

Article

Analyzing dialect (morpho)syntax in Austria: a non-aggregative dialectometric approach

Philip C. Vergeiner¹, Lars Bülow² and Stephan Elspaß³

¹Institute of German Philology, University of Munich, Munich, Germany, ²Institute of German Philology, University of Munich, Munich, Germany and ³Department of German Language and Literatures, University of Salzburg, Salzburg, Austria

Abstract

This paper aims to identify and analyze geographical patterns of (morpho)syntactic variation in traditional Austrian dialects using non-aggregative dialectometric methods (factor analysis). Based on a comprehensive dialect corpus obtained by direct dialect interviews including 163 speakers from 40 locations throughout Austria, our analyses of 79 variants of 30 (morpho)syntactic variables not only show geographical patterns in Austria's dialects, but also address the linguistic basis of the geographical structures revealed. In particular, the results show that variables at the morphology–syntax interface contribute most to geographical structuring. We argue that this finding is related to structural conditions of these variables and the historical development of the respective variants.

Keywords: (morpho)syntactic variation; Bavarian and Alemannic dialects in Austria; dialect syntax; dialectometry; factor analysis

1. Introduction

In traditional dialectology from the late nineteenth and early twentieth centuries, it was long assumed that (morpho)syntactic phenomena are spatially less structured than phonological, morphological, and lexical phenomena and do not show clear-cut regional distributions; see e.g. Glaser (2013) and Scheutz (2005) for critical assessments. Moreover, it was widely believed that morphosyntactic phenomena are very difficult to elicit through traditional dialectological methods (e.g. König et al., 2019:163). As a result, dialectologists have long neglected (morpho)syntax. However, the last few decades have seen increasing research interest in dialect (morpho)syntax. This becomes evident in many studies on single (morpho)syntactic phenomena (e.g. Schallert, 2014; Bülow et al., 2021b; Moser, 2021), as well as in large-scale dialect syntax atlas projects (e.g. the projects *Syntax of Hessian Dialects* (SyHD) and *Syntactic Atlas of German Speaking Switzerland* (SADS)). The results of these studies and projects confirm the existence of syntactic areas (see Birkenes & Fleischer, 2021). They show that “[i]n terms of areal variation ... dialect (morpho)syntax is not different, in principle, from what is known about areal variation in accents and dialect vocabularies” but “has a wider areal reach than phonological and lexical variation” (Kortmann, 2010:846). In contrast to the latter, syntactic variation is also “much subtler and less salient, less categorical, and in many cases a matter of statistical frequency” (Kortmann, 2010:846).

Despite the generally solid level of research into *individual* (morpho)syntactic features, it remains an open question as to

whether different (morpho)syntactic features *show patterns of co-occurrence* which result in different dialect areas, and if so, whether these areas correspond to areas defined by phonological, morphological, or lexical phenomena; see Birkenes & Fleischer (2021:159) for a brief discussion. It should be evident that such questions are challenging to answer by means of studies that focus exclusively on single (morpho)syntactic phenomena. Instead, they call for quantitative dialectometric analyses, abstracting away from individual variables to reveal general geographical structures.

However, dialectometric studies have mostly focused on phonology or lexis so far. Due to the lack of suitable data, the level of (morpho)syntax has only rarely been studied (Wieling & Nerbonne, 2015:256–257). Some of the exceptions include Spruit (2006) and Spruit et al. (2009) on Dutch dialects or Szmrecsanyi (2012) and Wolk (2014) on English and Scottish dialects. For German, recent studies have focused on dialects in Hesse (Birkenes & Fleischer, 2021) and, most importantly, on Swiss German dialects (e.g. Scherrer & Stoeckle, 2016; Derungs et al., 2020). In short, these studies identify geographical patterns of (morpho)syntactic variation which often differ from traditional accounts of dialect areas. Notably, since previous dialectometric studies on (morpho)syntactic variation build on aggregative methods developed in the (Romanistic) *Salzburg school* (e.g. Goebel, 1984) and/or the *Groningen school* (e.g. Wieling & Nerbonne, 2015), there are hardly any comprehensive analyses of which types of (morpho)syntactic variables or features contribute to the geographical patterns revealed. To address this issue, non-aggregative geolinguistic measures are perhaps more appropriate, because they preserve the variation of individual features while detecting areal structures (e.g. Grieve, 2014; Pickl & Pröll, 2019).

The aim of the present study is to uncover geographical patterns of (morpho)syntactic variation in traditional Austrian dialects

Corresponding author: Philip C. Vergeiner; Email: philip.vergeiner@germanistik.uni-muenchen.de

Cite this article: Vergeiner PC, Bülow L, Elspaß S. (2024) Analyzing dialect (morpho)syntax in Austria: a non-aggregative dialectometric approach. *Journal of Linguistic Geography* 12: 57–68, <https://doi.org/10.1017/jlg.2024.15>

using non-aggregative dialectometric methods. Two research questions are targeted:

- RQ1 Can geographical patterns for dialect (morpho)syntax in Austria be identified that are based on the co-occurrence of different variants of various variables, and if so, how do these patterns relate to traditional dialect areas?
- RQ2 Which variants of which variables form the linguistic basis of the geographical structures revealed? Are different (morpho)syntactic variables equally important for geolinguistic structuring, and if not, how can we explain the differences between them?

To address these questions, we draw on a comprehensive dialect corpus obtained by direct dialect interviews. The corpus includes data of 163 speakers from 40 locations throughout Austria. Our analyses are based on the frequency of 79 variants of 30 (morpho)syntactic variables. In order to arrive at generalizable statements beyond individual variables, we use factor analysis.

Section 2 provides a brief overview of previous research on dialect (morpho)syntax in Austria and beyond. In Section 3, we describe the data and methods of our study, and then present our results in Section 4. We conclude with a discussion and summary of our key findings in Section 5.

2. (Morpho)syntax of German dialects in Austria and beyond

Somewhat surprisingly, the increased interest in (morpho)syntactic phenomena in dialects within the last few decades originated in generative grammar (Scheutz, 2005:292–293) and not in traditional dialectology, although there is some earlier work on individual (morpho)syntactic features in dialects (e.g. Weise, 1907, 1917). Generativists “discovered” dialects in the 1980s and 1990s in order to address theoretical questions that revolved around specific syntactic phenomena, such as complementizers and negative concord (e.g. Bayer, 1990), while at the same time focusing on “microvariation,” i.e. using syntactic structures in dialects as a testing ground for syntax theories. Another reason why dialects became an attractive field of research for theoretical linguists is their status as “natural,” first-order (henceforth L1) languages (Weiß, 1998:3), i.e. L1 varieties that are immediate derivatives of internal languages, primarily acquired in family settings. This is not trivial in that the standard varieties in many regions in the German-speaking countries can be assumed to be only indirect derivatives of internal languages that do not meet the L1 criterion because they are taught and learned as L2 (second-order language) at school. This is especially true for the Upper German language area, which includes the Bavarian and Alemannic dialect regions in Germany, Austria, Switzerland, Liechtenstein, and South Tyrol.

The fact that the use of dialect in everyday communication is still widespread in the Upper German language area may also be one of the reasons why theoretical linguists have focused on microvariation in the (Central) Bavarian (e.g. Bayer, 1984, 1990; Weiß, 1998; Grewendorf & Weiß, 2014) and Alemannic dialects (e.g. Seiler, 2003; Brandner & Bräuning, 2013). At the same time, since the 1990s, linguists working within the dialectological, sociolinguistic, and typological research paradigms have increasingly addressed Bavarian and Alemannic dialect syntax (e.g. Patocka, 1997; Fleischer, 2004; Scheutz, 2005). However, while there are two major dialect syntax projects for Alemannic dialects—the above-mentioned *Syntactic Atlas of German*

Speaking Switzerland (SADS) and the project *Syntax of Alemannic* (SynAlm, focusing particularly on Alemannic dialects in Germany)—there has been no such project on Bavarian. Thus, despite the great interest in Bavarian dialect syntax, comprehensive variationist studies covering the entire Bavarian language area are still a major desideratum.

Regarding Austria, some comprehensive studies on individual (morpho)syntactic phenomena have been carried out in recent years within the context of the Special Research Program “German in Austria” (DiÖ, *Deutsch in Österreich*; see Budin et al., 2019) and the “Dictionary of Bavarian Dialects in Austria” (WBÖ, *Wörterbuch der bairischen Mundarten in Österreich*). For instance, several studies have been published on variation and change of subjunctive constructions in the Austrian language varieties (cf. contributions to Bülow et al., 2021a); Bülow et al. (2021b) explored structures of adnominal possession; Fingerhuth & Lenz (2020) and Vergeiner & Bülow (2021) studied complementizer agreement in Austria’s dialects; Vergeiner & Hartinger (2022) and Stöckle et al. (2021) focus on negative concord (NC); Bülow et al. (2023) investigated the variation of relativizers in attributive relative clauses; Vergeiner & Niehaus (2024) and Vergeiner (2024) examined the syntax of articles; and so on (see Lenz, 2019:333–338 for an overview of previous studies). Most of these studies found significant differences between the syntax of Bavarian and Alemannic dialects in Austria. For example, the use of the particle *was* to introduce attributive relative clauses (1a) is restricted to the Bavarian dialects in Austria, whereas the use of the particle *wo* (1b) is restricted to the Alemannic dialects.¹

- | | | | | | | | |
|-----|----|----------------|--------------|-----------------|----------|-----------------|------------|
| (1) | a. | des | haus | wos | i | kaff-t | hob |
| | | <i>DET.DEF</i> | <i>house</i> | <i>PTCL.REL</i> | <i>I</i> | <i>buy-PTCP</i> | <i>AUX</i> |
| | b. | des | huus | wo | i | koof-t | han |
| | | <i>DET.DEF</i> | <i>house</i> | <i>PTCL.REL</i> | <i>I</i> | <i>buy-PTCP</i> | <i>AUX</i> |
- ‘The house I have bought.’

Apart from the differences between Alemannic and Bavarian dialect (morpho)syntax, which will also prove to be a prominent factor in the results of our study (see Section 4.1), previous research has shown that the geographical patterns in dialect (morpho)syntax do not always correspond to traditional dialect classifications (see the dialect map in Figure 1 below). For a number of individual phenomena such as subjunctive and negative concord (NC), there is no north–south division in Bavarian, as shown in Figure 1, but only an east–west division. Regarding the use of NC, for example, negative spread constructions (2a) predominate in the western parts of Austria, namely in the Alemannic and South Bavarian dialect regions in Tyrol, whereas negative doubling constructions (2b) are limited to the central and eastern Bavarian dialect regions of Austria (Moser, 2021; Vergeiner & Hartinger, 2022).

- | | | | | | | |
|-----|----|--------------------------|-------------|--------------|--------------|--------------|
| (2) | a. | i | hun | nia | koa | geld |
| | | <i>I</i> | <i>have</i> | <i>never</i> | <i>no</i> | <i>money</i> |
| | | ‘I never have any money’ | | | | |
| | b. | i | hob | koa | gööd | ned |
| | | <i>I</i> | <i>have</i> | <i>no</i> | <i>money</i> | <i>not</i> |
| | | ‘I have no money’ | | | | |

In sum, there is an extensive and growing body of literature on individual (morpho)syntactic phenomena across traditional dialects in Austria on a broad empirical basis. However, there is a lack of studies that shift the focus away from case studies on

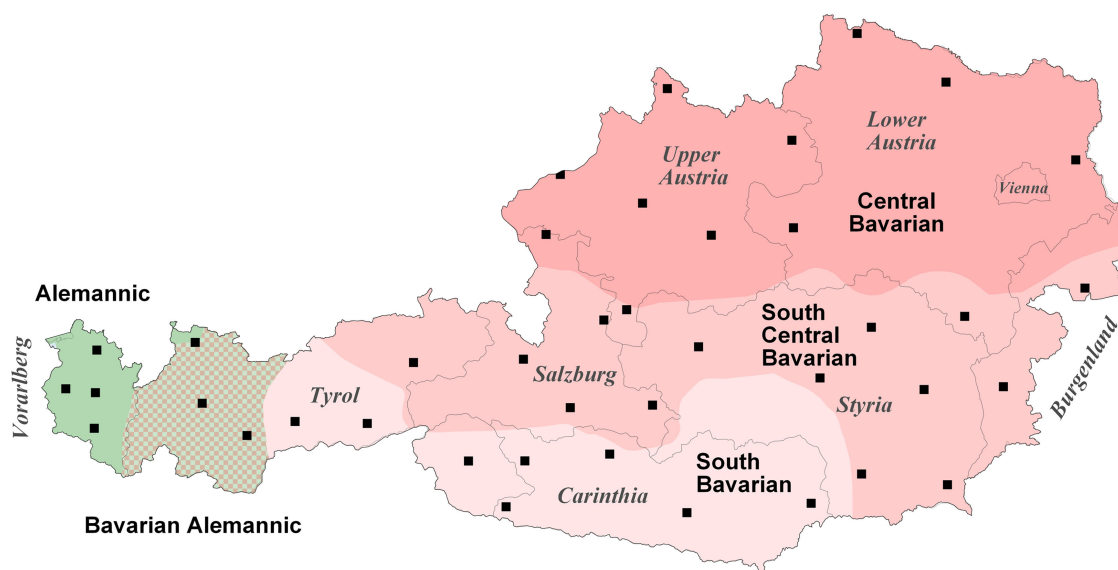


Figure 1. Research locations, federal states of Austria and dialect areas.

individual phenomena towards general areal patterns of (morpho)syntax in the Bavarian and Alemannic dialects in Austria. Moreover, there are no dialectometric studies on Austrian dialect syntax which could help to detect such patterns. This is where our study comes in, by addressing precisely these desiderata.

3. Data and methods

For our geolinguistic analyses of dialect (morpho)syntax in Austria, we draw on a dialect survey conducted within the framework of the project “Variation and Change of Dialect Varieties in Austria (in Real and Apparent Time).”² In what follows, we describe our project design, research locations, and participants (Section 3.1), before we elaborate on our variables (Section 3.2) and statistical methods (Section 3.3).

3.1. Research design, research locations, and participants

The data of the present study consist of dialect recordings obtained by trained fieldworkers. A survey was conducted in 40 small rural villages. Figure 1 shows the distribution of the research locations according to the most widely accepted classification of dialect areas in Austria (Wiesinger, 1983). Austria comprises a small Alemannic dialect area in the far west (Vorarlberg) and a much larger Bavarian dialect area with a Bavarian–Alemannic transition zone in between. The Bavarian area is divided into Central Bavarian, South-Central Bavarian, and South Bavarian dialects. Notably, this classification is based on a qualitative structuralist approach drawing mostly on phonology and partly also on (inflectional) morphology but ignores syntax completely (Wiesinger, 1983:813).

Four speakers of the traditional dialect from each location participated in our study.³ Two participants were chosen from an older (65+ years) and two from a younger (18–35 years) generation, with one male and one female per age group. In sum, the sample consists of 163 speakers. Traditional dialectological criteria for sampling were applied (see Chambers & Trudgill, 1998). The older speakers are typical NORM/Fs (= non-mobile, old, rural males/females). The younger participants can also be considered prototypical dialect speakers in that they have

been raised in local artisanal or agricultural networks and have not received higher education. Their parents were born and raised in the same location. Both their social and working lives are centered in the same local environments.

3.2. Variables

The interviews were conducted by trained fieldworkers using a traditional dialect questionnaire, which included several tasks such as translation tasks, cloze tasks, and picture naming tasks that were designed to elicit traditional dialect features on all linguistic levels. To analyze syntactic features, mainly translation tasks were employed. In these tasks, the participants had to translate sentences read out in standard German into their own dialects. While doing so, the participants were encouraged to use not only phonetic or morphological features but also syntactic features which they consider most natural in their everyday dialect. For example, there were twelve translation tasks in the questionnaire referring to a possessive relation, e.g. *Wo sind Mutters Schuhe?* ‘Where are mother’s shoes?’ or *Das ist Annas Fuß* ‘This is Anna’s leg’. To translate these sentences, most participants used either adnominal possessive dative constructions or competing variants such as *von* (‘of’) or genitive constructions (Bülow et al., 2021b).

Over the past few years, the data have been extensively analyzed in feature-based studies on (morpho)syntax (see Section 2). For the present study, we move away from individual features to detect more general patterns of variation. To do so, we selected a set of 30 variables, shown in Table 1. The variables include some of the most important characteristics of dialect (morpho)syntax in Austria, and as such they are discussed in feature-based studies and general overviews of this topic (e.g. Lenz, 2019:333–338). In addition, the variables can be assigned to different areas of (morpho)syntax, such as verbal and nominal syntax, government and agreement relations, conjunctions, and word order. Each variable was elicited by means of several different stimuli (mean = 5.1 stimuli per variable; median = 4.5), and in sum, our analyses include 24,140 responses (mean = 805 tokens per variable; median = 680.5).

The data for each variable are coded on a categorical scale, often of binary nature. In total, 79 variants were distinguished (mean = 2.6

Table 1. List of variables

#	Variables	Variants	n/stimuli	n/token
1.	past formation of <i>sein</i> ('be')	preterite perfect	5	796
2.	past formation of <i>wollen</i> ('want')	preterite perfect	4	536
3.	formation of subjunctive (past subjunctive/ <i>Konjunktiv II</i>)	strong synthetic weak synthetic periphrastic	21	3430
4.	infinitive complement with phase verb	bare infinitive infinitive with <i>z</i> infinitive with <i>zu</i> infinitive with <i>zum</i>	5	780
5.	adverbial infinitive construction	bare infinitive infinitive with <i>z</i> infinitive with <i>zu</i> infinitive with <i>zum</i> <i>dass</i> -clause	4	638
6.	IPP with <i>können</i> ('can')	infinitive (= IPP) participle (= no IPP)	7	1005
7.	IPP with <i>lassen</i> ('let')	infinitive (= IPP) participle (= no IPP)	3	444
8.	IPP with <i>müssen, wollen, dürfen</i> ('must, want, may')	infinitive (= IPP) participle (= no IPP)	6	902
9.	verb cluster order (V2 sentence) with perception verb	perception verb before dependent verb dependent verb before perception verb	3	454
10.	verb cluster order (V2 sentence) with modal verb	modal verb before full verb full verb before modal verb	7	983
11.	verb cluster order (verb-final sentence) with modal verb	auxiliary before modal verb before full verb modal verb before auxiliary before full verb full verb before auxiliary before modal verb full verb before modal verb before auxiliary	4	579
12.	complementizer agreement (singular)	complementizer agreement no complementizer agreement	3	473
13.	complementizer agreement (plural)	complementizer agreement no complementizer agreement ambiguous	3	479
14.	article before proper noun	article no article	5	971
15.	article before mass noun	article no article	5	1039
16.	weak article (feminine and plural)	weak article no weak article	6	960
17.	article doubling (indefinite article)	article doubling no article doubling	2	326
18.	prepositional dative construction	prepositional dative no prepositional dative	2	318
19.	adnominal possession construction	adnominal possessive dative construction prenominal genitive construction prenominal <i>von</i> -construction postnominal <i>von</i> -construction	12	1929
20.	inflection of attributive adjective (nominative)	inflection no (overt) inflection	3	476
21.	inflection of copredicative adjective	inflection no inflection	4	626
22.	comparative particle (equality relation)	<i>wie</i> <i>als wie</i>	3	441
23.	comparative particle (inequality relation)	<i>als</i> <i>wie</i> <i>als wie</i>	6	1005

(Continued)

Table 1. (Continued)

#	Variables	Variants	n/stimuli	n/token
24.	indefinite partitive pronoun (singular)	<i>ein-</i> <i>etwas</i> \emptyset	2	324
25.	indefinite partitive pronoun (plural)	<i>ein-</i> <i>ein paar</i> \emptyset <i>welch-</i>	3	475
26.	relativizer	<i>d-pronoun</i> <i>was</i> <i>d- was</i> <i>wo</i>	5	807
27.	inversion of personal pronoun	accusative before dative (= no inversion) dative before accusative (= inversion)	3	434
28.	pronominal adverb	simple pronominal adverb pronominal adverb with “short doubling” pronominal adverb with “long doubling” prepositional phrase for pronominal adverb	5	723
29.	negative spread	negative spread no negative spread	6	974
30.	negative doubling	negative doubling no negative doubling	5	813

variants per variable; median = 2).⁴ Since our corpus includes four speakers per location and at least two stimuli per variable, relative frequency distributions could be calculated for all variants per location. These relative frequencies are used in our subsequent analyses. Due to the less salient nature of syntactic data in comparison to phonological and lexical data (Kortmann, 2010:846), our approach seems to have significant advantages over previous work, which is often based either on the presence or absence of a feature (Spruit, 2006; Spruit et al., 2009) or on the choice of one dominant variant per location (Scherrer & Stoeckle, 2016). Another major difference to most previous work is our use of non-aggregative dialectometric methods, which are explained in the following section.

3.3 Statistical methods

Most dialectometrical studies are based on aggregation: they add up differences (or similarities) between every pair of locations to create a site-by-site distance matrix. This matrix is then used as input for further statistical analyses such as multidimensional scaling or cluster analysis to reveal geographical structures (e.g. Goebel, 1984; Wieling & Nerbonne, 2015). Although this approach has led to valuable insights, it has the main disadvantage that information on individual variant distributions gets lost, and the linguistic basis of aggregate dialect differences can only be added again in a *post hoc* fashion; for a discussion, see Wieling & Nerbonne (2015:248-250). This problem is avoided in non-aggregative measures such as factor analysis (FA) or principal component analysis (PCA), which have been used as alternative approaches (e.g. Shackleton, 2005; Nerbonne, 2006; Leino & Hyvönen, 2008; Szmrecsanyi, 2012; Grieve, 2014; Pickl, 2016; Pickl & Pröll, 2019; Pickl et al., 2019). For the present study, we apply a factor analysis because it is proved to be particularly suitable for identifying areal structures (e.g. Leino & Hyvönen, 2008; Vergeiner & Bülow, 2023).

A factor analysis (FA) is a multivariate statistical method for identifying patterns of variation in a data set. Based on a correlation matrix, it subsumes variants that are correlated with one another but

largely independent of other variants under a small set of underlying constructs, so-called “factors.” By means of this procedure, the variation in the data is reduced and restructured by identifying latent patterns behind the variants. In doing so, FA preserves as much information as possible from the original data set.

In the present study, FA is based on the interrelations between the research locations with regard to (morpho)syntactic variation. Notably, FA does not presuppose any information about the geographical positions of the research locations. Rather it reveals areal patterns only if there are sufficiently strong geographical signals in the linguistic data itself.

For the interpretation of the factor solution, two parameters are crucial.

- First, the *factor loadings* function as a measure of the relation between a factor and, in our study, the research locations. A loading close to 1 signals a high positive correlation between a factor and a location, indicating that the factor accounts for most of a location’s variance. In contrast, loadings close to 0 suggest no correlation, and loadings < 0 indicate a negative correlation.
- Second, the *factor scores* reveal how a particular variant ranks on a given factor. A high positive factor score signals that a given variant is (positively) associated with a factor while scores close to 0 indicate no association whatsoever. Negative associations result in negative factor scores.

For this study, FA were performed with the software IBM SPSS Statistics using principal axis factoring with varimax rotation. The estimation of the factor scores is based on the regression method. Notably, both the Kaiser–Meyer–Olkin Measure of Sampling Adequacy (= 0.759) and Bartlett’s Test of Sphericity ($X^2 = 6302.45$, $p < 0.000^{***}$) indicate that the data fit well to the analysis.

4. Results

In what follows, we present the results for a factor solution with four factors. This solution is based on the Kaiser–Guttman

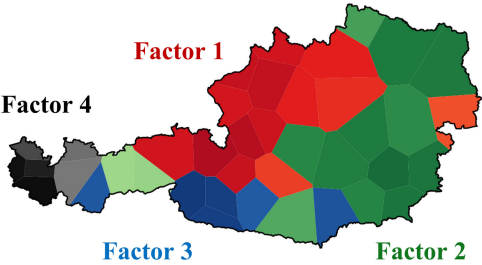


Figure 2. Dominant factor loadings.

Criterion (eigenvalues < 1). In total, the four factors account for not less than 82% of the variance in the data. Factor 1 (= F1) accounts for about 57.2%, F2 for about 11.4%, F3 for about 8.7%, and, finally, F4 for about 4.6%.

Section 4.1 focuses on RQ1; that is, it investigates the geographical structures. Section 4.2 addresses RQ2, dealing with the linguistic patterns underlying these structures.

4.1. Geographical patterns

Figure 2 provides a first overview of the geographical patterns. To enhance visibility, we generated an area-class map based on Voronoi partition (using the software REDE SprachGIS; <https://www.regionalsprache.de/SprachGIS/Map.aspx>). The colors in the map (Figure 2) display the dominant factors in each research location, i.e. the factors with the highest factor loadings, so that, for example, a red coloring indicates that F1 is dominant in the respective location. Green stands for F2, blue for F3, and black for F4. The shadings reveal the relative strength of the dominant factor loadings: the darker the shading, the higher the factor loadings. Hence, the darker shaded areas can be interpreted as core dialect areas and the lighter shaded areas as transition zones or “foothills.”

Figure 2 clearly indicates a strong geographical signal in our (morpho)syntactic data. Based on the dominant factor loadings, we find several coherent areas.

- F1 (red) dominates in the northwestern parts of Austria, in particular in the regions of Salzburg and Upper Austria, but also in northern Burgenland.
- F2 (green) is the most important factor in the eastern half of Austria and also in parts of Tyrol and Carinthia.
- F3 (blue) predominates in the south(west), i.e. in most parts of Carinthia and also in Tyrol.
- The loadings on F4 (black) are highest in the westernmost regions, i.e. in Vorarlberg and western Tyrol.

At least to a certain extent, the geographical patterns visible in Figure 2 resemble the traditional dialect classification of Austria (see Section 3.1). The area where F4 is dominant comprises the Alemannic dialects and also the Bavarian–Alemannic transition zone. In most South Bavarian locations, the factor loadings on F3 outweigh the other factor loadings. Interestingly, however, the differentiation between the Central and South-Central Bavarian dialects made on phonological maps (see Figure 1) does not materialize here for syntactic variation. Instead of a north–south division, Figure 2 displays an east–west division for the Central and South-Central Bavarian dialects. Although there is not a complete overlap, our results indicate that the geographical patterns in (morpho)syntactic

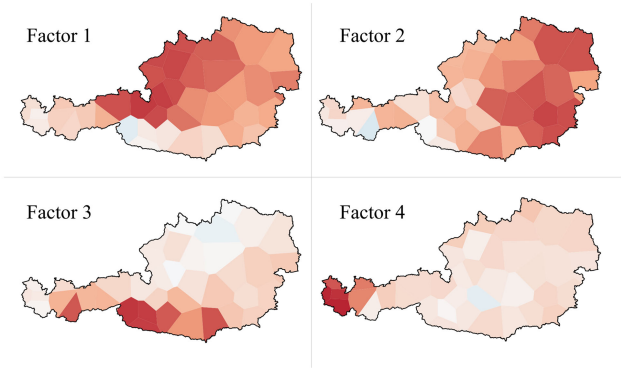


Figure 3. Individual factor loadings.

variation partially correspond to the traditional dialect classification explained in Section 3. However, mapping only the dominant factors “ignores the variation ‘below’ the threshold of dominance,” i.e. “the locally non-dominant parts of the globally dominant factors” (Pickl, 2016:91). To account for this fine-grained variation, we have to focus on one factor at a time, as reflected in the four individual heatmaps in Figure 3.

Figure 3 reveals that F1, although dominant mostly in the northwest, also has rather high loadings in the eastern parts of Austria. In contrast, there is a clear boundary to the southern and most western dialects. The geographical patterns for F2 are more scattered since there are moderately high loadings spread across Austria, except for some western and southwestern regions. In general, there is a broad continuum visible for F2, in both a northwestern and southwestern direction. For F3 and F4, the patterns are straightforward: F3 is a latent factor in several Tyrolean and Carinthian locations and shows minor loadings eastwards into Styria. For F4, a sharp boundary becomes apparent. Note that there is no spatial continuum for F4.

4.2 Linguistic patterns

RQ2 is concerned with the linguistic basis of the geographical structures discussed in Section 4.1. As explained in Section 3.3, this can be analyzed by examining the factor scores which indicate the association between a given factor and the linguistic variants. To this end, Table 2 shows the top three (positive) factor scores for each individual factor.

Regarding F1, the factor scores are highest for the usage of complementizer agreement in the singular (3a) as opposed to its non-occurrence (3b); the use of an indefinite article before mass nouns (4a) as opposed to the lack of an indefinite article (4b); the use of *eine* as a partitive pronoun in the singular (5a) as opposed to the use of *etwas* (5b) or a null morpheme (5c).

- (3)

a.

du

deaf-st

geh-n

wonn-st

wüü-st

you

may-2sg

go-INF

if-2sg

want-2sg
- b.

du

derf-sch

gia-n

wenn

will-sch

you

may-2sg

go-INF

if

want-2sg
- ‘You may go, if you want to.’

(4)

a.

ea

brauch-t

a

gööd

he

need-3sg

DET.INDEF

money

b.

ea

brauch-t

geld

he

need-3sg

money

‘He needs money.’

Table 2. Top three factor scores for the factors

F1	1.	complementizer agreement (sg.) (#12)	2.21
	2.	article before mass nouns (#15)	1.79
	3.	eine as partitive pronoun (sg.) (#24)	1.76
F2	1.	preterite of <i>sein</i> ('be') (#1)	2.89
	2.	overtly inflected attributive adjective (nominative) (#20)	2.13
	3.	IPP with <i>können</i> ('can') (#6)	1.9
F3	1.	prepositional dative (#18)	2.27
	2.	no weak article <i>d'</i> feminine/plural (#16)	2.25
	3.	no article doubling (#17)	1.9
F4	1.	no complementizer agreement (sg.) (# 12)	2.16
	2.	weak article <i>d'</i> feminine/plural (#16)	1.99
	3.	IPP with <i>können</i> ('can') (#6)	1.9

- (5) a. i hob koa gööd owa se hom oa-s
I have no money but they have PRO.PART-NOM.SG
 b. i hun koa geld owa se hom epes
I have no money but they have PRO.PART
 c. i han koi geld aba se han ø
I have no money but they have PRO.PART
 'I have no money, but they have some.'

F2 is strongly associated with the following features: preterite forms of *sein* ('be') (6a) as opposed to forms in the perfect tense (6b); overtly inflected adjectives in the nominative (7a) as opposed to adjective forms lacking overt inflection (7b); and finally, IPP constructions (*infinitivus pro participio*) with the modal verb *können* (8a) as opposed to forms without IPP (= constructions with a participle) (8b).

- (6) a. ea woa miad
he was tired
 b. ea is miad g-wes-n
he AUX tired PTCP-be-PTCP
 'He was tired.'
- (7) a. wo is da kloan-e eisl
where is DET.DEF little-NOM.SG donkey
 b. wo is da kloa eisl
where is DET.DEF little donkey
 'Where is the little donkey?'
- (8) a. se hom ned lees-n kenn-a
they AUX not read-INF can-INF
 b. se hom ned lees-n ge-kinn-t
they AUX not read-INF PTCP-can-PTCP
 'They could not read.'

Most characteristic for F3 is the prepositional dative as in (9a), in contrast to (9b); the use of *di/de* (10a) instead of the weak article *d* (10b) before feminine and plural nouns; and the lack of article doubling as in (11a), in contrast to (11b).

- (9) a. ea sog-t=s in di leit
he say-3SG=it PREP DET.DEF people
 b. ea sog-t=s di leit
he say-3SG=it DET.DEF people
 'He tells the people.'
- (10) a. di sau ghea-t eam
DET.DEF sow belong-to-3SG him
 b. d= sau ghea-t eam
DET.DEF= sow belong-to-3SG him
 'The sow is his.'
- (11) a. ea brauch-t a bissl salt
he need-3SG DET.INDEF bit salt
 b. ea brauch-t a bissl a soiz
he need-3SG DET.INDEF bit DET.INDEF salt
 'He needs a bit of salt.'

Finally, typical features for F4 are the non-occurrence of complementizer agreement in the singular (3b); the usage of the weak article *d* before feminine and plural nouns (10b); and IPP constructions with the modal verb *können* (8a).

The factor scores indicate how the variants are ranked on one particular factor. Note that variants can have (relatively) high positive factor scores for more than one factor, and they can have high negative factor scores for other factors. For example, the use of the prepositional dative construction (see (9a)) has high positive factor scores not only for F3 (2.27) but also for F1 (1.21), while it has rather high negative factor scores for F2 (−1.55) and F4 (−1.28). This indicates that the variance of this particular feature is well accounted for by the factor solution, and it can thus be regarded a key feature for the geographical patterns discussed above. On the other hand, several other variants have generally factor scores close to 0. For instance, the use of prepositional phrases (12a) instead of pronominal adverbs—either simple pronominal adverbs (12b), or pronominal adverbs with “short doubling” (12c) or “long doubling” (12d) (see Fleischer, 2002)—has factor scores close to 0 for all factors (F1, −0.3; F2, −0.2; F3, −0.28, F4, −0.12).

- (12) a. vo dem woß i nix
PREP PRO.DEM know I nothing
 b. davo woß i nix
about_that know I nothing
 c. da=davo woß i nix
ADV=about_that know I nothing
 d. da woß i nix davo
ADV know I nothing about_that
 'I know nothing about that.'

To examine which variants are most strongly—either positively or negatively—associated with all four factors, and thus best explained by the geographical structures revealed, we calculated a composite Factor Score Index for all variants. We did so by adding up the absolute values of all four factor scores per variant and dividing them by the total number of factors (= 4). Table 3 shows the results for the top three and lowest three Factor Score Index values.

Note that the Factor Score Index values are highest for the occurrence of the prepositional dative construction (see (9a)), the

Table 3. Top three and lowest three Factor Score Index values for the variants

1.	prepositional dative (#18)	1.58
2.	weak article <i>d'</i> feminine/plural (#16)	1.56
3.	no complementizer agreement (sg.) (# 12)	1.52
...
77.	comparative particle <i>als wie</i> (inequality rel.) (#23)	0.4
78.	<i>d</i> -pronoun as relativizer (#26)	0.28
79.	prepositional phrase for pronominal adverb (#28)	0.22

weak article *d* before feminine and plural nouns (see (10b)), and the absence of complementizer agreement in the singular (see (3b)). The values are lowest for the use of prepositional phrases instead of pronominal adverbs (see (12a)), for the use of the *d*-pronoun as a relativizer (13a) instead of *was* (13b), *d- was* (13c), or *wo* (13d), and the use of *als wie* (14a) instead of *als* (14b) or *wie* (14c) as a comparative particle in inequality relations. This comes as no surprise, since variants such as the prepositional dative construction or complementizer agreement are regionally restricted forms with limited areal distribution (see e.g. Vergeiner & Bülow, 2021 for complementizer agreement), while variants such as the *d*-pronoun as a relativizer are used throughout Austria with no clear areal patterns (e.g. Bülow et al., 2023 for relativizers).

(13)	a.	des	haus	des	i	kaff-t	hob
		DET.DEF	house	PRO.REL	I	buy-PTCP	AUX
	b.	des	haus	wos	i	koof-t	hob
		DET.DEF	house	PTCL.REL	I	buy-PTCP	AUX
	c.	des	haus	des	wos	i	kaff-t
		DET.DEF	house	PRO-REL	PTCL.REL	I	buy-PTCP
	d.	des	huus	wo	i	koof-t	han
		DET.DEF	house	PTCL.REL	I	buy-PTCP	AUX
		'the house I have bought'					
(14)	a.	se	is	kloan-a	ois	wia	ea
		she	is	small-COMP	PTCL.COMP	PTCL.COMP	he
		se	is	kloan-a	ois		ea
		she	is	small-COMP	PTCL.COMP		he
		se	is	kloan-a	wia		ea
		she	is	small-COMP	PTCL.COMP		he
		'She is smaller than him.'					

Notably, however, not all features with a clear geographical distribution show high factor score values or high Factor Score Index values. This can be demonstrated when comparing these values with the Moran’s I statistics for the variants. Moran’s I is the most common statistical measure for (global) spatial autocorrelation and clustering. It assesses whether neighboring locations tend to have above-average similarities or dissimilarities. Moran’s I values range from -1 to 1 , with values around 0 resulting from random distributions (= no spatial autocorrelation). When the data are spatially dispersed and neighboring locations are highly dissimilar, Moran’s I approximates -1 . In the case of strong spatial clustering, Moran’s I gets close to 1 ; for the use of spatial autocorrelation in dialectometry, see e.g. Grieve et al. (2011), Grieve (2014), and Szmrecsanyi (2012).

We computed Moran’s I for all 79 variants using the open-source software tool GeoDa 1.20 (<https://geodacenter.github.io/>). For defining spatial weights, we employed queen-based contiguity weights. Significance testing was based on Monte Carlo randomization tests with 9999 permutations. The significance level was set at pseudo $p < 0.0008$ (= $0.05/60$) with Bonferroni correction for multiplicity.⁵ Not surprisingly, all variants except two show positive values for Moran’s I, with a mean value of 0.37 (see also the histogram in Figure 4, left). In most cases (43 variants of 21 variables), the result is significant.

Interestingly, there is only an average correlation of $r = 0.4$ ($p < 0.001^{***}$) between the Moran’s I values and the Factor Score Index values (based on a Pearson correlation analysis) (see Figure 4, right). On the one hand, this can be attributed to the fact that the factor solution does not account very well for some variants with clear-cut regional distributions. For example, the use of adnominal possessive dative constructions (15a) instead of pronominal genitive constructions (15b) or pre- and postnominal *von*-constructions (15c, 15d) has a significant high Moran’s I value (0.7) but rather low factor scores ($F1, -0.1$; $F2, 0.7$; $F3, -0.4$; $F4, -1.1$) which indicates that there is only a minor overlap with the geographical patterns of other variants.

(15)	a.	wo	san	da	mama	ihr-e	schuach
		where	are	DET.DEF	mother	PRO.POSS-PL	shoes
	b.	wo	sein		mamas		schuach
		where	are		mother-GEN		shoes
	c.	wo	san	vo	da	mama	de
		where	are	PREP	DET.DEF	mother	DET.DEF
	d.	wo	san	de	schuach	vo	da
		where	are	DET.DEF	shoes	PREP	DET.DEF
		'Where are mother's shoes?'					

On the other hand, there are variants without clear-cut regional patterns but rather high factor scores. For instance, the occurrence of overtly inflected adjectives in the nominative (see (7a)) has a non-significant Moran’s I value (0.2) but rather high factor scores ($F1, 1$; $F2, -0.8$; $F3, 1.3$; $F4, -0.7$). This is in line with the finding of Szmrecsanyi (2012:153) that “even features that appear to be distributed non-geographically when analyzed in isolation may help to create geographically more or less focused layers of morphosyntactic variability in conjunction with other features.”

Finally, we also calculated mean Factor Score Index values for the variables by simply adding up the Factor Score Index values for all variants of a variable and dividing them by the total number of variants per variable. These mean values indicate which variables are, on average, most important for the geographical structures detected. Table 4 shows the results for all 30 variables, with the variables ranked by the mean Factor Score Index values. In addition, Table 4 also shows the mean Moran’s I values for the variables.

What is interesting about Table 4 is that, on average, variables at the syntax-morphology interface seem to have a higher Factor Score Index than features related to other aspects of syntax (e.g. word order variation). This applies, in particular, to features concerning nominal morphology, and, to a somewhat smaller degree, verbal morphology. Although it is difficult to draw a categorical distinction between the variables in that regard, we highlighted those variables in Table 4 which arguably have the closest links to morphology. As can be seen, these variables tend to

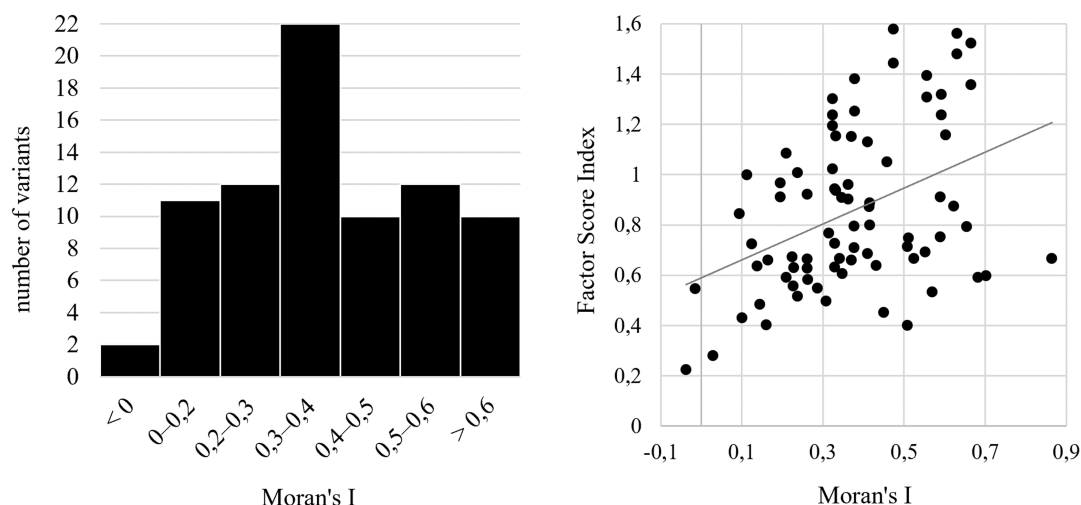


Figure 4. Histogram of the Moran's I values (left) and correlations between the Factor Score Index values and Moran's I (right).

be positioned in the upper half of the table. This result may be due to the fact that variants which are of a more morphological nature tend to co-occur more strongly and, thus, tend to form more coherent areas with one another than with the variants of other variables. We will discuss this aspect, among others, in Section 5.

5. Discussion and summary

The present study aimed at detecting areal patterns of (morpho)syntactic variation in traditional Austrian dialects using methods of non-aggregative dialectometry. In this final section, we address the two research questions and include a discussion of methodological and further linguistic implications of our findings.

In response to RQ1, our results indicate clear geographical patterns of (morpho)syntactic variation in Austria. The data from the factor analysis demonstrate a strong geographical signal in our (morpho)syntactic data. This was not necessarily expected, as it is a common assumption that syntax is geographically less structured than lexis or phonology; see e.g. Glaser (2013) and Kortmann (2010) for a discussion. It is noteworthy that these patterns could be shown with data gathered using more traditional methods of data collection, i.e. direct dialect interviews (see Section 3). Note also that it has often been questioned whether these methods allow us to detect geographical patterns in (morpho)syntax at all; see e.g. Scheutz (2005) for a discussion.

Compared to traditional dialect maps such as Wiesinger (1983), which are mostly based on phonological variation, the geographical patterns we identified show both similarities and differences. Most notably, our findings confirm the strong contrast between the Bavarian and Alemannic dialects as well as the peculiarities of South Bavarian; however, the north–south divisions between Central, South-Central, and South Bavarian which shape the dialect landscape of Austria on traditional maps (see Figure 1) are not reflected in our results. It could be assumed from these findings that the geographical structures depend on the level of linguistic architecture under investigation, i.e. that the syntactic patterns simply look different from the phonological or lexical ones. However, this assumption can be questioned by argument and empirical evidence. First, traditional classifications of German dialects are not based on patterns of a large number of variables,

but rather on boundaries defined by relatively few phonological isoglosses. Second, a study by Pickl et al. (2019), which used a similar method of factor analysis but drew on data from “colloquial” *lexis*,⁶ shows very similar geographical patterns. In particular, the three dominant factors in this study correspond to Factors 1, 2, and 4 in our study.

RQ2 asks about the strength of the contribution of individual variables to the formation of linguistic geographical structures. In our data, certain variants generally show high factor scores, others comparatively low scores. According to our findings, this is only partly related to the extent to which the variants have a stronger spatial distribution at all (see the measures of spatial autocorrelation in Section 4.2). Rather, variants that differ spatially only in their frequency distribution often seem to be important for structuring the language space; see Szmrecsanyi (2012) for similar findings for English. Such variants have often been overlooked in previous studies on dialect syntax (e.g. the space-structuring function of the past tense of the copula verb *sein* ‘to be’; see Factor 2) or they were featured as omnipresent in Upper German dialects in Austria (e.g. Lenz, 2019:336 on article doubling; see Factor 3).

In sum, we were able to show that certain variants co-occur more strongly with each other than others, which is reflected in higher factor scores. For instance, the usage—and non-usage (!)—of the prepositional dative construction (see (9a)), the weak article *d* before feminine and plural nouns (see (10b)), or complementizer agreement in the singular (see (3a)) shows clear patterns of co-occurrence with other variants, while variants such as the comparative particle *als wie* for inequality relations (see (14a)), the use of the *d*-pronoun as a relativizer (see (13a)), or the use of prepositional phrases for pronominal adverbs (see (12a)) do not exhibit such patterns.

Against the backdrop of this finding we obtained Mean Factor Score Index values for individual variables; that is, we determined which variables, on average, have the highest impact on the structure of the linguistic space under investigation. It emerges that variants of syntactic variables with an interface to morphology rank highest on the list (see Table 4). This suggests that these (morpho)syntactic variables are similar to each other on a (deep) structural and/or typological level (e.g. with respect to their nature as more analytic vs. more synthetic structures). The relevant similarities may be attributed to the fact that the formation

Table 4. Factor Score Index values and mean Moran's I values for all variables (variables at syntax-morphology interface highlighted in bold)

#	Variables	Mean Factor Score Index values	Mean Moran's I
16.	weak article (feminine and plural)	1.52	0.63
18.	prepositional dative construction	1.51	0.47
12.	complementizer agreement (singular)	1.44	0.67
17.	article doubling (indefinite article)	1.35	0.56
1.	past formation of <i>sein</i> ('be')	1.32	0.38
20.	inflection of attributive adjective (nominative)	1.28	0.59
6.	IPP with <i>können</i> ('can')	1.25	0.32
8.	IPP with <i>müssen, wollen, dürfen</i> ('must, want, may')	1.13	0.32
7.	IPP with <i>lassen</i> ('let')	1.04	0.33
21.	inflection of copredicative adjective	0.94	0.2
27.	inversion of personal pronoun	0.93	0.36
30.	negative doubling	0.91	0.37
13.	complementizer agreement (plural)	0.86	0.47
24.	indefinite partitive pronoun (singular)	0.85	0.44
14.	article before proper noun	0.84	0.21
25.	indefinite partitive pronoun (plural)	0.83	0.59
9.	verb cluster order (V2 sentence) with perception verb	0.83	0.33
10.	verb cluster order (V2 sentence) with modal verb	0.83	0.28
22.	comparative particle (equality relation)	0.77	0.26
29.	negative spread	0.76	0.44
3.	formation of subjunctive (past subjunctive/<i>Konjunktiv II</i>)	0.76	0.24
2.	past formation of <i>wollen</i> ('want')	0.75	0.38
4.	infinitive complement with phase verb	0.68	0.41
23.	comparative particle (inequality relation)	0.68	0.3
11.	verb cluster order (verb-final sentence) with modal verb	0.64	0.25
5.	adverbial infinitive construction	0.62	0.28
28.	pronominal adverb	0.6	0.16
15.	article before mass noun	0.56	0.51
19.	adnominal possession construction	0.55	0.42
26.	relativizer	0.51	0.44

processes of these variables are rooted in similar historical conditions. For instance, many constructions such as prepositional dative constructions, complementizer agreement, the weak article *d*, or zero inflection of the attributive adjective are all related to reduction or “weakening” processes. To get a clearer picture of these connections is a desideratum for future research.

In methodological terms, the results from this study show that dialectometric procedures can contribute to answering not only questions of linguistic geography but also structural questions. The method of factor analysis adapted here is innovative compared to earlier methods in dialectology (and aggregative methods in dialectometry, in particular) in that it assesses which variants and variables are most important for the formation of linguistic geographical structures. Moreover, it does not assume homogeneity at one location but allows for variation—or *orderly heterogeneity*—at specific data points in space. Thus, this approach is not based on the mere presence or absence of a feature or on the

dominance of a variant per location, as in previous work, but accounts for actual variation which may paint a more realistic picture of language use at individual locations.

Acknowledgments. The research for the present article was conducted within the framework of the project “Variation and change of dialect varieties in Austria (in real and apparent time)” by the Austrian Science Foundation (FWF), grant no. F 06002, as part of the Special Research Program “German in Austria” (SFB F 60).

Competing interests. The authors declare none.

Notes

1 The glossing of our examples is based on the Leipzig Glossing Rules (<https://www.eva.mpg.de/lingua/resources/glossing-rules.php>). Note that for reasons of economy and readability we omitted glossing morphological features which are invisible on a formal level.

- 2 The project is based at the University of Salzburg and the Acoustics Research Institute (ARI) in Vienna, see Acknowledgments.
- 3 In the village of Ulrichsberg, seven speakers were recorded as part of an in-depth investigation into sound change in real time.
- 4 To avoid a distortion of our results, we removed outliers, i.e. variants with a general frequency < 5%.
- 5 We applied one test per variant, except for binary variables, where the statistics are identical for both variants.
- 6 The study was based on 191 variables from the *Atlas zur deutschen Alltagssprache* (AdA), collected via online surveys between 2010 and 2014.

References

- AdA = Elspaß, Stephan & Robert Möller. 2003ff. *Atlas zur deutschen Alltagssprache* (AdA). Retrieved from <https://www.atlas-alltagssprache.de> (January 5, 2025).
- Bayer, Josef. 1984. COMP in Bavarian syntax. *The Linguistic Review* 3(3). 209–274.
- Bayer, Josef. 1990. What Bavarian negative concord reveals about the syntactic structure of German. In Joan Mascaró & Marina Nespó (eds.), *Grammar in progress: GLOW essays for Henk van Riemsdijk*, 13–23. Dordrecht: Foris.
- Birkenes, Magnus Breder & Jürg Fleischer. 2021. Syntactic vs. phonological areas: A quantitative perspective on Hessian dialects. *Journal of Linguistic Geography* 9(2). 142–161.
- Brandner, Eleonore & Iris Bräuning. 2013. Relative *wo* in Alemannic: Only a complementizer? *Linguistische Berichte* 234. 131–169.
- Budin, Gerhard, Stephan Elspaß, Alexandra N. Lenz, Stefan M. Newerkla & Arne Ziegler. 2019. The Research Project (SFB) “German in Austria”: Variation—Contact—Perception. In Lars Bülow, Kristina Herbert & Ann-Kathrin Fischer (eds.), *Dimensions of linguistic space: Variation—Multilingualism—Conceptualisations*, 7–35. Vienna: Lang.
- Bülow, Lars, Stephan Elspaß & Philip C. Vergeiner (eds.). 2021a. *Formen und Funktionen des Konjunktivs in Österreich aus variationslinguistischer Perspektive*. *Linguistik Online* 114(2). Available at <https://bop.unibe.ch/linguistik-online/issue/view/1168>.
- Bülow, Lars, Philip C. Vergeiner & Stephan Elspaß. 2021b. Structures of adnominal possession in Austria’s traditional dialects: Variation and change. *Journal of Linguistic Geography* 9(2). 69–85.
- Bülow, Lars, Anja Wittibschlager & Alexandra N. Lenz. 2023. Variation and change of relativizers in Austria’s German varieties. *Sprachwissenschaft* 48(3). 243–280.
- Chambers, Jack K. & Peter Trudgill. 1998. *Dialectology*. Cambridge: Cambridge University Press.
- Derungs, Curdin, Christian Sieber, Elvira Glaser & Robert Weibel. 2020. Dialect borders: Political regions are better predictors than economy or religion. *Digital Scholarship in the Humanities* 35(2). 276–295.
- Fingerhuth, Matthias & Alexandra N. Lenz. 2020. Variation and dynamics of “complementizer agreement” in German: Analyses from the Austrian language area. *Linguistic Variation* 21(2). 1–48.
- Fleischer, Jürg. 2002. Preposition stranding in German dialects. In Sjef Barbiers, Leonie Cornips & Susanne van der Kleij (eds.), *Syntactic microvariation*, 116–151. Amsterdam: Meertens Instituut.
- Fleischer, Jürg. 2004. A typology of relative clauses in German dialects. In Bernd Kortmann (ed.), *Dialectology meets typology: Dialect grammar from a cross-linguistic perspective*, 211–243. Berlin/New York: De Gruyter.
- Glaser, Elvira. 2013. Area formation in morphosyntax. In Peter Auer et al. (eds.), *Space in language and linguistics: Geographical, interactional and cognitive perspectives*, 195–221. Berlin: De Gruyter.
- Goebel, Hans. 1984. *Dialektometrische Studien: Anhand italo-romanischer, rätoromanischer und galloromanischer Sprachmaterialien aus AIS und ALF*. Tübingen: Max Niemeyer.
- Grewendorf, Günther & Helmut Weiß. 2014. *Bavarian syntax: Contributions to the theory of syntax*. Amsterdam/Philadelphia: John Benjamins.
- Grieve, Jack. 2014. A comparison of statistical methods for the aggregation of regional linguistic variation. In Benedikt Szmrecsanyi & Bernhard Wälchli (eds.), *Aggregating dialectology, typology, and register analysis: Linguistic variation in text and speech*, 53–88. New York: De Gruyter.
- Grieve, Jack, Dirk Speelman & Dirk Geeraerts. 2011. A statistical method for the identification and aggregation of regional linguistic variation. *Language Variation and Change* 23. 193–221.
- König, Werner, Stephan Elspaß & Robert Möller. 2019. *dtv-Atlas zur deutschen Sprache*, 19th edn. Munich: dtv.
- Kortmann, Bernd. 2010. Areal variation in syntax. In Peter Auer & Jürgen E. Schmidt (eds.), *Language and space: An international handbook of linguistic variation*, vol. 1, *Theories and methods*, 837–864. Berlin/New York: De Gruyter.
- Leino, Antti & Saara Hyvönen. 2008. Comparison of component models in analysing the distribution of dialectal features. *International Journal of Humanities and Arts Computing* 2(1–2). 173–187.
- Lenz, Alexandra N. 2019. Bairisch und Alemannisch in Österreich. In Joachim Herrgen & Jürgen E. Schmidt (eds.), *Sprache und Raum: Ein internationales Handbuch der Sprachvariation*, vol. 4, *Deutsch*, 318–363. Berlin/Boston: De Gruyter.
- Moser, Anne-Marie. 2021. *Negationskongruenz in den deutschen Dialekten*. Stuttgart: Steiner.
- Nerbonne, John. 2006. Identifying linguistic structure in aggregate comparison. *Literary and Linguistic Computing* 21(4). 463–476.
- Patocka, Franz. 1997. *Satzgliedstellung in den bairischen Dialekten Österreichs*. Frankfurt: Lang.
- Pickl, Simon. 2016. Fuzzy dialect areas and prototype theory: Discovering latent patterns in geolinguistic variation. In Marie-Hélène Côté, Remco Knooihuizen & John Nerbonne (eds.), *The future of dialects: Selected papers from Methods in Dialectology XV*, 75–97. Berlin: Language Science Press.
- Pickl, Simon & Simon Pröll. 2019. Ergebnisse geostatistischer Analysen areal-sprachlicher Variation im Deutschen. In Joachim Herrgen & Jürgen Erich Schmidt (eds.), *Sprache und Raum: Ein internationales Handbuch der Sprachvariation*, vol. 4, *Deutsch*, 861–879. Berlin/Boston: De Gruyter.
- Pickl, Simon, Simon Pröll, Stephan Elspaß & Robert Möller. 2019. Räumliche Strukturen alltagssprachlicher Variation in Österreich anhand von Daten des “Atlas zur deutschen Alltagssprache (AdA)”. In Lars Bülow, Ann Kathrin Fischer & Kristina Herbert (eds.), *Dimensions of linguistic space: Variation—Multilingualism—Conceptualisations*, 39–60. Berlin: Lang.
- Schallert, Oliver. 2014. IPP-constructions in Bavarian and Alemannic in comparison. In Günther Grewendorf & Helmut Weiß (eds.), *Bavarian syntax: Contributions to the theory of syntax*, 249–304. Amsterdam: John Benjamins.
- Scherrer, Yves & Philipp Stoeckle. 2016. A quantitative approach to Swiss German: Dialectometric analyses and comparisons of linguistic levels. *Dialectologia et Geolinguistica* 24. 92–125.
- Scheutz, Hannes. 2005. Perspektiven einer neuen Dialektsyntax. In Eckhard Eggers, Jürgen Erich Schmidt & Dieter Stellmacher (eds.), *Moderne Dialekte—Neue Dialektologie*, 291–312. Stuttgart: Steiner.
- Seiler, Guido. 2003. *Präpositionale Dativmarkierung im Oberdeutschen*. Stuttgart: Steiner.
- Shackleton, Robert G. Jr. 2005. English–American speech relationships: A quantitative approach. *Journal of English Linguistics* 33(2). 99–159.
- Spruit, Marco René. 2006. Measuring syntactic variation in Dutch dialects. *Literary and Linguistic Computing* 21(4). 493–506.
- Spruit, Marco René, Wilbert Heeringa & John Nerbonne. 2009. Associations among linguistic levels. *Lingua* 119(11). 1624–1642.
- Stöckle, Philipp, Christina Hemetsberger & Manuela Stütz. 2021. Die WBÖ-Belegdatenbank als Quelle für syntaktische Analysen: Möglichkeiten, Grenzen, Perspektiven. *Wiener Linguistische Gazette* 89. 579–626.
- Szmrecsanyi, Benedikt. 2012. *Grammatical variation in British English dialects: A study in corpus-based dialectometry*. Cambridge: Cambridge University Press.
- Vergeiner, Philip C. 2024. Formen und Funktionen reduzierter Definitartikel: Zur Artikelsyntax in den österreichischen Dialekten. *Zeitschrift für Dialektologie und Linguistik* 91(2). 199–227.
- Vergeiner, Philip C. & Lars Bülow. 2024. Complementizer Agreement in den Dialekten Österreichs: Eine variationslinguistische Perspektivierung. *Wiener Linguistische Gazette* 89. 107–134.

- Vergeiner, Philip C. & Lars Bülow. 2023. Geolinguistic structures of dialect phonology in the German-Speaking Alpine region. A dialectometric approach using crowdsourcing data. *Open Linguistics* 9(1). 20220252.
- Vergeiner, Philip C. & Marlene Hartinger. 2022. Zur Mehrfachnegation in den Dialekten Österreichs: Aktuelle Befunde zu außer- und innersprachlichen Faktoren. In Philip C. Vergeiner, Stephan Elspaß & Dominik Wallner (eds.), *Struktur von Variation zwischen Individuum und Gesellschaft*, 129–147. Stuttgart: Steiner.
- Vergeiner, Philip C. & Konstantin Niehaus. 2024. Article use in Upper German: A “radical” stage of grammaticalization? New findings from Austria. *Linguistic Variation* 24(1). 37–76.
- Weise, Oskar. 1907. Die sogenannte Flexion der Konjunktionen. *Zeitschrift für Deutsche Mundarten* 2. 148–166.
- Weise, Oskar. 1917. Die Relativpronomina in den deutschen Mundarten. *Zeitschrift für Deutsche Mundarten* 12. 64–71.
- Weiß, Helmut. 1998. *Syntax des Bairischen: Studien zur Grammatik einer natürlichen Sprache*. Tübingen: Narr.
- Wieling, Martijn & John Nerbonne. 2015. Advances in dialectometry. *Annual Review of Linguistics* 1(1). 243–264.
- Wiesinger, Peter. 1983. Die Einteilung der deutschen Dialekte. In Werner Besch et al. (eds.), *Dialektologie*, 807–900. Berlin/New York: De Gruyter.
- Wolk, Christoph. 2014. *Integrating aggregational and probabilistic approaches to language variation*. PhD dissertation, University of Freiburg. Retrieved from <https://freidok.uni-freiburg.de/data/9656>.