Language Variation and Change (2024), **36**, 307–331
doi:10.1017/S0954394524000188

CAMBRIDGE UNIVERSITY PRESS

Spatial, occupational, and age-related effects on reported variation in colloquial German

Mason A. Wirtz[1], Simon Pickl[1], Konstantin Niehaus[1], Stephan Elspaß[1] and Robert Möller[2]

[1]Department of German Language and Literatures, University of Salzburg, Salzburg, Salzburg, Austria and [2]Department of Modern Languages, University of Liège, Liège, Wallonia, Belgium
Corresponding author: Mason A. Wirtz. Email: mason.wirtz@plus.ac.at

Abstract
While dialectal variation is often investigated from a geographical angle, there exists substantial variation both within the community and individual. The aim of the present article is to investigate the extent to which spatial, occupational, and age-related factors are associated with the diversity of linguistic variants reported per informant at a given locality. Drawing on colloquial language data from the *Atlas zur deutschen Alltagssprache* 'Atlas of Colloquial German,' we found that informants from southeastern Germany and Austria reported familiarity with more variants. Moreover, we multifactorially operationalize *occupational complexity*, a variable that can capture the effects of different communicative, technical, and physical skills required in a job (via the Dictionary of Occupational Titles). Bayesian multilevel modeling revealed that informants in occupations involving physical precision work and communicative complexity reported less familiarity with variants, and that younger informants were familiar with a wider range of variants.

Keywords: occupational complexity; colloquial language; age-related effects; spatial autocorrelation; variationist sociolinguistics

Introduction

Dialectal variation in traditional atlas projects is most often investigated from a geographical angle, the goal being to illustrate variation between regions and map out linguistically (dis-)similar areas. However, as has been attested (e.g., Steiner, Jeszenszky, Stebler, & Leemann, 2023; Stoeckle, 2016), there may be substantial variation *within* both localities and individual speakers (see also Bülow & Pfenninger, 2021). This necessitates not only approaches that can capture within-person variation, but also begs the question as to what may explain this individual-level variation. To address the former, we explore the diversity of reported variants per informant at a given locality by drawing on crowdsourcing data collected via the *Atlas zur deutschen Alltagssprache* 'Atlas of Colloquial German' (hereinafter: AdA).

Previous work has found that the diversity of reported variants[1] at the *locality* level differs according to region (Stoeckle, 2016) and generation (Wirtz, Pickl, Niehaus, Elspaß, & Möller, under review). The current study expands on these insights by investigating whether the diversity of variants reported by *individual* informants—operationalized in what follows as a standardized measure of the number of variants reported for a linguistic variable—is subject to spatial and age-related effects as well. What is more, exploring the diversity of reported variants at the individual level allows us to incorporate person-level differences lost at the locality level, such as an individual's occupation. Specifically, we draw on the "occupational complexity" measure included in the Dictionary of Occupational Titles (DOT; U.S. Department of Labor, 1977) in order to multifactorially operationalize occupation. The DOT is a source of occupational information in the United States in which occupations have been defined with respect to three complexity dimensions: (i) complexity with people, (ii) complexity with data, and (iii) complexity with things. Broadly speaking, these reflect the premise that each occupation requires a worker to function in relation to these three complexity dimensions. For example, higher complexity in working with people involves occupations associated with more monitoring and supervising, and requires extensive communicative competence. Complexity with data, contrariwise, is characterized by operations of analyzing and computing, and complexity with things is associated with occupations requiring more physical precision work (Smart, Gow, & Deary, 2014).

The paper begins with an overview of variation as source of information and of career as a predictor of language variation. The atlas data, participants, and dependent and independent variables used in the current study are detailed in the methods section. Subsequently, the effects of region, occupational complexity, and age are presented, following which the identified effects are discussed.

Theoretical background

On the utility and meaningfulness of variation

In many traditional dialectological surveys, linguistic variation has been explored primarily from a geographical angle (e.g., Nerbonne, 2010), the reason being that most atlas projects aimed to document old base dialects and their geographical distributions (e.g., Stoeckle, 2016). What is more, many of these atlas projects have only considered one or a few select speakers at each location (e.g., NORMs/NORFs [i.e., *nonmobile, older, rural males/females*]). This, perhaps inadvertently, encourages a view toward *one place = one variety*. As Stoeckle (2016:196) pointed out, however, this view "does not correspond to linguistic reality" in which inter- and intra-speaker variation are the norm rather than the exception (see Bülow & Pfenninger, 2021).

The concept of variation can be a crucial source of information about the individual and/or locality, and this idea is cultivated across disciplines. For example, in cognitive developmental research, larger intraindividual variation—that is "differences in the level of a developmental variable within individuals and between repeated measurements" (van Geert & van Dijk, 2002:341)—in cognitive resources is associated with vulnerability or impairment and is often taken to be indicative of lower cognitive functioning (e.g., Fagot, Mella, Borella, Ghisletta, Lecerf, & De Ribaupierre, 2018).

https://doi.org/10.1017/S0954394524000188 Published online by Cambridge University Press

Relatedly, greater linguistic variation either at the level of the individual or of a locality may be an indicator of language change (e.g., Elspaß, 2018), though it is not necessarily clear whether variation is a precondition of language change or a consequence thereof (e.g., Glaser, 2014). Given the idea that variation may carry information about impending cases of language change, recent years have seen an uptake in attempts to integrate the concept of variation into, for example, syntactic theory (e.g., Cornips & Corrigan, 2005). The idea here is that variation may be an inherent part of an individual's or a locality's grammar and provide the basis for dynamism and language change (e.g., Seiler, 2008), as has also been argued in sociolinguistics since the field's conception (e.g., Labov, 1966).

Stoeckle (2016) investigated interpersonal variation in the *Syntactic Atlas of German-speaking Switzerland* to explore which regions of German-speaking Switzerland evince the most variation, and whether higher variability may be regarded as an indicator of dynamism and linguistic change. To this end, he introduced a "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 most informants, at each location). Localities in which all individuals reported only a single variant (i.e., the dominant variant) received a value of 0 indicative of no variation within the locality, whereas localities in which individuals reported multiple variants received higher scores indicative of higher locality-level variation. Stoeckle concluded from his analysis that the higher variability in certain dialect areas is a result of variants dominant in one dialect area becoming more widespread—in other words, certain local variants may be in the process of becoming more supra-regional variants.

Wirtz et al. (under review) employed a similar measure of variation, the "Variation Intensity Index," which takes into account frequency data concerning all variants and not just the dominant one, and conducted an apparent-time analysis in order to determine whether the degree of variation differs intergenerationally. Concerning generational differences in variation intensity, Wirtz et al.'s apparent-time analysis lends weight to the assumption that the younger generation of informants reported more familiarity with linguistic variants of a variable and thus a higher degree of variation than did the older generation. Specifically, these findings indicate that young adult informants consider a broad repertoire of variants to be common in their locality as compared to individuals from older age groups. These age effects necessitate analyses focusing on measures of variability as the response variable to consider, in addition to other predictor–outcome relationships, (chronological) age as a potential covariate. Additionally, their Variation Intensity Index was calculated at the locality level and disaggregated by two age cohorts—whether related measures of variation at the *informant* level (e.g., the diversity of linguistic variants an informant reports for a variable at a given locality) follows similar age-related patterns, and also between which age group(s) the diversity of reported linguistic variants is prone to substantial change, remains to be seen. Additionally, the aforementioned authors strictly explored the spatial clustering patterns in variation intensity of 15 variables, a comparatively small sample for an aggregate analysis. It thus remains an outstanding question whether the diversity of reported variants computed across a larger sample of variables is subject to spatial autocorrelation.

Career as a meaningful predictor for language variation

The idea of career as a potentially influential factor has not necessarily been left unattended, but empirical results attempting to disentangle how differential aspects of speakers' primary occupation impact on their patterns of variation do not run rampant in the sociolinguistic literature. Traditionally, occupation has been treated as a class-related, economic variable, at least in Western societies. For example, Labov's (1966) study of five phonological variables in the Lower East Side of New York City employed a class model that followed the model developed in 1961 by the Mobilization for Youth project. This included conceptualizing class as a 10-category linear scale (Warner & Lunt, 1942; Warner, Meeker, & Eells, 1949), which combined education level, family income, and occupational groupings. Such an operationalization was considered advantageous, as it would simultaneously tap into several dimensions of socioeconomic status (e.g., Labov, 2001). Along similar lines, Trudgill (1974) drew on a combined class scale but weighted occupation as the most important component. This is based on the notion that class identity alongside the associated (linguistic) behaviors tends to remain constant, at least in Great Britain during the late 20th century, such that "even the most affluent manual workers retain the values, ideas, behaviour patterns and general culture of the working class, and there has been little embourgeoisement of the British working class" (Trudgill, 1974:34). Others (e.g., Macaulay, 1977) rely exclusively on occupational groupings to operationalize a class scale, the rationale being ease of data collection and the fact that further class indicators such as income and education are typically strongly correlated with occupational groupings (see also Dodsworth, 2009).

Recently, starker occupation-centered scales have also emerged. Prediger (2016), for example, classified the occupational status of his participants (second language learners of German) on a 5-point scale based on the degree of manual labor required, the goal being to approximate the intensity of exposure to vernacular speech in the workplace. Advocating for a more multidimensional approach to occupation, Wirtz (2024) proposed a measure of occupational complexity based on the DOT (U.S. Department of Labor, 1977). The goal herewith is to take a multifactorial perspective to the differential communicative and manual requirements of an occupation and how these are intertwined with patterns of varietal behavior. His rationale for a multifactorial approach to occupational status, explicitly not drawing on class-related covariates, was to circumvent spuriously drawing categorical lines or placing clear-cut thresholds on occupations. As Wirtz (2024) noted, occupations house diverse sociological phenomena and complexity dimensions (e.g., communicative, technical, and/or physical skills), which are interwoven with the extent to which speakers are exposed to diverse contextual environments and, by extension, locality- and population-level sociolinguistic variation. Thus, adhering to operational definitions of different complexity dimensions on *continuous* scales allows us to more faithfully capture the granularity in the effects of occupation.

Of course, as Sankoff & Laberge (1978) also noted, even the most carefully conceptualized scales and operationalized variables fall short at capturing the sociohistorical features of occupational factors. Sankoff & Laberge (1978) proposed focusing on the linguistic market, that is, a domain in which (non-)standard language forms are

https://doi.org/10.1017/S0954394524000188 Published online by Cambridge University Press

regarded as useful and desirable capital for particular economic roles. However, as Dodsworth (2009:1318) noted, the index measured standard language competence and ignored possible alternative linguistic markets (as well as alternative cultural settings with different or no linguistic standardization) which do not orientate toward a prestigious standard variety. Given this, we must consider approaches that take into account the influence of occupational complexity on a speaker's linguistic repertoire. Additionally, as lifespan perspectives (e.g., Riverin-Coutlée & Harrington, 2022) have exemplified, shifts in job-related complexity may be strong predictors of differences in linguistic variation, in addition to classic sociolinguistic factors (e.g., Labov, 1966, 2001; Trudgill 1974). This necessitates the additional inclusion of changes in job complexity and/or age as potential moderating covariates when examining the relationship between occupation and linguistic outcomes (e.g., the linguistic repertoire may change more for people who switch careers often, or occupation may have a stronger impact on the language use of young adults recently joining the workforce as opposed to working midlife and older adults).

Finally, whereas previous investigations have explored whether economic factors and differences in occupational status impact, for example, speakers' (socio-stylistic) differential use of sociolinguistic variables, there are no investigations attending to whether the diversity of reported linguistic variants per informant at a given locality relates to career-specific variables such as occupational complexity. This is a desideratum insofar as the range of an informant's reported variants can provide meaningful information on an individual's knowledge of linguistic variation, and whether this diversity can be predicted by career-related complexity metrics.

Research questions

The main goal of this article is to analyze the association between the diversity of an informant's reported variants and occupational complexity, and whether age moderates this relationship—while also attending to potential issues of spatial autocorrelation. To this end, we address the following research questions:

(1) To what extent is the diversity of individual informants' reported variants spatially clustered?
(2) To what extent does occupational complexity with data, people, and things predict differences in the diversity of reported variants (when controlling for locality-related variation as a potential confound)?
(3) To what extent is the relationship between occupational complexity and the diversity of reported variants moderated by chronological age?

As detailed above, occupational complexity with data, people, and things may play a role in influencing the range of an individual informant's reported variants, but empirical evidence along these lines is lacking. From a variationist sociolinguistic angle, the hypothesis stands that the career of a speaker may indeed hold high predictive power when it comes to, for instance, their own linguistic behavior (e.g., Steinegger, 1998). This is because an individual's career is complexly interwoven with the extent to which they are exposed to the linguistic behavior of others, a diversity in contextual

https://doi.org/10.1017/S0954394524000188 Published online by Cambridge University Press

environments, and, consequently, the locality- and community-level (socio-)linguistic variation. Given the exploratory nature of the present analyses, we refrain from any *a priori* hypotheses concerning the (directionality of the) effects of occupational complexity. In light of Wirtz' et al. (under review) findings, however, it does appear plausible that there may be age-related effects on informants' familiarity with different variants at a given locality—specifically, with younger adults reporting familiarity with more variants. What is more, since occupational complexity may change (and likely increase, at least in certain domains) throughout the lifespan (e.g., in relation to work-related milestones such as taking on management positions and/or engaging in more cognitively complex careers due to increased experience), it seems additionally necessary to incorporate age as a moderating covariate of occupational complexity.

Data and methodology

Atlas of Colloquial German

The data for the present study result from the *Atlas zur deutschen Alltagssprache* 'Atlas of Colloquial German' (AdA; Elspaß & Möller, 2003; see also Pickl, Pröll, Elspaß, & Möller, 2019), the largest and longest running linguistic atlas of contemporary colloquial German in the German-speaking world. The data are collected in approximately annual to biannual intervals via online surveys in German-speaking regions (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 variants in lexis, pronunciation, grammar, phraseology, and pragmatics, and in some cases also to declare how common/uncommon a certain variant or construction is for the "everyday colloquial speech" in their respective locality. The informants are usually presented with a brief concept in the form of a picture and/or description along with a list of potential variants from which they are then requested to choose the expression(s) "normally" used in their locality, or otherwise to provide a variant not listed (for examples from the current round, see https://www.atlas-alltagssprache.de). In the AdA, participants act as *informants* and are thus requested not to indicate their own personal use, but are rather instructed to name variants that are used in the colloquial speech of their hometown/city, that is, the kind of speech one would normally hear, be it more closely oriented toward a dialect or standard German. Importantly, given the heterogeneity in potential varietal spectra in the German-speaking realm (e.g., diaglossic and diglossic settings, see Auer, 2005), the concept of "everyday colloquial speech" can range from local dialects (e.g., in German-speaking Switzerland) over intermediate varieties (e.g., in some areas of Bavarian-speaking Austria) to different regional forms of standard German varieties (e.g., in northern Germany and metropolitan areas).

In the AdA, the localities are not predefined, and participation is not constrained by any social-specific factors, which facilitates a more realistic portrait of colloquial language variation by way of capitalizing on informants' "expertise of the local language use" (Pickl et al., 2019:41, our translation). Such crowdsourcing data collection methods necessarily result in socially very diverse datasets (see Leemann, Derungs, & Elspaß, 2019), particularly as concerns variables relating to, for example, age, gender, mobility, socioeconomic status and, importantly for this contribution, occupational status.

Informants

The present study draws on data from Round 11 of the AdA, and we concentrate on Germany, Austria, Switzerland, and German-speaking northern Italy, as these countries evinced the highest response rates. In terms of data cleaning, we excluded participants who (a) did not provide a postal code and/or place of residence, and thus could not be categorized in terms of regional location; (b) did not provide information on their age; (c) did not provide information on their occupation; or (d) for whom no occupational complexity measure could be determined. For 9,173 informants—the main sample pool for the present study—it was possible to link their main occupation to an occupational complexity measure in accordance with the DOT (see the section, Occupational Complexity, for more details on the categorization process). As concerns the nationality distribution of informants, we received the most responses from individuals in Germany ($n = 7563$), followed by Austria ($n = 1156$), Switzerland ($n = 390$), and the German-speaking parts of northern Italy ($n = 64$). These quantities are approximately proportional to the German-speaking populations of the respective countries and thus the four countries were sampled at a consistent rate.

Participation was comparatively balanced in terms of gender (men = 4276, women = 4897), but there was an uneven distribution of age cohorts (10–19 = 743; 20–29 = 3276; 30–39 = 2168; 40–49 = 1324; 50–59 = 1,117; 60+ = 545), with adolescents and older adults being the most underrepresented age cohorts. Note that data on informants' chronological age were collected categorically, which is why age is entered a categorical rather than as a numeric predictor in the following analyses.

Following Wirtz et al. (under review), localities were determined by using 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. This resulted in 114 localities and allowed us to examine a geographically large area, but which was condensed enough to still capture areal variation (note that, for the AdA maps published on the internet, the data are grouped according to a fixed network of locality points).

Quantifying the diversity of reported linguistic variants

In order to capture the diversity of reported variants in the AdA data quantitatively, we employed a standardized measure of the number of reported variants of a variable for each informant, averaged across the 62 variables detailed in the following section. Specifically, a variable with n variants was assigned a value of $1 - (1/n)$ (e.g., a single variant reported yields a measure of 0, two variants reported yield a measure of 0.5, etc.). In other words, the integer value of the number of variants informants report being familiar with was scaled to the interval $[0, 1)$ and the scores were subsequently averaged across 62 variables. While from a purely computational perspective it makes no difference to use the scaled measure of diversity of reported variants as opposed to the number of reported variants for each informant, the scaled measure functions as

https://doi.org/10.1017/S0954394524000188 Published online by Cambridge University Press

an intuitive standardized measure of variation in terms of the informants' familiarity with linguistic variants. Additionally, since this measure of variation was adapted from Wirtz et al.'s (under review) Variation Intensity Index, the aforementioned scaling procedure also ensures that measures of variation at the informant-level and locality-level can be more readily compared across studies. Furthermore, the aggregate approach (i.e., a single measure of informants' familiarity with linguistic variants) is in line with operationalizations of intraindividual variation, which often encapsulate task-specific individual-level variation in a single measure (e.g., from developmental psychology the "intra-individual standard deviation" score [Fagot et al., 2018]).

Table 1 provides the descriptive statistics for the scaled measure of the diversity of reported variants across the six age cohorts (see also Figure A2 in the online Appendix for a visual overview, and also Figure A3 for a visual overview in relation to gender).

Table 1. Descriptive statistics of the scaled measure of the diversity of reported variants across age cohorts

Age cohort	Mean	SD	Range
10–19	0.050	0.035	0–0.151
20–29	0.049	0.035	0–0.195
30–39	0.040	0.034	0–0.142
40–49	0.034	0.032	0–0.129
50–59	0.028	0.029	0–0.134
60+	0.025	0.029	0–0.149

Variables

The scaled measure of the diversity of reported variants was calculated based on 62 variables from Round 11 of the AdA (for the variables in Round 11, see https://www.atlas-alltagssprache.de/elfte-runde/). Items requiring informants to declare whether a certain variant or construction is common/uncommon were excluded from analysis, as these inherently result in single-answer responses and thus leave no room for informants to indicate familiarity with multiple variants of the variable. The variables included were mainly lexical in nature, though select morphological, phonological, and syntactic variables were also included. Table 2 provides several examples of the variables from which the scaled measure of the diversity of reported variants was derived.

Table 2. Examples of variables from which the scaled measure of the diversity of reported variants was derived

Linguistic domain	Variable	Potential variants
Lexical	*Apfelrest* 'apple core'	*Kerngehäuse*, (*Apfel-*)*Butzen*, (*Apfel-*) *Griebsch*, etc.
Morphological	Past participle of *aufhängen* 'to hang up'	*aufgehängt, aufgehangen*
Phonological	Word accent of *Büro* 'office'	First syllable, second syllable
Syntactic	Relative clause pronoun/particle	*die, die wo, wo, die was*

https://doi.org/10.1017/S0954394524000188 Published online by Cambridge University Press

Figure A1 in the online Appendix illustrates the raw number of reported variants by individual informants for each of the 62 variables and illustrates that participants typically did not report more than two, three, or four variants of a variable.

Note that the different linguistic domains mentioned prior were collated in the final measure of the diversity of reported variants. This is in line with previous dialectometric work which emphasizes aggregation as a way to smooth over the noise of individual variables and linguistic architecture and produce a more representative measure. For instance, in his dialectometric analysis of the *Sprachatlas von Bayerisch-Schwaben* ('Language Atlas of Bavarian Swabia'), Pröll (2015) examined different linguistic domains both separately and in an aggregated manner and found that, on the whole, the overall picture was more representative and more than the sum of the individual parts (i.e., linguistic domains). Moreover, in their pilot study, Wirtz et al. (under review) found no difference between the locality-level variation intensity in lexical, morphological, and phonetic variables.

Occupational complexity

In each AdA round, informants were asked to provide their current primary occupation. Main occupation was then matched with the best fitting category listed in the fourth edition of the United States DOT (U.S. Department of Labor, 1977), which comprises more than 12,000 occupations that have been evaluated based on observations by job analysts. Of the 9-digit classification code listed for each occupation, the middle three digits represent occupational complexity with data, people, and things respectively. This measure reflects the notion that each occupation requires workers to function in relation to these three complexity dimensions.

Table 3 indicates the dimensions used to classify occupations. To facilitate interpretive ease, scores were recoded so that for each dimension a higher value was indicative of higher occupational complexity (ranges for complexity with data: 0–6; for complexity with people: 0–8; and for complexity with things: 0–7, and in the statistical models, occupational complexity was z-scored). For example, the occupation "Automobile Mechanic" is assigned scores of 4 for complexity with data, 2 for complexity with

Table 3. Dimensions used in the rating of occupations into complexity of working with data, people, and things

Data		People		Things	
0	Comparing	0	Taking instructions-helping	0	Handling
1	Copying	1	Serving	1	Feeding-offbearing
2	Computing	2	Speaking—signaling	2	Tending
3	Compiling	3	Persuading	3	Manipulating
4	Analyzing	4	Diverting	4	Driving-operating
5	Coordinating	5	Supervising	5	Operating-controlling
6	Synthesizing	6	Instructing	6	Precision working
		7	Negotiating	7	Setting up
		8	Mentoring		

people, and 6 for complexity with things. This indicates that the occupation involves "analyzing" data (i.e., "examining and evaluating data; presenting alternative actions in relation to the evaluation is frequently involved"), "speaking-signaling" people (i.e., "talking with and/or signaling people to convey or exchange information"), and "precision working" with things (i.e., "using body members and/or tools or work aids to work, move, guide, or place objects or materials in situations [...] [in which the] selection of appropriate tools, objects, or materials, and the adjustment of the tool to the task require exercise of considerable judgment"). As another example, the occupation "Pastor" is assigned scores of 5 with data, 8 with people, and 0 with things, which implies that the job components involve "coordinating" data (i.e., "determining time, place, and sequence of operations or action to be taken"), "mentoring" people (i.e., "dealing with individuals in terms of their total personality in order to advise, counsel, and/or guide them"), and "handling" things (i.e., "involves little or no latitude for judgment with regard to attainment of standards or in selecting appropriate tool, object, or materials") (for more extensive descriptions of the occupational complexity metrics, see the *Explanation of Data, People, and Things*: https://www.dol.gov/agencies/oalj/PUBLIC/DOT/REFERENCES/DOTAPPB). Importantly, the occupational complexity measure does come with several caveats, the most notable one being that no occupational complexity measure for "Student" (neither school pupils nor university

Table 4. Descriptive statistics of occupational complexity with data, people, and things across age cohorts

Age cohort	Occupational complexity dimension	Mean	SD	Min.	Max.
10–19	Data	3.02	0.39	0.00	6.00
	People	2.04	0.48	0.00	7.00
	Things	0.18	0.94	0.00	6.00
20–29	Data	3.44	0.97	0.00	6.00
	People	2.52	1.52	0.00	8.00
	Things	0.68	1.81	0.00	7.00
30–39	Data	4.29	1.29	0.00	6.00
	People	3.55	2.24	0.00	8.00
	Things	1.67	2.52	0.00	7.00
40–49	Data	4.23	1.39	0.00	6.00
	People	3.46	2.26	0.00	8.00
	Things	2.08	2.64	0.00	6.00
50–59	Data	4.17	1.42	0.00	6.00
	People	3.61	2.32	0.00	8.00
	Things	2.12	2.64	0.00	7.00
60+	Data	4.22	1.25	0.00	6.00
	People	4.08	2.38	0.00	8.00
	Things	1.74	2.55	0.00	7.00

https://doi.org/10.1017/S0954394524000188 Published online by Cambridge University Press

students) exists. However, based on the guidelines for occupational complexity classification outlined in the DOT, we assigned a value of 3 to the complexity measure with data (i.e., "gathering, collating, or classifying information about data, people, or things"), and following Wirtz (2024), a value of 2 for complexity with people and a value of 0 for complexity with things. This was rationalized by the fact that occupations with arguably similar complexity, activities, and cognitive requirements such as administrative assistants (complexity = 320) and researchers (complexity = 420) are coded in the DOT with similar complexity values. Importantly, an individual's occupation was only coded for the aforementioned complexity measures when a direct translation of the occupation was evident in the DOT (apart from students). This was done to ensure comparability between (a) occupations in the German-speaking countries and the DOT measures developed for the US marketplace and (b) modern occupations and complexity measures developed for occupations in the late 20th century.

While the *O*Net* program (https://www.onetonline.org/) is now the primary source of occupational information in the United States, the occupational complexity measure remains in use in, for example, cognitive psychology (e.g., Smart et al., 2014), as it is nevertheless useful for broadly classifying occupational requirements.

Table 4 lists the descriptive statistics for the three complexity dimensions in informants' primary occupation disaggregated by age cohort. As Figure 1 additionally illustrates, the younger cohort unsurprisingly evinced comparatively homogeneous

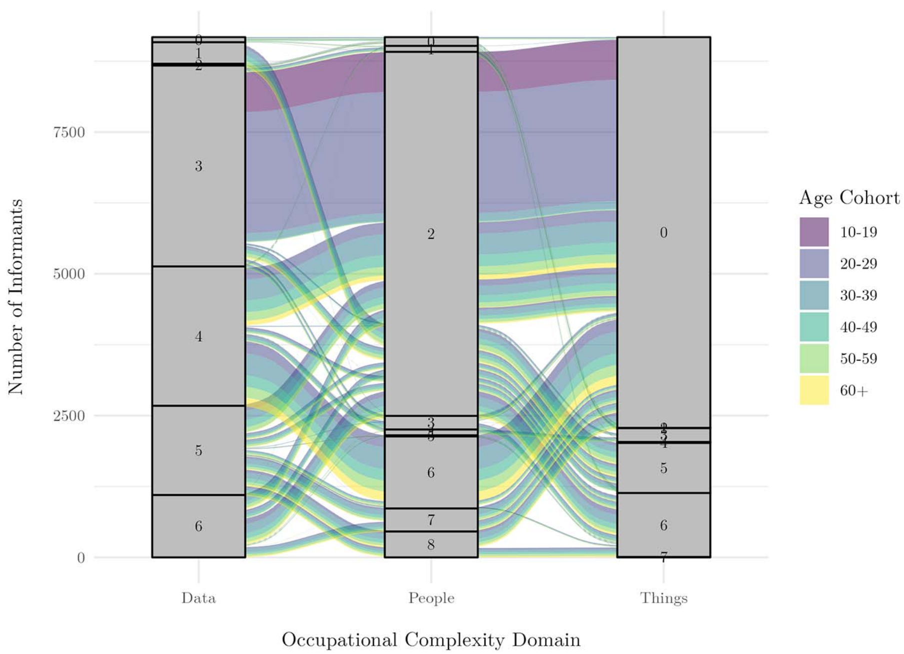

Figure 1. Distributions of occupational complexity measures across the domains of data, people, and things.

https://doi.org/10.1017/S0954394524000188 Published online by Cambridge University Press

complexity scores across data, people, and things, largely given their status as students (either as school pupils or university students). The remaining age cohorts were more heterogeneous in terms of their respective occupational complexity. As concerns complexity with things, we note the surplus in zeros, that is, informants practicing occupations with little to no manual labor demands, which is likely an artifact of the crowdsourcing method and its tendency toward sampling bias (Leemann, Kolly, Purves, Britain, Glaser, 2016), for instance, in terms of socioeconomic status and education. Along similar lines, the people domain of occupational complexity evinces an excess of twos ("Speaking-Signaling"), illustrating that occupations involving an expressly high degree of communicative competence are also a minority in this sample.

Statistical analyses

Given the scaling of the diversity of reported variants as the outcome variable—that is, strictly bounded by [0, 1)—and additionally in light of the surplus of zeros in the data (i.e., informants who reported familiarity with a single variant across all variables), we employed zero-inflated beta models (Markl, 2023). Zero-inflated beta models are tailored to handle datasets that (a) contain many zeros and (b) are bounded by [0, 1), which makes them an optimal choice for current data distribution of the outcome variable.

The zero-inflated beta models were fitted using a Bayesian approach with the *brms* package (version 2.20.4, Bürkner, 2017) in *R* (version 4.2.2, R Core Team, 2020). Bayesian models can estimate generalized (non)linear multivariate models using the probabilistic programming language *Stan* (version 2.32.3, Carpenter, Gelman, Hoffman, Lee, Goodrich, Betancourt, Brubaker, Guo, Li, & Riddell, 2017). The Bayesian framework allows us to investigate the absence of "null effects." The focus of the analyses is directed toward the distributions of the inquired effects (i.e., the posterior distributions) rather than on point estimates. By doing this, we effectively avoid asking questions strictly relating to whether there is an effect of a variable (null hypothesis significance testing) but rather ask what the most probable direction and magnitude of an effect is. (For variationist accounts of the conceptual advantages of Bayesian methods, we refer interested readers to Markl [2023], and for tutorials on Bayesian inferential statistics for the language sciences, see Vasishth, Nicenboim, Beckman, Li, and Jong Kong [2018]).

In this paper, we fitted two models to investigate the relationship between occupational complexity with data, people, and things (note that the occupational complexity measures were z-scored before being entered into the respective models), and the diversity of reported variants as a dependent variable. The first model was fitted with solely the occupational complexity predictors, which provides an unadjusted estimate of the effect of complexity of primary occupation on the diversity of individual informants' reported variants. In this model, we also included an interaction effect between occupational complexity with people and things. Given that previous research has shown that individuals in manual professions tend toward dialectal variants, while individuals in communicatively oriented professions tend toward standard language (e.g., Chambers & Trudgill, 1998:57–59; Niebaum & Macha, 2014:211–212), individuals in occupations requiring a combination of higher complexity in both domains may be more likely to report familiarity with a wider range of variants. In line with our interest in age as a potential moderator variable—age may affect the strength of the relationship between diversity of reported variants and occupational complexity—we fitted a second model including age as an interaction effect with the three occupational complexity measures. The age cohort 20–29 being the largest informant group was entered as the reference level, such that the model compares all other age groups with the 20–29 cohorts. Due to reasons of computational feasibility and comprehensibility, we did not include a three-way interaction effect between occupational complexity with people and things and age cohort—we found this to be rational also because the interaction effect in the first model, though theoretically justified, did not robustly predict differences in informants' familiarity with a wider range of variants, as the Results section will show.

Given that the diversity of reported variants per variable was averaged across the 62 variables in Round 11 of the AdA for each informant—resulting in a single value per informant—there was no need to include by-person random intercepts. However, in order to control for spatial autocorrelation as a potential confounder (see the first Results section for the rationale) when not being entered into the model as a fixed effect, we introduced by-locality random intercepts into the models. The model formulas are detailed in the online Appendix.

In interpreting the effects, we established a region of practical equivalence (ROPE) of ± 0.05 around the point null value 0, which is a region/interval that is practically equivalent to zero (Kruschke, 2018). The underlying idea of the ROPE is that although the coefficient may not precisely equal zero, the effect size may potentially be so small that it holds little practical importance. We judge there to be compelling evidence for a given effect when 95% of the highest density interval (HDI, a type of credible interval, basically the Bayesian analog to the frequentist confidence interval) of the posterior predictive distribution for a parameter β falls outside the ROPE, and when the maximum probability of effect (MPE, the proportion of the posterior distribution that is of the median's sign, indicating whether the probability of the effect is positive or negative) is close to 1 (i.e., 100%). In other words, when the ROPE = 0% and MPE = 100%, we judge the respective effect to be significant.

Model convergence was assessed using the *Rhat* statistic, which was at the ideal value of 1 for each parameter. For each parameter, the effective sample sizes, which are an estimate of the number of independent draws from the posterior distribution, were also at sufficiently high values (>1000, i.e., larger than 10% of the total number of post-warmup draws [Vasishth et al., 2018]).

Results

This section is divided into three main parts corresponding to the research questions. First, we investigate whether the diversity of reported variants in this set of variables is subject to spatial autocorrelation, and if so, in which areas of the German-speaking regions. Second, we explore the effects of occupational complexity with data, people, and things. Finally, we examine the extent to which the diversity of reported variants as a function of complexity in primary occupation is moderated by age.

RQ1: Spatial autocorrelation

In order to determine whether the diversity of reported variants is spatially clustered, that is, subject to spatial autocorrelation, we computed Moran's I at the aggregated locality level. Results indicated slight autocorrelation (Moran's $I = 0.08$, $SD = 0.02$, $p < 0.001$). Specifically, as Figure 2 illustrates, informants in Austria and in south-eastern Germany (i.e., Bavaria) reported familiarity with a wider range of variants as opposed to informants in Switzerland, who reported the lowest diversity of variants. This result can be interpreted to mean that in diglossic Switzerland and Vorarlberg, the westernmost province of Austria, everyday language is characterized by the local dialects in a relatively stable manner, while in all other areas in Austria and in south-eastern Germany that can be characterized as diaglossic, everyday colloquial language oscillates between dialectal and regiolectal varieties, depending on the age of the speakers (see below).

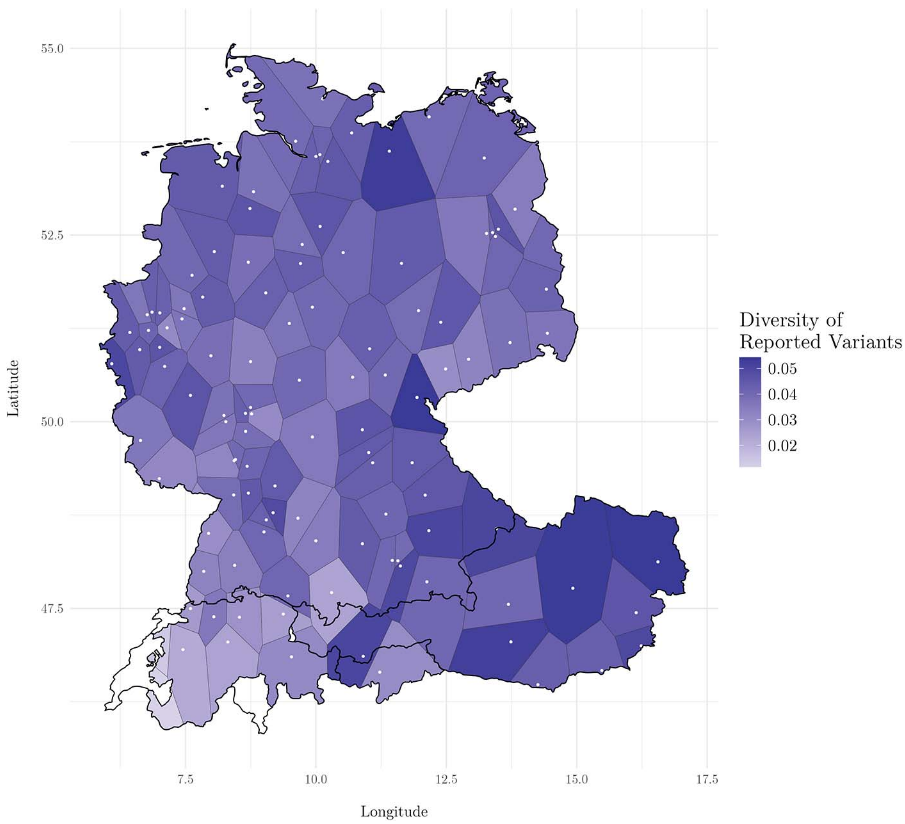

Figure 2. Distribution of the diversity of reported variants (total $n = 114$).

https://doi.org/10.1017/S0954394524000188 Published online by Cambridge University Press

RQ2: Effects of occupational complexity

As regards occupational complexity metrics as predictor variables, we find several notable trends (for the visual model summary, see Figure 3). To start, there is no evidence of an effect of complexity with data on the diversity of reported variants ($\beta = -.004$, HDI = [−.02, .01], ROPE = 100%, MPE = 70%). Interestingly, there is evidence to suggest that the directionality of the effects of occupational complexity with people ($\beta = -.045$, HDI = [−.06, −.03], ROPE = 73.1%, MPE = 100%) and things ($\beta = -0.054$, HDI = [−0.07, −0.04], ROPE = 29.7%, MPE = 100%) is negative—that is, as Figure 4 illustrates, the model predicted informants in occupations with higher communicative demands, as well as in occupations requiring more handling and physical precision work, to report fewer variants overall. Indeed, this rationalizes the inclusion of the interaction effect as well, which, granted, was predicted to be largely positive in directionality ($\beta = .01$, HDI = [.00, .03], ROPE = 100%, MPE = 94.5%), but still included zero and fell within the ROPE. This indicates that the interaction effect did not robustly predict inter-individual changes in the diversity of reported variants. Importantly, whereas the *directionality* of these effects is comparatively clear, and several effects (e.g., complexity with people and things) indeed do not include zero, we underscore that all effect sizes fall within the ROPE, that is, in the interval functionally

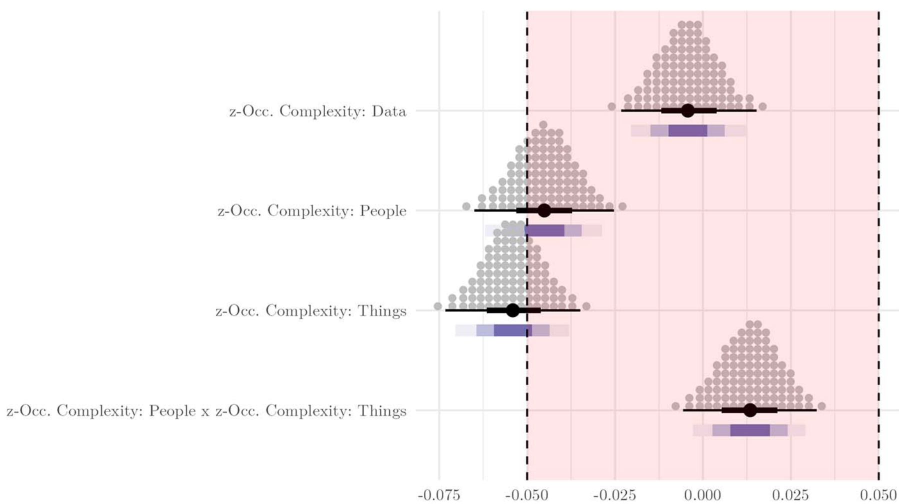

Figure 3. Visual model summary for the effects of occupational complexity on the diversity of reported variants (total $n = 9173$; random intercepts for locality = 114). Quantile dotplots visualize the height, shape, and range of the posterior probability distribution of the predictor variable's effect size (here, in log-odds). Each dot represents a 1% likelihood of a given value. The bars below the dots indicate (from darker to lighter) the 50%, 80%, and 95% HDIs. The black point with bars is the posterior mean (the point), the 98% (thin bar) and 66% (thicker bar) HDIs. The shaded area around point null is the ROPE set at ±.05. Effects that fall within the ROPE, indicating non-sufficient evidence for an effect, are shaded lighter.

https://doi.org/10.1017/S0954394524000188 Published online by Cambridge University Press

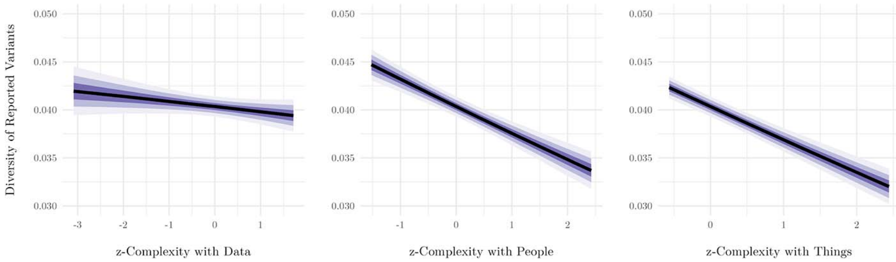

Figure 4. Conditional effects plots for the diversity of reported variants as a function of occupational complexity. The differential shading (from darker to lighter) represents the 50%, 80%, and 95% credible bands around the conditional effects (z-scored) trend line, which represent uncertainty around the population-level averages (black lines).

equivalent to zero. Thus, the magnitude of the effect sizes is simply too small to consider there to be compelling evidence for these effects.

RQ3: Age as a moderator variable

The second model we computed included age cohort as an interaction effect with occupational complexity to determine its potentially moderating role (see Figure 5 for the visual model summary). Again, note that age was entered as a categorical variable with reference level "20–29 age cohort," and thus comparisons in Figure 5 are made with this group. On the whole, there appear to be clear age-related trends in the diversity of reported variants. Specifically, the older age cohorts 40–49 ($\beta = -.20$, HDI = [−.25, −.15], ROPE = 0%, MPE = 100%), 50–59 ($\beta = -0.29$, HDI = [−0.35, −0.24], ROPE = 0%, MPE = 100%), and 60+ ($\beta = -0.33$, HDI = [−0.42, −0.25], ROPE = 0%, MPE = 100%), were estimated to report fewer variants of a variable as compared to the 20–29 age cohort.

The conditional effects plots in Figure 6 visualize the effects of occupational complexity on the diversity of reported variants as mediated by age cohort. The conditional effects are particularly useful in illustrating the directionality of the respective trends, irrespective of the reference level of age cohort. A simple effect analysis did not reveal that any of these trends fall outside of the ROPE interval, suggesting that age did not significantly moderate the effects of occupational complexity on the diversity of reported variants. There were, however, several notable effects in which the directionality was comparatively clear. For example, among the 20–29 age cohort, there was evidence that the directionality of the effects of occupational complexity with people ($\beta = -.04$, HDI = [−.08, −.01], ROPE = 63.8%, MPE = 99.5%) and things ($\beta = -.05$, HDI = [−.09, −.02], ROPE = 39.5%, MPE = 100%) is negative. Similarly, among the 50–59 age cohort, the effect of complexity with people on the diversity of reported variants is negative ($\beta = -.06$, HDI = [−.10, −.01], ROPE = 40.6%, MPE = 99.1%).

On the whole, the Bayesian zero-inflated models provide compelling evidence for age as a predictor of the diversity of reported variants, and Figure 7 displays the age-related differences. Holding the occupational complexity metrics constant at their means, the diversity of reported variants is predicted to decrease from younger to older age cohorts, with the middle age brackets evincing particularly pronounced

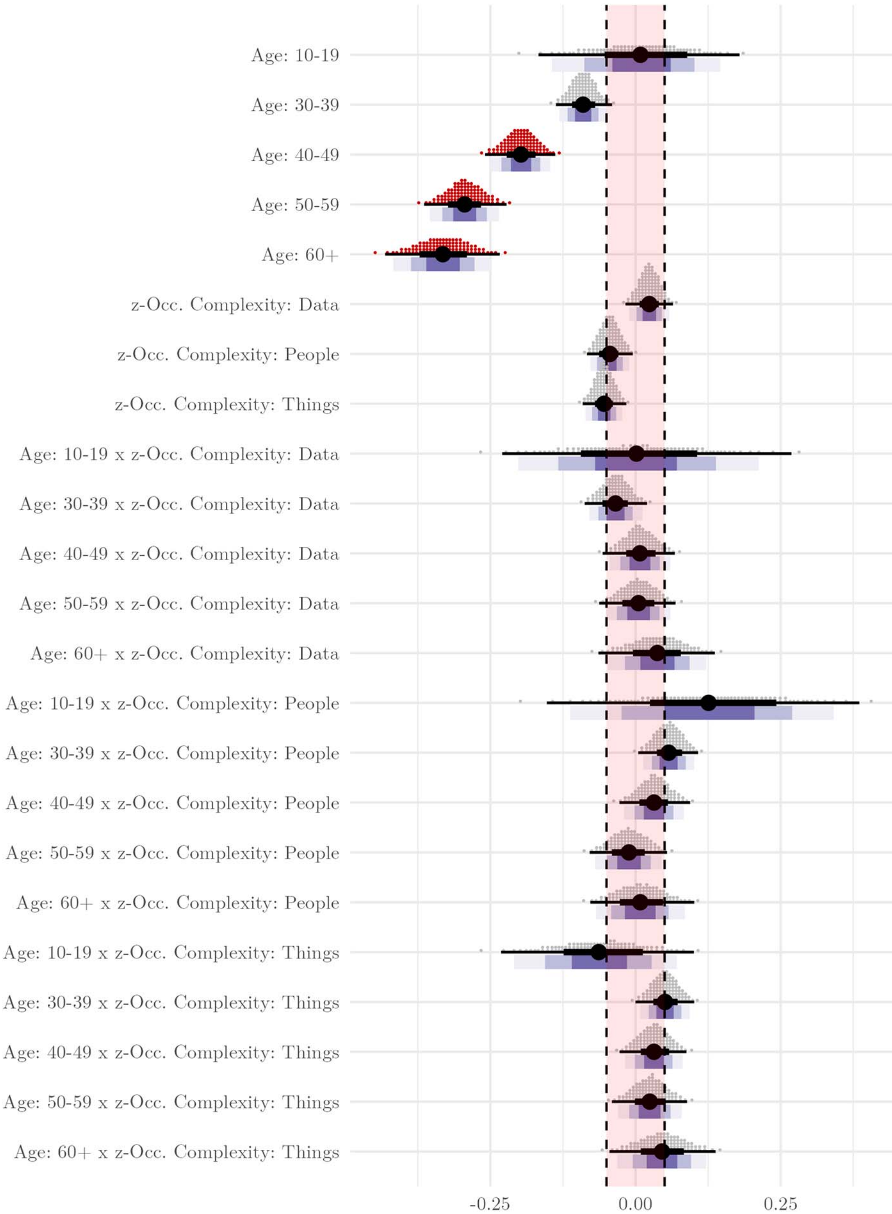

Figure 5. Visual model summary for occupational complexity * age cohort interaction effects (total $n = 9173$; random intercepts for locality = 114).

differences. That said, it would not appear that chronological age moderates the relationship between occupational complexity and the diversity of reported responses to any significant degree.

https://doi.org/10.1017/S0954394524000188 Published online by Cambridge University Press

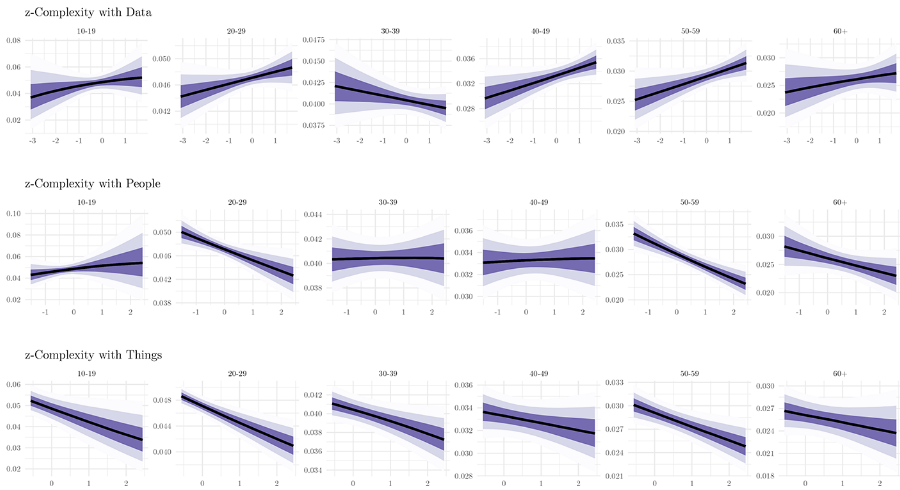

Figure 6. Conditional effects plots for the diversity of reported responses as a function of occupational complexity moderated by age cohort.

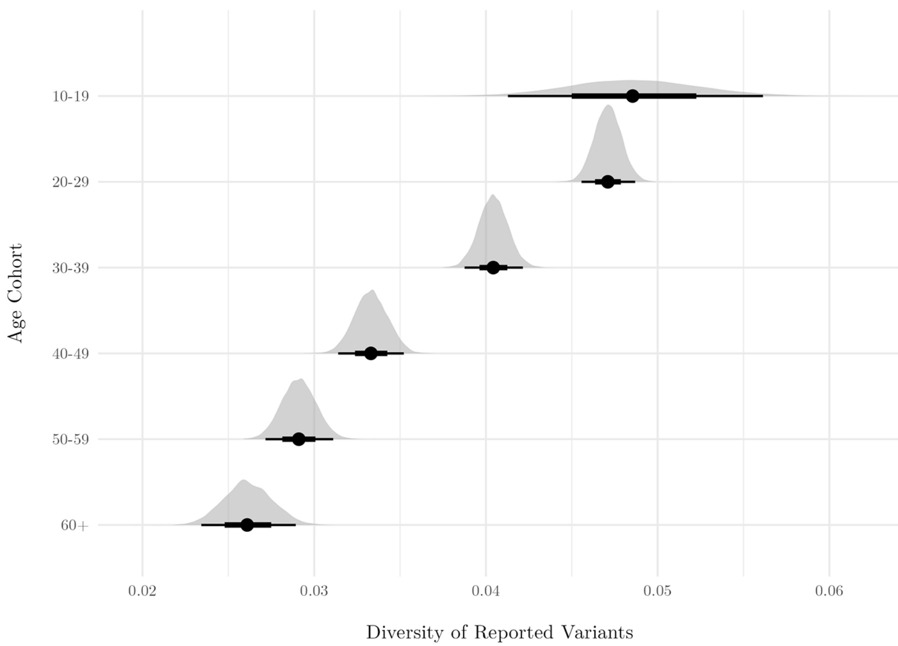

Figure 7. Conditional effects plot for the diversity of reported responses as a function of age holding the z-scored occupational complexity measures constant at their mean (i.e., 0).

Discussion

At any age, in any occupation, there is a pronounced degree of variability among individuals. The aim of this study was to explore which variables (region, occupational

https://doi.org/10.1017/S0954394524000188 Published online by Cambridge University Press

complexity, and the moderating effects of age) are associated with the diversity of reported variants, operationalized here using a scaled measure of the number of reported variants of a variable.

Spatial, occupational, and age-related effects on the diversity of reported variants

We found that the diversity of reported variants aggregated across 62 variables to be subject to spatial autocorrelation. The smaller range of reported variants in Switzerland and the western part of Austria and the higher degree of diversity in all other areas in Austria and in south-east Germany were particularly striking. We suggested that these results adequately reflect the diglossic versus diaglossic language situation in these regions (cf. Auer, 2005 for a typology of such dialect/standard constellations). If the everyday colloquial language in a diglossic environment is characterized by the local dialect, the informants will generally only provide one or at most two variants that are commonly used in the dialect. In a diaglossic constellation, on the other hand, in which the local everyday language can be influenced by dialects, regiolects and even regional standard varieties, it is more likely that the informants in the locality will come into contact with and perceive a broader range of variants.

Slight (though nonsignificant) negative relationships between complexity with people and things and the diversity of reported variants were found, indicating that individuals who practice occupations involving more handling and precision work, alongside informants in positions requiring higher interpersonal competence, were predicted to report *fewer* variants of a variable. Even if this may seem contradictory at first glance, this trend is perhaps not all that surprising. For example, dialectological studies have emphasized that people in manual labor positions tend toward dialectal speech (e.g., Chambers & Trudgill, 1998:57–59; Niebaum & Macha, 2014:186–187). In our case, this may result in informants with high scores for occupational complexity with things (and, by virtue of the regression model, average scores for complexity with people) indicating dialectal and/or regional variants, but abstaining from reporting supra-regional or standard language variants, which likely explains the lower number of reported variants. Conversely, the negative effect of occupations with higher scores for complexity with people (and average scores for complexity with things) on the diversity of linguistic variants reported by informants may relate to higher educational attainment and thus a more pronounced affinity for the standard language. In other words, the fact that individuals with high occupational complexity with people and individuals with high occupational complexity with things (and with average scores for the other occupational complexity covariates) achieve similar values in terms of their familiarity with variants does not suggest that they have *similar* but rather *opposite* repertoires (i.e., individuals with high occupational complexity with people or things do not appear likely to mix varieties, which would explain the fewer reported variants for individuals with high scores on occupational complexity with people or things).

The most robust effects for predicting differences in the diversity of reported variants were found in relation to age cohort. The age effects identified suggest that younger informants, specifically adolescents (10–19) and young adults (20–29), report

https://doi.org/10.1017/S0954394524000188 Published online by Cambridge University Press

familiarity with a wider range of variants whereas informants from the older age cohorts were predicted to report familiarity with fewer variants. These results, based on further data and more variables, confirm Wirtz' et al. (under review) apparent-time analysis between younger and older informants and their hypothesis that younger informants consider a broader repertoire of variants to be common in their locality, composed of both regional and supra-regional variants. Our findings moreover conform with the hypothesis that adolescents (potentially alongside young adults) are leaders of change (see Sankoff, 2018) and that younger generations can be a prime source of information concerning language change (e.g., Eckert, 1997; Kerswill, 1996). As Wirtz et al. (under review) showed, older informants tend to report regional variants, while younger informants reported both regional and supra-regional variants, resulting in overall higher variation among younger individuals. In other words, the fact that there is a difference in the number of variants reported from the older and the younger generation gives rise to the hypothesis of change under the apparent-time assumption. Because younger people reported higher numbers of variants, we assume that in this case the directionality of change, especially in diaglossic regions such as Austria and southern Germany, is toward a larger range of variants, provided the apparent-time hypothesis holds. At first glance, one might be inclined to expect the opposite trend, specifically a decrease of variation in a context of dialect attrition. However, an increase of variation is a necessary phase in a transition from more local to more supra-regional repertoires. In terms of models of language change, this process can also be conceptualized using the S-curve model of language change (e.g., Denison, 2003). At the local level, repertoires were traditionally made up of local/regional variants. These come to be replaced gradually by supra-regional variants, leading to increased variation at the local level. Only when this process nears completion will local variation decrease again. Thus, one hypothesis resulting from both the aforementioned authors' and our results that would need further empirical validation is whether younger informants' tendency to indicate both regional and supra-regional variants is a marker of early-stage long-term change from regional to supra-regional variants, which may potentially result in an overall reduction in linguistic variation.

Given that occupational complexity may change throughout the lifespan (e.g., increases in complexity upon taking on management positions and/or engaging in more cognitively complex careers due to increased experience), we incorporated age as a potential moderator variable for occupational complexity and diversity of reported variants. We did not, however, find that age cohort significantly moderated the relationship between occupational complexity and the diversity of reported variants. Arguably, this may reflect Trudgill's (1974) justification to allocate the most weight to occupation in his combined-class scale, in that, despite the possibility of economic mobility (across the life-course), class identity and the associated behaviors tend to remain constant (see also Dodsworth, 2009).

Limitations and perspectives for future research

While measures of (within- or between-person) variability have been described across disciplines (e.g., the intraindividual standard deviations in developmental psychology [Fagot et al., 2018], or the variation index in dialectology [Stoeckle, 2016]), and thus

https://doi.org/10.1017/S0954394524000188 Published online by Cambridge University Press

the measure of diversity of reported variants employed here is by no means a novel idea conceptually, there are several aspects of the measure that need to be critically assessed. Perhaps the most notable is that analyses employing such operational measures of variability inherently answer different research questions than do traditional analyses of sociolinguistic variation. For example, while many sociolinguistic analyses explore the contextual rationales for the differential use of a standard versus a vernacular variant, measures such as the diversity of reported variants model the degree to which different variants are encoded in a speaker's linguistic repertoire. Thus, in our context the status of a variant as standard or vernacular is of little importance. Moreover, rather than generating information about what speakers *actually* say, the AdA data provide information about informant's exposure to variation at a certain moment in time, and about which variants informants' perceive as potential realization possibilities. By virtue of the AdA questionnaire data, we are therefore likely capturing the largest number of variants encoded in an informant's repertoire (i.e., the highest number of variants with which an informant is familiar). Given this, statements based on comparisons between analyses of classic sociolinguistic variables and measures of variation such as our use of the number of reported variants, and similarly between factors that predict these two measures, are to be made with caution.

In addition to age and region (two factors often in the limelight of sociolinguistic and dialectological work), we also explored the effects of occupational complexity on the diversity of reported variants. In light of this novel measure, it is necessary to acknowledge some of its limitations. Perhaps the most prevalent is that the DOT measures were developed for the United States workforce and last updated during the 1990s. Thus, some modern occupations may not be listed, and potential differences in job complexity metrics between countries, Western or otherwise, are necessarily lost. Additionally, given that AdA collects only very few social variables, correlating occupational complexity with more traditional class-related factors such as education was not possible. This presents an interesting avenue for future endeavors, however.

Finally, it is notable that, despite the fact that the zero-inflated beta distributional family is most appropriate for our data, the posterior predictive checks revealed discrepancies between the observed data and the simulated posterior draws (see Figures A4 and A5 in the Appendix). While other diagnostics (i.e., *Rhat*) indeed indicated that the models successfully converged, the posterior predictive checks suggested there may be much more variation (e.g., relating to individual differences) not captured by the models. In other words, there are likely to be additional variables moderating the relationship between occupational complexity and the diversity of reported variants (in addition to age), but which were not measured in this study. Given the inherent multidimensionality involved in exploring the effects of occupation, it appears necessary to incorporate a more diverse set of individual differences that are related to occupational complexity and thus can be hypothesized to moderate or mediate the relationship between occupation and informants' familiarity with linguistic variants. Future work exploring the role of career on patterns of variation would do well to incorporate measures of cognitive resources given their close relationship with occupation (Coe, von Gaudecker, Lindeboom, & Maurer, 2012); for example, there is the issue of reverse causation between these two variables, that is, individuals with higher levels of cognitive ability may select more cognitively complex work. This similarly necessitates the incorporation of further individual differences, such as educational attainment, personality factors (Steiner et al., 2023), among many others in order to more wholly capture the (also statistically adjusted) relationship between occupational complexity and the diversity of linguistic variants reported by informants.

Conclusion

The present study examined the potential effects of region, occupational complexity, and the moderating role of age on the diversity of variants reported by German-speaking informants. As concerns occupational complexity, our findings highlight slight negative relationships between complexity with people and things and the diversity of reported linguistic variants, though the effect sizes are notably small and must thus be interpreted with caution. Additionally, we found clear effects of age on the diversity of reported variants, such that, echoing Wirtz et al.'s (under review) preliminary results, younger participants generally seem to report familiarity with a wider range of variants—a trend which, in a gradient way, is subject to decline across older generations.

As we see it, the factor occupational complexity has some notable implications for sociolinguistics. In drawing on such a predictor variable in sociolinguistic lines of inquiry, it should be possible to more directly investigate the distinct contribution that career-related factors such as occupational complexity have on patterns of variation. For instance, macro-definitions of socioeconomic class/status (e.g., aggregated education level, family income, and occupational rank, see Warner and Lunt [1942] and Warner et al. [1949]) are oftentimes employed as primary explanatory variables, but these arguably only function as *proxies* for occupation/occupational status. While there may be overlapping properties between different socioeconomic status indicators, it is likely that operational definitions of class- and occupation-related variables may also regress different variance and thus have differential effects on a linguistic outcome variable of interest. In a similar way as, for example, Darin-Mattsson, Fors, & Kåreholt (2017) explored differences in socioeconomic status indicators in capturing variation in health outcomes, sociolinguistics may benefit substantially from exploring in a more nuanced way how, whether, and the extent to which different class- and occupation-related variables predict differences in inter- and intra-individual patterns of (socio)linguistic variation and change. In so doing, it would be possible to more clearly tease apart the differential effects of person-related (socio)economic and occupational traits—an influx of studies drawing on occupational complexity as an *occupation*-distinct variable would present a preferable step in this direction.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0954394524000188.

Acknowledgements. 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. We also wish to express our gratitude to Lara Wlcek for her meticulous work in coding the occupational complexity data, and to the anonymous reviewers for their constructive and very helpful comments on earlier versions of this manuscript. Any remaining errors are, of course, our own.

Competing interests. The author(s) declare none.

Note

1. Note that we use phrases such as "diversity or range of reported variants" as a shortened version of "diversity/range of reported linguistic variants of a variable in a given locality," which more accurately captures the fact that, in the online questionnaire, informants indicated the variants they believe are used *in their respective locality*, rather than indicating all potential variants they know.

References

Auer, Peter. (2005). Europe's sociolinguistic unity, or: A topology of European dialect/standard constellations. In Nicole Delbecque, Johan van der Auwera & Dirk Geeraerts (Eds.), *Perspectives on Variation* (Trends in Linguistics Studies and Monographs 163). Berlin & New York: de Gruyter. 7–42.

Bülow, Lars, & Pfenninger, Simone E. (2021). Introduction: Reconciling approaches to intra-individual variation in psycholinguistics and variationist sociolinguistics. *Linguistics Vanguard* 7(2s):1–10.

Bürkner, Paul-Christian. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software* 80(1):1–28.

Carpenter, Bob, Gelman, Andrew, Hoffman, Matthew D., Lee, Daniel, Goodrich, Ben, Betancourt, Michael, Brubaker, Marcus A., Guo, Jiqiang, Li, Peter, & Riddell, Allen. (2017). Stan: A Probabilistic Programming Language. *Journal of Statistical Software* 76(1):1–32. https://doi.org/10.18637/jss.v076.i01.

Chambers, Jack K., & Trudgill, Peter. (1998). *Dialectology*. Cambridge: Cambridge University Press.

Coe, Norma B., von Gaudecker, Hans-Martin, Lindeboom, Maarten, & Maurer, Jürgen. (2012). The Effect of Retirement on Cognitive Functioning. *Health Economics* 21(8):913–927. https://doi.org/10.1002/hec.1771.

Cornips, Leonie, & Corrigan, Karen P. (eds.) (2005). *Syntax and variation. Reconciling the biological and the social*. Amsterdam & Philadelphia: John Benjamins.

Darin-Mattsson, Alexander, Fors, Stefan, & Kåreholt, Ingemar. (2017). Different indicators of socioeconomic status and their relative importance as determinants of health in old age. *International Journal for Equity in Health* 16:173. https://doi.org/10.1186/s12939-017-0670-3.

Denison, David. 2003. Log(ist)ic and simplistic S-curves. In *Motives for language change*. Cambridge: Cambridge University Press. 54–70.

Dodsworth, Robin. 2009. Modeling socioeconomic class in variationist sociolinguistics. *Language and Linguistics Compass* 3(5):1314–1327. https://doi.org/10.1111/j.1749-818x.2009.00167.x.

Eckert, Penelope. (1997). Age as a sociolinguistic variable. In Florian Coulmas (Ed.), *The handbook of sociolinguistics*. Oxford: Blackwell. 151–167.

Elspaß, Stephan. (2018). Sprachvariation und Sprachwandel. In E. Neuland & P. Schlobinski (Eds.), *Handbuch: Sprache in sozialen Gruppen*. Berlin & Boston: de Gruyter. 87–107.

Elspaß, Stephan, & Möller, Robert. (2003). Atlas zur deutschen Alltagssprache (AdA). https://www.atlas-alltagssprache.de.

Fagot, Delphine, Mella, Nathalie, Borella, Erika, Ghisletta, Paolo, Lecerf, Thierry, & De Ribaupierre, Anik. (2018). Intra-individual variability from a lifespan perspective: A comparison of latency and accuracy measures. *Journal of Intelligence* 6(1):1–18.

Glaser, Elvira. (2014). Wandel und Variation in der Morphosyntax der schweizerdeutschen Dialekte. *Taal en Tongval* 66(1):21–64.

Kerswill, Paul. (1996). Children, adolescents and language change. *Language Variation and Change* 8(2):177–202.

Kruschke, John K. (2018). Rejecting or Accepting Parameter Values in Bayesian Estimation. *Advances in Methods and Practices in Psychological Science* 1(2):270–280.

Labov, William. (1966). *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.

Labov, William. (2001). *Principles of linguistic change, vol. II: Social factors*. Oxford: Wiley-Blackwell.

Leemann, Adrian, Derungs, Curdin, & Elspaß, Stephan. (2019). Analyzing linguistic variation and change using gamification web apps: The case of German-speaking Europe. *PLOS ONE* 14(12):1–29. https://doi.org/10.1371/journal.pone.0225399.

Leemann, Adrian, Kolly, Marie-José, Purves, Ross, Britain, David, & Glaser, Elvira (2016). Crowdsourcing Language Change with Smartphone Applications. *PLOS ONE* 11:e0143060. https://doi.org/10.1371/journal.pone.0143060.

Macaulay, Ronald K. S. (1977). *Language, social class, and education: a Glasgow study*. Edinburgh: Edinburgh University Press.

Markl, Nina. (2023). "I can't see myself ever living any[w]ere else": Variation in (HW) in Edinburgh English. *Language Variation and Change* 35(1):1–27. https://doi.org/10.1017/S0954394523000078.

Nerbonne, John. (2010). Measuring the diffusion of linguistic change. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365(1559):3821–3828. https://doi.org/10.1098/rstb.2010.0048.

Niebaum, Hermann, & Macha, Jürgen. (2014). *Einführung in die Dialektologie des Deutschen*. De Gruyter. https://doi.org/10.1515/9783110338713.

Pickl, Simon, Pröll, Simon, Elspaß, Stephan, & Möller, Robert. (2019). Räumliche Strukturen alltagssprachlicher Variation in Österreich anhand von Daten des "Atlas zur deutschen Alltagssprache (AdA)." In L. Bülow, A. Fischer & K. Herbert (Eds.), *Dimensions of Linguistic Space: Variation – Multilingualism Conceptualisations. Dimensionen des sprachlichen Raums: Variation – Mehrsprachigkeit – Konzeptualisierung*. Berlin et al.: Peter Lang. 39–60.

Prediger, Alexander. (2016). *Erwerb und Gebrauch alemannischer Dialektmerkmale durch russische Muttersprachler*. Doctoral dissertation, Universität Freiburg.

Pröll, Simon. (2015). *Raumvariation zwischen Muster und Zufall. Geostatistische Analysen am Beispiel des Sprachatlas von Bayerisch-Schwaben*. Stuttgart: Franz Steiner Verlag.

R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/.

Riverin-Coutlée, Josiane, & Harrington, Jonathan. (2022). Phonetic change over the career: a case study. *Linguistics Vanguard* 8(1):41–52. https://doi.org/10.1515/lingvan-2021-0122.

Sankoff, David, & Laberge, Suzanne. (1978). The linguistic market and the statistical explanation of variability. In D. Sankoff (Ed.), *Linguistic variation: models and methods*. New York: Academic Press. 239–50.

Sankoff, Gillian. (2018). Language change across the lifespan. *Annual Review of Linguistics* 4(1):297–316.

Seiler, Guido. (2008). Syntaxgeographie und die Plastizität der Grammatik. In M. Valentin & H. Vinckel (Eds.), *Akten des XI. Internationalen Germanistenkongresses Paris 2005*. Bern: Peter Lang. 49–58.

Smart, Emily L., Gow, Alan J., & Deary, Ian J. (2014). Occupational complexity and lifetime cognitive abilities. *Neurology* 83(24):2285–2291. https://doi.org/10.1212/WNL.0000000000001075.

Steinegger, Guido. (1998). *Sprachgebrauch und Sprachbeurteilung in Österreich und Südtirol, Ergebnisse einer Umfrage*. Frankfurt a. M. et al.: Peter Lang.

Steiner, Carina, Jeszenszky, Péter, Stebler, Viviane, & Leemann, Adrian. (2023). Extraverted innovators and conscientious laggards? Investigating effects of personality traits on language change. *Language Variation and Change* 51(1):1–28. https://doi.org/10.1017/S0954394523000091.

Stoeckle, Philipp. (2016). Horizontal and vertical variation in Swiss German morphosyntax. In M. Côté, R. Knooihuizen & J. Nerbonne (Eds.), *The future of dialects*. Berlin: Language Science Press. 195–215.

Trudgill, Peter. (1974). *The social differentiation of English in Norwich*. Cambridge: Cambridge University Press.

U.S. Department of Labor (1977 [revised 1991]). Dictionary of occupational titles (4th ed.). Washington. Retrieved December 10, 2023 from https://occupationalinfo.org/appendxb_1.html

Van Geert, Paul, & van Dijk, Marijn. (2002). Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behavior and Development* 25(4):340–374. https://doi.org/10.1016/S0163-6383(02)00140-6.

Vasishth, Shravan, Nicenboim, Bruno, Beckman, Mary E., Li, Fangfang, & Jong Kong, Eun. (2018). Bayesian data analysis in the phonetic sciences: A tutorial introduction. *Journal of Phonetics* 71(4):147–161.

Warner, W. Lloyd, & Lunt, Paul. (1942). *The social life of a modern community*. New Haven: Yale University Press.

Warner, W. Lloyd, Meeker, Marchia, & Eells, Kenneth. (1949). *Social class in America*. Chicago: Science Research Associates.

Wirtz, Mason A. (2024). On the relationship between complexity of primary occupation and L2 varietal behavior in adult migrants in Austria. *Linguistics Vanguard*. 1–15. https://doi.org/10.1515/lingvan-2023-0069.

Wirtz, Mason A., Pickl, Simon, Niehaus, Konstantin, Elspaß, Stephan, & Möller, Robert. (under review). Reconciling the social and spatial: An apparent-time analysis based on variation intensity.

Cite this article: Wirtz, Mason A., Simon Pickl, Konstantin Niehaus, Stephan Elspaß and Robert Möller. (2024). Spatial, occupational, and age-related effects on reported variation in colloquial German. *Language Variation and Change* **36**:307–331. https://doi.org/10.1017/S0954394524000188