

Article

Reconciling the social and spatial: An apparent-time analysis of variation intensity in colloquial German

Mason A. Wirtz¹, Simon Pickl², Konstantin Niehaus², Stephan Elspaß² and Robert Möller³

¹English Department, University of Zurich, Zurich, Switzerland, ²Department of German Language and Literatures, University of Salzburg, Salzburg, Austria and

³Department of Modern Languages, University of Liège, Liège, Wallonia, Belgium

Abstract

In this article, we aim to reconcile the interrelated role of social and spatial influences on colloquial German by conducting an apparent-time analysis of 15 lexical, morphological, and phonetic variables from the *Atlas of Colloquial German* (AdA). Specifically, we introduce a “change in variation intensity” (CVI) measure meant to gauge the degree of change in colloquial speech between younger and older cohorts. Using this measure, we investigate (a) whether the CVI within and across variables is prone to spatial patterns and (b) the extent to which the CVI differs systematically between variable types. Results from visual–spatial analyses indicate clear spatial patterns, though the geographical distributions are largely item-specific, and a linear mixed-effects model revealed no effect of variable type. More broadly, the CVI measure represents an important advancement in how we address language variation and change from a methodological–statistical perspective, specifically when dealing with heterogeneous, crowdsourced data.

Keywords: apparent-time analysis; *Atlas zur deutschen Alltagssprache*; language change; variation intensity; ergodicity

1. Introduction

Traditional aggregative dialectometry often operates under the premise that, in order to “neutralize” or at least generalize over variable-level noise and item-related discrepancies in variation patterns, it is an arguable necessity to “aggregate the differences in many linguistic variables” (Nerbonne, 2010:3822; see also Nerbonne, 2009). In so doing, the goal is to sketch global, overall spatial tendencies of language variation, and myriad atlases have been dedicated to investigating the way different variants and aggregates thereof are distributed across space (and time). These notably range from geographically constrained, small-scale areas (e.g. Pröll, Elspaß & Pickl, 2021) to larger regions and entire countries, and can transcend national boundaries (e.g. Pickl et al., 2019). In a complementary vein, variationist approaches to sociolinguistics address questions concerning how speaker-internal (see e.g. Werth et al., 2021; Vergeiner & Elspaß, 2025) alongside external social (e.g. Montemagni & Wieling, 2017; Wirtz, 2022; Steiner, Jeszenszky & Leemann, 2023) and spatial/distance-related (e.g. Heeringa & Nerbonne, 2001; Gooskens, 2005) factors influence variable patterns of language use, and how biological–cognitive variables (Tamminga, MacKenzie & Embick, 2016) and career-related life-course transitions (Riverin-Coutlée & Harrington, 2022) affect (changes in) linguistic behavior.

That said, the scope of work attempting to reconcile dialectometric analyses with variationist-informed approaches to

how the social plays into the spatial remains comparatively limited (though see e.g. Steiner, Jeszenszky & Leemann, 2023). This is because areal language variation is often argued to be a function of geographical distance (see e.g. Séguy, 1971; Nerbonne, 2010; Stanford, 2012), which is why most major atlases attempt to cover large areas of space while at the same time constraining the social scope (thus the focus on NORM/NORF [i.e. *nonmobile, older, rural male/female*] speakers). As Pröll et al. (2021:244) put it, the overwhelming focus on geographical variation has strengthened the notion that “social variation would distract from the spatial picture and obscure the spatial signal.”

Notably, recent years have seen a rise in more socially diverse datasets capitalizing increasingly on crowdsourcing and gamification-based data collection methods (see e.g. Kolly & Leemann, 2015; Leemann, Derungs & Elspaß, 2019; Steiner, Jeszenszky & Leemann, 2023). For example, in atlases such as the *Atlas zur deutschen Alltagssprache*—*Atlas of Colloquial German* (hereinafter AdA), from which the data for the present article are analyzed, the localities are not predefined, and participation is not constrained by any social-specific factors. The goal herewith is to facilitate a more realistic portrait of colloquial language variation by capitalizing on informants’ “expertise of the local language use” (Pickl et al., 2019:41, our translation).

In considering a socially more diverse set of informants inherent to AdA’s aforementioned design, however, we are likely to capture a higher degree of linguistic variation within (and likely also between) localities compared to other atlas projects that consider only one or a few select NORM/NORF speakers per locality. The question then arises: How do we measure the degree of variation present within a locality? To address this issue,

Corresponding author: Mason A. Wirtz; Email: mason.wirtz@es.uzh.ch

Cite this article: Wirtz MA, Pickl S, Niehaus K, Elspaß S, Möller R. (2025). Reconciling the social and spatial: An apparent-time analysis of variation intensity in colloquial German. *Journal of Linguistic Geography* <https://doi.org/10.1017/jlg.2025.10001>

Stoeckle (2016) introduced the “Variation Index,” which operationalizes variation as the degree of agreement between the informants regarding the locally dominant variant (i.e. the variant that was provided by the highest number of informants, at each location). We build on Stoeckle’s index, proposing in this contribution the “variation intensity” measure, which quantifies the degree of variation by taking into account not only the dominant variant, but rather the total number of variants used and their relative frequencies at each respective locality.

What is more, Stoeckle (2016:212) raises the question regarding to what extent the observed degree of variation within a specific locality is stable or subject to ongoing change, claiming that “it will be helpful to take a closer look at sociolinguistic variables [in order to] see whether there are indicators for apparent-time change.” In line with Stoeckle’s suggestion, and also with our broader agenda to reconcile geolinguistic and variationist sociolinguistic approaches, we quantify the difference in variation intensity between two age cohorts (i.e. a younger and older generation) within a locality, which we refer to as change in variation intensity. This measure allows us to capture not only the degree to which apparent-time change in variation intensity in colloquial German can be observed, but also the *directionality* (i.e. with more extensive variation among either younger or older informants) as well as the *spatial patterns* thereof. In what follows, we investigate change in variation intensity on the basis of 15 lexical, morphological/morphosyntactic, and phonetic variables collected in the AdA.

Additionally, we take this article in a methodological direction. To start, we address the “group-to-individual generalizability issue” in applied linguistics, which focuses on notions of how the individual—or, in our case, individual linguistic *variables*—measure up to group-level trends (i.e. aggregate analyses). From a complexity theory perspective, Larsen-Freeman & Cameron (2008:238) underscored that “variability in data is not noise to be discarded when averaging across events or individuals, or the result of measurement error,” and this raises an important issue in classic dialectometric and variationist approaches concerning whether apparent-time change at the aggregate level is also present at the level of different linguistic variables (see also Pickl, 2013b:16–20). In other words, is the average “grand sweep” (Larsen-Freeman, 2006:594) representative of any of the individual variables used to compute said group-level effect (see also the ergodicity hypothesis, e.g. Molenaar & Campbell, 2009; Molenaar, 2015; Bülow & Vergeiner, 2021; Wirtz & Pfenninger, 2023)? Such a “variant-based” or “bottom-up” approach facilitates more nuanced insights into how the change of variation intensity may differ by variable (see also Pickl, 2013a), and ultimately allows for a more concentrated view of “variation among the spatialities of linguistic features” (Pickl & Rumpf, 2012:212).

Our second methodologically oriented goal of this contribution is to demonstrate how the change in variation intensity measure can be integrated into spatial-independent analyses, specifically by way of mixed-effect regression modeling. Related to the group-to-individual generalizability issue is the question of *how* individual distributions differ from each other, assuming that those differences are not “noise,” and that variation patterns are distributed in a non-random way across different parts of the dataset, for instance across different linguistic domains such as phonetics, morphology, and lexis. We aim to explore group- and item-level effects while at the same time controlling for locality-related idiosyncrasy (via random effects). With regard to synchronic variation, Pröll (2015:133–149), for instance, found that the spatial distributions of variants were less homogeneous

and more complex in lexis than in phonetics or morphology in a southern German dialect atlas, but also that the separately aggregated spatial patterns of the individual domains evinced differences both from each other and from the joint aggregate. Since synchronic variation and diachronic change are interlinked, the question arises of whether similar differences can be found in diachronic change in variation across different linguistic domains.

At the broadest level, this contribution can be interpreted as a pilot study seeking to make methodological strides in bridging research questions geared towards exploring both social and spatial variation. We demonstrate the versatility of the (change in) variation intensity measure, our goal being to open new ground upon which geolinguistic and variationist lines of inquiry can operationalize and methodologically capture different forms of variation.

2. Methodology and data

2.1. Aim and scope

Based on the research desiderata outlined in Section 1, we pursue the following exploratory research questions:

- (1) Can we identify any spatial patterns in the change of variation intensity in colloquial German (i.e. on the basis of the individual variables) between younger and older informants?
- (2) How can the aggregated change of variation intensity across 15 variables of lexical, morphological (including morphosyntactic), and phonetic nature be characterized in terms of age effects and spatial factors? (For example, do younger or older participants report more variation, and is this effect spatially conditioned?)
- (3) To what extent does the change of variation intensity in colloquial German differ systematically between lexical, morphological (including morphosyntactic), or phonetic variables?

Importantly, we caution against overinterpretation of the following analyses given that we consider only 15 variables (see also Section 2.3 for the rationale behind this choice). As mentioned, from a methodological perspective, this contribution seeks to highlight the versatility of the (change in) variation intensity measure, discussing *where* and in which contexts this measure may be particularly useful, and what (new) questions we can answer by drawing on this measure. In order to maintain comprehensibility, especially regarding the issue of group-to-individual generalizability, we limit this investigation to 15 core variables.

2.2. Atlas of Colloquial German

The data for the present study stem from the *Atlas zur deutschen Alltagssprache* (AdA; Elspaß & Möller, 2003–; see also Möller & Elspaß, 2014, 2015). The AdA is a linguistic atlas of contemporary colloquial German, the data for which are collected in approximately (bi-)annual intervals via online surveys in German-speaking regions of central Europe (i.e. Germany, Austria, Switzerland, Liechtenstein, Luxembourg and the German-speaking parts of northern Italy, eastern Belgium and the Alsace and Lorraine regions in eastern France).

In the AdA questionnaires, informants are asked to identify local variant(s) in lexis, pronunciation, grammar, phraseology, and pragmatics, and in some cases to declare how common/uncommon a certain variant or construction is for the “everyday colloquial

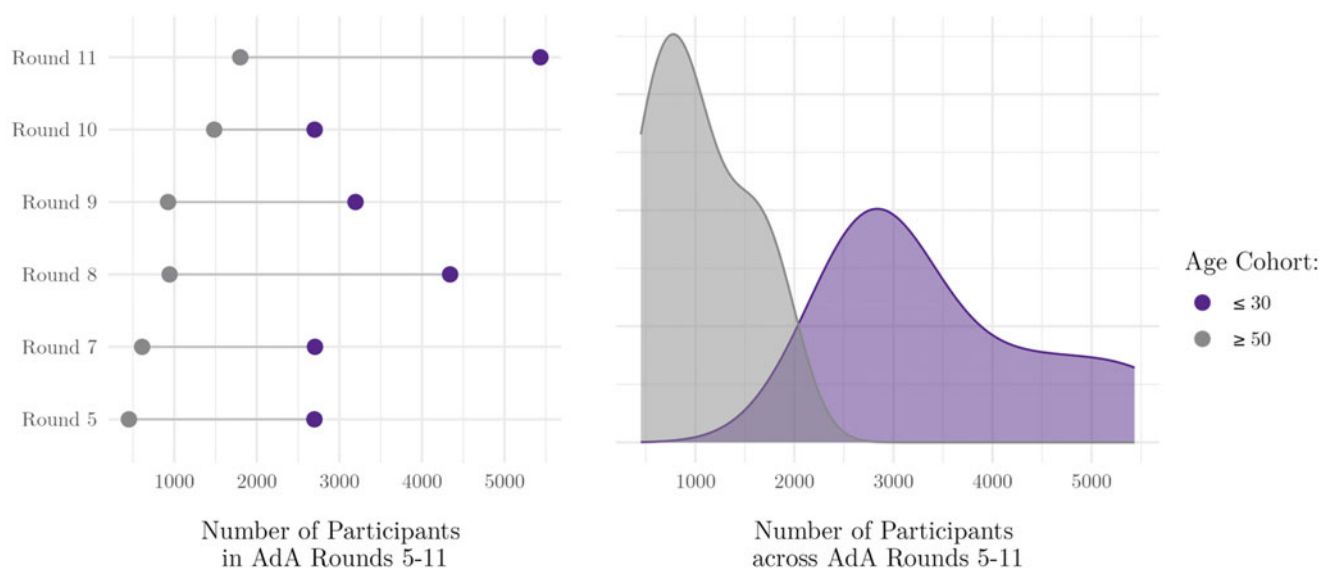


Figure 1. Age cohorts within and across AdA rounds. The Cleveland dot plot (left) illustrates the total number of participants in each round by age (≤ 30 shaded in purple, ≥ 50 shaded in gray). The density plot (right) shows the density distribution of the number of participants from the two age cohorts across AdA rounds 5–11 (excluding round 6) in an aggregated form. Note that the full area below the density line equals 1, i.e. 100% of the participants in the two age cohorts respectively. Exclusively informants from Germany, Austria, German-speaking Switzerland, and German-speaking northern Italy are illustrated in this plot.

speech” in their hometown. Specifically, informants are presented with a brief concept (e.g. a short description and/or picture) and are provided with a list of potential variants from which they are to choose the expression(s) “normally” used for the respective concept in their town/city, or otherwise provide a variant not listed (for examples from the current round, see <https://www.atlas-alltagssprache.de>). Importantly, participants in the AdA survey rounds act as *informants* and are explicitly requested not to indicate their own personal variant preferences. Rather, participants are instructed to name variants used in the colloquial speech of their (home-)towns and cities, specifically, the kind of speech “one would normally hear” in these places, “be it more dialect or more standard German.” In light of the heterogeneity in potential varietal spectra in the German-speaking realm (e.g. diaglossic and diglossic settings), what is used as “everyday colloquial speech” can range from different regional forms of standard German varieties (e.g. in northern Germany and metropolitan areas) to intermediate varieties (e.g. in some areas of Bavarian-speaking Austria) to local dialects (e.g. in German-speaking Switzerland). Supplementary Figure 1 (SF1) provides an example of the spatial distribution of variants of the comparison particle variable (*er ist größer als ich*, *er ist größer wie ich*, *er ist größer als wie ich*, *er ist größer wan ich* ‘he is bigger than me’) as collected in round 9 of AdA.

2.3. (Change in) variation intensity

Variation intensity (VI) is defined as the probability of two randomly chosen cases/answers/instances in a single locality (or, more generally, in any area or domain for which VI is to be calculated) representing differing variants with respect to one linguistic variable, calculated on the basis of relative variant frequencies. Essentially, this represents an intuitive measure of the degree of variation within a location. Variation intensity is calculated as the sum of a variant’s relative frequency multiplied by one minus the same value of overall variants per linguistic variable per locality, resulting in a measure bounded by [0, 1).

$$VI_X = \sum_{x \in X} w_x \cdot (1 - w_x)$$

The variation intensity *VI* of a variable *X* in a given context (e.g. at a particular location) is thus determined by the individual relative frequencies (w = weights) of all variants x belonging to *X* in that context. *VI* is thus not only dependent on the context under investigation (e.g. an individual, a location, or an area), but also on the granularity of the variant categorization. *VI* values for one and the same variable are thus comparable across different contexts (e.g. locations) and variables if the same variant classification standards are employed. It must be pointed out that the data used here are not yet optimal in this respect: for practical reasons, the variant differentiation in the AdA questionnaires was retained, which besides systematicity also aims at recognizability of the variants for the informants and comparability with older surveys.

In cases when categorical independent variables (e.g. social factors such as gender, age group, and so forth) are of empirical interest, these can be employed as additional grouping variables at the locality level (and beyond). In our case, given our intent to investigate apparent-time change in the degree of variation in colloquial German within and across localities, informant age was included as an additional grouping variable at the locality level. Specifically, we computed a variation intensity value for two generations (informants ≤ 30 and ≥ 50) at each locality. To demonstrate apparent-time change, we computed the respective difference in variation intensity between older and younger informants per locality, which we term “change in variation intensity” (CVI). In the CVI measure, negative values indicate more variation among the younger generation and positive values the opposite.

2.4. Informants

During the computation of the CVI measure (see Section 2.3), the data were arranged into subsets of informants ≤ 30 years of age (see Figure 1, purple shading) and ≥ 50 years of age (see Figure 1,

gray shading), i.e. with 20 years between the two cohorts (the approximate minimum amount for one generation). As is evident from the density distribution, our data show a substantial sampling bias towards younger participants. There also seem to be gender discrepancies across the AdA rounds 5–11, though reversed for the two cohorts, such that younger females and older males were sampled more (≤ 30 , 11964 female, 9107 male; ≥ 50 , 2907 female, 3302 male).

Importantly, we note that participants' role as "informants" of their locality's respective language use may influence the results, particularly when subdividing informants into age cohorts for the purposes of apparent-time insights. That said, previous real- and apparent-time analyses (e.g. Leemann et al., 2019; Möller, 2021) indeed provide evidence that in the apparent-time comparison between the reports of the older and the younger AdA informants, tendencies become visible that also show up in real-time comparisons of older and younger survey data (especially change towards more spacious variants). This can be explained by the fact that the respective impression of the local usage is of course strongly dependent on the usage of the most frequent interlocutors, especially in the case of the younger informants with respect to their respective peer group.

2.5. Variables

We investigate 15 lexical, morphological (including morphosyntactic), and phonetic variables¹ (five of each) collected in the AdA survey rounds 5 through 11.² In order to scrutinize the effect of item-specific idiosyncrasies and differences in spatial dispersion on the CVI measure, we selected variables with different geographical distributions (see also Leemann et al., 2019) and with different numbers of total variants, taking care to approximately balance these across linguistic domains (see Table 1; see also Supplementary Table 2 [ST2] for all variants of each variable). It is, however, important to note that the observed patterns at the aggregate level may (in part) be influenced by the non-representative selection of variables for the present dataset. Thus, the aggregate analysis of the change in variation intensity is to be taken first and foremost as a conservative estimate of the spatial and social trends that emerge from the data.

2.6. Localities

Responses to the 15 variables under investigation here were collected from a total of 6,396 locations (i.e. postal codes). As mentioned previously, the data structure of the AdA is not dictated by a clearly defined network of locations,³ but is rather designed to accommodate a growing number of informants and locations with each survey round. Consequently, not every location is represented in each AdA round, and the distribution of response patterns differs per location as well as per variable (informants are not required to answer every question). To ensure that the CVI measure was computed using a sufficient number of responses, the total number of locations was reduced (Supplementary Table 1 [ST1] outlines the number of localities represented across AdA rounds 5 through 11, and Figure 2 indicates the largely low response rates per locality, i.e. < 10 informants, which necessitates some form of geographical aggregation). In this contribution we concentrate on Germany, Austria, Switzerland, and German-speaking northern Italy, as these countries evinced the highest response rates allowing for intergenerational comparisons. As concerns the reduction of the individual locations, we used the first two digits of the participant-reported postal code in Germany, and,

in light of their drastically smaller geographical size, the first digit of the participant-reported postal code in Austria, Switzerland, and northern Italy, as the grouping variables. This resulted in geographically sparser areas (which we refer to as "localities"), but which were condensed enough to capture areal variation (as opposed to focusing on individual locations). Since each AdA round comprises a distinct number of localities once completed, the number of condensed areas for each round can also vary slightly.

3. Results

The following section is divided into three parts corresponding to the three research questions (RQs):

- CVI in lexical, morphological, and phonetic variables (RQ1);
- aggregated CVI across the aforementioned 15 variables (RQ2); and
- the degree to which CVI may systematically differ in function of variable type (RQ3).

3.1. Change of variation intensity in individual variables

To demonstrate step-by-step the computation of the CVI measure, Figure 3 illustrates the variation intensity for both age cohorts with respect to the phonetic form of the indefinite article *ein*, Engl. 'a' (see also Figure 8 in Section 3.1.3 for the CVI map). Here we find that the variation intensity in the northern regions of Germany remains consistently near zero, whereas central and southern Germany (most notably the southwest) and Switzerland are subject to a higher degree of variation. In addition to these regional patterns, we find more pronounced variability in the younger generation (Figure 3, left) than in the older one (Figure 3, right), suggesting intergenerational differences in the degree of within-locality variability.⁴ To encapsulate this information in a single measure, we display in the following maps the difference in variation intensity between older and younger informants per locality (i.e. the CVI measure), where negative values indicate higher variability among the younger generations, and positive values the opposite.

3.1.1. Change of variation intensity in lexical variables

The item-level analysis of the lexical variables indicates a considerable amount of variance in the spatial patterns of the CVI measure, though we did identify several reoccurring themes.

Figure 4 and Supplementary Figure 2 (SF2) ('mashed potatoes' and 'pancake' respectively) display a substantial amount of variance in the CVI measure. The directionality of the CVI measures in these two maps show that both younger and older participants reported variants not reported by the respective other age cohort. For example, as Figure 4 shows, younger informants in Austria reported variants not or only marginally reported by older adults, e.g. a high number of reports of *Kartoffelpüree* in addition to *Erdäpfelpüree* (i.e. the variant dominant in the older informant cohort). SF2 illustrates marked generational differences in northern and southwestern Germany, with the younger informants reporting a wider diversity of variants. For example, younger informants reported higher rates of variants such as *Pfannenkuche(n)* (in southwestern Germany) and *die Plinse/der Plinz* (in eastern Germany) in addition to variants such as *Eierkuchen*, which the older generation predominately reported. We find a similar trend for SF3 ('hicups'), such that the degree of variation in colloquial German appears subject to inter-regional

Table 1. Variables chosen for the apparent-time analysis of colloquial German

Variable	Example variants	# variants	Type	AdA map
'Mashed potatoes'	<i>Erdäpfelpüree, Erdäpfelbrei</i>	11	lexical	https://www.atlas-alltagssprache.de/kartoffelbrei/
'Pancakes'	<i>Palatschinken, Eierkuchen</i>	11	lexical	https://www.atlas-alltagssprache.de/runde-7/f01a/
'Hiccups'	<i>Hickser, Schluckauf</i>	18	lexical	https://www.atlas-alltagssprache.de/r10-f3c/
'To chat'	<i>tratschen, plaudern</i>	11	lexical	https://www.atlas-alltagssprache.de/runde-7/f08b/
'Slippers'	<i>Patschen, Schlapfen</i>	9	lexical	https://www.atlas-alltagssprache.de/runde-7/f09a-b/
Possessive construction	<i>der Schlüssel von Anna, der Anna ihr Schlüssel</i>	7	morphological/ morphosyntactic	https://www.atlas-alltagssprache.de/attribut/
Infinitival constructions	<i>(es fängt...) an zu regnen, zum regnen an</i>	3	morphological/ morphosyntactic	https://www.atlas-alltagssprache.de/runde-7/f13b-d/
<i>sich etwas nicht zu sagen trauen</i> 'not to dare to say sth.' construction	<i>mich sagen getraut, mich zu sagen trauen</i>	10	morphological/ morphosyntactic	https://www.atlas-alltagssprache.de/runde-7/f06b/
1st person singular reflexive pronoun in <i>ich tu mich schwer mit ...</i> 'I find it difficult to ...'	<i>mich, mir</i>	2	morphological	https://www.atlas-alltagssprache.de/r8-f4j-k-2/?child=runde
Plural of <i>Balkon</i> 'balcony'	<i>Balkon, Balkone, Balkons</i>	4	morphological	https://www.atlas-alltagssprache.de/r10-f9c/
Phonetic form of <i>habe ich</i> 'have I'	<i>hab i, han isch</i>	13	phonetic	https://www.atlas-alltagssprache.de/r11-f8a/
Phonetic form of indefinite article <i>ein</i> 'a'	<i>'n, a, e</i>	7	phonetic	https://www.atlas-alltagssprache.de/runde-5/f18a-b/
Phonetic form of <i>heute um vier</i> 'today at four'	<i>heut um viere, heute um vier</i>	7	phonetic	https://www.atlas-alltagssprache.de/runde-7/f08g/
Phonetic form of <i>Milch</i> 'milk'	<i>Milch, Milli</i>	9	phonetic	https://www.atlas-alltagssprache.de/r10-f7d/
Phonetic form of <i>darin</i> 'inside'	<i>drin, drinnen</i>	7	phonetic	https://www.atlas-alltagssprache.de/r10-f6abc/

change, though the directionality does tend to differ: for example, in Austria, the widely reported variant *Schnackerl* is accompanied by young-informant reports of the more inter-regional variant *Schluckauf*.

Additionally, we observed a pronounced change in variation intensity in Figure 5 and SF4 ('to chat' and 'slippers' respectively), particularly in Germany. For example, in central, southern and (north)eastern Germany, younger informants reported a wider diversity of variants for the variable 'to chat', e.g. *klatschen, plauschen, plaudern, schnacken, schwatzen*, in addition to the variants which the older informants predominately reported, such as *quatschen* and *tratschen*.

3.1.2. Change of variation intensity in morphological/morphosyntactic variables

With respect to the morphological variables, SF5 and Figure 6 ('possessive construction' and 'not to dare to say sth.' construction respectively) suggest that younger informants particularly throughout most parts of Germany and, though notably to a lesser extent, Austria report more diverse variants. In east central Germany, it would appear that, in addition to more dialectal morphosyntactic variants such as the possessive dative *der Anna ihr Schlüssel*, standard German variants such as the genitive construction *Annas Schlüssel* and the *von*-construction *der*

Schlüssel von Anna have taken hold among younger individuals (see SF5). Figure 6 illustrates that younger participants in northern and southern Germany reported variants with infinitive 'to' and (dialectally colored) dative reflexive pronoun *mir* (e.g. *mir zu sagen getraut* and *mir getraut zu sagen*) along with standard German variants such as infinitive 'to' and accusative *mich* (e.g. *mich zu sagen trauen*), in addition to variants without infinitive 'to' and accusative *mich* (e.g. *mich sagen getraut* and *mich sagen trauen*) which were widely reported by both age cohorts.

Conversely, SF6, SF7, and Figure 7 (infinitive 'to', 1st person singular reflexive pronoun, and the plural form of 'balcony') illustrate that, with respect to some variables, the older cohort demonstrates a higher degree of variation, notably in northern and central Germany and central Austria (SF6), southern Germany (SF7), and throughout the German-speaking areas (Figure 7). Interestingly, when the older cohort demonstrated a higher degree of variability, this was found for variables for which fewer variants were reported overall.

3.1.3. Change of variation intensity in phonetic variables

We found that younger informants reported more diverse variants of certain phonetic variables. For example, the younger cohort reported inter-regional variants such as *drinnen* (SF8), *'n* (Figure 8), *heute um vier* and *heute um viere* (Figure 9), but also

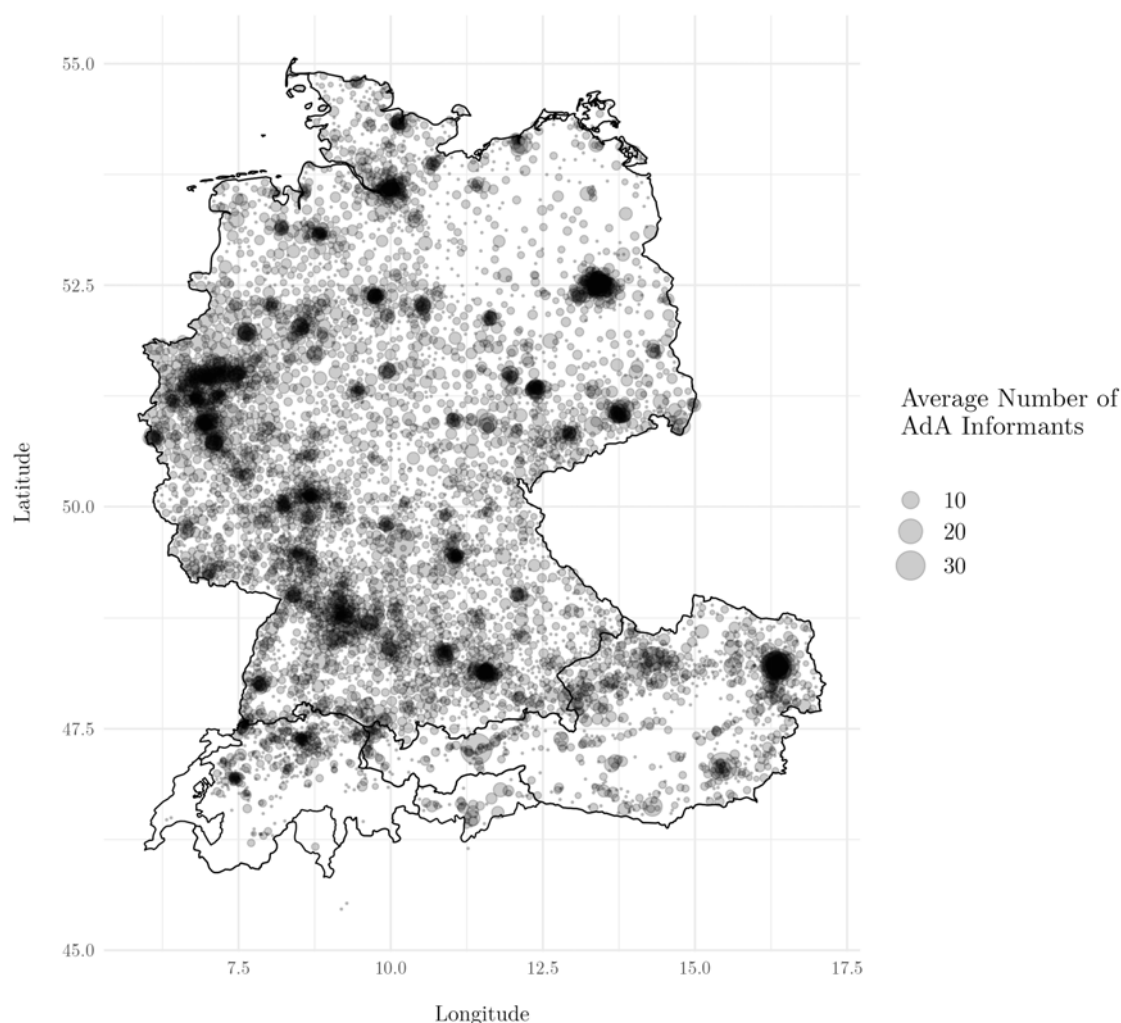


Figure 2. Number of AdA informants per individual location across AdA rounds 5–11, aggregated across all age groups. Each respective bubble indicates an unaggregated location in the AdA rounds 5–11 (excluding round 6, as this round comprised linguistic similarity judgements), and the size visualizes the number of informants per location per round.

dialectal variants such as *ho i* (SF9) and *heut um vier* (Figure 9). For example, in southern Germany and Austria, these variants were indicated in addition to the variants consistently reported among both the older and younger cohorts (i.e. *drin* [SF8], *hob i* [SF9], *a* [Figure 8], *heut um viere* [Figure 9]).

We also identified areas in which we found no or only very minor generational differences in the degree of variation reported, for instance in central and northern Germany, as illustrated in Figures 8, and to a lesser extent, SF8 and SF10.

3.2. Aggregated change of variation intensity across lexical, morphological, and phonetic variables

Figure 10 presents the mean CVI value at each locality, computed across the 15 variables under investigation.

We found inter-generational differences in variation intensity in most parts of Germany (except the northwest), in South Tyrol, and in parts of western Austria, with the younger cohort demonstrating a higher degree of variation. Conversely, in northwestern Germany, Switzerland, and eastern Austria we observed little to no apparent-time change. What does stand out, however, is the directionality of apparent-time change across regions: While the CVI value is certainly more pronounced in

some regions than in others, the overall trend is that younger informants report more overall variants than do informants from the older age cohort.

3.3. Change of variation intensity between lexical, morphological/morphosyntactic, and phonetic variables

In order to determine whether the CVI measure differed systematically between lexical, morphological, and phonetic variables—regardless of any spatial dependencies—we computed a linear mixed-effects model with CVI as the outcome variable and variable type (three levels: *Lexical*, *Morphological*, *Phonetic*; reference level: *Lexical*) as the sole predictor. To account for spatial- and item-related idiosyncrasies, we included random intercepts for the country postal code (e.g. for Salzburg, “5020_AT,” given that some postal codes exist in multiple countries) and the 15 items.

ST2 indicates that, given the data and data-generating model, the CVI in the morphological and phonetic variables did not systematically differ from the lexical variables. Figure 11 visualizes the conditional effects of the three variable types, illustrating (a) that the CVI measure did not differ based on the variable type, and (b) that the younger participants (i.e. < 30)

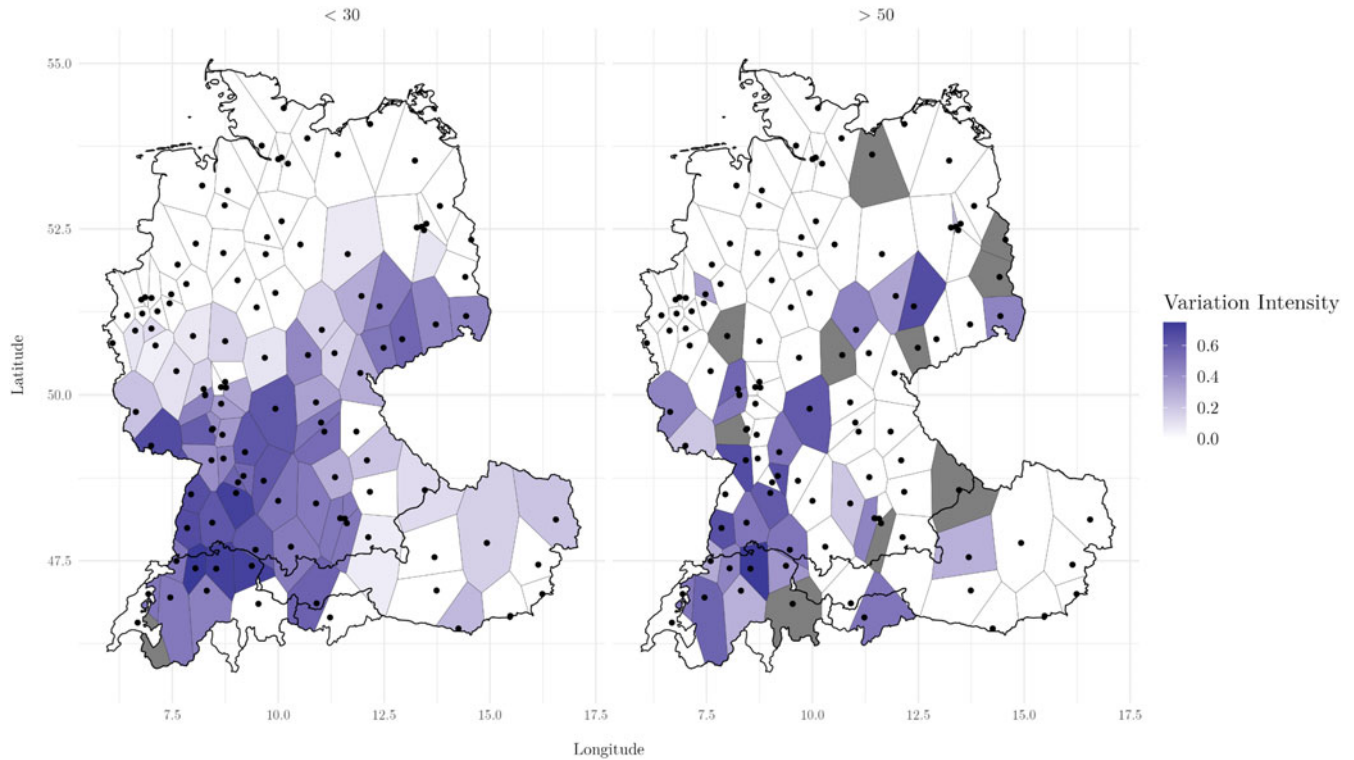


Figure 3. Variation intensity in maps for the phonetic form of indefinite article *ein*, Engl. 'a'. The two panels illustrate the computed variation intensity measure for each aggregate locality, with darker shading visualizing higher values in the respective variation intensity. Gray Voronoi cells indicate lack of responses from the respective cohort.

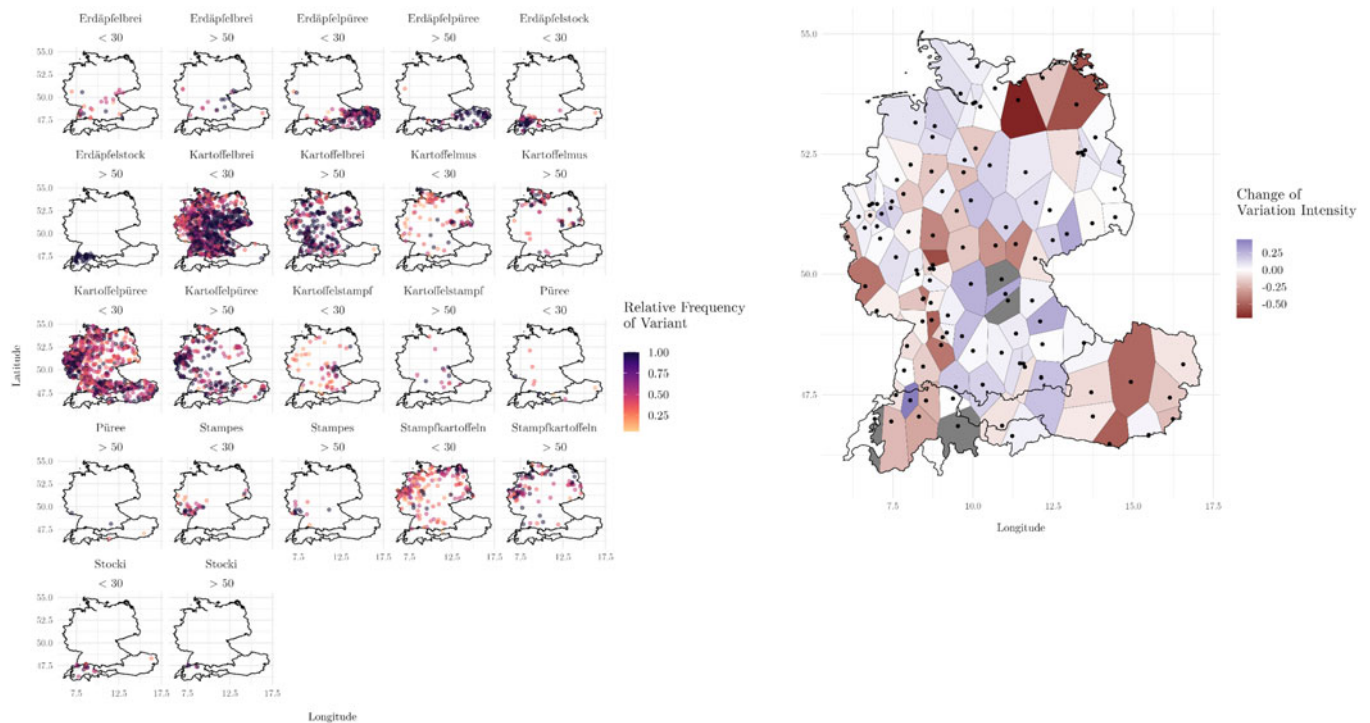


Figure 4. Variant and change of variation intensity maps for 'mashed potatoes'. The left-hand panel shows the raw data of the relative frequency of each respective variant at the unaggregated localities at which they were reported. This allows insights into which variants may be at the forefront of change. The right-hand panel indicates the CVI, with red shading (negative values) indicating more variance in the young cohort as compared to the older adults, and blue shading (positive values) vice versa. Gray Voronoi cells indicate that no comparison between the younger and older generations was possible (e.g. lack of responses from both cohorts in the respective aggregated locality).

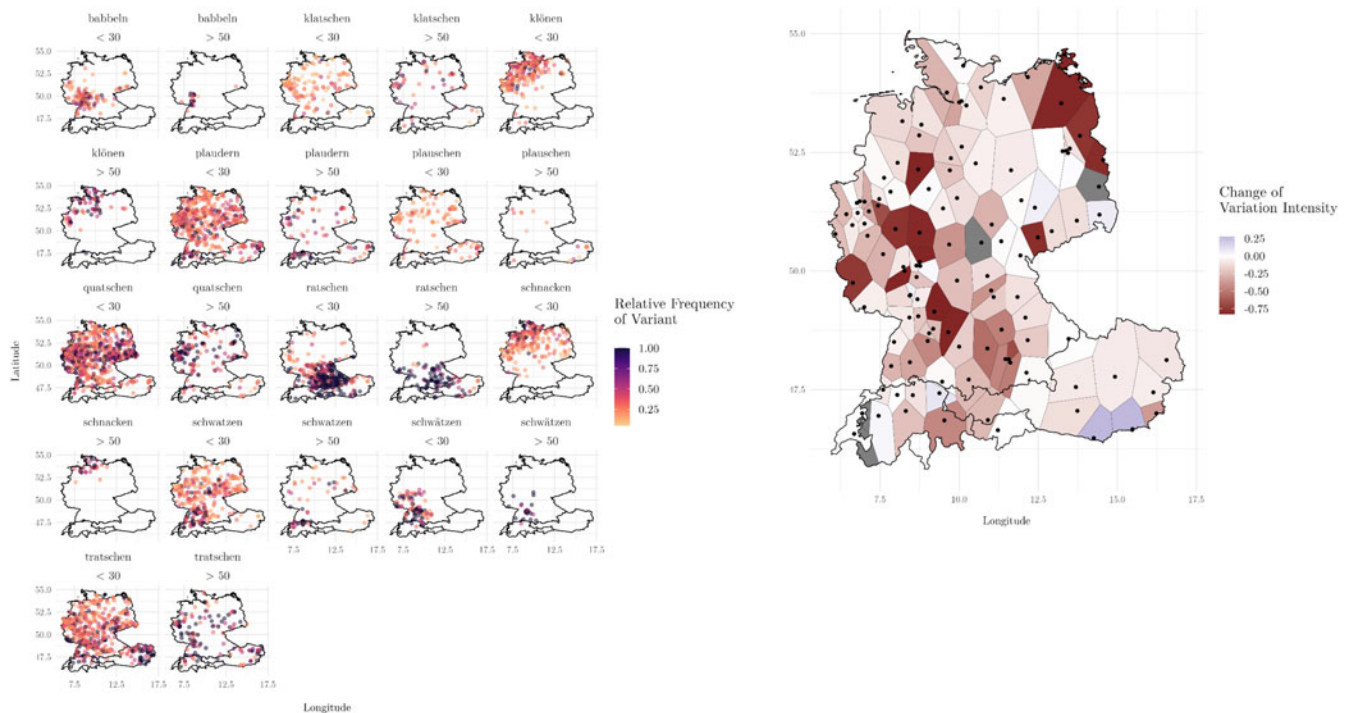


Figure 5. Variant and change of variation intensity maps for ‘to chat’.

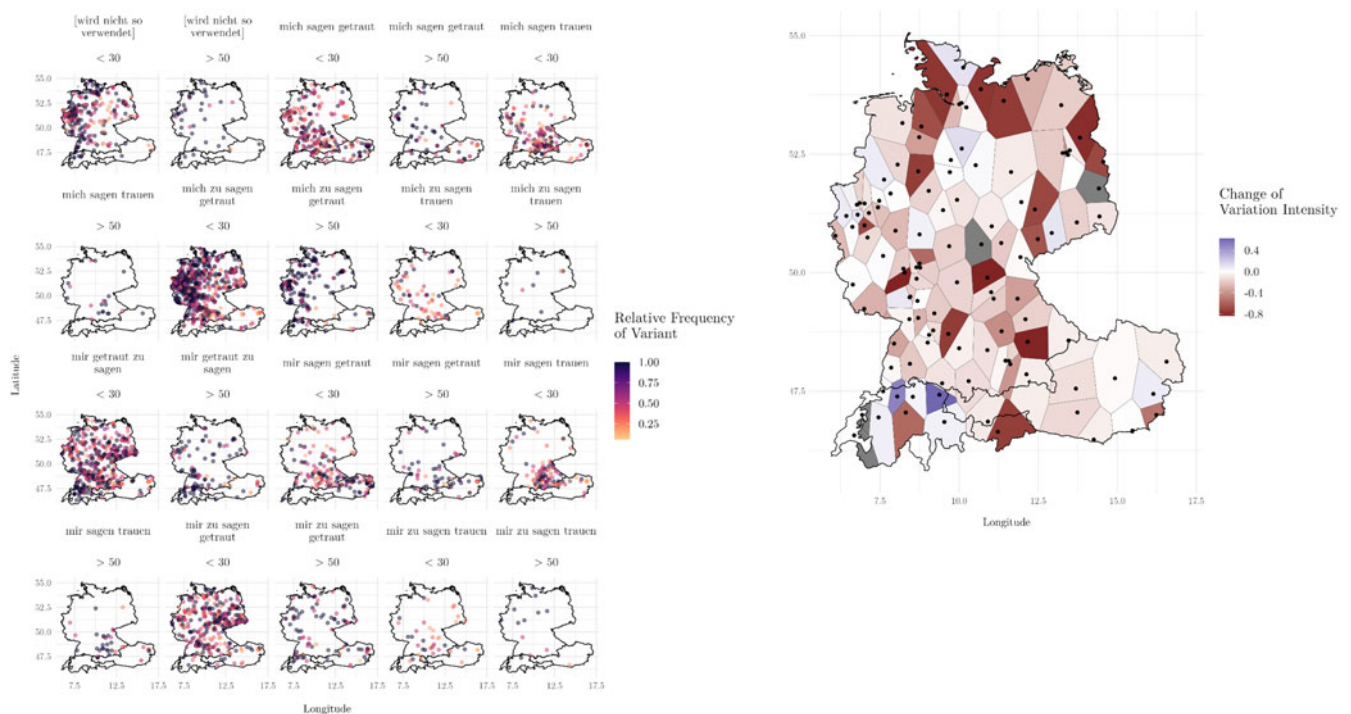


Figure 6. Variant and change of variation intensity maps for the ‘not to dare to say sth.’ construction.

reported a wider range of variants than did the older informants. This complements the findings from the previous section: not only is the degree of variation conditioned spatially, but we can also identify *spatial-independent* apparent-time effects, where younger informants—regardless of region—tend to report a more diverse set of variants.

Importantly, we do note marked item-level variance in these data (controlled for via the random effects in the regression analysis). Figure 12 illustrates the spatial-independent item-level variance in the form of violin plots. Overall, the item effects indicate what our previous analyses have shown, i.e. a generally higher degree of variation among the younger cohort. That said,

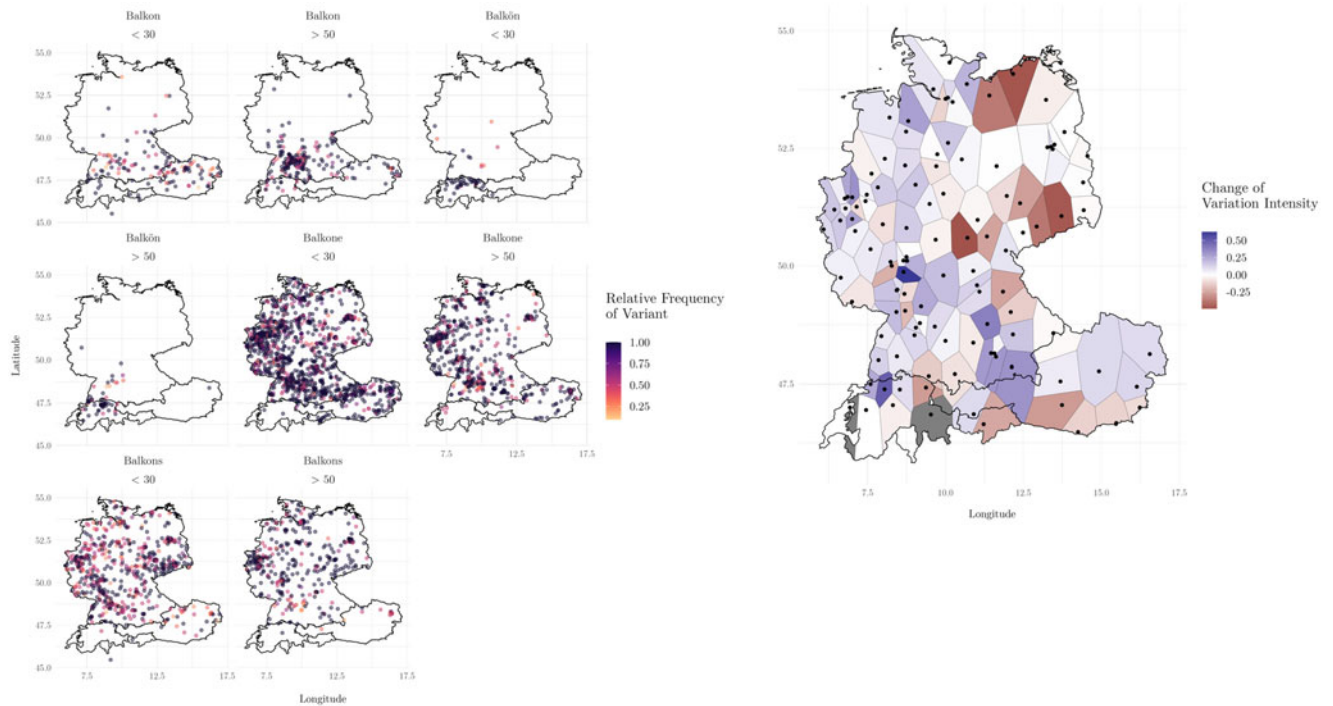


Figure 7. Variant and change of variation intensity maps for the plural form of 'balcony'.

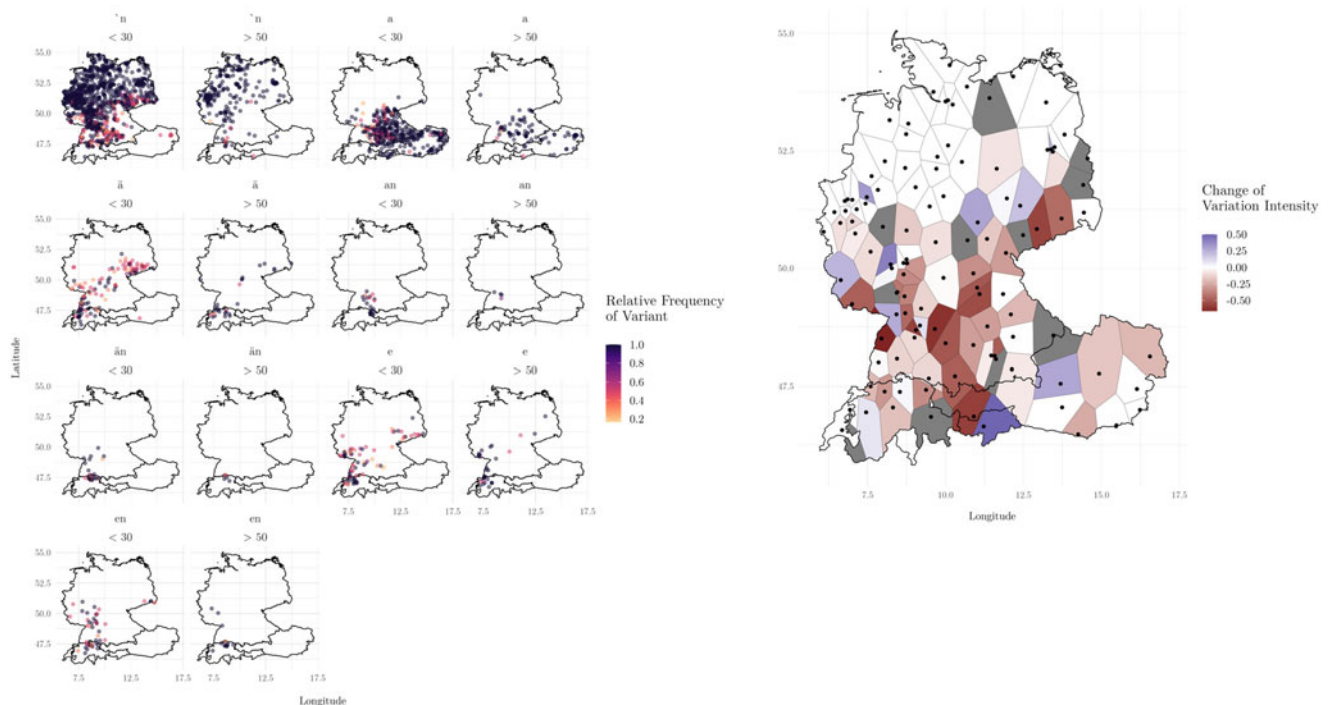


Figure 8. Variant and change of variation intensity maps for the phonetic form of indefinite article *ein*, Engl. 'a'.

there were only a few select variables for which all three quartiles fell below zero (the variables 'slippers', 'to chat', and 'not to dare to say sth.', and the upper quartiles of the morphological variable 1st person singular reflexive pronoun and the phonetic variable indefinite article *ein*, Engl. 'a', hovered at the zero mark). This indicates that only for these select variables did the

younger informants inter-individually, and spatial-independently, report more variants of a variable than the older generation. In other words, while our previous results highlight a *tendency* for younger participants to report a more diverse set of variants, there are indeed individual cases of disparity at the level of the variables.

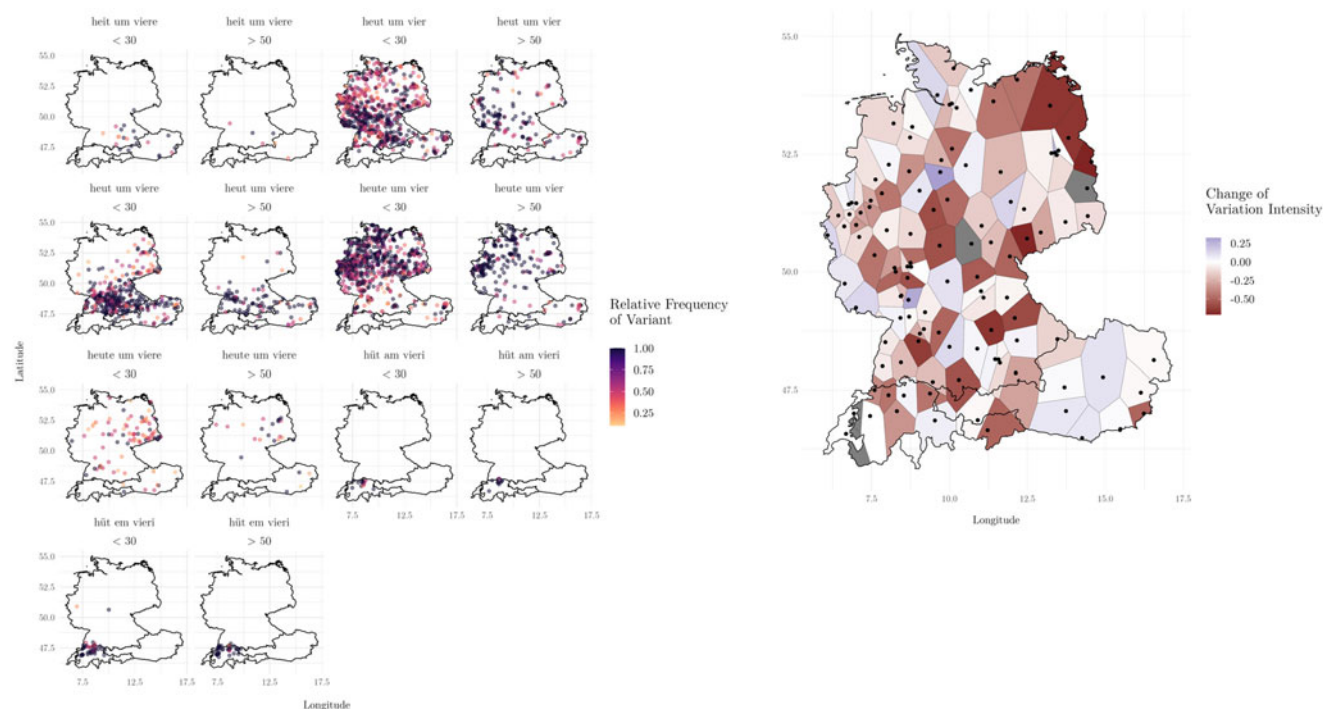


Figure 9. Variant and change of variation intensity maps for the phonetic form of *heute um vier* 'today at four'.

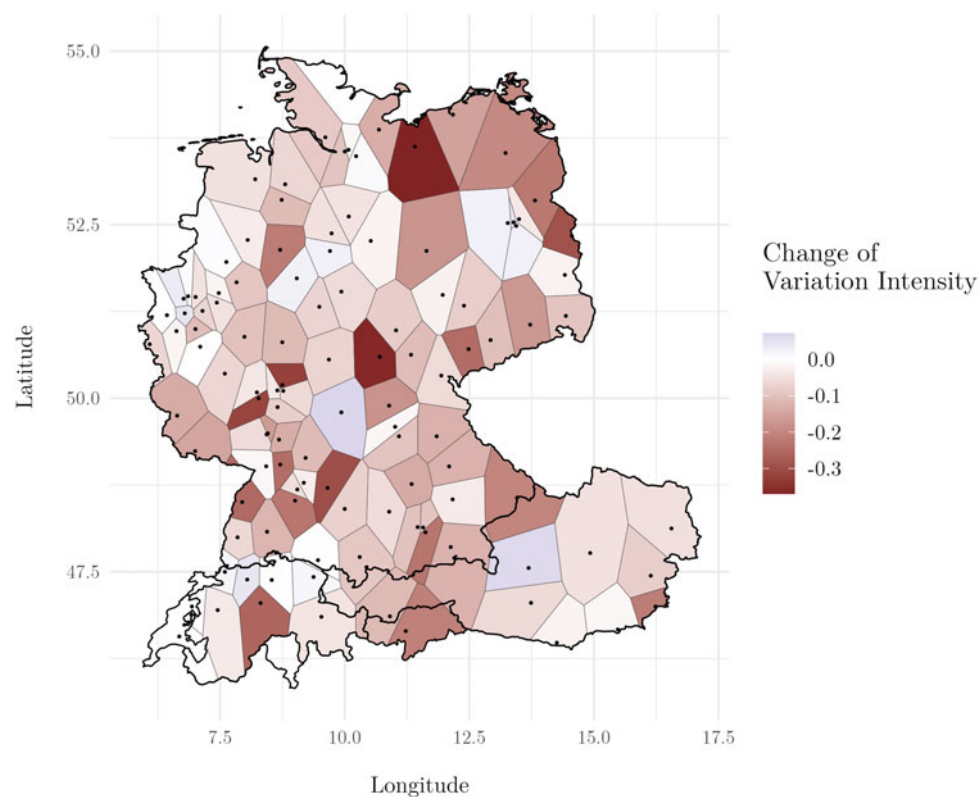


Figure 10. Mean change of variation intensity aggregated across 15 lexical, morphological, and phonetic variables.

4. Discussion

The goal of the present contribution was to introduce a novel measure, the change in variation intensity (CVI), which intuitively captures and quantifies generational differences in the *degree* of variation within a given locality (taking into account the number of

variants used and their relative frequencies). We found that in the northwestern and southern parts of Germany, as well as in Austria and Switzerland, the younger cohort did not demonstrate much more variability in terms of the diversity of variants reported than the older cohort. That said, there were indeed individual cases of

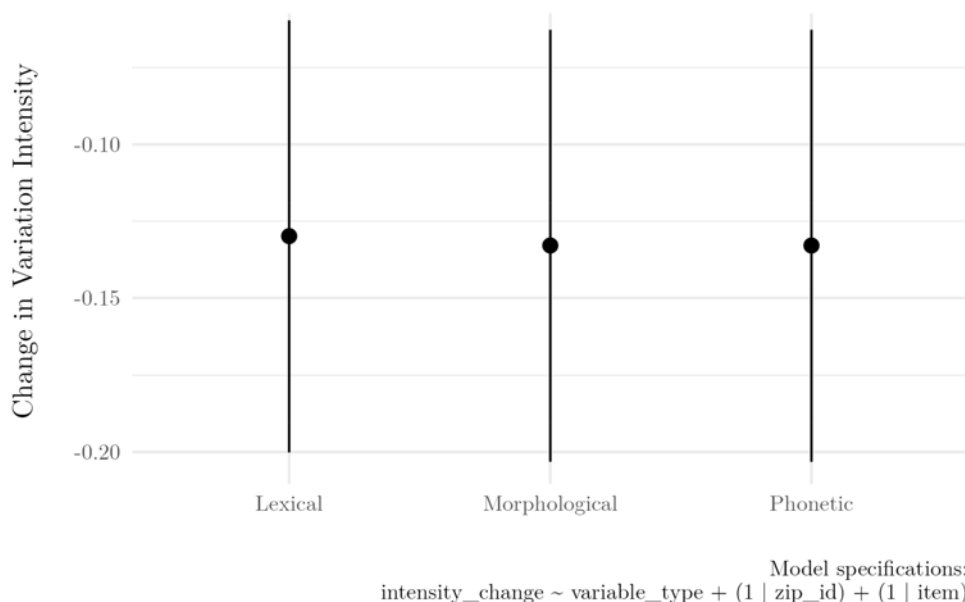


Figure 11. Linear mixed-effects regression model for the investigated change of variation intensity as a function of variable type (i.e. lexical, morphological/morphosyntactic, phonetic variables).

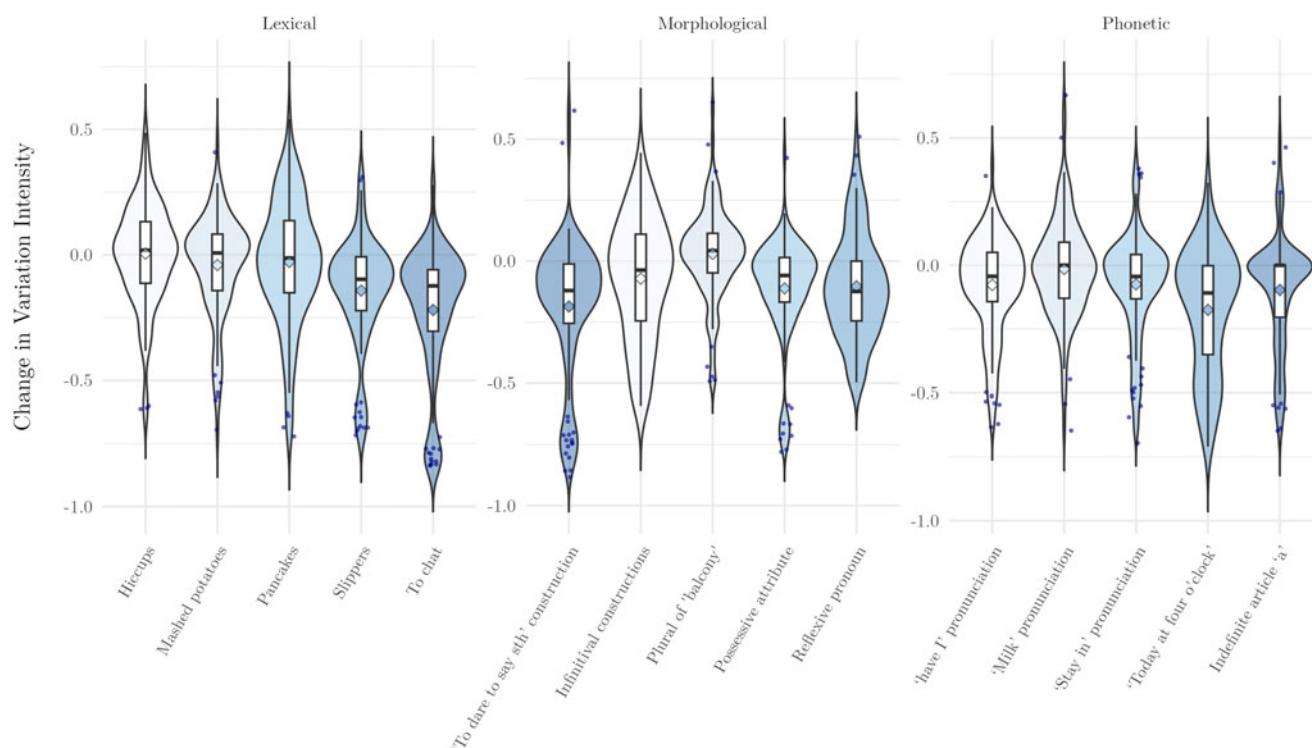


Figure 12. Item-level idiosyncrasies in the spatial-independent change of variation intensity measures. Violin plots show the probability density of the data at different values. Box plots indicate the median and the respective quartiles. The rhombus at the center illustrates the mean and the dots indicate outliers, which are defined as values above 1.5 interquartile range from the median.

disparity in the variable-level analyses indicating patterns inconsistent with the aforementioned observations. For instance, we identified inter-generational stability in terms of variation intensity in both the center and north as well as in the south for the variable ‘not to dare to say sth.’ (Figure 6), and for the variable ‘plural form of *Balkon*’ (‘balcony’) (Figure 7) we even found that older adults were slightly more variable than the younger adults, a trend which stands in contrast to the results of the aggregate analysis. We take this as a call to action for geolinguistic

approaches not to strictly disregard more nuanced item-level analyses in favor of the aggregate. As Pickl (2013a:67) put it, exclusively aggregate-based investigations render it “impossible to make any statements about the relations between the distributions of single variables or their alignment with extra-linguistic constraints, as the aggregation reduces the many dimensions of diatopic variation to only one that accounts for the overall dialectal variation but not for the inter-feature variability within the collection of variables.”

Complexivists (see inter alia Larsen-Freeman & Cameron, 2008; Lowie & Verspoor, 2019) also maintain that individual/item-level variation, and by extension (inter-)feature variability, can indeed provide meaningful information as concerns issues of group-to-individual generalizability (see e.g. Bülow et al., 2018; Lowie & Verspoor, 2019; Wirtz & Pfenninger, 2023). In addition to exploring overall, group-level spatial distributions on the basis of mean trends, it is equally necessary to inquire about the extent to which the aggregate actually represents the individual variation in the respective items. Importantly, item-level and aggregate-level approaches need not be mutually exclusive, but should rather be seen as complementary: one examines the (potentially functional) meaningfulness of idiosyncrasy and the extent to which the variable relates to the aggregate, the other whether aggregates can generalize over item-related noise in order to project the overarching tendencies (see also Pickl & Rumpf, 2012). In reconciling these two approaches, we can address issues relating to the degree to which findings at the variable and aggregate level are *mutually inferable*, which, in turn, may allow for more nuanced insights into *which* social factors generalize and, importantly, *how* they generalize—i.e. perhaps only for certain subgroups of individuals, only for certain regions, etc. (see also Peng et al., 2022).

We offer these current findings not as generalizations for colloquial German and for the influence of social factors such as age alone, but rather as conceptual and methodological approaches to *reconciling* the spatial and social on the basis of dialectometric and variationist sociolinguistic methods. As mentioned earlier, this contribution, given its methodological density, did not set out to perform a large-scale aggregate analysis or generalize over potentially interesting item-level idiosyncrasies. Rather, our main intent was to broadly illustrate the versatility and usefulness of measures gauging the *degree* of variation present within a locality, and whether this variation is subject to change in apparent time. Employing such variation-based measures, indeed, brings with it new questions for future variationist sociolinguistic and geolinguistic agendas. For example, expanding further on the CVI measure than what we have empirically demonstrated in this article, future applications may include exploring the degree to which formal and/or functional linguistic characteristics of variables have an impact on their variability in the spirit of Pickl (2013a) and Franco et al. (2019). It is to be carefully noted, however, that analyses using the variation intensity metric inherently answer different research questions than more traditional geolinguistic approaches. Specifically, variation intensity can be used to model the (spatial) distribution of *variation* rather than *variants*, thus allowing us to address patterns of language variation and change from another angle.

Additionally, our current analysis drew on variation intensity at the locality level, focusing on generational differences therein. While this facilitated a sufficient degree of variance to conduct spatial and regression analyses, this approach necessarily neglects the *individual* informant and the potential idiosyncrasies inherent to individuality. Thus, future work using measures of VI may consider analyzing variation intensity across a wide breadth of linguistic variables for each individual informant. In so doing, one could conduct between-person correlational analyses of individual-level variation intensity (e.g. the diversity of variants in individual reports or produced in speech) and individual differences in, for example, gender, age, occupation, length of local residence, degree of multilingualism, or personality factors (e.g. Steiner, Jeszenszky, Stebler & Leemann, 2023).

While future research must decide whether our findings are spurious in light of the few variables considered here, the present results do lend weight to the notion that the *younger* generation of informants (both at the aggregate level and in many of the individual items as well) reported more variants and thus evinced a higher degree of variation than did their older counterparts. This suggests that young informants have a broad range of knowledge concerning the linguistic variation in their locality, arguably more so than do informants from the older age cohort; however, productive data may not necessarily capture or reflect this notion (see e.g. Kaiser & Bülow, 2022, for data on processes of language change in children's production data). This can be taken as provisional evidence to hypothesize that we may be observing impending processes of language change in the younger generations, and more specifically that younger generations indeed represent a prime source of information about linguistic change and the role of language in social and societal practice (e.g. Eckert, 1997; Kerswill, 1996). More specifically, it is possible that the inter-generational results are a first indication of what may lead to variation loss in the long run, if the additional variants in younger speakers' responses are supraregional or even standard variants that, in a first stage, represent additions to originally regionally constituted repertoires, but may eventually supplant their more traditional, regionally confined alternative forms.

Finally, in order to provide further accounts of the versatility of measures of variation intensity in terms of spatial-independent analyses while at the same time controlling for locality and thus spatial idiosyncrasy, we addressed by way of mixed-effects regression modeling the issue of whether generational differences in variation intensity systematically varied between the linguistic domains of phonetics, morphology, and lexis. We did not find any credible differences between the three domains, as Pröll (2015) did in his analyses of synchronic variation; however, this may also be an artifact of the small subset of selected variables. In order to determine whether these results are spurious or not, future research needs to draw on a larger set of variables.

5. Conclusion

Across the social and behavioral sciences, it has been suggested that aggregative approaches are advantageous in order to neutralize or generalize over item-related idiosyncrasies (Nerbonne, 2010:3822; see also Nerbonne, 2009), the key interest being overall spatial tendencies of language variation. Moreover, the notion that “social variation would distract from the spatial picture and obscure the spatial signal” (Pröll et al., 2021:244) has been pervasive in the dialectometric literature for decades, and to date only a very few studies have attempted to juxtapose the relative effects of, and the interaction between, social *and* spatial factors on linguistic variation (but see e.g. Wieling et al., 2014; Gilles, 2023; Vergeiner et al., 2025; Wirtz et al., 2025). This remains a research desideratum in need of more careful attention, especially in light of the recent uptake in crowdsourcing data collection methods that necessarily generate socially very diverse datasets, such as those in the AdA.

What is more, these socially diverse sets of informants are likely to generate a higher degree of linguistic variation both within and across localities as opposed to more traditional atlases focusing on the language reports of a few select NORM/NORF speakers, and this begs the question: How can we measure the *degree* of variation within a locality? And is the degree of within-locality variation subject to change, at least in apparent time? To address these issues,

we proposed in this contribution a variation intensity measure which intuitively operationalizes the degree of variable-specific variation in a given context. Such variation-based measures open the doors to many new questions, for example where is variation most pronounced, and which social factors contribute to increased variation? Providing answers to such questions will undoubtedly complement the more traditional issues concerning which variants are reported where.

Supplementary material. To view supplementary material for this article, please visit <https://doi.org/10.1017/jlg.2025.10001>

Acknowledgments. This project was funded by Land Salzburg: Kultur und Wissenschaft (State of Salzburg, Austria: Culture and Science Department, reference number 20204-WISS/262/9-2021), which is hereby gratefully acknowledged. Any remaining errors are our own.

Notes

1 We focus on lexical, morphological, and phonetic variables since these are the most commonly collected variables in AdA. AdA also includes variables on phraseology (e.g. greetings, leave-taking formulas, idioms), syntax (e.g. verb order in subordinate clauses), and, to a degree, pragmatics (e.g. socio-situational use of T- vs. V-forms of address). To facilitate interpretability of the results and contain the scope of this contribution, however, domains other than the three aforementioned were not considered for analysis.

2 Given that the AdA allows not only for multiple-choice responses but also for individualized responses, exclusively variants with ≥ 40 occurrences were considered in order to ensure the legitimacy of the variants reported.

3 However, for the AdA maps published on the internet, the data are aggregated according to a fixed network of locality points.

4 In order to determine whether certain variables underlie spatial clustering effects, we initially computed Moran's I (see ST3), which calculates the degree of spatial autocorrelation of the CVI across the entirety of the research area. That said, if there is only spatial clustering in small areas or regions, the Moran's I effect may be consequently smaller, which necessitates visualizations of the CVI measure for each variable.

References

- Bülow, Lars & Philip C. Vergeiner. 2021. Intra-individual variation across the lifespan: Results from an Austrian panel study. *Linguistics Vanguard* 7(s2). 1–11.
- Bülow, Lars, Kees de Bot & Nanna Hilton. 2018. Zum Nutzen der Complex Dynamic Systems Theory (CDST) für die Erforschung von Sprachvariation und Sprachwandel. In Helen Christen, Peter Gilles & Christoph Purschke (eds.), *Räume, Grenzen, Übergänge: Akten des 5. Kongresses der Internationalen Gesellschaft für Dialektologie des Deutschen (IGDD)*, 45–69. Stuttgart: Franz Steiner.
- Eckert, Penelope. 1997. Age as a sociolinguistic variable. In Florian Coulmas (ed.), *The handbook of sociolinguistics*, 151–167. Oxford: Blackwell.
- Elspaß, Stephan & Robert Möller. 2003–. Atlas zur deutschen Alltagssprache (AdA). <https://www.atlas-alltagssprache.de> (July 1, 2023).
- Franc, Karlien, Dirk Geeraerts, Dirk Speelman & Roeland Van Hout. 2019. Concept characteristics and variation in lexical diversity in two Dutch dialect areas. *Cognitive Linguistics* 30(1). 205–242.
- Gilles, Peter. 2023. Regional variation, internal change and language contact in Luxembourgish: Results from an app-based language survey. *Taal en Tongval* 75(1). 29–57.
- Gooskens, Charlotte. 2005. Travel time as a predictor of linguistic distance. *Dialectologia et Geolinguistica* 13. 38–62.
- Heeringa, Wilbert & John Nerbonne. 2001. Dialect areas and dialect continua. *Language Variation and Change* 13(3). 375–400.
- Kaiser, Irmtraud & Lars Bülow. 2022. Der Dialekt der jüngsten Generation: Wandel, Remanenz und Abbau dialektaler Merkmale bei Kindern im westmittelbairischen Raum Salzburgs. In Philip C. Vergeiner, Stephan Elspaß & Dominik Wallner (eds.), *Struktur von Variation zwischen Individuum und Gesellschaft: Akten der 14. Bayerisch-Österreichischen Dialektologietagung*, 437–467. Stuttgart: Franz Steiner.
- Kerswill, Paul. 1996. Children, adolescents and language change. *Language Variation and Change* 8(2). 177–202.
- Kolly, Marie-José & Adrian Leemann. 2015. Dialäkt Äpp: Communicating dialectology to the public—crowdsourcing dialects from the public. In Adrian Leemann, Marie-José Kolly, Stephan Schmid & Volker Dellwo (eds.), *Trends in phonetics and phonology*, 271–285. Bern: Peter Lang.
- Larsen-Freeman, Diane. 2006. The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics* 27(4). 590–619.
- Larsen-Freeman, Diane & Lynne Cameron. 2008. *Complex systems and applied linguistics*. Oxford: Oxford University Press.
- Leemann, Adrian, Curdin Derungs & Stephan Elspaß. 2019. Analyzing linguistic variation and change using gamification web apps: The case of German-speaking Europe. *PLoS ONE* 14(12). 1–29.
- Lowie, Wander M. & Marjolijn H. Verspoor. 2019. Individual differences and the ergodicity problem. *Language Learning* 69(S1). 184–206.
- Molenaar, Peter C. M. 2015. On the relation between person-oriented and subject-specific approaches. *Journal for Person-Oriented Research* 1(1–2). 34–41.
- Molenaar, Peter C. M. & Cynthia G. Campbell. 2009. The new person-specific paradigm in psychology. *Current Directions in Psychological Science* 18(2). 112–117.
- Möller, Robert. 2021. Alltagssprachlicher Gebrauch heute und morgen: Die Daten einiger Karten des “Atlas zur deutschen Alltagssprache” im Apparent-Time-Vergleich. *Alltag im Rheinland* 2021. 120–126.
- Möller, Robert & Stephan Elspaß. 2014. Zur Erhebung und kartographischen Darstellung von Daten zur deutschen Alltagssprache online: Möglichkeiten und Grenzen. In Fabio Tosques (ed.), *20 Jahre digitale Sprachgeographie*, 121–131. Berlin: Humboldt-Universität, Institut für Romanistik.
- Möller, Robert & Stephan Elspaß. 2015. Atlas zur deutschen Alltagssprache. In Roland Kehrein, Alfred Lameli & Stefan Rabanus (eds.), *Regionale Variation des Deutschen: Projekte und Perspektiven*, 519–540. Berlin/Boston: De Gruyter.
- Montemagni, Simonetta & Martijn Wieling. 2017. Exploring the role of extra-linguistic factors in defining dialectal variation patterns through cluster comparison. In Martijn Wieling, Martin Kroon, Gertjan van Noord & Gosse Bouma (eds.), *From semantics to dialectometry: Festschrift in honor of John Nerbonne (Tributes 32)*, 241–251. College Publications.
- Nerbonne, John. 2009. Data-driven dialectology. *Language and Linguistics Compass* 3(1). 175–198.
- Nerbonne, John. 2010. Measuring the diffusion of linguistic change. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365(1559). 3821–3828.
- Peng, Hongying, Wander Lowie & Sake Jager. 2022. Unravelling the idiosyncrasy and commonality in L2 developmental processes: A time-series clustering methodology. *Applied Linguistics* 43(5). 891–911.
- Pickl, Simon. 2013a. Lexical meaning and spatial distribution: Evidence from geostatistical dialectometry. *Literary and Linguistic Computing* 28(1). 63–81.
- Pickl, Simon. 2013b. *Probabilistische Geolinguistik: Geostatistische Analysen lexikalischer Variation in Bayerisch-Schwaben* (Zeitschrift für Dialektologie und Linguistik, Beihefte 154). Stuttgart: Franz Steiner.
- Pickl, Simon & Jonas Rumpf. 2012. Dialectometric concepts of space: Towards a variant-based dialectometry. In Sandra Hansen, Christian Schwarz, Philip Stoeckle & Tobias Streck (eds.), *Dialectological and folk dialectological concepts of space*, 199–214. 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 Fischer & Kristina Herbert (eds.), *Dimensions of Linguistic Space: Variation—Multilingualism—Conceptualisations/Dimensionen des sprachlichen Raums: Variation—Mehrsprachigkeit—Konzeptualisierung* (Schriften zur deutschen Sprache in Österreich 45), 39–60. Berlin: Peter Lang.
- Pröll, Simon. 2015. *Raumvariation zwischen Muster und Zufall: Geostatistische Analysen am Beispiel des Sprachatlas von Bayerisch-Schwaben* (Zeitschrift für Dialektologie und Linguistik, Beihefte 160). Stuttgart: Franz Steiner.

- 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 in variationist sociolinguistics* (Studies in Language Variation 27), 227–252. Amsterdam & Philadelphia: John Benjamins.
- Riverin-Coutlée, Josiane & Jonathan Harrington. 2022. Phonetic change over the career: A case study. *Linguistics Vanguard* 8(1). 41–52.
- Séguy, Jean. 1971. La relation entre la distance spatiale et la distance lexicale. *Revue de Linguistique Romane* 35(139–140). 335–357.
- Stanford, James N. 2012. One size fits all? Dialectometry in a small clan-based indigenous society. *Language Variation and Change* 24(2). 247–278.
- Steiner, Carina, Péter Jeszenszky & Adrian Leemann. 2023. Variation and change in Swiss German agreement morphology: Spatial, social, and attitudinal effects. *Journal of Linguistic Geography* 11(1). 8–24.
- Steiner, Carina, Péter Jeszenszky, Viviane Stebler & Adrian Leemann. 2023. Extraverted innovators and conscientious laggards? Investigating effects of personality traits on language change. *Language Variation and Change* 35(1). 1–28.
- Stöckle, Philipp. 2016. Horizontal and vertical variation in Swiss German morphosyntax. In Marie-Hélène Côté, Remco Knooihuizen & John Nerbonne (eds.), *The future of dialects*, 195–215. Berlin: Language Science Press.
- Tamminga, Meredith, Laurel MacKenzie & David Embick. 2016. The dynamics of variation in individuals. *Linguistic Variation* 16(2). 300–336.
- Vergeiner, Philip C. & Stephan Elspaß. 2025. Infinitival constructions in the German dialects of Austria: On variation, constraints, and change of a prominent syntactic feature. *Journal of Germanic Linguistics* 37(1). 1–34.
- Vergeiner, Philip C., Lars Bülow & Stephan Elspaß. 2025. The social versus the regional: A multivariate analysis of (morpho-)syntactic variation in Austria's rural dialects. In James M. Stratton & Karen V. Beaman (eds.), *Expanding variationist sociolinguistic research in varieties of German*, 35–60. New York: Routledge.
- Werth, Alexander, Lars Bülow, Simone E. Pfenninger & Markus Schiegg (eds.). 2021. *Intra-individual variation in language* (Trends in Linguistics Studies and Monographs 363). Berlin & Boston: De Gruyter.
- Wieling, Martijn, Simonetta Montemagni, John Nerbonne & R. Harald Baayen. 2014. Lexical differences between Tuscan dialects and standard Italian: Accounting for geographic and sociodemographic variation using generalized additive mixed modeling. *Language* 90(3). 669–692.
- Wirtz, Mason A. 2022. Discourse completion tasks meet virtual reality: A pilot study on virtual reality as an elicitation instrument. *Research Methods in Applied Linguistics* 1(3). 100029.
- Wirtz, Mason A. & Simone E. Pfenninger. 2023. Variability and individual differences in L2 sociolinguistic evaluations: The GROUP, the INDIVIDUAL and the HOMOGENEOUS ENSEMBLE. *Studies in Second Language Acquisition* 45(5). 1186–1209.
- Wirtz, Mason A., Simon Pickl, Konstantin Niehaus, Stephan Elspaß & Robert Möller. 2025. Gebrauchsstandard in der deutschen Alltagssprache: Eine integrative Modellierung räumlicher und sozialer Variation. *Zeitschrift für Germanistische Linguistik* 53(1). 97–125.