogeneous.

4. **Specificity of meaning**

   Specificity of meaning is situated in the mental lexicon, where the connections between linguistic forms and their conceptual representation are stored. These connections can be straightforward or rather fuzzy.[32] They can also vary from dialect to dialect and from speaker to speaker. What is more, the semantic vagueness can have a massive impact on data elicitation, as the surveyed items may not be represented by a uniform concept in the informants' mental lexicons, resulting in insecure and semantically inconsistent answers, and very often multiple responses. This clearly reduces the homogeneity of a lexical variable's variant areas drastically.[33] Complexity, on the other hand, should remain unaffected by this mechanism, as it is independent of geography and exerts its effect mainly on the level of data collection.

These four categories can be used to account for the variation in spatial distributions as outlined in Section 6. Each lexical variable has certain intrinsic characteristics (e.g. social relevance) that can have an effect on the spatial distribution of its variants, depending on which level a specific characteristic plays a role. An overview of the expected effects as outlined above is provided in Table 2.

Table 2 Factors for the distributional indices complexity (C) and homogeneity (B). The presumed effects of the factors were extracted from the interpretations of the statistical tests and their results in the previous sections. A plus-sign indicates an effect that leads on average to an increase of the respective index's values, a minus-sign stands for a decreasing effect (cf. Table 1).

			C	*B*
Innovation affinity	high	e.g. expressivity, emotionality, taboo	+	–
	low	e.g. obsolete vocabulary	–	+
Arbitrariness	free		+	–
	restricted	e.g. disposition for onomatopoeia	–	+
Diffusion affinity	high	e.g. mobility, social relevance, phatic component	–	+
	low	e.g. immobility, taboo	+	–
Specificity of meaning	clear			+
	fuzzy	e.g. unclear differentiation from other concepts, ambiguity		–

The findings from Section 6 can now be viewed anew in the light of this classification. In Table 3, it is shown how different characteristics inherent in subject areas (such as a tendency to terminological fuzziness or emotionality) trigger one (or more) of the four mechanisms outlined above, which likewise lead to higher or lower values of C and B. It is quite obvious that the strength of the effect is not predictable with this simple model, at best its direction. Moreover, C and B are not affected to the same extent, and in some cases the effect might go unnoticed as the values are too small to be significant.

Table 3 Topics in the SBS, significant mean index values and possible explanation. A plus-sign indicates an increasing effect on the respective index, a minus-sign stands for a decreasing effect. The highlighted cells indicate that the effect was found to be statistically significant (cf. Table 1).

community		*C*	*B*
social	diffusion affinity: high	–	+

the farmhouse		*C*	*B*
place-bound	diffusion affinity: low	+	–
terminological fuzziness	specificity of meaning: fuzzy		

weather		*C*	*B*
emotional potential	innovation affinity: high	+	–
referential variance	specificity of meaning: fuzzy		

children's games		C	B
place-bound	diffusion affinity: low	+	−
emotionality, play instinct, imagination	innovation affinity: high		

household		C	B
place-bound	diffusion affinity: low	+	−

farmers and agricultural workers		C	B
size of agriculture: emotional potential	innovation affinity: high	+	−
conceptual fuzziness	specificity of meaning: fuzzy		

time & greeting		C	B
socially relevant, phatic	diffusion affinity: high	−	+

poultry & bee keeping		C	B
onomatopoeia	arbitrariness: restricted	−	+
obsolete vocabulary	innovation affinity: low		

fertilisation		C	B
obsolete vocabulary	innovation affinity: low	−	+

wood & timber		C	B
obsolete vocabulary	innovation affinity: low	−	+

8. Conclusion

To sum up, it can be said that a clearly non-random distribution of significant complexity and heterogeneity values is observable in the semantic groups of the SBS. In both cases, seven out of thirty-one groups, or 22.5 %, show significant values (cf. Table 1), which is remarkably more than the 10 % of significances (5 % for each end of the scale) that were to be expected as type I errors. Individual details of the interpretations suggested in Section 6 might be arguable, but the model introduced in Section 7 provides a consistent and plausible explanatory framework. On the whole, it is indisputable that the meaning of a lexical variable can have a notable effect on the spatial distributions of its variants. This result also demonstrates the power of the (geo)statistical approach taken.

Funding and Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft (DFG), which funded the joint research project "New Dialectometry Using Methods from Stochastic

Image Analysis" of the Department of German Linguistics (University of Augsburg) and the Institute of Stochastics (Ulm University). My cordial thanks go to all of my colleagues who have contributed to this work, as well as to the editors and reviewers of this issue.

References

Bach, A. (1950). *Deutsche Mundartforschung. Ihre Wege, Ergebnisse und Aufgaben. Mit 58 Karten.* Heidelberg.

Francis, W. N. (1983). *Dialectology. An Introduction.* London.

Geeraerts, D., Grondelaers, S., and Bakema, P. (1994). *The Structure of Lexical Variation. Meaning, Naming, and Context.* Cognitive Linguistics Research, 5. Berlin/New York.

Geeraerts, D. (2010). Lexical Variation in Space. In Auer, P. and Schmidt, J. E. (eds), *Language and Space. An International Handbook of Linguistic Variation. Volume 1: Theories and Methods.* Handbooks of Linguistics and Communication Science, 30.1. Berlin/New York, 821–837.

Geeraerts, D. and Speelman, D. (2010). Heterodox concept features and onomasiological heterogeneity in dialects. In: Geeraerts, D., Kristiansen, G. and Peirsman, Y. (eds), *Advances in Cognitive Sociolinguistics.* Berlin/New York, pp. 23–40.

Goebl, H. (1984). *Dialektometrische Studien. Anhand italoromanischer, rätoromanischer und galloromanischer Sprachmaterialien aus AIS und ALF. Band 1.* Tübingen.

Goossens, J. (1969). *Strukturelle Sprachgeographie. Eine Einführung in Methodik und Ergebnisse.* Sprachwissenschaftliche Studienbücher, Zweite Abteilung. Heidelberg.

Grieve, J. (2009). *A Corpus-Based Regional Dialect Survey of Grammatical Variation in Written Standard American English.* Ph.D. Dissertation. Northern Arizona University.

Grieve, J., Speelman, D., and Geeraerts, D. (2011). A statistical method for the identification and aggregation of regional linguistic variation. *Language Variation and Change* 23, 193–221.

Heeringa, W. (2004). *Measuring Dialect Pronunciation Differences using Levenshtein Distance.* Groningen Dissertations in Linguistics, 46. Groningen.

Hildebrandt, R. (1983). Typologie der arealen lexikalischen Gliederung deutscher Dialekte aufgrund des Deutschen Wortatlasses. In Besch, W., Knoop, U., Putschke, W., and Wiegand, H. E. (eds), *Dialektologie. Ein Handbuch zur deutschen und allgemeinen Dialektforschung. Zweiter Halbband.* Handbücher zur Sprach- und Kommunikationswissenschaft, vol. 1.2. Berlin/New York, pp. 1331–1367.

Kranzmayer, E. (1956). *Historische Lautgeographie des gesamtbairischen Dialektraums. Mit 27 Laut- und 4 Hilfskarten in besonderer Mappe.* Wien.

Kronenfeld, B. J. (2005). Gradation and Map Analysis in Area-Class Maps. In Cohn, A. G. and Mark, D. M. (eds), *Spatial Information Theory. International Conference, COSIT 2005. Ellicottville, NY, USA, September 14–18, 2005. Proceedings.* Lecture Notes in Computer Science, 3693. Berlin.

Kronenfeld, B. J. (2007). Triangulation of Gradient Polygons: A Spatial Data Model for Categorical Fields. In Winter, S., Duckham, M., Kulik, L., and Kuipers, B. (eds), *Spatial Information Theory. 8th International Conference, COSIT 2007. Melbourne, Australia, September 2007. Proceedings.* Lecture Notes in Computer Science, 4736. Berlin.

Lee, J. and Kretzschmar, W. A. Jr (1993). Spatial analysis of linguistic data with GIS functions. *International Journal of Geographical Information Systems* 7(6): 541–560.

Leinonen, T. (2010). *An Acoustic Analysis of Vowel Pronunciation in Swedish Dialects.* Groningen Dissertations in Linguistics, 83. Groningen.

Lötscher, A. (2005). Areale Diversität und Expressivität im Dialektwortschatz. In Krämer-Neubert, S., Wolf, N. R. (eds), *Bayerische Dialektologie. Akten der Internationalen Dialektologischen Konferenz. 26.–28. Februar 2002.* Heidelberg, pp. 303–311.

Lötscher, A. (2006). Wortgestalt und areale Diversität. In Klausmann, H. (ed.), *Raumstrukturen im Alemannischen. Beiträge der 15. Arbeitstagung zur alemannischen Dialektologie. Schloss Hofen, Lochau (Vorarlberg) vom 19.–21.9.2005.* Schriften der Vorarlberger Landesbibliothek, vol. 15. Graz, pp. 141–150.

Lötscher, A. (2010). Der Flickenteppich als areales Muster in Dialektkarten. Alltagskommunikation und Diffusionstypen im Dialekt. Paper given at the 40th Annual Meeting of the Gesellschaft für Angewandte Linguistik e.V. (GAL), “SprachRäume”, 15.–17. September 2010, Leipzig.

Meschenmoser, D. and Pröll, S. (forthcoming). Fuzzy clustering of dialect maps using the empirical covariance. Grouping maps in dialect corpora based on spatial similarities. *International Journal of Corpus Linguistics*.

Meschenmoser, D. and Pröll, S. (submitted). Automatic Detection of Radial Structures in Dialect Maps: Determining Diffusion Centers.

Nerbonne, J. (2010). Mapping aggregate variation. In Lameli, A., Kehrein, R., and Rabanus, S. (eds), *Language and Space. An International Handbook of Linguistic Variation. Volume 2: Language Mapping.* Handbooks of Linguistics and Communication Science, 30.2. Berlin/New York, pp. 476–495.

Nerbonne, J. and Kleiweg, P. (2003). Lexical Distance in LAMSAS. *Computers and the Humanities* 37: 339–357.

Pickl, S. (2012). *Probabilistische Geolinguistik. Geostatistische Analysen lexikalischer Distributionen in Bayerisch-Schwaben.* Ph.D. thesis, University of Augsburg.

Pickl, S. and Rumpf, J. (2011). Automatische Strukturanalyse von Sprachkarten. Ein neues statistisches Verfahren. In Glaser, E., Schmidt, J. E., and Frey, N. (eds), *Dynamik des Dialekts – Wandel und Variation. Akten des 3. Kongresses der Internationalen Gesellschaft für Dialektologie des Deutschen (IGDD)*. Stuttgart, pp. 267–285.

Pickl, S. and Rumpf, J. (forthcoming). Dialectometric Concepts of Space. Towards a Variant-Based Dialectometry. In Hansen, S., Schwarz, C., Stoeckle, P., and Streck, T. (eds), *Dialectological and folk dialectological concepts of space.* Berlin.

Pröll, S. (same issue). Detecting structures in linguistic maps. Fuzzy clustering for pattern recognition in Geostatistical Dialectometry.

Rumpf, J., Pickl, S., Elspaß, S., König, W., and Schmidt, V. (2009). Structural analysis of dialect maps using methods from spatial statistics. *Zeitschrift für Dialektologie und Linguistik* 76(3): 280–308.

Rumpf, J., Pickl, S., Elspaß, S., König, W., and Schmidt, V. (2010). Quantification and Statistical Analysis of Structural Similarities in Dialectological Area-Class Maps. *Dialectologia et Geolinguistica* 18: 73–100.

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

Speelman, D. and Geeraerts, D. (2008). The Role of Concept Characteristics in Lexical Dialectometry. *International Journal of Humanities and Arts Computing* 2(1–2): 221–242.

Spruit, M. R., Heeringa, W., and Nerbonne, J. (2009). Associations among linguistic levels. *Lingua* 119(11): 1624–1642.

Viereck, W. (1986). Dialectal Speech areas in England: Orton's lexical evidence. In Kastovsky, D. and Szwedek, A. (eds). *Linguistics across Historical and Geographical Boundaries. In Honour of Jacek Fisiak on the Occasion of His Fiftieth Birthday.* Trends in Linguistics: Studies and Monographs, 32. *Volume 1: Linguistic Theory and Historical Linguistics.* Berlin/New York/Amsterdam, 725–740.

Wenzel, W. (1930). *Wortatlas des Kreises Wetzlar und der umliegenden Gebiete.* Marburg.

Wiesinger, P. (2005). Historische Grundlagen und Voraussetzungen der gegenwärtigen deutschen Wortgeographie. In Cruse, D. A., Hundsnurscher, F., Job, M., and Lutzeier, P. R. (eds), *Lexicology. An International Handbook on the Nature and Structure of Words and Vocabularies.* Handbooks of Linguistics and Communication Science, 21.2. Berlin/New York, pp. 1108–1137.

---

1 For a more detailed discussion of the relationship between lexical variables, variants and concepts, cf. Pickl 2012, 42–43.

2 Only Lötscher (2005; 2010) arrives at a theoretical framework by explaining his observations in terms of innovation and diffusion.

3 Cf. e.g. Croft (2000, pp. 44–46).

4 'arealer Diversität im Ursprung' (Lötscher 2005, p. 310).

5 Lee and Kretzschmar (1993, p. 554).

6 Cf. also Geeraerts and Speelman (2010).

7 Cf. Goossens (1969, pp. 53 and 69); Francis (1983, 20); Viereck (1986, 725).

8 Cf. SBS (vol. I, pp. 20–24).

9 For further information, cf. SBS (vol. I).

10 For details of the classification process, cf. Pickl (2012, 76–80).

11 Cf. Spruit *et al.* (2009, p. 1624).

12 Cf. Leinonen (2010, p. 38).

13 Cf. Rumpf *et al.* (2009; 2010).

14 The methods referred to and the results presented in this article were obtained in the joint research project "New Dialectometry Using Methods from Stochastic Image Analysis" (funded by the Deutsche Forschungsgemeinschaft) of the Department of German Linguistics (University of Augsburg) and the Institute of Stochastics (Ulm University).

15 Cf. Pickl and Rumpf (forthcoming).

16 A somewhat related geostatistical approch is taken by Jack Grieve, who uses the spatial autocorrelation measures Moran's *I* and Getis-Ord G_i* to test if significant spatial clustering is found in the data (cf. e.g. Grieve 2009; Grieve, Speelman, and Geeraerts 2001).

17 For the exact mathematical procedure, cf. Rumpf *et al.* (2009, pp. 285–290); Pickl and Rumpf (2011, pp. 272–275), Pickl (2012, pp. 85–101).

18 So-called graded area-class maps reproduce the "fuzziness" of a location's affiliation with the areas, making it possible to visualize transition zones between areas and allowing locations to have partial membership of areas (cf. Kronenfeld 2005; 2007).

19 For the definition of a "lexical distance measure" and an application of the procedure, cf. Pickl (2012, 101–111). Other (related) lexical distance measures are introduced by Nerbonne and Kleiweg (2003, 346–349) and Speelman and Geeraerts (2008, 227–228). For phonetic data, the so-called Levenshtein distance is very popular (cf. e.g. Heeringa 2004; Nerbonne 2010, pp. 480–481). One of the most well-known general distance measures is probably the method proposed by Goebl (1984, pp. 75–77).

20 The Voronoi tessellation that has been used for the visualization assigns each point in the plain to the closest data point.

21 Francis (1983, p. 5).

22 Cf. Rumpf *et al.* (2010); Pickl (2012); Pröll (this volume), Meschenmoser and Pröll (forthcoming; submitted).

23 Sharp transitions are, however, to be expected along major linguistic boundaries. Therefore, the effect can be reduced by using linguistic instead of geographical distances for the intensity estimation, which has been done with the lexical data of the SBS. For technical reasons, and in order to be concise, however, I will only use maps generated with geographical distances in the present article.

24 It can also be worthwhile to have a closer look at categories other than semantic ones. It is, for example, possible to test the effect of parts of speech (cf. Pickl 2012, pp. 142–144). For independent variables other than categorical ones (e.g. frequency of use) other test procedures like regression analysis have to be used.

25 A type I error occurs when a sample's test statistic is determined to be significant although its value is only a result of random variation. The probability of a type I error is exactly α, i.e. in this case 5 %.

26 'in der Sache nicht immer genau zu unterscheiden wußten und daß die vorhandenen Bezeichnungen in ihrem Bedeutungsumfang nicht genau einzugrenzen sind' (SBS, vol. VIII, p. 24).

27 'referentielle Varianz' (Hildebrandt 1983, pp. 1333, 1364).

28 Cf. Lötscher (2006, pp. 147–148).

29 Cf. Wenzel (1930, p. 13); Bach (1950, p. 174).

30 Also Lötscher (2010) emphasized the role of innovation and diffusion for the areal distribution of variants.

31 Speelman and Geeraerts (2008, pp. 230–233) have already established the influence of 'negative effect' on the heterogeneity of a variable's geographical distribution (cf. also Geeraerts and Speelman 2010, pp. 34–35).

32 For a more detailed discussion of problems of conceptual differentiation, cf. Geeraerts *et al.* (1994) and Geeraerts (2010). The influence of the independent variables 'lack of familiarity' and 'non-uniqueness', which are also associated with fuzzy conceptualisations, on the distributional heterogeneity of lexical variables has been shown by Speelman and Geeraerts (2008) and Geeraerts and Speelman (2010).

33 In principle the same observation has been made by Hildebrandt (1983, pp. 1333, 1364). The relationship between the prototypicity of a concept and the areal diversity of its linguistic representations according to Lötscher (2005, pp. 304–305, 310) also follows a similar line of reasoning.