

## Chapter 3

# The “Atlas of colloquial German in Salzburg”

Julian Blaßnigg<sup>a</sup>, Irmtraud Kaiser<sup>a</sup>, Peter Mauser<sup>a</sup> &  
Konstantin Niehaus<sup>a</sup>

<sup>a</sup>University of Salzburg

This article is concerned with the application of geolinguistic methods onto messy big data in dialectology. It presents data from an on-going project, the “Atlas of Colloquial German in Salzburg” (*Atlas zur Salzburger Alltagssprache*, ASA), which collected data on language use via a large online survey. The questionnaire consisted of 76 items, each of them given in a situational context. The goal of the project is to provide a first detailed description of the everyday German in the federal state of Salzburg, ranging from local dialect to regional dialect (regiolect) to near-standard and standard. The present study provides first results and identifies areas with similar diatopic variation in Salzburg via a cluster analysis of a total of approx. 10,000 data sets (collected via online questionnaires). Regional Spatial Planning Areas (*Planungsregionen*) prove to be a relevant category in order to account for diatopic variation in the state of Salzburg, with clusters of more traditional, dialect-orientated areas in the south, a regiolectal central area, and a northern urban area (encompassing the capital Salzburg City) orientated towards standard German. The article shows that even with anonymous and uncontrolled data collections, the cluster analysis offers a suitable approach towards larger dialect/regiolect areas as well as smaller transition areas. Ultimately, these methods may prove useful not only for the dialectology of (Austrian) German but also for other languages.



## **1 Introduction**

### **1.1 Motivation**

In this contribution, we want to demonstrate from the perspective of German dialectology how recent quantitative geolinguistic methods can be applied to relatively uncontrolled approaches of data collection (compared to traditional survey scenarios in dialectology), in our case crowd sourcing. To this end, we will introduce the project “Atlas of colloquial German in Salzburg” (ASA), which serves as the data basis for our analysis. The ASA arose from the idea of going beyond established dialectological approaches for a central linguistic area of Austria and to present the status quo of current language use in the federal state of Salzburg by collecting data on a broad scale. Compared to earlier dialectological approaches, the focus does not lie on selected prototypical speakers but on the colloquial language use of as many different people as possible. High diversity in terms of age, gender, education, profession etc. should go hand in hand with a dense network of places. We achieved this, as we received answers from nearly all municipalities in Salzburg and by very different people in terms of socio-demographic background. This should allow for statements about small-scale variation as well as it should also render a more representative picture of language use in the region, not least because of the large data source. Before going into more detail about data collection, we start with some general information about Salzburg and the current state of research on diatopic variation in this part of Austria.

### **1.2 General information about Salzburg**

Salzburg is one of the nine federal states of Austria. It is located in the central-western part of Austria. The capital is the city of Salzburg, which is located in the north of the state, right next to the German-Austrian border (cf. Figure 1). Salzburg connects the western part of Austria (Vorarlberg and Tyrol) with the eastern part and borders on four other Austrian federal states (Tyrol, Upper Austria, Styria and Carinthia).

In addition, parts of Salzburg’s border constitute a national border, stretching over 174 km, and mostly shared with the Federal Republic of Germany (districts of Berchtesgadener Land and Traunstein), i.e., 164 km. Only 10 km in the south-west border on the Tauferer Ahrntal in (German-speaking) South Tyrol (Italy) (cf. Landesstatistik Salzburg 2022).



Figure 1: Location of Salzburg in Austria (self created image)

The federal state of Salzburg covers an area of 7,154.6 km<sup>2</sup> and has a population of 562,606 inhabitants (01.01.2022).<sup>1</sup> Salzburg consists of six districts, one of which is the city of Salzburg. Perhaps more interestingly from a geolinguistic perspective, Salzburg is also divided into *Außergebirg* (literally ‘outside the mountains’) in the north and *Innergebirg* (‘in the mountains’) in the south of the Northern Limestone Alps (cf. Figure 2). This geographical division is accompanied by differences in terms of transport connections and infrastructure and therefore influences the dialectal situation in Salzburg. From a dialectological view, the South-Central Bavarian dialects cover most of Salzburg, a smaller part in the north belongs to the Central Bavarian group. Salzburg is therefore a central space in more than one respect: on the one hand, it is a purely geographical centre, due to the many geopolitical borders involved, and on the other hand, it also constitutes a centre of (dialect) variation, due to its transitional position between different dialect regions.

<sup>1</sup>For comparison: Austria has a population of approx. 9 million (01.01.2023).

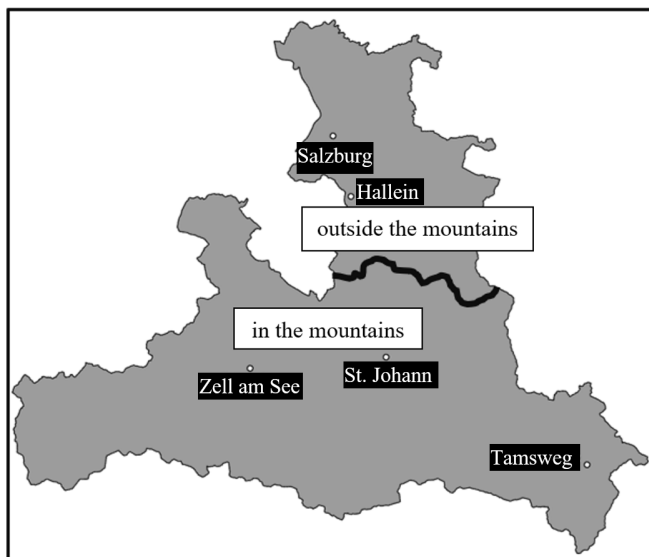


Figure 2: Division between ‘in the mountains’ and ‘outside the mountains’. The border corresponds to the district border between Tennengau and Pongau. N.B.: Populous areas in the South are mostly restricted to valleys. (self-created image)

### 1.3 State of research and research gaps

Dialectological and variational research has often been conducted only for selected locations in the state of Salzburg. Most studies in Salzburg dialectology (Reiffenstein 1955, Scheutz 2007, 2022, Scheutz et al. 2019, Mauser 2021, 2022) focus exclusively on the base dialect. There are also non-Salzburg-specific research projects which include data from the region like the Task-Cluster B of SFB *Deutsch in Österreich* (DiÖ) ‘German in Austria’, *Wortatlas der deutschen Umgangssprachen* ‘Lexical atlas of colloquial German varieties’ (Eichhoff 1977, 1978, 1993, 2000) and *Atlas zur deutschen Alltagssprache* (AdA) ‘Atlas of colloquial German’ (Elspaß & Möller 2003ff.).<sup>2</sup> However, these research projects often only include data on a few locations and usually cannot capture variation within the individual regions of Salzburg.

There are three main aspects which are missing from previous research: first, data that go beyond the local base dialect; second, up-to-date variational data, not least for variation of regiolectal (regional dialect) and near-standard as well as standard lexis; third, more exhaustive and detailed data not restricted to sub-regions or individual locations in Salzburg. Hence, our goal was to conduct a

<sup>2</sup>For a popular science publication based on AdA data, cf. Leemann et al. (2018).

large online survey which draws from the method of the AdA. Its focus on language use includes any point on the vertical variational spectrum in order to grasp colloquial German in Salzburg regardless of its varietal status. In the next section we will elaborate in more detail on what is understood by *colloquial German* and how exactly the data collection took place.

## 2 Data and methods

### 2.1 Definition of colloquial language

Dialectological research in Salzburg typically concentrates on the base dialect, often archaic variants produced by NORMs/NORFs (the oldest form possible). The focus of this project, however, is the colloquial German language of the majority of speakers in Salzburg. So what is meant by the term *colloquial German*?

The project conforms with the AdA’s socio-pragmatic definition of colloquial language:

[D]iejenige Sprachform, die in informellen alltäglichen Situationen spontan und routiniert gesprochen wird (Elspaß 2010: 419)<sup>3</sup>

This criterion applies regardless of whether the language used would be categorised linguistically as standard language, regiolect, or (base) dialect. As a consequence, the project targets the whole spectrum of varieties, including the regiolectal and near-standard variants neglected so far.

### 2.2 Data collection

For the data collection, the project co-operated with the largest regional newspaper *Salzburger Nachrichten* in order to reach as many speakers as possible. All survey rounds were advertised and published by the *Salzburger Nachrichten*. Items used in the AdA questionnaire were replicated or modified. Some items were adapted to the Salzburg situation, for example in the case of expressions for ‘female child’ several variants in regional dialect were added (round 1, question 2). Additionally, regional variables expected to be more specific to Salzburg were created (e.g. variation of prepositions: *an der Bushaltestelle* vs. *bei der Bushaltestelle* ‘at the bus stop’). Sometimes genuine ASA items were created, i.e., variables that are specifically relevant to the (south-central) Bavarian language area,

---

<sup>3</sup>‘That form of a language which is spoken spontaneously and routinely in informal everyday situations’.

e.g. subordinate clauses introduced by *dass* versus introduced by *um ... zu* (round 4, question 6) (cf. Bayer 2020). The survey participants were asked to select their individual preferred variant from the options given.

Four rounds of questionnaires were distributed from October to December 2019, covering 76 variables of lexis, grammar, phonetics and pragmatics. We also collected multiple types of metadata for a quantitative variational analysis: place of residence, age, gender, highest educational qualification, occupational group, duration of residence and origin of parents. More than 10,000 speakers from Salzburg participated in the survey, of which 9,668 responses have been analysed. Some answers had to be excluded from the analysis due to missing data.

The questionnaire was created using LimeSurvey. We mostly relied on items with single answers but also used multiple answers in some cases, e.g. with variants of greeting and saying farewell. We tried to contextualise the item as much as possible and used visual material to explain meaning and context without additional language input. Each round consisted of approximately 20 thematic items, questions on socio-demographic metadata, and a free-comment section at the end of the questionnaire. The following welcome text provided respondents with an explanatory introduction to the survey:

Bitte geben Sie bei den folgenden Fragen an, was man in Ihrem Ort normalerweise hört – egal, ob eher Mundart, eher Hochdeutsch oder etwas dazwischen. Es gibt keine richtigen und falschen Antworten.

Am Ende des Fragebogens können Sie (unter „Ihre Anmerkungen“) beschreiben, ob und wie einzelne Ausdrücke unterschiedlich gebraucht werden (z.B. wie häufig, mit welcher unterschiedlichen Bedeutung usw.).

Versuchen Sie bitte, sich alle Beispiele in Ihrer ‚normalen‘ Aussprache vorzustellen. (ASA, round 4)<sup>4</sup>

Respondents' average age (40.7 years) roughly corresponds to that of the federal state (43.1 years). There is a slight overrepresentation of participants with higher educational qualifications (A-levels or degree). However, this is not unusual in surveys originating from the university sector – even if they are disseminated via the regional newspaper (for more details see Niehaus et al. 2022:

---

<sup>4</sup>In the following questions, please indicate what is normally heard in your location – whether it is more dialect, more High German or something in between. There are no right and wrong answers.

At the end of the questionnaire, you can describe (under “Your comments”), whether and how individual expressions are used differently (e.g. how often, with what different meanings, etc.). Please try to imagine all examples in your “normal” pronunciation.

N.B.: “Hochdeutsch” here is used in its lay meaning “Standard German”.

112–114). Moreover, there are significantly more female than male participants in our survey, the proportion of men is only 33.1%.

34.4% of the respondents have always lived in the place for which they provided the information, 18% for more than 30 years, 37.1% for more than 10 years, and only 9.2% have lived there for less than 10 years (no information: 1.3%). Thus, we can conclude that most respondents are sufficiently familiar with their place of residence and the local language.

## 2.3 Data analysis

### 2.3.1 Methods used

Data was collected and cleaned up, i.e. responses with missing data records were deleted. The remaining total of 9,668 responses from all four rounds were then examined more closely and in detail via SPSS. Statistical analyses were conducted for each district, planning region (see next section) and municipality in Salzburg. At the same time, the metadata were utilised in a quantitative variational analysis, e.g. with regard to gender<sup>5</sup> or age. Furthermore, social factors of regional language use were examined, e.g. whether young women from the city chose significantly more standard (or near-standard) variants than old men from rural areas (who chose variants of traditional dialect more often).

In order to determine whether the results are only due to individual phenomena or indeed similarities across all variables, the data were subjected to a factor analysis with GeoLing as well as a cluster analysis with R. To cluster data we used hierarchical cluster analyses with Ward’s algorithm. For variational linguistic applications, the Ward algorithm seems to be relatively useful as it does not overestimate smaller differences and fluctuations in the data but strongly sanctions larger differences. More information on cluster analysis in general can be found in Section 3.3.1.

Each item as well as the results of the factor and cluster analyses are finally visualised using ArcGis-Pro, which has proven to be well suited for the cartographic representation of geolinguistic data.

### 2.3.2 Basic diatopic levels of analyses

Popular belief in Salzburg may refer to several extra-linguistic criteria which index diatopic differences within the federal state. In terms of physical geography, the division into Innergebirg (in the south) and Außergebirg (in the north)

---

<sup>5</sup>In fact, men and women gave very similar answers to almost all questions. This is the reason why our results are not devalued by the somewhat imbalanced gender ratio.

is also of some importance. Innergebirg includes the districts of Zell am See (Pinzgau), St. Johann (Pongau) and Tamsweg (Lungau) while the Außergebirg includes the district of Hallein (Tennengau), Salzburg-Land (Flachgau) and the city of Salzburg (cf. Figure 2).<sup>6</sup>

The areas covered by the individual districts differ considerably, for example the largest district in size – the Pinzgau – occupies a cadastral area of 2,641.1 km<sup>2</sup> and is thus the third largest district in Austria, whereas the city of Salzburg measures only 65.7 km<sup>2</sup> and is thus significantly smaller than many municipalities in the federal state (see Figures 3 and 4). The opposite is true with regard to population size, where the city of Salzburg has the highest number of inhabitants. The data on which the ASA is based was generally analysed at the municipality level. A total of 109 of the 119 municipalities in Salzburg are included in all four rounds of our survey and were thus included in the analyses. Since approx. 92% of the Salzburg municipalities are represented in all survey rounds, it is possible to conduct fine-grained analyses at the municipal level. At the same time, some municipalities are represented by only a few responses, which means that an analysis of the results in larger areas seems to be useful in order to avoid generating a distorted picture of the actual language use by overrating a few individual answers.

Districts, on the other hand, often proved too coarse for our variational as well as geolinguistic analyses. For this reason, recourse was made to a level between municipalities and districts: the so-called *planning regions*. The planning regions are a unit of supra-local spatial planning, they are administratively anchored in the national spatial planning programme and also reflect historically-grown regional identities in many cases, which makes them ideal as a basis for geolinguistic analyses in Salzburg. The planning regions offer a good compromise and therefore lie at the core of our quantitative approach and cartography. In an extract from the *Salzburg Regional Development Programme* (revision 2003, published by the Department of Spatial Planning of the Salzburg Federal State Government) the planning regions are defined as follows:

Nach dem Raumordnungsgesetz ist es eine der ausdrücklich festgelegten Aufgaben des Landesentwicklungsprogramms, das Land in Planungsregionen zu gliedern. [...] Zu Regionen wurden darin nach Anhörung der Gemeinden Gebiete zusammengefasst, die strukturell und funktional zusammengehören und entsprechend den Erfordernissen der Raumordnung als

---

<sup>6</sup>This is reflected, for example, by dialect lexicons compiled by amateurs, which usually attempt to capture the dialect of a district, e.g. *Pinzgauer Mundart Lexikon* (Schwaiger & Höck 2010).



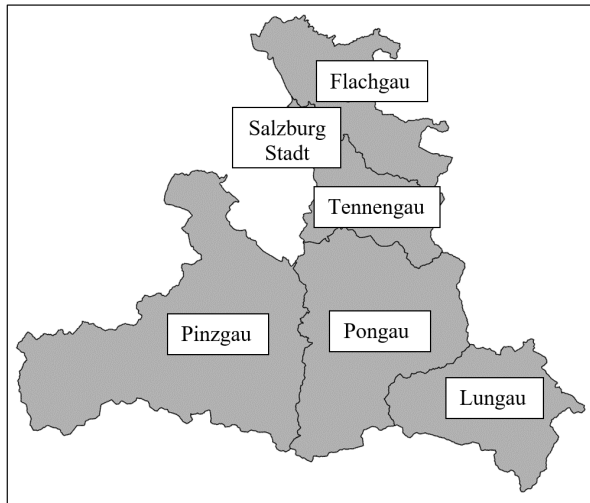


Figure 3: Districts of Salzburg (self created image)

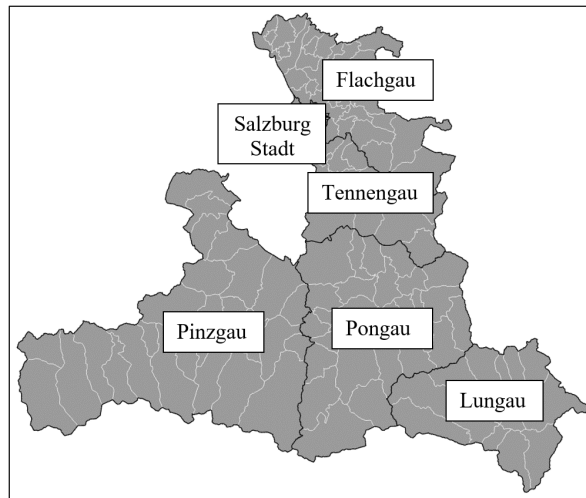


Figure 4: Municipalities in Salzburg (self created image)

Einheit entwickelt werden sollen, wobei auch die „Identifikation“ der Gemeinden mit einer bestimmten Region berücksichtigt wurde. (Mair 2003: 90)<sup>7</sup>

As can be seen, the planning regions are based on national planning as well as regional identities.

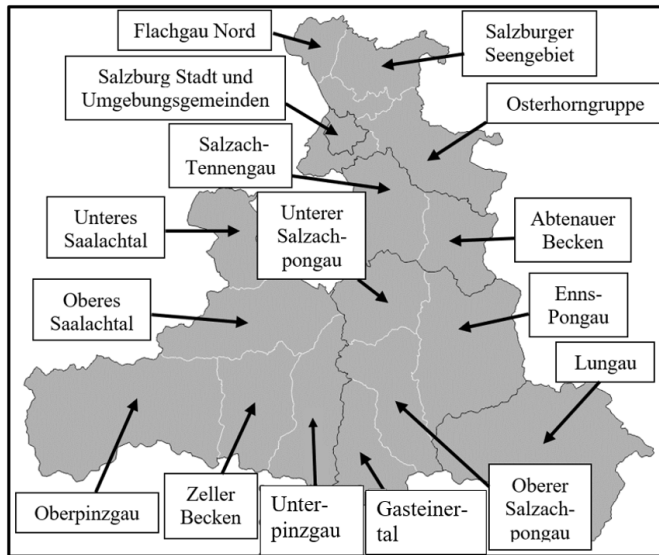


Figure 5: Official regional planning regions (cf. Mair 2003) – (self created image)

Each of the 16 planning regions in the province of Salzburg consists of an average of seven municipalities. As far as the number of respondents per planning region is concerned, numbers range from 102 (round 4, Zeller Becken) to 1003 (round 1, Lungau) per 100,000 inhabitants. These numbers ensure a realistic approximation to inner-regional language use in Salzburg while at the same time much more detailed insights can be gained than if only the district level were considered. Additionally, analyses were carried out at district and municipality level in order to avoid pitfalls during interpretation; most importantly, we applied multidimensional scaling (MDS) on both levels, which allows for a calculation of linguistic similarities.

<sup>7</sup>According to the Spatial Planning Act, one of the expressly defined tasks of the Development Programme is to divide the federal state into Regional Spatial Planning Areas [“Planungsregionen”]. [...] After consultation with the municipalities, areas that belong together structurally and functionally and are to be developed as a unit in accordance with the requirements of spatial planning were grouped together in these regions, whereby the “identification” of the municipalities with a particular region was also taken into account’.

## 3 Findings

### 3.1 Mapping: General approach

The following maps depict one variable each, i.e. dominant variants in a combined heat map. For this purpose, each dominant variant is given its own colour, while the relative frequency of the dominant variant is represented by colour intensity. Put simply: the more intense the colour, the stronger the prevalence of a variant. All sub-divisions on the maps are based on the planning regions, with district boundaries and the respective district capitals visualised for better orientation. At the same time, many maps show that the distribution of individual phenomena is often not identical with district boundaries, which increases the additional informative value of analyses at the level of planning regions. To illustrate this, two maps will be showcased below and their results will be presented as examples. The presented maps (Figures 6 and 7, page 68) should be understood such that the intensity of the color – either red or blue – correlates with the prevalence of the variant. The stronger the color shading, the more widespread the variant.

### 3.2 Showcase results

#### 3.2.1 Example ‘female child’

The designations for ‘female child’ are a classic in German-speaking dialectology. The German dialects in general, including those spoken in Austria, boast a multitude of expressions and phonetic variants of these expressions for referring to the category of ‘female child’. Our results for Salzburg show that two expressions are particularly common here: *Mölz* and *Dirndl*.

The south-western region of the Pinzgau is strongly dominated by *Mölz* except for the Unteres Saalachtal (‘Lower Saalach Valley’), an area in the northern Pinzgau. This demonstrates the strength of working with planning regions, for the popular assumption that *Mölz* is the preferred term for ‘female child’ in the entire Pinzgau cannot be confirmed by our data. Our survey would also have yielded this very result if the districts alone had been the basis of our cartographic representation. By using planning regions, however, it becomes apparent on the one hand, that there is a region in the Pinzgau where *Mölz* is not at all the dominant variant and, moreover, that the frequency of this particular variant differs within the district, specifically that *Mölz* is used most frequently in the far west. Similar observations can be made for *Dirndl*: Although it is the most frequently used variant throughout the rest of the federal state, usage is particularly frequent in the

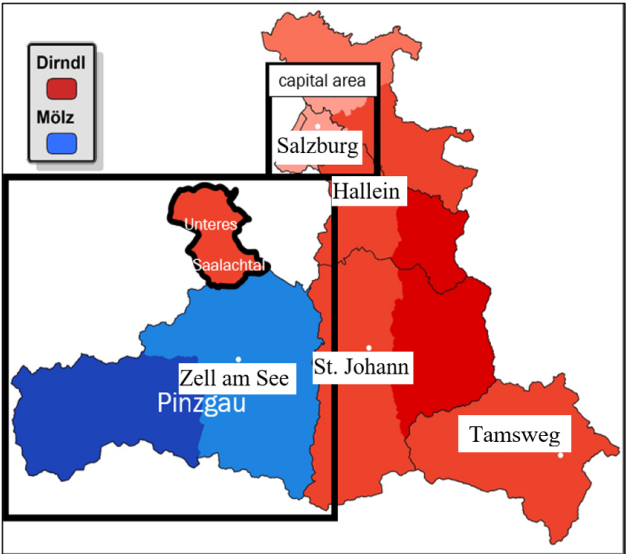


Figure 6: Names for ‘female child’ (self created image)

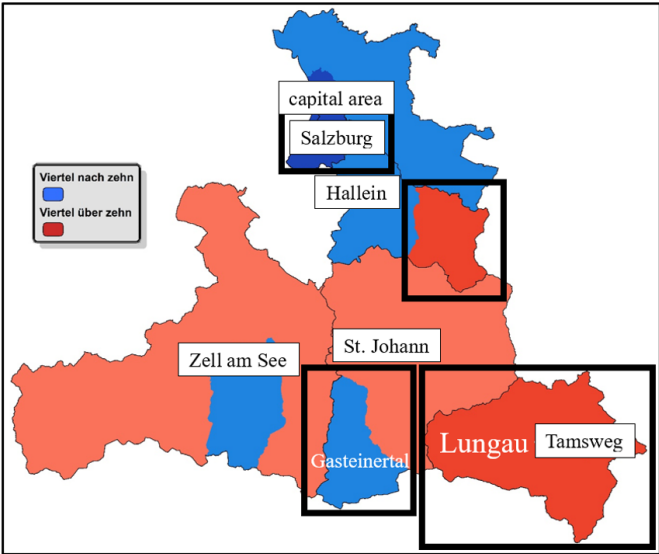


Figure 7: Verbalisation of 10:15 (self created image)

east of the federal state. In the capital area around the city of Salzburg, *Dirndl* is only just the most frequently mentioned designation for ‘female child’: Although it is still dominant, there is a higher percentage of standard variants – and more variation in general (which does not come as a surprise and applies to many variables). In addition to *Dirndl*, the standard German *Mädchen* (approx. 17%) and near-standard *Mädl* (approx. 36%) are also frequently used in and around the city of Salzburg. Hence, *Dirndl* is not a quasi-absolute default variant in most of the federal state but a relatively dominant variant most prevalent in the east.

#### 3.2.2 Example 10:15

A particularly suitable item for our demonstration of the data is the verbalisation of a certain time of day, in (standard) English *ten-fifteen* or *quarter past ten*. To avoid verbal stimuli, this question only showed a picture of a clock and asked for the phrase to name the corresponding time. The German variants in the state of Salzburg are *Viertel nach zehn* ‘quarter past ten’ (in blue) and *Viertel über zehn* ‘quarter over ten’ (in red). The map (Figure 7) demonstrates a pattern which surfaces often in our data, i.e. an opposition between the more urban north (capital area) and the more rural southeast (Lungau), notwithstanding a few exceptions such as the northern Abtenauer Becken (‘Abtenau basin’), which often joins the language use of the south – and vice versa the southern Gasteinertal (‘Gastein valley’), which often conforms with the language use of the northern capital area. Within the individual districts clear-cut areas with different degrees of variant use become apparent.

### 3.3 Results of cluster analyses

#### 3.3.1 General approach

Our data have a nominal scale form (as is customary in variational linguistics) and can easily be spread in a table: Each row of the table contains the data of exactly one informant. The columns represent the variables of the survey, the individual cells contain the respective variant that the informant has given for the variable. As already mentioned, most of the variables are of a purely linguistic nature, with the exception of the metadata information on the respondents as well as the free comments section. These are also documented individually for each informant.

This was the basis on which we evaluated the most prevalent variants for each planning region. However, our interest is not only in the results for each individual item but we also wanted to identify inter-local and inter-regional similarities,

which only become apparent with a sufficient amount of data combined. Factor and cluster analyses are suitable for such approaches (cf. Goebel 2005: 509, Pröll et al. 2021: 234–235, note 14). In the following, the results of the cluster analysis will be presented.

The basis for these further analytical steps is no longer the original table of nominally scaled, individual raw data, but a table with the relative frequencies of the individual variants per survey location. These calculations were conducted for all our diatopic levels of analysis, i.e. the districts, the planning regions, and single municipalities. In the present paper, we will only take a closer look at two cluster maps which are based on the planning regions. A distance matrix of the planning regions to each other is calculated from the relative frequencies of the linguistic variants per planning region. In order to make this high-dimensional variation less complex and interpretable, cluster analysis was used: The individual planning regions were grouped based on their similarity or dissimilarity. These groups (= clusters) should be as homogeneous as possible within themselves and at the same time as different as possible from each other. Ultimately, therefore, one obtains a concrete allocation of the individual regions to groups, which in turn can be represented and interpreted cartographically. These groupings can be represented at different levels of granularity, e.g. as a 2-cluster, 3-cluster or 4-cluster solution.

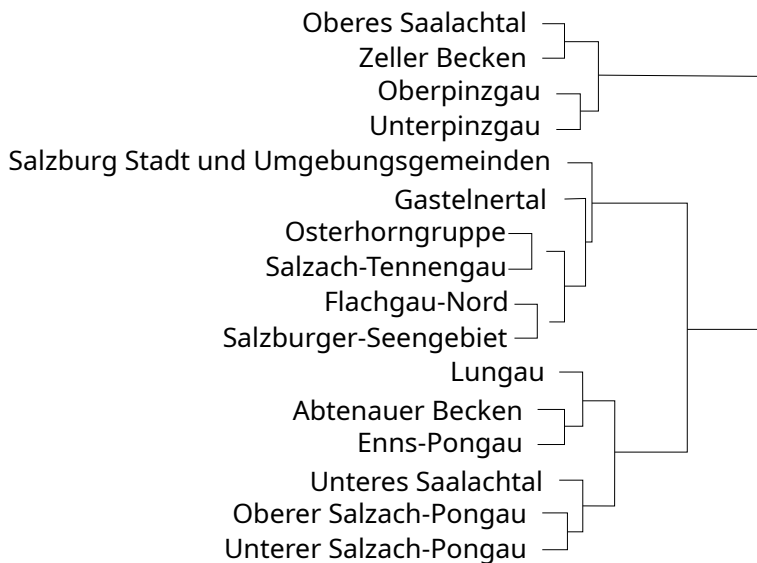


Figure 8: Cluster analysis dendrogram (self created image)

The allocation of the planning regions to individual clusters is best represented in the form of a dendrogram (cf. Figure 8). Depending on where the cut is made to distinguish clusters, solutions with different numbers of clusters are obtained. Since the length of the horizontal lines provides information about the linguistic distance between the individual clusters, the dendrogram can also be used to determine which cluster solutions ultimately represent plausible groupings.

As can be seen from the dendrogram, mainly 2-, 3- and 4-cluster solutions are suitable. We will briefly discuss a 3- and a 4-cluster solution.

### 3.3.2 3-cluster solution

Popular belief as to how the federal state of Salzburg could be divided into dialect areas include the idea that there is a clear linguistic division between the dialects *Außergebirg* (‘outside the mountains’, i.e. in the north) and *Innergebirg* (‘in the mountains’, i.e. in the south). In addition, political districts are ascribed their own regional dialects such as *Pinzgauerisch* in the south-western part of the state and the like. The 3-cluster map (cf. Figure 9) shows that the actual linguistic situation is clearly more complex.

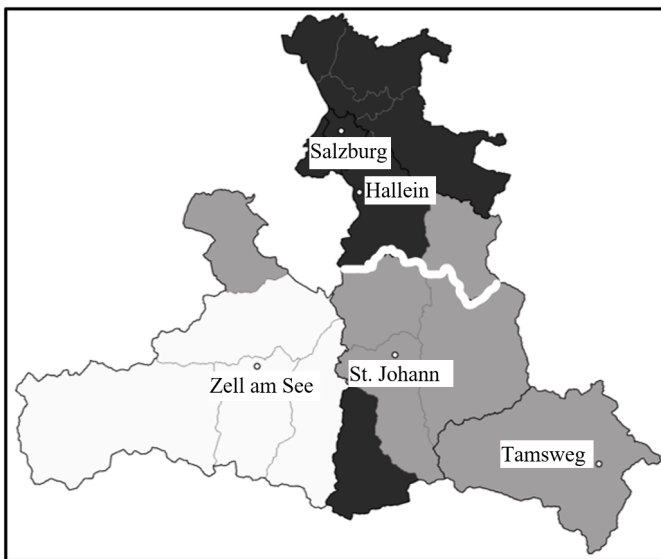


Figure 9: 3-cluster solution (self created image)

In addition to the boundaries of the planning regions and the district capitals, we also drew the border between Innergebirg and Außergebirg (bold white line).

The border runs between the two districts of Tennengau and Pongau. From the perspective of physical geography, the lay-linguistic notion is understandable: Some Salzburgers from the Innergebirg travelling to the capital will immediately notice a change, when the Salzach valley widens outside the Pass Lueg and the rugged mountains give way to gentle hills, and, of course, natural obstacles can be a factor of linguistic variation. However, linguistic difference may subjectively be perceived as much more salient and relevant than the distance measured by linguistic means. For example, the Abtenauer Becken (in grey, north of the white line – a basin in terms of geology) clearly belongs to the south in terms of linguistic distance in this cluster solution. On the other hand, the cluster resembling the Außergebirg area (black) also includes the Gasteinertal in the south, which is geographically clearly a mountain valley of the south. On top of that, the Innergebirg does not add up to a single cluster but instead consists of three clusters, mainly two large clusters in white and in grey in the (south-)west and in the (south-)east respectively. Not even the Pinzgau district (white) consists of a single cluster, as again the northern part, Unteres Saalachtal, belongs to a different cluster – a phenomenon which could already be observed with the lexical variants of ‘female child’ (see Section 3.2.1, Figure 6). The map for ‘10:15’ (see Section 3.2.2, Figure 7) again shows very well the outliers also evident in the 3-cluster solution, namely the Abtenauer Becken and the Gasteinertal, which in many maps stand out with regard to other linguistic variables from their surroundings. In sum, it can be said that neither the linguistic relevance of political districts nor the popular binary distinction of varieties *in* and *outside* the mountains can be confirmed through our cluster analysis. Rather, it seems that other factors play a crucial role, such as the infrastructural proximity to the city of Salzburg. At the same time, the cluster which mainly covers the south-east and centre-most areas seems to form a transition area between the urban north around the city of Salzburg and the rural south(-west). The fact that regional centres such as St. Johann im Pongau are relatively well-connected to the city of Salzburg in terms of infrastructure seems to be at work here. Even from the Unteres Saalachtal – via a short route through German territory – you can reach the city of Salzburg relatively fast, from some places (e.g. Unken) even faster than the district’s own capital, Zell am See.

### **3.3.3 4-cluster solution**

The 4-cluster solution (see Figure 10) can refine the 3-cluster picture a little more: Differences to the 3-cluster solution are especially evident in the transition area, which now breaks up into two clusters.



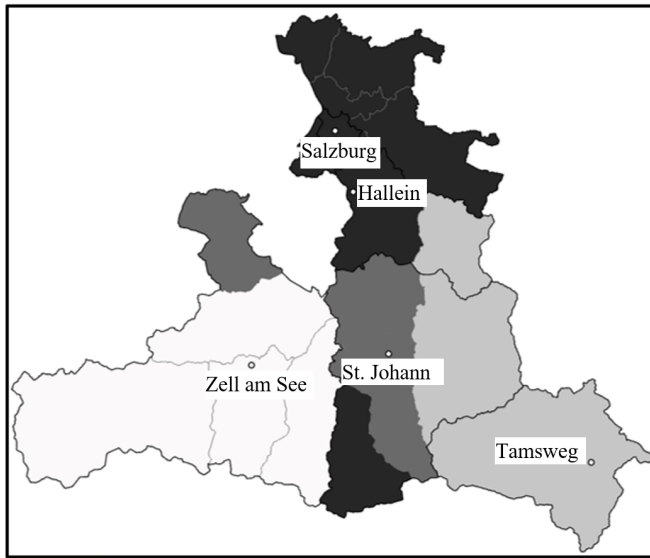


Figure 10: 4-cluster solution (self created image)

Similar to the 3-cluster solution, there is a pseudo-Außergebirg cluster (in black), with a tendency towards either standard variants or Bavaro-Austrian regiolectal variants (that is, variants which are also common in other parts of Austria and in south-eastern Germany/Bavaria). In the Innergebirg, the large eastern cluster of the 3-cluster solution breaks apart into smaller clusters, one transition area more similar to the northern cluster (in dark grey) and the easternmost rural area, which maintains base dialect variants more strongly (in light grey).

The differences between the dark grey and the light grey clusters can be illustrated by numerous examples: the Pongau central area (in the west of the district of St. Johann in dark grey) more often tends to use variants that are common in the city of Salzburg and its surroundings, e.g. the grammatical gender of email, neutr. *das E-Mail* (Enns-Pongau) vs. fem. *die E-Mail* (Pongau central area). The West (in white) is obviously orientated towards dialect variants similar to those of adjacent west Austrian dialects, such as Tyrolean, a relatively stable result regardless of the number of clusters.<sup>8</sup>

The assumption that the infrastructural proximity to the city could play a decisive role is corroborated by this 4-cluster solution. Those regions of the original transition cluster in the 3-cluster solution which are particularly well connected

<sup>8</sup>Cf. data in the *Tiroler Dialektarchiv* (Archive of Tyrolean Dialects): <https://www.tiroler-dialektarchiv.at>. For instance, *Mölz* appears as *Mötz* in Tyrol, cf. [https://wiki.uibk.ac.at/tda/index.php/Bedeutung\\_14](https://wiki.uibk.ac.at/tda/index.php/Bedeutung_14).

to the city of Salzburg now fall into the more urban, regiolectal<sup>9</sup> transition cluster (dark grey), while the eastern and southeastern regions (light grey) are closer to the rural-dialectal forms of the southwest (white). This interpretation is corroborated by an MDS-analysis.

Once again, however, the Gasteinertal in the south (black), which is located *in* the mountains, joins the northern urban cluster. Infrastructural proximity to the city can hardly account for this. Yet, a long-standing history of (health and spa) tourism and subsequent migration (e.g. German or Viennese retirement residences in the Gasteinertal, going back as far as the late 19th century) may explain the similarities of the Gasteinertal and the northern regions: It entails either the tendency to converge towards wider regional, non-local variants or to use standard variants. Historical migration through jobs in tourism and retirement could have left a linguistically sustainable mark by changing the composition of the population. We will pursue this case further in the future, among other factors that still await analysis.

## 4 Conclusion and outlook

In addition to traditional dialect areas and popular divisions (Innergebirg/Außergebirg), which are only partially confirmed by our data, various factors influence patterns of language variation in the federal state of Salzburg.

Our analysis has demonstrated that infrastructural proximity to the city of Salzburg seems to be a major factor which may have contributed to creating a transitional regiolect. Infrastructure seems to be at least as important as mere geographical distance, sometimes it even trumps the latter. In other cases, this results in groups of diatopic variants which can only be explained by consulting both factors: The Abtenauer Becken, for example, is geographically closer to the city of Salzburg than, for example, St. Johann. However, St. Johann is better connected to the city of Salzburg in terms of infrastructure due to the nearby motorway and the railway connection. The MDS-analysis mentioned above shows that St. Johann is in fact linguistically closer to the city of Salzburg than the Abtenauer Becken.

Smaller regional centers also play a certain role in many cases (in addition to the city of Salzburg as the overall centre of the federal state). These are often also the target of commuter flows and usually have a better infrastructure than many

---

<sup>9</sup>Common standard linguistic variants would be, for example, *Mädchen* 'female child' as opposed to regiolectal *Dirndl* and dialectal *Mölz/Tättl*, but also pronunciation variants such as [ˈmɪɕ] *Milch* 'milk' as opposed to regiolectal [ˈmyːɕ] and dialectal [ˈmiːɕ] as well as subjunctive II constructions such as (*wir*) *bräuchten* 'we would need' as opposed to regiolectal *täten/dadn* ... *brauchen* and dialectal *brauchatn* (cf. also Niehaus et al. 2022).

rural communities in the mountain regions. At the same time, these regional centres often show linguistic characteristics that bring them closer to the regionally influenced urban language of the capital than to the rural base dialects which surround these regional centres. In sum, the situation is strongly reminiscent if not a prime example of Trudgill’s (1974) gravity model where linguistic change spreads from one (urban) centre to another and only then diffuses into the surrounding areas. Accordingly, there appears to be a preservation of traditional dialect variants outside these smaller centres, i.e. in the westernmost and easternmost rural parts of Salzburg, in contrast to more dynamic areas in the north.

This insight has been gained by a combination of quantitative analyses which are sufficiently fine-grained and could be complemented by qualitative studies to account for the data in future approaches. As the case of the Gasteinertal has shown, even with comprehensive statistics for a large amount of data one must take into account the specific histories of locations which can leave their mark linguistically, irrespective of geographic and/or infrastructural circumstances.

Finally, social factors such as age and profession are expected to play a significant role for the regional language use in Salzburg (e.g. young urban vs. old rural). Analyses on these factors will be presented in forthcoming research.

## References

- Bayer, Josef. 2020. What’s unique in Bavarian syntax? Thoughts on the occasion of a performance of Bach’s St Matthew Passion. In Tomoya Tsutsui Tanaka & Masashi Hashimoto (eds.), *Linguistic research as an interdisciplinary science*, 27–43. Tokyo: Hituzi Publishers.
- Eichhoff, Jürgen. 1977. *Wortatlas der deutschen Umgangssprachen*, vol. 1. Bern: Franke.
- Eichhoff, Jürgen. 1978. *Wortatlas der deutschen Umgangssprachen*, vol. 2. Bern: Franke.
- Eichhoff, Jürgen. 1993. *Wortatlas der deutschen Umgangssprachen*, vol. 3. München: Saur.
- Eichhoff, Jürgen. 2000. *Wortatlas der deutschen Umgangssprachen*, vol. 4. Bern: Saur.
- Elspaß, Stephan & Robert Möller. 2003ff. *Atlas zur deutschen Alltagssprache*. <https://www.atlas-alltagssprache.de>.
- Elspaß, Stephan. 2010. Alltagsdeutsch. In Hans-Jürgen Krumm, Christian Fandrych, Britta Hufeisen & Claudia Riemer (eds.), *Deutsch als Fremd- und Zweitsprache. Ein internationales Handbuch*, 2nd edn., vol. 1, 418–424. Berlin, New York: de Gruyter.

- Goebel, Hans. 2005. Dialektometrie. In Reinhard Köhler, Gabriel Altmann & Rajmund G. Piotrowski (eds.), *Quantitative Linguistik: Ein internationales Handbuch* (Handbücher zur Sprach- und Kommunikationswissenschaft), 498–531. Berlin, New York: de Gruyter.
- Landesstatistik Salzburg. 2022. *Gemeindeportraits*. <http://www.salzburg.gv.at/themen/statistik/gp-statistik-daten-gemeindeportraet>.
- Leemann, Adrian, Stephan Elspaß, Robert Möller & Timo Grossenbacher. 2018. *Grüezi, Moin, Servus*. Reinbek bei Hamburg: Rowohlt.
- Mair, Friedrich (ed.). 2003. *Salzburger Landesentwicklungsprogramm*. Gesamtüberarbeitung. Salzburg: Amt der Salzburger Landesregierung.
- Mauser, Peter. 2021. *Wiarach ba ins ret: Das Lungauer Sprachbuch*. Tamsweg: Pfeifenberger.
- Mauser, Peter. 2022. *Wiarach ba ins ret: Die sprechende Landkarte zum Buch*. <https://www.pfeifenberger.at/wiarach/landkarte/>.
- Niehaus, Konstantin, Irmtraud Kaiser & Peter Mauser. 2022. Der Konjunktiv II in Salzburger Varietäten: Grammatik, Gebrauch, soziale Faktoren. *Linguistik Online* 114(2). 99–128.
- Pröll, Simon, Stephan Elspaß & Simon Pickl. 2021. Areal microvariation in German-speaking urban areas (Ruhr area, Berlin, and Vienna). In Arne Ziegler, Stefanie Edler & Georg Oberdorfer (eds.), *Urban matters: Current approaches of international sociolinguistic research*. (Studies in Language Variation 27), 227–252. Amsterdam, Philadelphia: Benjamins. DOI: 10.1075/silv.27.10pro.
- Reiffenstein, Ingo. 1955. *Salzburgische Dialektographie. Die südmittelbairischen Mundarten zwischen Inn und Enns* (Beiträge zur deutschen Philologie 4). Gießen: Wilhelm Schmitz Verlag.
- Scheutz, Hannes (ed.). 2007. *Drent und herent: Dialekte im salzburgisch-bayerischen Grenzgebiet*. Mit einem sprechenden Dialektatlas auf CD-ROM. Unter Mitarbeit von Sandra Aitzetmüller und Peter Mauser. Salzburg: Salzburg – Berchtesgadener Land – Traunstein EU Regio.
- Scheutz, Hannes. 2022. *Salzburger Sprachatlas*. <https://www.sprachatlas.at/salzburg/index.html>.
- Scheutz, Hannes, Aitzetmüller Sandra & Peter Mauser. 2019. *Drent und herent: Dialekte im salzburgisch-bayerischen Grenzgebiet*. <https://www.sprachatlas.at/drentherent/>.
- Schwaiger, Alois & Leonhard Höck. 2010. *Pinzgauer Mundart Lexikon*. Leogang.
- Trudgill, Peter. 1974. Linguistic change and diffusion: Description and explanation in sociolinguistic dialect geography. *Language in Society* 2(2). 215–246. DOI: 10.1017/S0047404500004358.