To appear: Language Variation and Language Change Across the Lifespan: Theoretical and Empirical Perspectives from Panel Studies. K. V. Beaman and I. Buchstaller (eds). New York: Routledge.

Deconfounding the effects of competition and attrition on dialect across the lifespan: A panel study investigation of Swabian

R. Harald Baayen, University of Tübingen and University of Alberta
Karen V. Beaman, Queen Mary University and University of Tübingen
Michael Ramscar, University of Tübingen[1]

Abstract

The dialectical changes seen across the course of individual lives are typically thought to reflect the attritional influence of standard languages on native dialects. However, the distributional properties of natural languages, which guarantee that lexical knowledge continuously increases across the lifespan, suggest these changes might simply reflect the broadening and diversification of individual vocabularies, not the loss of dialect itself. Consistent with this proposal, speech analyses from 20 speakers of the southwestern German dialect Swabian, recorded in 1982 and again in 2017, reveal that across their lifetimes, these speakers did not suffer a significant loss of dialect, but rather gained a vast amount of non-dialectal vocabulary, a pattern of change that was promoted or constrained by local orientation and personal identity. The analyses show that dialect words were actually used with similar frequencies across the two recording periods, indicating that speakers' dialectal knowledge remains largely intact, while in the later recordings low-frequency words from the standard language were used at increased rates, reflecting gains in non-dialectal vocabulary across the lifespan. These results suggest an alternative account of the changes in individual speech patterns in which the changes observed in lexical choice across the lifespan primarily reflect the increased influence of later acquired, usually non-dialect, lexical knowledge, and not necessarily the "loss" of dialect itself.

---

[1] Authors are presented in alphabetical order to represent the collaborative nature of this research: Beaman provided that Swabian corpus and sociolinguistic framework, Baayen the models and statistical methods for investigating changes in lexical richness, and Ramscar the hypotheses on aging and language change across the lifespan. All three authors collaborated on the writing and interpretation of the results. The authors would like to thank Jenny Cheshire, John Nerbonne, and Ulrich Reubold for their comments on an early draft. Any deficiencies remaining are, of course, our own.

Keywords: lexical frequency, lifespan change, language change, ageing, dialect contact, dialect attrition, standard language, dialect identity.

1. Introduction

Rising mobility, increasing levels of education, and intensifying immigration are bringing more diverse people into more frequent contact, more prolonged interactions (Auer 2007, Britain 2013, 2016, Britain and Trudgill 1999, Dodsworth 2017, Trudgill 1992). These factors, coupled with continuing globalisation and ubiquitous social media, are pushing standard languages into the forefront of people's experience and relegating non-standard varieties to the background. As a consequence, a growing body of research suggests that many dialects, i.e., non-standard language varieties, are receding across the globe (Britain 2009, Schilling-Estes and Wolfram 1999, Smith and Durham 2012), and nowhere is this more evident than in Europe, notably in Germany (Auer 2005, 2018, Auer, Baumann, and Schwarz 2011, Auer and Spiekermann 2011, Kehrein 2012, Pedersen 2005, Schmidt 2011, Streck and Auer 2012).

Dialectologists generally focus on changes in lexical items and the use of dialect-specific words (e.g., Swabian *Grombiere* versus standard German *Kartoffel* 'potato'), variationists tend to target changes in the frequencies of various phonological, grammatical, and discourse-pragmatic forms (e.g., Alemannic *Fescht* [fɛʃt] versus standard German *Fest* [fɛst] 'party'), and corpus/computational linguists often look at the competition between grammatical forms and changes in frequencies between different word forms (e.g., colloquial *geh* versus standard German *gehe* 'go'). Cumulatively, these metrics reveal that where different language varieties come into contact, accommodation occurs, and most commonly, it is the dialect variants that lose out and the more prestigious, standard variants that "win" (Britain and Trudgill 1999, Giles, Taylor, and Bourhis 1973, Trudgill 1986, Wieling, Nerbonne, and Baayen 2011). Indeed, this pattern is also seen in individuals: as they age, adult speakers appear to lose dialect as they gain greater experience with the standard language, gained through their participation in various educational, commercial, and public institutions (Britain 2010, Eckert 1997, Labov 1964, Sankoff and Laberge 1978).

The idea of *dialect attrition* is the dominant way of interpreting these patterns of language development. This interpretation assumes that standard languages encroach on dialects, such that, at the lexical level dialect words are replaced by their standard language counterparts,

resulting in the attrition of individual dialect vocabularies. There are, however, problems inherent in this assumption that are particularly relevant to lifespan studies of dialect usage. First, the lexical distributional properties of natural languages (Baayen 2001) ensure that the lexical knowledge of healthy individuals increases continuously across their lifespan. These same distributional properties also guarantee that the majority of lexical types any individual knows are relatively rare and that many of these types will be shared only with subsets of the wider community. As people age, their knowledge expands as they gain new experiences (e.g., in schools, on the job, at leisure), face various new life events (e.g., graduation, marriage, childbirth), and tackle new challenges (e.g., driving a fork-lift, climbing Kilimanjaro). In the course of these undertakings, speakers encounter new words and add them to their vocabularies. Many of these new words are specific to particular areas of knowledge, such as medicine, plumbing, or linguistics, and are not in the vocabularies of other speakers in the community. In an increasingly technology-driven world, this increased lexical knowledge may involve words for new inventions and technologies (e.g., *cell phone, fax, emoji*). Importantly, it is likely that many of the specialisation-specific words, as well as words for cultural innovations, have the same form in both the dialect and the standard language.

The second challenge in investigating dialect attrition across the lifespan relates to the differing social settings in which the use of standard language or dialect is appropriate. A local dialect is lexically strong for discussing traditional methods of farming and socially appropriate for informal interactions with family and friends in the local community. The standard language comes into its own for interactions with speakers from different backgrounds or to cover topics for which the dialect does not offer the relevant specialised words. These two considerations thus suggest an alternative account of the changes in speech patterns as individuals age: specifically, lexical change across the lifespan does not necessarily represent skill loss, as has often been claimed (e.g., Köpke and Schmid 2004), but rather reflects the fact that experience tends to make individuals *more* skilled when measured in terms of their ability to communicate about an expanding repertoire of topics. Thus, many changes in speech patterns merely reflect the increased influence of later acquired, standard language lexical knowledge, and not necessarily a substantial loss of the dialect itself.

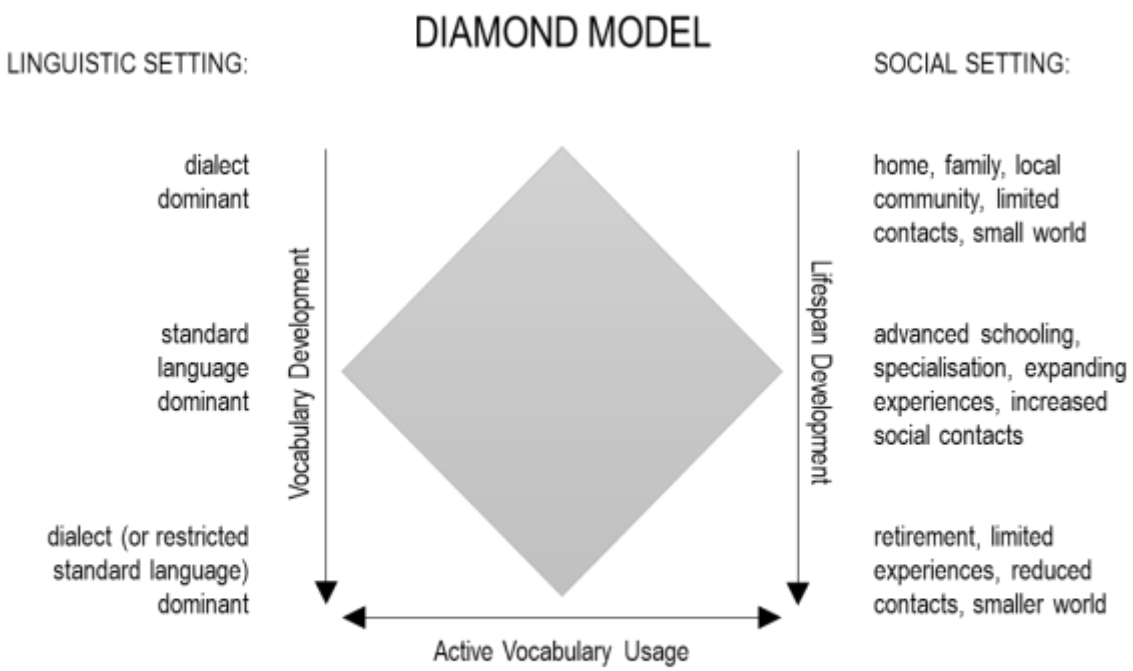

Figure 1. Diamond Model of vocabulary development over the lifespan

Furthermore, as we will argue, earlier acquired dialect forms are likely to be more deeply embedded in speakers' repertoires than later acquired standard forms and, therefore, are more likely to be reactivated as individuals age. Our view of lexical change across the lifespan can be visualised as a diamond. In their youth, speakers' experiences are naturally quite limited, and the breadth of their active vocabulary usage reflects this narrower range of experiences, which are primarily with the local community in the native dialect (the upper point of the diamond). As individuals move into the workforce, their vocabularies expand along with their experiences, to express a widening range of interests and activities (the wider midpoint of the diamond). Then, in later life, particularly in retirement, the range of activities and the breadth of social contacts slowly decrease, resulting in talk about a smaller subset of topics (the lower point of the diamond). Our Diamond Model of vocabulary development over the lifetime (see Figure 1) proposes that dialects are typically the primary medium for communicating in the early and the later stages of life (the upper and lower parts of the diamond), with standard languages playing a greater role during mid-life when individual experiences and specialisations are most varied (the widest part of the diamond, cf. 'linguistic market' (Sankoff and Laberge 1978)). As a result, if

there is active vocabulary loss[2] across the lifespan, as individuals advance in age, narrow their circle of contacts, and reduce their exposure to standard language settings, we expect it to be primarily visible in the standard language, following the principle of "use it or lose it" (Shors et al. 2012).

1.1. The Hypotheses

The central hypothesis for this study is thus that rather than losing dialect as a result of myriad experiences throughout their lifetime, speakers actually gain a massive amount of new lexical knowledge that is not dialect (i.e., the expanding of the diamond). The standard language is the medium *par excellence* for diversification in all fields of specialisation, an aspect of human language development that is particularly acute during mid-life. Thus, it follows that, so long as speakers continue to participate in those situations where speaking dialect is appropriate, the dialect will remain strong. As more experiences accumulate outside the sphere in which the dialect is the primary mode of communication, we expect to see an increase in the use of the standard language and a relative decrease in the use of dialect. As the breadth of experiences begin to subside in later life and speakers' worlds become smaller (i.e., the contracting of the diamond), we expect the trend to reverse, revealing a decrease in standard language usage and an increase in dialect (with an exception, of course, for individuals who have moved outside of the dialect sphere, such as, to other localities or with non-dialect-speaking partners.

How broadly or narrowly individuals' dialect and standard vocabularies expand or contract over the course of their lifetime will also be heavily dependent, amongst other things, on their personal orientation and identity with the local dialect and community. Studies have shown that individuals who identify more with the local community and place high value on their culture and traditions are more likely to retain more dialect forms, while individuals who orient themselves beyond the local community and manifest broader and more diverse world views are more likely to use more standard language forms (Beaman 2020, Beaman and Tomaschek, this volume). Hence, our second hypothesis predicts that the degree of dialect loss or maintenance is modulated by speakers' local orientation and identity with their homeland.

[2] It is important to point out that a loss in active vocabulary use does not necessarily imply a loss of knowledge; vocabulary knowledge may be passive and certain words may simply not have been observed in the current sample.

Related to the supposition that speakers lose dialect forms as they age is whether dialect forms are used with the same intensity (i.e., frequency) as standard language forms across the lifespan. Prior research has shown that the encroachment of the standard language on the dialect is most successful for the lowest frequency words (drawn from the CELEX lexical database) (Wieling, Nerbonne, and Baayen 2011). Lower-frequency words tend occur in contexts that are ever-more specific, in which they are used far more frequently than their average probabilities would otherwise predict (Katz 1996). An empirical consequence of this is that the lower average frequency any given word has, the harder it will become to disentangle its loss from its not having been relevant to any given context observed. This problem is further confounded by an inevitable consequence of our first hypothesis, simply because any growth in knowledge and use of standard language vocabulary items must inevitably lead to a decrease in the frequency at which dialect items are used. That is, when Swabians add *gemelli* and *ravioli* to their vocabularies, and when they eat *gemelli* and *ravioli* on days when they might previously have eaten *Spätzle* and *Maultaschen*, the average frequencies at which they use *Spätzle* and *Maultaschen* must inevitably decrease, along with the number of contexts in which they are used in. Given that these changes will not reflect an individual's loss of *Spätzle* and *Maultaschen* so much as the increase in the breadth and specialization of their vocabulary as a result of their extra experiences, our third hypothesis predicts that where dialect vocabulary items do appear to be lost across the lifespan, these apparent losses will actually reflect reduced intensity of use and not the loss of the knowledge of individual word types, which we expect to be most prominent in the lowest frequency words.

1.2. The Current Study

This study investigates lexical richness in dialect and standard language word usage across the lifespan. Our investigation is positioned at the intersection of the fields of dialectology (dialect contact and attrition studies), sociolinguistics (longitudinal variationist and identity studies), psycholinguistics (lexical frequency studies), and psychology (ageing and cognition studies). We first describe the corpus we used and explain the methodology we employed, followed by a presentation of the analysis and results. We conclude with a discussion on the importance of considering lexical distributions and the nature of lifetime learning in studies of language change across the lifespan.

2. Data and Methods

2.1. Swabian Corpus

The corpus for this investigation comprises Labovian-style, semi-structured sociolinguistic interviews (Labov 1984) with 20 speakers of Swabian, a high-Alemannic dialect spoken in southwestern Germany by just over 800,000 people or one percent of the German population (Eberhard, Simons and Fennig 2019). Each speaker was interviewed twice, once in 1982 and again in 2017, for approximately an hour, although the interviews in 2017 tended to be somewhat longer. The interviews were typically conducted in the speakers' homes over coffee and cake, using the same template in 1982 and in 2017, covering questions about the speakers' childhood, games, leisure activities, family, friends, and the Swabian language and culture. If speakers wandered off topic, the interviewer did not interrupt them, with the aim of obtaining natural, unmonitored speech. Local native Swabian speakers were selected as interviewers, matched in 1982 and 2017 for similar social characteristics (i.e., same age group, gender, educational level). Although the overall goal was to create similar interview situations for all speakers in both time periods, in reality, the 'Gap Effect' in longitudinal panel studies is unavoidable (Cukor-Avila and Bailey 2017; Wagner and Tagliamonte 2017). A key difference between the two interview periods is the 'Interviewer Closeness', that is, the degree to which the interviewer and informant are familiar with each other. In 1982, the interviewees were close family and friends of the interviewer; in 2017, except for two speakers, the interviewers and interviewees were strangers, reintroduced after a 35-year break as a 'friend-of-a-friend-of-a-friend' (Milroy 1987). This difference in the relationship 'closeness' has the effect of creating a slightly more formal situation in 2017, hence, we expect to see somewhat greater use of the standard language in the later recordings.

The corpus comprises two different speech communities, providing the opportunity to investigate changes in language use in both an urban and a semi-rural setting. Stuttgart is a large urban centre with over one million inhabitants and one of the most diverse populations in the country with almost twice as many "foreigners" (individuals with at least one parent who immigrated) as in Germany overall (Auer 2020).[3] In contrast, Schwäbisch Gmünd is a considerably smaller, semi-rural town of 60,000 inhabitants. Seven speakers are from Stuttgart,

[3] *Statistisches Amt, Landeshauptstadt Stuttgart*, https://statistik.stuttgart.de/statistiken/tabellen/7392/jb7392.php

four men and three women, and 13 from Schwäbisch Gmünd, seven men and six women. Most speakers are of the same age group, 18-25 in 1982 and 53-60 in 2017, and socioeconomic status (middle class); four speakers, parents of the younger speakers, were in their late 40's to early 50's in 1982, and hence in their 80's in 2017; 14 of the 20 speakers completed their *Abitur*, the 'German college preparatory exam'. As mentioned above, in 1982 both communities exhibited many, dense, multiplex social relationships between family and close friends, whereas by 2017, community ties among the members had weakened and social connections become considerably more dispersed, particularly in Stuttgart.

Transcriptions were completed in ELAN (Wittenburg et al. 2006) by native German speakers, students at the University of Tübingen. A standard orthography was developed for easily and distinctly transcribing relevant Swabian dialect forms. All transcripts were verified by the principal investigator (Beaman) to ensure that standards were followed and to neutralise transcriber bias. All words (delineated by punctuation marks or blanks, Hay 2018) were extracted, and forms were automatically identified as Swabian, colloquial or standard using a bespoke Swabian-German Lexicon (SGL), which contains over 10,000 dialect and standard variants. Swabian-specific forms were tagged with a code indicating one of 32 linguistic variables under investigation (see Table 7 in the Supplementary Materials[4] for a description of the variables). Colloquial forms were identified as any form differing from the standard German form. For example, with the verb *haben* 'to have', *habe* is identified as the Standard form, *hab* as the colloquial variant (with the reduction of the final 'e'), and *han* as the Swabian variant (an irregular verb in the dialect). Because we are primarily interested in vocabulary growth and attrition across the lifespan, we first group the colloquial and Swabian-specific forms together (henceforth called, "dialect") in order to contrast them with the standard German forms. Subsequently, we zoom in on an analysis of word frequency for the 32 identified Swabian-specific features versus their standard German counterparts.

For the lexical richness analysis, the corpus was divided into two subsets, dialect words (n=22,401 in 1982 and n=20,795 in 2017) and standard words (n=50,149 in 1982 and n=69,619 in 2017). Dialect words made up less than a third (30.9%) of the speakers' active vocabularies in 1982, dropping to less than a quarter (23.0%) in 2017. For the linguistic feature analysis, the

[4] Supplemental Materials can be found at: https://osf.io/nhjxk/

corpus comprises two groups, Swabian variants (n=13,876 in 1982 and n=13,839 in 2017) and standard language variants (n=12,118 in 1982 and n=23,537 in 2017). Swabian variants made up more than half (53.4%) of the speakers' lexical choice in 1982, dropping to just over a third (37.0%) in 2017. As our results will later show, this level of dialect attrition stands in stark contrast with the exceptional growth of the standard vocabulary between 1982 and 2017: speakers used 15.8% more standard vocabulary in 2017 than in 1982, even when talking about the same general topics in both years.

2.2. Social Predictors

Five social factors were considered in this study: (1) two recording years, i.e., 1982 and 2017; (2) two speech communities, i.e., Stuttgart and Schwäbisch Gmünd; (3) two genders, i.e., male and female (as self-reported by the informants)[5]; (4) two education levels, i.e., with *Abitur* 'German college preparatory exam' and without; and (5) a composite index to assess speakers' level of identification and belonging to the Swabian community (see Table 1 for a summary of the speakers and social predictors). The Swabian Orientation Index (SOI),was adapted from Hoffman and Walker's (2010) Ethnic Orientation (EO) model based on speakers' answers to 16 questions posed during the interview covering their attitudes to the Swabian culture and language, knowledge of Swabian icons and markers, participation in Swabian events, and the nature of linguistic interactions with Swabian and non-Swabian friends and family (see Table 5 in the Supplementary Materials as well as Beaman 2020 for details). Speakers' answers were evaluated on a five-point scale, from one for the lowest to five for the highest Swabian orientation, and averaged to create an overall score for each speaker in each year.

[5] Tests on gender differences between dialect and standard language use were not significant, and hence have been eliminated from further discussion.

				1982		2017	
Pseudonym	Community	Sex	Abitur	Age	SOI	Age	SOI
Alfried	Schwäbisch Gmünd	M	Yes	23	4.5	58	4.2
Angela	Schwäbisch Gmünd	W	Yes	18	4.5	52	4.2
Anneliese	Schwäbisch Gmünd	W	Yes	21	3.5	56	3.6
Berdine	Schwäbisch Gmünd	W	Yes	21	3.9	56	3.3
Elke	Schwäbisch Gmünd	W	No	22	4.2	57	4.3
Herbert	Schwäbisch Gmünd	M	No	51	4.2	86	4.2
Jurgen	Schwäbisch Gmünd	M	Yes	19	3.8	55	3.3
Louise	Schwäbisch Gmünd	W	No	53	4.3	88	4.0
Markus	Schwäbisch Gmünd	M	Yes	22	4.3	56	2.8
Rachael	Schwäbisch Gmünd	W	No	47	4.4	83	4.3
Rupert	Schwäbisch Gmünd	M	Yes	23	4.0	58	2.6
Siegfried	Schwäbisch Gmünd	M	Yes	21	4.2	57	4.8
Theo	Schwäbisch Gmünd	M	Yes	18	4.0	53	3.7
Bertha	Stuttgart	W	No	18	3.6	53	3.3
Egbert	Stuttgart	M	Yes	24	4.0	59	3.6
Ema	Stuttgart	W	No	48	4.2	83	4.2
Helmut	Stuttgart	M	Yes	22	3.3	57	2.0
Manni	Stuttgart	M	Yes	23	3.7	59	2.8
Pepin	Stuttgart	M	Yes	25	3.4	60	3.8
Ricarda	Stuttgart	W	Yes	18	3.5	53	2.1
Overall Average				27	4.0	62	3.6

Table 1. Socio-demographic statistics and composite indices for 20 Swabian panel speakers in 1982 and 2017 (sorted alphabetically by speaker pseudonym).

2.3. Types and Tokens

We use WORD TYPE to refer to any unique word, a string of letters delineated by spaces or punctuation marks, and WORD TOKEN to refer to any instance of a specific WORD TYPE that occurs or reoccurs in the transcript regardless of its identity. For each transcript, TEXT LENGTH is measured by the number of WORD TOKENS, while VOCABULARY SIZE is measured by the number of WORD TYPES.[6] In 1982, the 20 interviews consisted of 17,707 TYPES and 72,560 TOKENS, and in 2017, the 20 transcripts contained 17,134 TYPES and 90,414 TOKENS. Of the roughly 17,000 WORD TYPES in each recording year, more than half (11,688 in 1982 and 11,337 in 2017) occurred only once, emphasising that many of the words speakers use are indeed quite rare. In

[6] No lemmatisation was carried out; thus, all inflected forms of verbs, e.g., *ich gehe* 'I go', *du gehst* 'you go', *sie geht* 'she goes', and all declined forms of nouns, e.g., *Buch* 'book', *Bücher* 'books', are counted as separate lemmas. Both content words (e.g., nouns, verbs, adjectives) and function words (e.g., determiners, prepositions) were included.

our corpus, however, because the same interview questions are used for all speakers and across both time periods, there is considerable overlap in the topics spoken about, e.g., hobbies, favourite books and films, making *Spätzle* 'Swabian egg noodles' and *Maultaschen* 'Swabian ravioli', and local activities and festivals.

2.4. Vocabulary Growth

The most straightforward measure for investigating differences in word use between texts is the size of the vocabulary (Baayen 2001). However, vocabulary size is dependent on text length, which, for the present study, is the length of the interview. Quite naturally, the longer the interview, the greater the opportunity for the speaker to utter a new word. Simple ways to sidestep this problem are to either base the analysis on a comparison of texts that are the same length or to plot interpolated VOCABULARY GROWTH CURVES side-by-side for texts of differing lengths (Baayen 2001, 2008). Due to the nature of our spontaneously spoken sociolinguistic interviews, we chose the latter approach. VOCABULARY GROWTH CURVES are projected by counting the number of TOKENS within equally spaced measurement points throughout the text (referred to as TOKEN TIME) and graphing the corresponding count of WORD TYPES. This curve depicts how vocabulary increases throughout the text, which is typically quite steep at first and then flattening as more and more different WORD TYPES are encountered. By plotting two VOCABULARY GROWTH CURVES side-by-side, core properties of the different dynamics between TYPES and TOKENS become available for visual inspection and statistical evaluation.

2.5. Statistical Methods

For the analysis of frequency of use as one progresses through a text or corpus, statistical methods based on the urn model (Johnson and Kotz 1977) have the disadvantage that they build on the assumption that words are used independently in text. As shown by Baayen (1996), topical cohesion in discourse can lead to a substantial divergence between model prediction and actual vocabulary development. In the present study, we therefore opted for using a randomisation-based method. To avoid artefactually enhancing vocabulary growth estimates that would arise by within-text randomisation and its concomitant destruction of topical structure, we opted for randomising the order of complete interviews. This is a natural choice, as the set of interviews does not have any intrinsic order, and is not governed by an overall cohesive narrative. For a given analysis, we permuted the sequence of entire interviews 50 times. For each

of the 50 permutations, we calculated the vocabulary size at ten equally-spaced measurement points, called TEXT CHUNKS (due to the varying lengths of the interviews, we used 100 TEXT CHUNKS for dialect and 200 for the standard language). For each TEXT CHUNK, we applied the Wilcoxon test to evaluate whether vocabulary sizes at a given TOKEN TIME differed significantly between 1982 and 2017. We also added outer polygons to the permutation-based vocabulary sizes to provide non-parametric confidence intervals indicating the uncertainty regarding vocabulary size. The following section presents our analysis and the results.

3. Analysis and Results

Our investigation into lexical frequency effects in dialect usage in Swabian cover four areas: vocabulary growth across the lifespan, Swabian orientation and lexical choice, individual patterns of lexical change, and changes in frequency of use of standard and dialect variants across the lifespan.

3.1. Vocabulary Growth

Recall that our central hypothesis proposes that, rather than lose dialect, speakers actually gain substantial numbers of new standard words as the range of their experience grows over the course of their lifetime. Figure 2 depicts the VOCABULARY GROWTH CURVES for our 20 speakers for the two time periods. Dialect vocabulary growth is pictured on the left and standard vocabulary growth on the right; VOCABULARY SIZE (in TYPES) is shown on the vertical axis and TEXT LENGTH (in TOKENS) on the horizontal axis. Grey illustrates speakers' vocabulary growth curve in 1982, and black portrays their growth curve in 2017. The results of the randomisation process are displayed via a polygon that surrounds the outer boundary (technically, the convex hull) and encircles all of the points. The dots represent vocabulary sizes for the 50 interview permutations. The asterisks ("*") at the top of each plot signify a significant difference in vocabulary size between 1982 and 2017 at the corresponding TEXT CHUNK according to a Wilcoxon test ($p<0.0001$).

From the left panel, we see a significant growth in dialect vocabulary from the first text chunk. The overlapping polygons reveal that there has been relatively little change in the extent to which speakers use dialect words over the 35-year timespan: speakers appear to use around a hundred fewer dialect WORD TYPES in 2017 than they did in 1982. In contrast, on the right panel, the larger black polygon shows that speakers have considerably enriched their standard language

vocabulary by 2017 (compared with the grey polygon for 1982). Their conversations made use of some 3,000 more WORD TYPEs in 2017 than in 1982. These findings provide support for our hypothesis that, rather than using fewer dialect forms, in fact, speakers actually gained an immense amount of additional lexical knowledge that is not dialect, making it appear as if dialect forms have been lost. These results replicate numerous other studies that show vocabulary size increases with age (Keuleers et al. 2015, McCabe et al. 2010, Park et al. 2002). Keuleers et al. (2015:1685) claim that "age is by far the most important variable in predicting vocabulary size…. every day lived represents an opportunity for acquisition of vocabulary and that existing vocabulary is not forgotten." As the *Diamond Model* predicts, the knowledge gained through additional experience is manifested in the standard language rather than in the dialect.

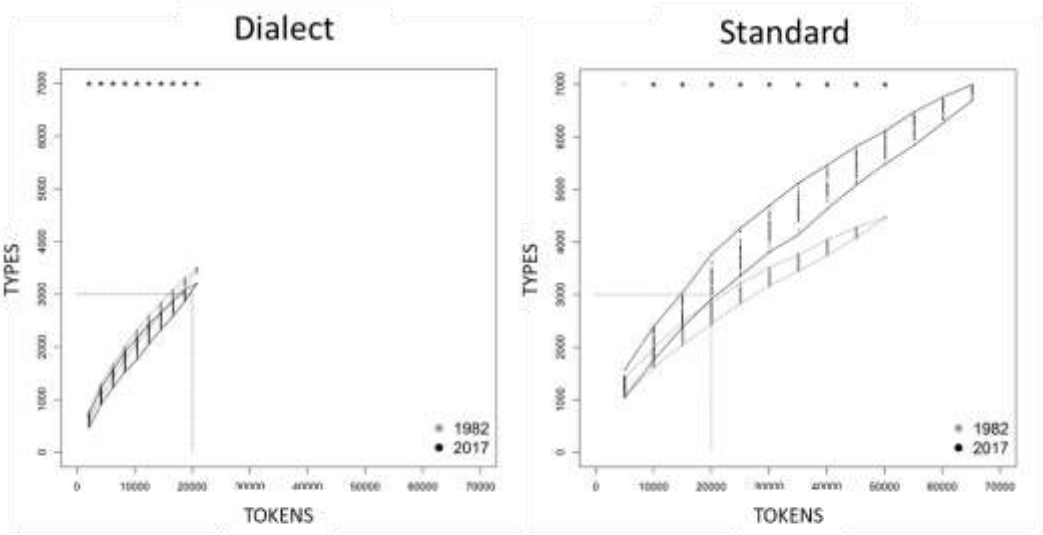

Figure 2. Vocabulary growth curves (dialect and standard) for 20 Swabian panel speakers over a 35-year time span, using 50 permutations of interview orders, for ten equally-spaced measurement points. Grey dots represent 1982 vocabulary and black dots 2017. Asterisks ("*") at the top of each plot indicate significant differences between TEXT CHUNKS based on a Wilcoxon test ($p<0.0001$).

It is interesting to note that the dialect vocabularies in 1982 and 2017 (left panel) are quite similar, which can be observed in how the polygons overlap for most of the trajectory. The two *active* vocabularies, i.e., the counts of different words used by the speakers in the interviews, only begin to disassociate about three quarters into the curve and are not entirely disassociated until the last interval. However, for the standard *active* vocabulary (right panel), the two trajectories disassociate much earlier, almost from the beginning, signifying that the standard language vocabularies in 1982 and 2017 are considerably more dissimilar. This difference can be explained with the *Diamond Model* and the premise that the domains and contexts in which

dialect is spoken have changed little over the years, whereas the spheres and settings in which the standard language is encountered are vast and multifarious. In 1982 most of the speakers were students in their 20's at the university or starting their first jobs, with naturally quite limited life experiences. As they completed their education, travelled, moved away from home, entered the workforce, and made new friends, they encountered novel and diverse experiences, experiences it appears that strengthened their standard language.

Figure 3 shows similar VOCABULARY GROWTH CURVES by community, Stuttgart on the top and Schwäbisch Gmünd on the bottom, dialect on the left and the standard language on the right. As we would expect, more dialect is spoken in the semi-rural community of Schwäbisch Gmünd than in the urban centre of Stuttgart, in fact, almost double: there are close to 3,000 dialect WORD TYPES in our sample from Schwäbisch Gmünd and only 1,500 dialect WORD TYPES in our sample from Stuttgart. We also note that speakers' *active* dialect vocabulary has declined somewhat in Stuttgart between 1982 and 2017 (by around 500 TYPES), yet remains more constant over the 35 years in Schwäbisch Gmünd (a difference of about 100 TYPES, yet significant).

The right panels of Figure 3 establish quite clearly that speakers' *active* standard language vocabulary has expanded substantially over the 35 years in both communities. Looking at a fixed TEXT LENGTH, say 20,000 TOKENS and 3,000 TYPES about 30 minutes into the interview (indicated by the dotted box in the lower left of each plot marking this fixed text length), we see that speakers in both communities actively use a similar number of standard forms; however, active dialect usage is considerably greater for speakers in Schwäbisch Gmünd than in Stuttgart. In addition, the lower panels in Figure 3 show that speakers from Schwäbisch Gmünd are considerably more chatty than those from Stuttgart: they produce more TOKENS (ca. 40,000 standard TOKENS and 15,000 dialect TOKENS in Schwäbisch Gmünd versus ca. 30,000 standard and 7,000 dialect TOKENS in Stuttgart) and more WORD TYPES (ca. 4,800 standard and 2,800 dialect TYPES in Schwäbisch Gmünd versus ca. 4,500 standard and 1,200 dialect TYPES in Stuttgart). Based on our ethnographic investigations of the speakers in these communities, we know that people from Schwäbisch Gmünd place a high value on their dialect, which is strengthened in the social setting via intense and frequent communication with friends and family. In the urban centre of Stuttgart, social connections are weaker and looser. Time appears to be of the essence; hence, communication is briefer and to the point.

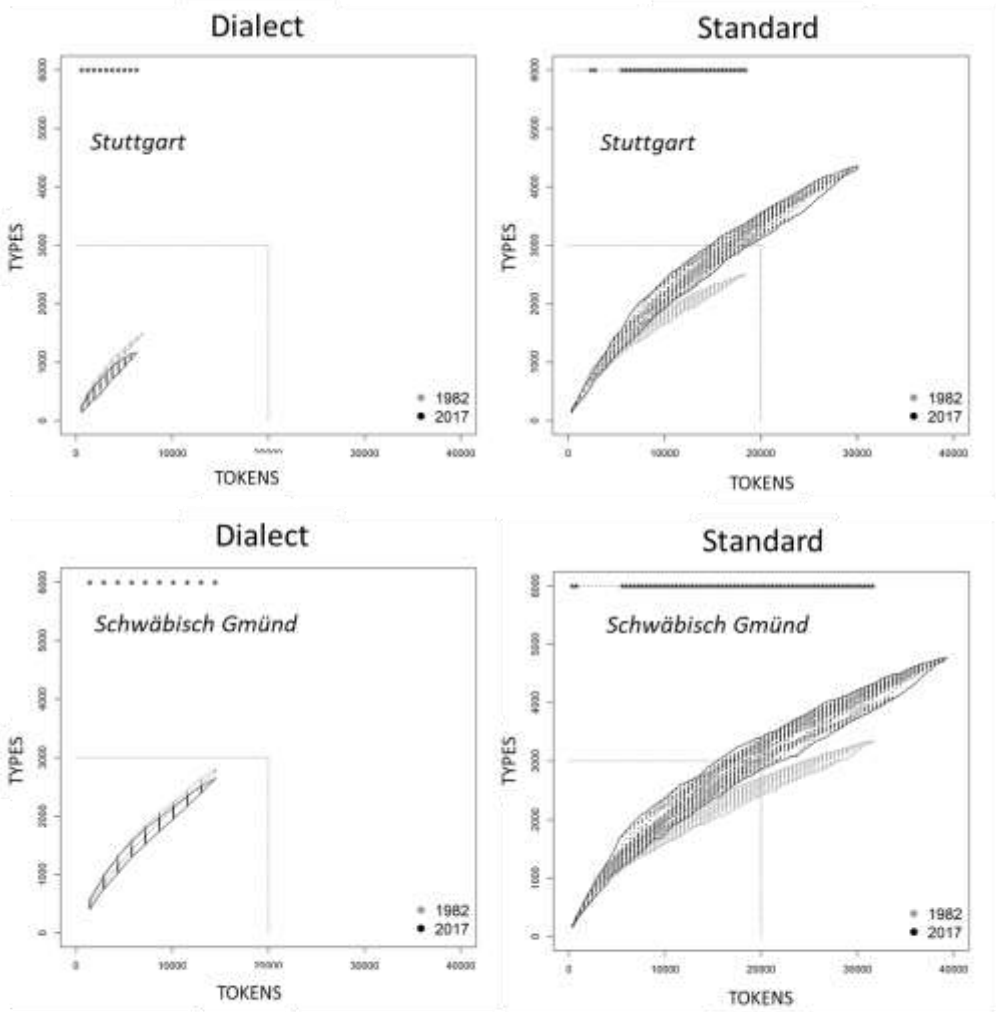

Figure 3. Vocabulary growth curves by community (Stuttgart and Schwäbisch Gmünd) for 20 Swabian panel speakers (7 from Stuttgart, 13 from Schwäbisch Gmünd) over a 35-year time span, using 50 permutations of interview orders, for ten equally-spaced measurement points. Grey dots represent 1982 vocabulary and black dots 2017. Asterisks ("*") at the top of each plot indicate significant differences between TEXT CHUNKS based on a Wilcoxon test ($p<0.0001$).

The following citation from Angela[7] in 2017 illustrates the fact that speakers in Schwäbisch Gmünd manifest a strong orientation to Swabia, with the dialect providing a conduit for bonding with the people around them:

ich bin ein sehr kommunikativer Mensch	'I am a very communicative person'
ich schwätz gern	'I like chit-chatting'
de Schwertkampf vom meinr Kinder	'my children's sword-fighting classes'
da bin i mit einige Lait befraindet	'I have some friends there'
mr rufet uns au mal ã	'sometimes we call each other'
oder wenn ôiner e Sorge hat	'or if one person has a concern'

[7] All names have been replaced by pseudonyms in order to protect the speakers' identities and maintain their privacy.

dann ruft er de andere ã	'then he calls the others'
un mã kã des dann bespreche	'and then you can talk about it'
i bin au gern mit dene zsamme	'I like being together with them'
ôifach so zum schwätze	'simply to chit-chat'

Figure 4 presents a third perspective on the lexical growth picture in Swabia by exploring the speakers' VOCABULARY GROWTH RATE by level of education. Speakers who did not complete an *Abitur*, the 'German college preparatory exam', are shown on the top and speakers with an *Abitur* on the bottom. From the left panels, there is little change in the use of dialect based on educational attainment: both groups of speakers have retained most of their dialect words over the years, a finding which is significant through the entire text. However, on the right panels, there is considerable growth in the active standard language vocabulary for both groups of speakers, those with and without an *Abitur*.

While both the high and less highly educated groups have increased their standard vocabulary over the 35 years, we see a striking surge in 2017 for speakers with higher education (lower right panel). Knowledge with its accompanying vocabulary naturally increases with education and diversification of experience, both of which have introduced new and novel ideas into speakers' vocabularies. Increased standard language vocabulary reflects the contact and involvement that the more educated group has with the standard language (Hart and Risley 1995). The more highly educated speakers are also more loquacious in the standard language, presumably because they have encountered a broader range of experiences in the standard language, which in turn offers a broader vocabulary for expressing their thoughts and experiences.

Our findings confirm the central hypothesis of this research that rather than lose an extensive amount of dialect vocabulary as they age, speakers actually acquire substantially more non-dialect forms through the myriad experiences of their lives (i.e., the expanding diamond). Further, this growth of non-dialect has a cumulative and confounding effect when measuring vocabulary use which is influenced by aging (earlier versus later recordings), community (urban versus semi-rural), level of education (college preparation or not), and, as we will see in the next section, by their orientation to the local community.

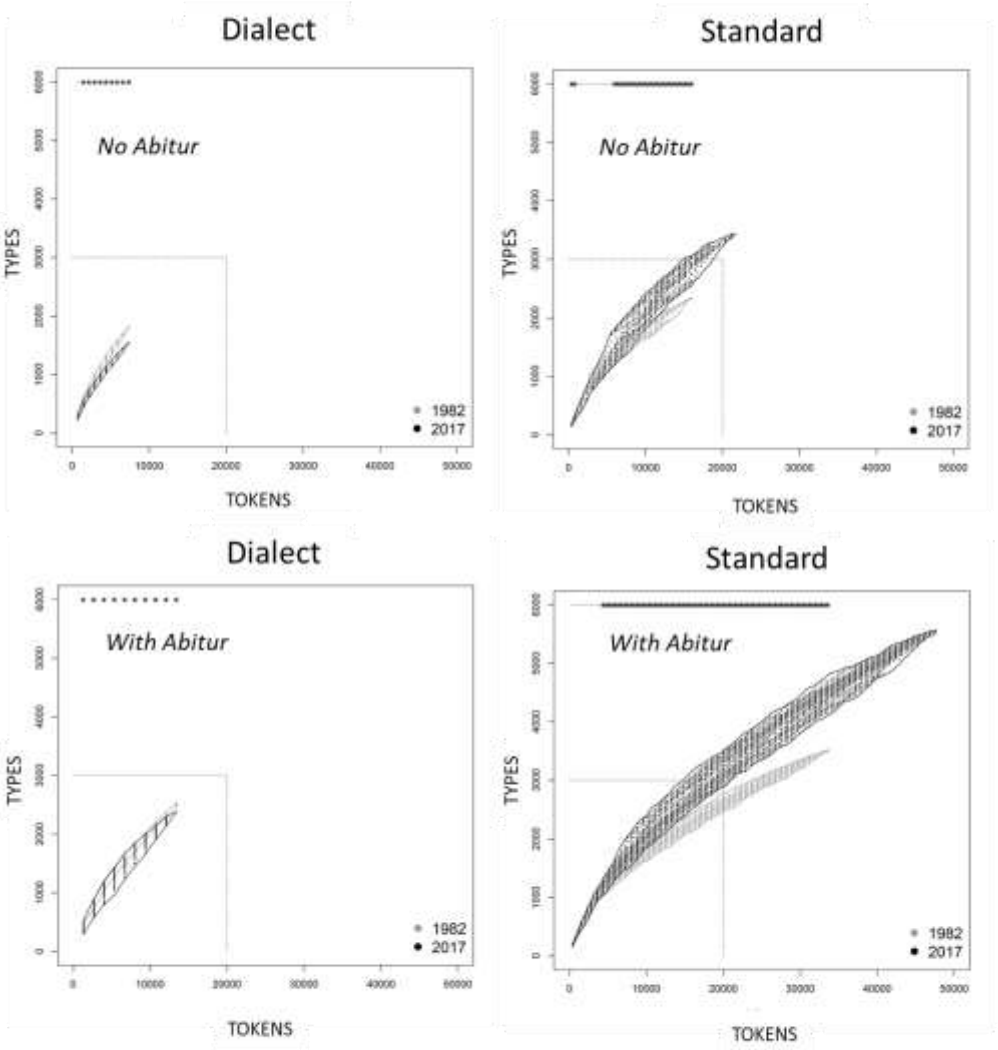

Figure 4. Vocabulary growth curves by educational achievement for 20 Swabian panel speakers (14 with *Abitur*, 'German college preparatory exam', and 6 without *Abitur*) over a 35-year time span, using 50 permutations of interview orders, for ten equally-spaced measurement points. Grey dots represent 1982 vocabulary and black dots 2017. Asterisks ("*") at the top of each plot indicate significant differences between TEXT CHUNKS based on a Wilcoxon test ($p<0.0001$).

3.2. Orientation and Lexical Choice

Many studies have shown that speakers' linguistic choices are influenced by their orientation or personal affinity towards the dialect or the standard language (Cheshire et al. 2008, Coupland 2007, Eckert 1989, Hoffman and Walker 2010, Horvath and Sankoff 1987, Labov 1963, 1966, Schilling-Estes 2004), which can, of course, change across the lifespan. Figure 5 depicts the changing prominence of Swabian orientation over the 35 years in the two communities, exposing two critical effects of Swabian orientation on society. The left panel brings to light the powerful role that the Swabian orientation played in 1982 (average 4.0); by contrast, by 2017, Swabian

orientation scores have fallen (average 3.6) and now stretch out over a much broader range. The right panel shows that Stuttgart has a noticeably lower overall Swabian orientation index (3.5) than Schwäbisch Gmünd (4.3), which is not unexpected: the mid-sized, semi-rural town of Schwäbisch Gmünd is a much smaller, tighter-knit community than the vast urban metropolis of Stuttgart. Figure 5 makes it evident that the role of Swabian identity has changed dramatically over the years, especially for Stuttgart.

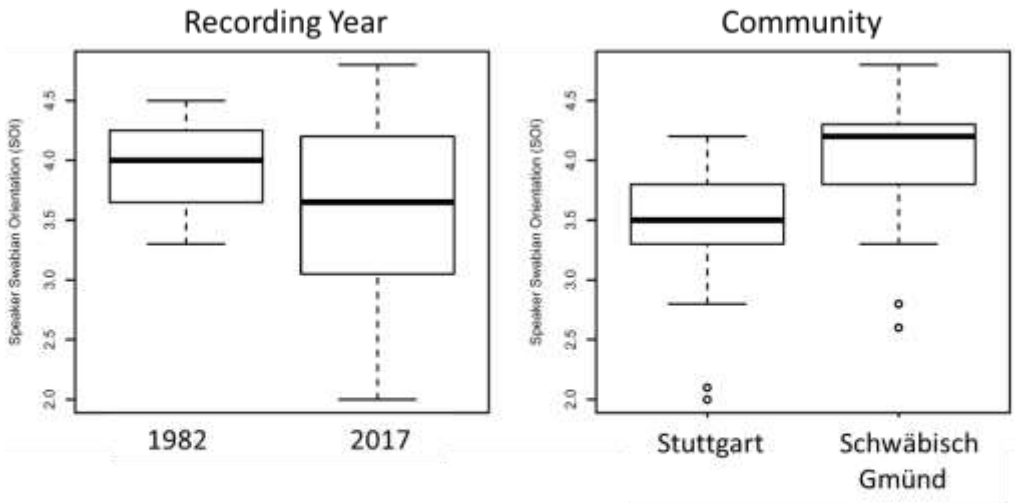

Figure 5. Swabian Orientation Index (SOI) by recording year (1982 and 2017) and community (Stuttgart and Schwäbisch Gmünd) for 20 panel speakers. SOI is calculated from speakers' answers to 16 questions posed in the interview, evaluated on a five-point scale, from 1 for the lowest to 5 for the highest Swabian orientation.

We now turn to the effect that Swabian orientation has on individual speakers and their propensity to use dialect or standard vocabulary. Figure 6 plots active dialect vocabulary size (number of WORD TYPES at the fourth TEXT CHUNK, about 20 minutes into the interview) and Swabian orientation for the 20 speakers, 1982 on the left and 2017 on the right. The Stuttgart speakers are denoted by open circles and the Schwäbisch Gmünd speakers by filled squares. Our first observation is that dialect vocabularies in Stuttgart and Schwäbisch Gmünd were much more homogeneous in 1982 than they have become in 2017. By 2017, for many speakers, Swabian orientation has declined concomitantly with dialect vocabulary (demonstrated in the right panel by the dots spreading down and to the left). Still, we see a number of speakers, particularly from Schwäbisch Gmünd, who have retained their high Swabian orientation and dialect vocabulary (illustrated by more black squares clustering in the upper right corner of the scatterplots). The trend is clear: the higher the Swabian orientation, the larger the active dialect

vocabulary; and conversely, the lower the speakers' orientation scores, the smaller the active dialect vocabulary. These findings are confirmed with a linear regression analysis (see Table 2) which shows orientation not to be a significant predictor of dialect vocabulary in 1982 ($\hat{\beta}$ = 13.285, p = 0.141, Adjusted R^2 = .067), whereas it has become highly significant in 2017 ($\hat{\beta}$ = 27.82, p = 0.001, Adjusted R^2 = .455). In 1982, Swabians simply spoke more dialect independent of their individual orientation. We also see that orientation had no effect on the size of the standard vocabulary in 1982 ($\hat{\beta}$ = -9.802, p = 0.768, Adjusted R^2 = -.050), however a significant effect has emerged in 2017 $\hat{\beta}$ = -49.56, p = 0.040, Adjusted R^2 = .170).

Figure 6 corroborates the findings from Figure 2 and confirms our second hypothesis that the extent of dialect loss and standard language gain over the lifespan is heavily influenced by the speakers' orientation to Swabia. According to the *Diamond Model*, the expanding and contracting of words in the standard language as individuals age, like the widening and narrowing angles of the diamond, vary according to the speakers' social setting and their individual orientation to the local community, language, and culture. This individualistic approach raises a question: which speakers have changed their vocabulary the most, and what are the reasons behind this change? To offer some insight into this, the next section examines the lifespan change patterns of individual speakers.

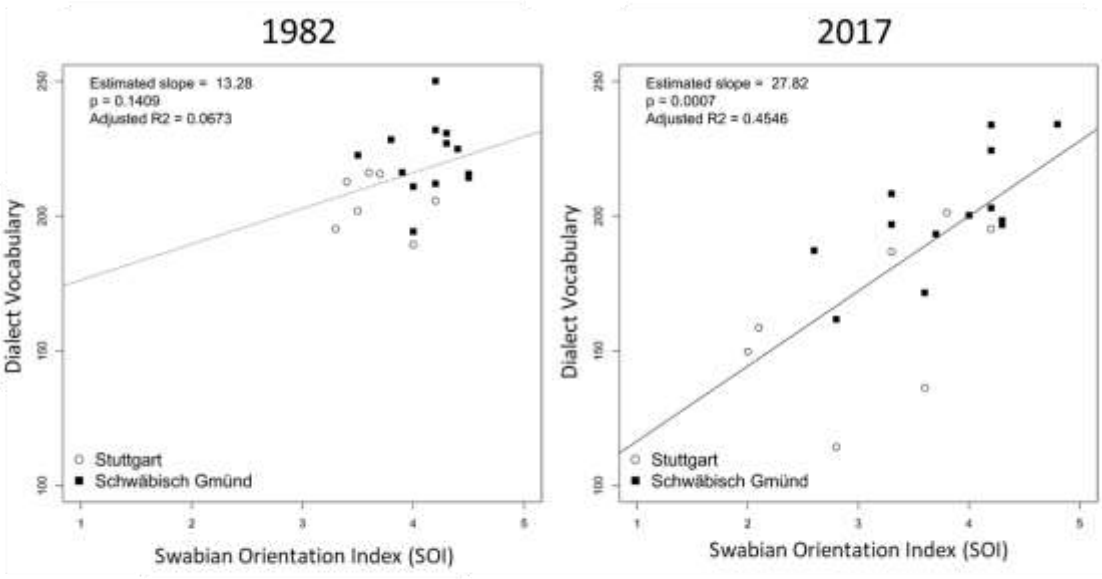

Figure 6. Dialect vocabulary size (at the fourth TEXT CHUNK, about 20 minutes into the interview) and Swabian Orientation Index (SOI) for two recording years (1982 and 2017) and two communities (Stuttgart and Schwäbisch Gmünd).

```
1982 SOI/DIALECT VOCABULARY CHANGE:
lm(Sw_vocab1982~soi1982, data=dfr3.SWG)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  162.994     34.432   4.734 0.000166 ***
soi1982       13.285      8.626   1.540 0.140940
Multiple R-squared:  0.1164,    Adjusted R-squared:  0.06734
```

```
2017 SOI/DIALECT VOCABULARY CHANGE:
lm(Sw_vocab2017~soi2017, data=dfr3.SWG)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    88.75      24.64   3.602 0.002038 **
soi2017        27.82       6.78   4.104 0.000667 ***
Multiple R-squared:  0.4833,    Adjusted R-squared:  0.4546
```

```
1982 SOI/STANDARD VOCABULARY CHANGE:
lm(St_vocab1982~soi1982, data=dfr3.SWG)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  708.441    130.907   5.412 3.84e-05 ***
soi1982       -9.802     32.796  -0.299    0.768
Multiple R-squared:  0.004938, Adjusted R-squared:  -0.05034
```

```
2017 SOI/STANDARD VOCABULARY CHANGE:
lm(St_vocab2017~soi2017, data=dfr3.SWG)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   845.31      81.37  10.389 4.95e-09 ***
soi2017       -49.56      22.39  -2.214     0.04 *
Multiple R-squared:  0.214,     Adjusted R-squared:  0.1703
```

Table 2. Summary of linear regression models by year for dialect and standard vocabulary size (at the fourth TEXT CHUNK, about 20 minutes into the interview) as a function of Swabian orientation.

3.3. Individual Patterns of Change

Individual patterns of linguistic change have been shown to complement and enhance insights gained from overall community change (Sankoff 2006, Wagner and Buchstaller 2017). To assess these individual effects, we modelled dialect and standard vocabulary differences as a function of Swabian orientation and speaker age[8] in 2017, using generalised additive mixed models (GAMMs) (*gam* function in the *mgcv* R package, version 1.8-27) (see Table 3). The results corroborate the preceding findings that Swabian orientation is a significant factor for both dialect and standard language vocabulary change ($\hat{\beta}$ = 23.59 for dialect and $\hat{\beta}$ = -42.61 for standard, p = 0.0057), while speaker age is significant only for the standard language ($\hat{\beta}$ = -11.99, p = 0.0149), showing only marginal significance for dialect vocabulary change ($\hat{\beta}$ = -4.713, p = 0.0640).

[8] The four older speakers (82 to 88 years old) were excluded from the age analysis due to skewing from the large age gap between the two generations; this resulted in 16 speakers between 52 and 60 years of age.

```
DIALECT VOCABULARY CHANGE:
Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  155.204    131.932   1.176  0.26053
soi2017       23.593      7.135   3.307  0.00567 **
age2017       -4.713      2.329  -2.024  0.06404 .

R-sq.(adj) =  0.452   Deviance explained = 52.5%
GCV = 588.58  Scale est. = 478.22    n = 16
```

```
STANDARD VOCABULARY CHANGE:
Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  824.184    242.176   3.403  0.00471 **
soi2017      -42.613     13.097  -3.254  0.00628 **
age2017      -11.991      4.274  -2.805  0.01487 *

R-sq.(adj) =  0.538   Deviance explained = 59.9%
GCV = 1983.2  Scale est. = 1611.3    n = 16
```

Table 3. Summary of generalised additive models *(gam* function in *mgcv* R package, version 1.8-28) for dialect and standard vocabulary size as a function of Swabian orientation and speaker age in 2017 (n = 16 speakers aged 52 to 60; four speakers between 82 and 88 were excluded due to the large age gap).

In order to visualise individual speaker vocabulary change across the lifespan as a function of Swabian orientation, Figure 7 (upper panels) plots the 20 speakers according to their vocabulary change between 1982 and 2017 and their Swabian orientation score in 2017. The left panel depicts dialect vocabulary change, and the right panel portrays standard vocabulary change (vocabulary size is calculated at the fourth TEXT CHUNK, 20 minutes into the interview). On the vertical axis, vocabulary gain is shown by positive numbers and loss by negative numbers. The upper left plot reveals two speakers, Angela and Siegfried, who actively use more dialect in 2017 than they did in 1982 (both have Swabian orientation scores greater than 4). Theo is on the cusp, which means that he continues to use the same amount of dialect in 2017 as in 1982. Toward the bottom of this plot are Manni and Markus, the two speakers who have lost the most active dialect vocabulary over the years (both with Swabian orientation scores below 3). The majority of speakers cluster between 0 and -30 in their dialect vocabulary loss, indicating a modest amount of loss over the years.

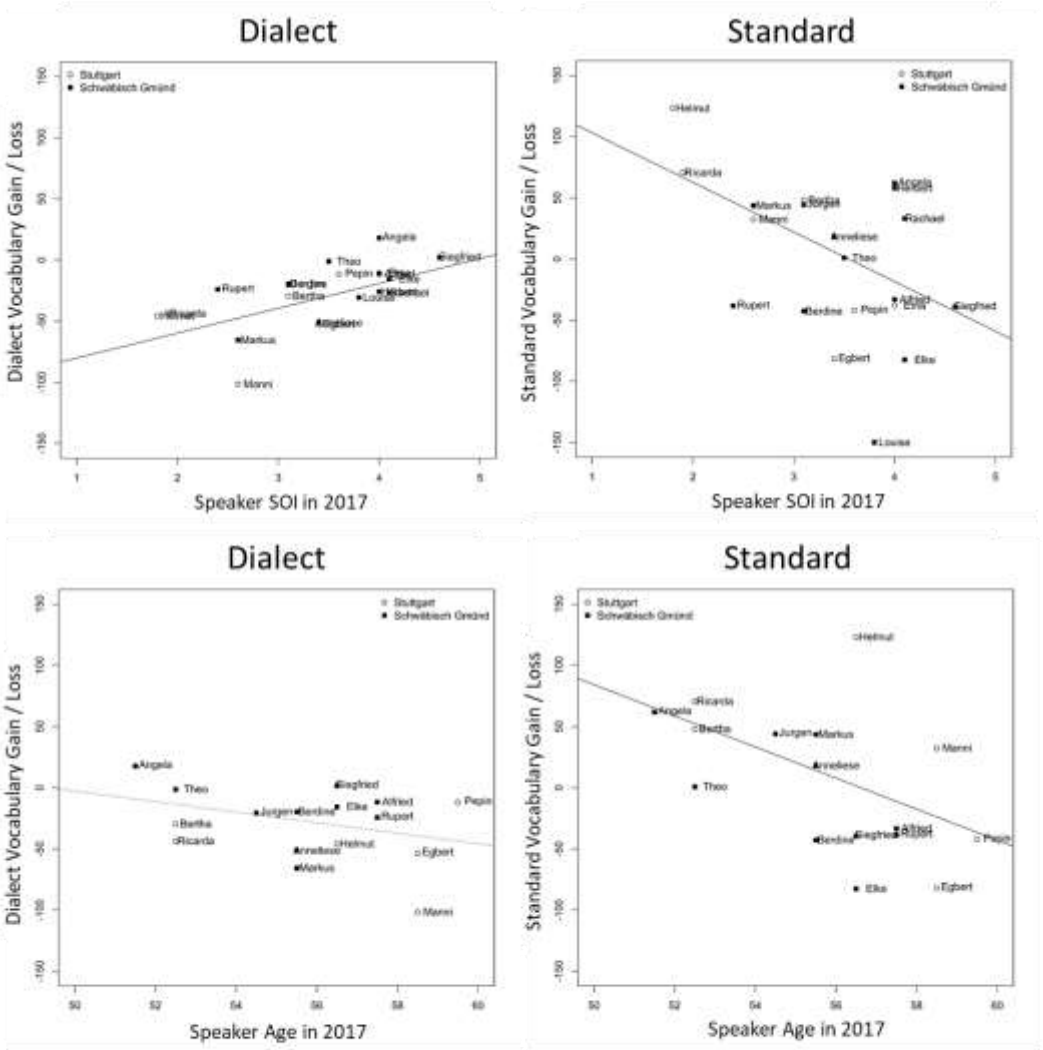

Figure 7. Dialect and standard language vocabulary change between 1982 and 2017 (calculated at the fourth TEXT CHUNK, 20 minutes into the interview) as a function of speakers' Swabian Orientation Index (SOI) (upper panels) and speakers aged 52 to 60 (lower panels) in 2017. Positive numbers show vocabulary gain, negative numbers vocabulary loss. Regression lines based on models presented in Table 3.

Consistent with our previous findings, the upper right panel of Figure 7 confirms that many speakers have gained considerable standard vocabulary over the years. Yet, a decrease in active standard vocabulary is clearly visible for a large number of high SOI speakers, establishing that Swabian orientation is also a critical predictor of standard vocabulary change. As speakers' Swabian orientation scores increase, their standard vocabularies decrease. The upper right plot reveals that Louise, the oldest speaker in our sample (88 years old), has the greatest standard vocabulary loss (-150 WORD TYPES) yet an average amount of dialect loss (-30 WORD TYPES), suggesting a prototypical example of standard vocabulary contraction due to aging as predicted by the *Diamond Model*.

The lower panels in Figure 7 visualise dialect and standard vocabulary change by speaker age for the 16 speakers between 52 to 60 years old. The lower right panel illuminates the *Diamond Model* in action*:* there is a steady, but significant decline in active standard language vocabulary as individuals age, an effect that is not readily apparent with the dialect vocabulary (lower left panel). While it is tempting to contribute this change in standard vocabulary to cognitive decline, our speakers in 2017 are pre-retirement age, and the literature has shown retirement itself to be the primary risk factor for cognitive decline (Dufouil et al. 2014, Nikolov and Adelman 2019, Xue et al. 2018). In addition, if this loss of standard language vocabulary were related to cognitive decline, we would expect to see the same effect in the dialect, which is questionable. The *Diamond Model* presumes that as speakers age and get closer to retirement, their social interactions slowly narrow, and they no longer use the standard language with as much variety as they did earlier (see Table 3 for regression coefficients and adjusted R-squared values).

Table 4 provides a summary of the 20 panel speakers ranked by their degree of dialect vocabulary change, illustrating a continuum that reflects Sankoff's (2006) three types of individual change: at the top, LIFESPAN CHANGE, individuals moving in the direction of the overall community change by using less dialect and more standard language forms; in the middle, SPEAKER STABILITY, individuals continuing to use a similar amount of dialect across the years; and at the bottom, RETROGRADE CHANGE, speakers moving in the opposite direction of the general community change by using more dialect in 2017 than they did in 1982.

It is interesting to see that there are more Stuttgart speakers at the top of the table, signalling greater loss of the dialect across the lifespan in the large urban centre than the semi-rural community of Schwäbisch Gmünd. The prominence of the urban-rural divide can also be seen when comparing the lifespan change of speakers with similar socio-demographics, so-called *social twins* (Nordberg and Sundgren 1998, Sankoff and Blondeau 2013). Egbert is a middle school teacher in Stuttgart, and Siegfried is a middle school teacher in Schwäbisch Gmünd; similarly, Ricarda is an elementary school teacher in Stuttgart, and Elke is an elementary teacher in Schwäbisch Gmünd. Both teachers from Stuttgart have moved to more standard language usage across their lifespan than the two teachers from Schwäbisch Gmünd, providing additional support for the dialect levelling occurring in the urban centre in contrast to the vital ongoing role that dialect retains in the Swabian countryside.

		2017				Vocabulary Change		
Pseudonym	Community	Age	Sex	Abi	SOI	Dialect	Standard	Net
Manni	Stuttgart	59	M	Yes	2.8	-101	33	-69
Markus	Gmünd	56	M	Yes	2.8	-65	44	-21
Egbert	Stuttgart	59	M	Yes	3.6	-53	-82	-135
Anneliese	Gmünd	56	W	Yes	3.6	-51	18	-33
Helmut	Stuttgart	57	M	Yes	2.0	-46	124	78
Ricarda	Stuttgart	53	W	Yes	2.1	-43	71	28
Louise	Gmünd	88	W	No	4.0	-30	-150	-180
Bertha	Stuttgart	53	W	No	3.3	-29	48	19
Rachael	Gmünd	83	W	No	4.3	-26	33	7
Herbert	Gmünd	86	M	No	4.2	-26	58	32
Rupert	Gmünd	58	M	Yes	2.6	-24	-38	-62
Jurgen	Gmünd	55	M	Yes	3.3	-20	45	24
Berdine	Gmünd	56	W	Yes	3.3	-19	-42	-62
Elke	Gmünd	57	W	No	4.3	-15	-82	-97
Pepin	Stuttgart	60	M	Yes	3.8	-11	-42	-53
Alfried	Gmünd	58	M	Yes	4.2	-11	-33	-44
Ema	Stuttgart	83	W	No	4.2	-10	-38	-48
Theo	Gmünd	53	M	Yes	3.7	-1	1	0
Siegfried	Gmünd	57	M	Yes	4.8	2	-39	-37
Angela	Gmünd	52	W	Yes	4.2	16	62	78

Table 4. Change in speaker vocabulary size across the years (1982 versus 2017), indicating three types of dialect change across the lifespan (sorted by change in dialect vocabulary size).

Table 4 brings out several key patterns of individual change across the lifespan. First, Manni and Markus, at the top of the table, have experienced the most dialect loss and gained a typical amount of standard vocabulary across their lifespan. Manni is a consulting architect for the Stuttgart airport, and Markus is a marketing executive who travels to Munich each week. Both are businessmen in their late fifties, actively participating in the linguistic market (Bourdieu 1977, Eckert 1997, Sankoff and Laberge 1978, Wagner 2012) and frequently interacting with speakers from other dialect areas, hence it is not surprising they would use fewer dialect words at this stage in their lives. Helmut, a radio moderator, also uses fewer dialect words and has gained the most standard words compared to his cohorts. Clearly his profession in the news media introduces considerable standardisation pressures and high expectations of in-group conformance (Bell 1991; Coupland 2001).

In contrast, at the bottom of the table are Angela and Siegfried who have not lost any of their dialect; with two of the highest Swabian orientation scores in 2017, these speakers illustrate the importance of considering the social setting and the personal identities of the speakers in understanding lifespan change. Angela, a medical doctor, living near Schwäbisch Gmünd and

commuting to Stuttgart for work each day, is a stark advocate for Swabian. In 1982, when asked what she thought of the Swabian language, she provocatively exclaimed, *das beschte Daitsch wo es gib!* 'the best German that there is!' Responding to the same question in 2017, she responded similarly, yet more thoughtfully:

> *Schwäbisch isch fe mi kôi Daitsch*
> 'Swabian is for me not German'
>
> *sondern des isch mei Muttersprache*
> 'rather it is my mother tongue'
>
> *in so fern isch se zentral fe mich*
> 'in that respect it is crucial for me'

Close to Angela in his sheer love for the Swabian dialect is Siegfried, a middle school teacher who has spent his entire life in Schwäbisch Gmünd. In 2017, he expressed his sorrow about the loss of Swabian, remarking:

> *viele Schwââbe erziehet ihre Kinder jetzt als net-Schwââbe*
> 'many Swabians raise their kids now as non-Swabian'
>
> *weil se willet, dass se Hochdeutsch schwätzet*
> 'because they want them to speak standard German'
>
> *dâ kommet se an dr Uni besser zrecht ond was-wôis-i,*
> 'then they do better at the university und whatever'
>
> *dâ gheer i net dazu,*
> 'I don't belong to [that group]'
>
> *i bin, wenn du so willsch, e stolze Schwââbe*
> 'I am, if you will, a proud Swabian'
>
> *on i find es schade, dass die Sprââch verlore gâht*
> 'and I think it's a shame, that Swabian is being lost'

In contrast to the composite diagram in Figure 2, the individual analysis of vocabulary change reveals that vocabulary richness, both dialect and standard, has diminished over the 35-year timespan for some speakers and increased or stayed largely the same for others. These results underscore the importance of incorporating individual lifespan analyses into general trend

studies if critical influences that are otherwise hidden in community-wide averages are to be teased apart.

3.4. Word Frequency

We now turn to our third hypothesis which predicts that if there is some loss of dialect vocabulary as individuals age, then this loss will be observed primarily in the low frequency range. Studies have reported that high-frequency words are more resistant in giving way to the standard language than low-frequency words (Bybee 2002, Keuleers et al. 2015, Wieling et al. 2014, 2011), which leads us to expect that higher frequency dialect words should be more deeply embedded in speakers' vocabularies. Testing this hypothesis on our data, however, is not straightforward, as one-hour interviews are unlikely to capture truly low-frequency words. What we can investigate, however, are differences in the use of words in the frequency ranges from our sample. Within these ranges, it seems likely that the lower frequency dialect words are the ones that are most vulnerable to replacement by standard language equivalents.

Figure 8 presents the results of our frequency analysis of Swabian and standard words for the 20 panel speakers in 1982 and 2017. In this analysis we separated words containing specific Swabian-specific features from other dialect forms by coding 32 Swabian linguistic variables (n=63,370), creating a binary distinction between Swabian and standard German forms (see Table 7 in the Supplementary Materials for a description of the variables). We look first at the two left panels. The leftmost panel depicts word frequency for the Swabian variants, and the middle panel shows word frequency for the standard language variants (log transformed after backing off from zero by adding 1). The horizontal axes plot word frequency in 1982, and the vertical axes plot word frequency in 2017. Both plots indicate a slight non-linear trend[9] (via GAMM). Words are scattered roughly around the diagonal, with greater scatter for lower frequencies, visually revealing a large number of words in the low-frequency range. The shaded area represents a 95% confidence area for log frequency in 2017 as predicted from log word frequency in 1982.

[9] Note that a line in a log-log plane indicates a power (non-linear) relation between the two variables.

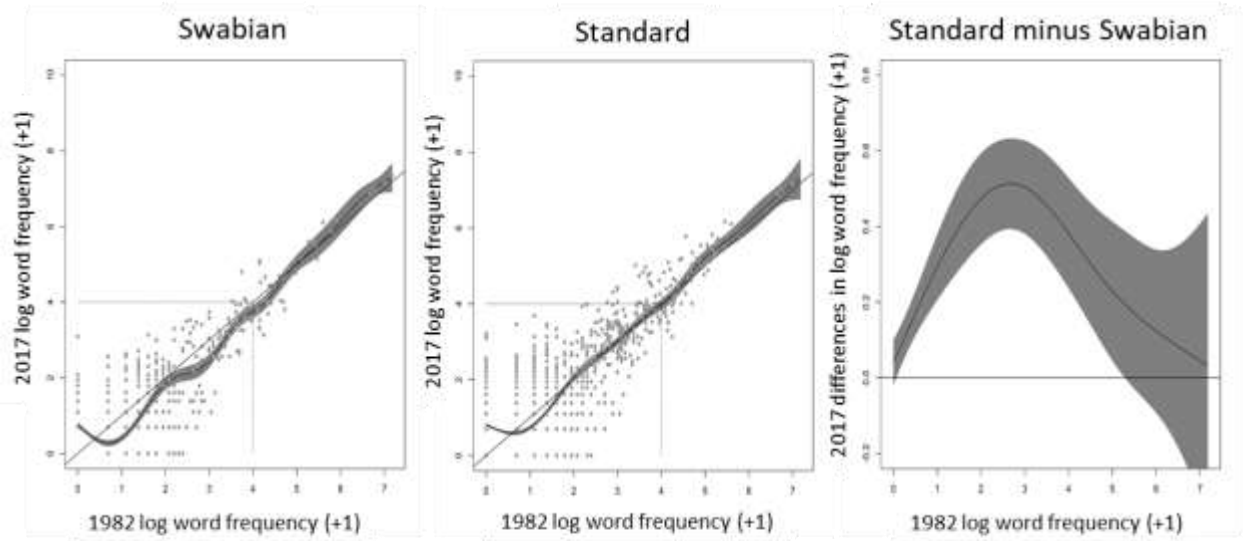

Figure 8. Lexical frequency effects between 1982 and 2017 (log transformed +1) based on 32 selected Swabian linguistic features coded for either standard or dialect variant (n=63,370). Left panel: Swabian word frequencies; middle panel: Standard language word frequencies; right panel: difference between standard and Swabian word frequencies.

Looking first at words with frequency of 1 in 1982 (i.e., log frequency = 0 on the horizontal axis), we see a number of novel words (i.e., the dots lined up along the vertical axis), both Swabian and standard, that were used in 2017 but were not used in 1982, implying that there is a repository of forms that people know but did not use in the first interview. Quite obviously, with one-hour interviews, a vast amount of knowledge is not sampled. Most important for our hypothesis, we see that Swabian forms with log frequencies less than 4 were used more frequently in 1982 than in 2017, revealed by the curvy shaded line appearing largely below the diagonal in the leftmost plot. For the standard forms, we see a slight (non-significant) effect of low-frequency which peters out after log frequency 2, demonstrating a much stronger effect of low-frequency for Swabian forms than for standard forms. No such significant difference is visible in the use of the highest frequency words, i.e., words with log frequency greater than 4, as seen by the shaded area remaining tight along the diagonal for both Swabian and standard forms.

The rightmost panel of Figure 8 depicts this difference in frequency of use between the standard and Swabian forms between 2017 and 1982 (i.e., log frequency of standard variants minus log frequency of Swabian variants). This curve clearly illustrates a change in the relationship in word frequency differences across the two recording periods, exhibiting a peak around the 1982 log frequency of 3. Specifically, the difference between the two years (y-axis) increases as the log frequency in 1982 (x-axis) increases, but only for words in the low-

frequency range (i.e., log frequency < 3), while differences in word frequency approach zero for words in the high-frequency range (i.e., log frequency > 3). In other words, there is a greater difference between the years in usage that favours the standard variants for low-frequency words than for high-frequency words. This finding supports our hypothesis that higher frequency words are more resistant in yielding to the standard language.

In sum, our frequency analysis of changes in use of standard and Swabian forms shows: (1) the large number of Swabian words used in 2017 but not in 1982 imply there is a repository of Swabian forms available to the speaker that were simply not actively sampled in our one-hour recordings; (2) low- to mid-frequency Swabian forms (log frequencies < 4) were used more often in 1982 than in 2017, indicating some loss; however, high-frequency Swabian words (log frequency > 4) have retained their use, indicating that the attack from the standard language is effective only in the low- to mid-frequency range; (3) low- and mid-frequency words from the standard language were re-used more often in 2017, indicating the advance of the standard language; and, (4) Swabian and standard forms not used in 1982 are used at the same rate in 2017 (i.e., there is no significant difference for frequencies of 1 (log(1)=0)), signifying the ongoing active production of Swabian forms. Although Swabian appears to be holding its own, it is still evident that the standard language is encroaching on the dialect in the low- to mid-frequency range, a finding consistent with our hypothesis.

4. Concluding Remarks

Conventional views of dialect change across the lifespan are typically seen as reflecting attrition due to the encroachment of the standard language on the native dialect. However, the findings from our investigation of 20 panel speakers of Swabian indicate that a substantial part of change in active vocabulary use across the lifespan reflects the increased influence of later acquired, non-dialect, lexical knowledge, with some minimal loss of dialect forms through attrition. Our findings demonstrate that change in vocabulary usage across the lifespan can best be explained with reference to the *Diamond Model*: in the initial phase of life-cycle development (from youth through adolescence to early adulthood), the dialect is the primary medium of communication; as speakers move through life, rather than substantial reduction of well-established dialect words, their standard language vocabularies actually expand, as they acquire a different register – the standard language – in conjunction with their exposure to the rich variety of experiences that

comprise a lifetime (confirming our first hypothesis). We have also seen that speaker vocabularies are highly influenced by their personal orientation to the language and culture of the local community: higher levels of community orientation produce lower levels of dialect loss, while lower levels of orientation promote greater gain in the standard language (substantiating our second hypothesis). Finally, we found that the encroachment of the standard language may be most active in the lower word frequency ranges, establishing that high-frequency words are more resistant change in intensity of use (supporting our third hypothesis).

Our findings underscore the importance of accounting for the dynamic properties of lexical distributions in interpreting language development across the lifespan. Ramscar and colleagues have recently shown how many changes in cognitive performance across the lifespan that are taken to reflect decline look very different once we control for the inevitable interactions that occur between lexical distributions and the nature of lifetime learning (Baayen, Tomaschek et al. 2016, Ramscar et al. 2013, 2014, 2017). Although we identified some changes in some of our speakers that are consistent with some degree of attrition in dialect use, to a large degree, our data support the more likely scenario that many of the changes in the balance between standard and dialect lexical choice observed across the lifespan reflect change in speakers' lifestyles and growth in their knowledge of the standard language in conformance with the *Diamond Model*. Although the standard language is expanding into new domains that speakers encounter over their lifetimes, the Swabian dialect, dependent on speakers' local orientation and dialect identity, is holding its traditional ground well.

References

Auer, Peter. 2005. "Europe's Sociolinguistic Unity, or: A Typology of European Dialect/Standard Constellations." In *Perspectives on Variation: Sociolinguistic, Historical, Comparative*, eds. Nicole Delbecque, Johan van der Auwera, and Dirk Geeraerts. Berlin: Mouton de Gruyter, 7–42.

———. 2007. "Mobility, Contact and Accommodation." In *The Routledge Companion to Sociolinguistics*, eds. Carmen Llamas, Louise Mullany, and Peter Stockwell. London: Routledge, 109–15.

———. 2018. "Dialect Change in Europe -- Leveling and Convergence." In *The Handbook of Dialectology*, eds. Charles Boberg, John Nerbonne, and Dominic Watt. Oxford: Wiley-Blackwell, 159–76.

———. 2020. "Dialect (Non-)Acquisition and Use by Young People of Migrant Background in

Germany.” *Journal of Multilingual and Multicultural Development*: 1–15.

Auer, Peter, Peter Baumann, and Christian Schwarz. 2011. “Vertical vs. Horizontal Change in the Traditional Dialects of Southwest Germany: A Quantitative Approach.” *Taal en Tongval* 63(1): 13–41.

Auer, Peter, and Helmut Spiekermann. 2011. “Demotisation of the Standard Variety or Destandardisation? The Changing Status of German in Late Modernity (with Special Reference to Southwestern Germany).” In *Standard Languages and Language Standards in a Changing Europe*, eds. Tore Kristiansen and Nikolas Coupland. Oslo: Novus Press, 161–76.

Baayen, R. Harald. 1996. “The Randomness Assumption in Word Frequency Statistics.” In *Research in Humanities Computing 5*, ed. G. Perissinotto. Oxford: Oxford University Press, 17–31.

———. 2001. *Word Frequency Distributions*. Dordrecht: Kluwer Academic Publishers.

———. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R (Second Edition)*. Cambridge: Cambridge University Press.

Baayen, R. Harald, Fabian Tomaschek, Susanne Gahl, and Michael Ramscar. 2016. “The Ecclesiastes Principle in Language Change. Kindle Edition.” In *The Changing English Language: Psycholinguistic Perspectives*, eds. Marianne Hundt, Sandra Mollin, and Simone E. Pfenninger. Cambridge: Cambridge University Press, 50–86.

Beaman, Karen V. 2020. “Identity and Mobility in Linguistic Change across the Lifespan: The Case of Swabian German.” In *Urban Matters. Current Approaches of International Sociolinguistic Research (Studies in Language Variation)*, eds. Arne Ziegler, Stefanie Edler, Nina Kleczkowski, and Georg Oberdorfe. Amsterdam: John Benjamins.

Beaman, Karen V., and Fabian Tomaschek. 2020. “Loss of Historical Phonetic Contrast Across the Lifespan: Articulatory, Lexical, and Social Effects on Sound Change in Swabian.” In *Language Variation and Language Change Across the Lifespan: Theoretical and Empirical Perspectives from Panel Studies*, eds. Karen V. Beaman and Isabelle Buchstaller. New York: Routledge.

Bell, Allan. 1991. *The Language of News Media*. Oxford: Blackwell Publishing.

Bourdieu, Pierre. 1977. “Economics of Linguistic Exchanges.” *Social Science Information* 16(6): 645–68.

Britain, David. 2009. “One Foot in the Grave? Dialect Death, Dialect Contact, and Dialect Birth in England.” *International Journal of the Sociology of Language* (196–197): 121–55.

———. 2010. “Supralocal Regional Dialect Levelling.” In *Language and Identities*, eds. Carmen Llamas and Dominic Watt. Edinburgh: Edinburgh University Press, 193–204.

———. 2013. “Space, Diffusion and Mobility.” In *The Handbook of Language Variation and Change*, eds. J. K. Chambers and Natalie Schilling. Malden MA: Wiley-Blackwell, 489–519.

———. 2016. “Sedentarism and Nomadism in the Sociolinguistics of Dialect.” In *Sociolinguistics: Theoretical Debates*, ed. Nikolas Coupland. Cambridge: Cambridge University Press, 217–41.

Britain, David, and Peter Trudgill. 1999. “Migration, New-Dialect Formation and Sociolinguistic Refunctionalisation: Reallocation as an Outcome of Dialect Contact.” *Transactions of the Philological Society* 97(2): 245–56.

Bybee, Joan L. 2002. “Word Frequency and Context of Use in the Lexical Diffusion of Phonetically Conditioned Sound Change.” *Language Variation and Change* 14(3): 261–90.

Cheshire, Jenny, Sue Fox, Paul Kerswill, and Eivind Torgersen. 2008. “Ethnicity, Friendship Network and Social Practices as the Motor of Dialect Change: Linguistic Innovation in London.” *Sociolinguistica* 22(1): 1–23.

Coupland, Nikolas. 2001. “Dialect Stylization in Radio Talk.” *Language in Society* 30(3): 345–75.

———. 2007. *Style: Language Variation and Identity. Kindle Edition*. Cambridge: Cambridge University Press.

Cukor-Avila, Patricia, and Guy Bailey. 2017. “The Effect of Small Ns and Gaps in Contact on Panel Survey Data.” In *Panel Studies of Variation and Change*, eds. Suzanne Evans Wagner and Isabelle Buchstaller. New York: Routledge, 181–212.

Dodsworth, Robin. 2017. “Migration and Dialect Contact.” *Annual Review of Linguistics* 3: 331–46.

Dufouil, Carole et al. 2014. “Older Age at Retirement Is Associated with Decreased Risk of Dementia.” *European Journal of Epidemiology* 29(5): 353–61.

Eberhard, David M., Gary F. Simons, and Charles D. Fennig. 2019. “Swabian: A Language of Germany.” *Ethnologue: Languages of the World. Twenty-second edition. SIL International*. http://www.ethnologue.com.

Eckert, Penelope. 1989. *Jocks and Burnouts: Social Categories and Identity in the High School*. New York: Teachers College Press.

———. 1997. “Age as a Sociolinguistic Variable.” In *The Handbook of Sociolinguistics*, eds. Carol Myers-Scotton and Florian Coulmas. United Kingdom: Blackwell Publishing, 151–67.

Giles, Howard, Donald M. Taylor, and Richard Y. Bourhis. 1973. “Towards a Theory of Interpersonal Accommodation through Language: Some Canadian Data.” *Language in*

Society 2(2): 177–92.

Hart, B., and T. Risley. 1995. *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore: Paul H. Brookes Publishing.

Hay, Jennifer. 2018. "Sociophonetics: The Role of Words, the Role of Context, and the Role of Words in Context." *Topics in Cognitive Science* 10: 696–706.

Hoffman, Michol F., and James A. Walker. 2010. "Ethnolects and the City: Ethnic Orientation and Linguistic Variation in Toronto English." *Language Variation and Change* 22: 37–67.

Horvath, Barbara M., and David Sankoff. 1987. "Delimiting the Sydney Speech Community." *Language in Society* 16(2): 179–204.

Johnson, Norman L., and Samuel Kotz. 1977. *Urn Models and Their Application. An Approach to Modern Discrete Probability Theory*. New York: John Wiley & Sons.

Katz, Slava M. 1996. "Distribution of Content Words and Phrases in Text and Language Modelling." *Natural Language Engineering* 2(1): 15–59.

Kehrein, Roland. 2012. *Regionalsprachliche Spektren Im Raum: Zur Linguistischen Struktur Der Vertikale. Zeitschrift Für Rezensionen Zur Germanistischen Sprachwissenschaft. Beihefte 152*. Stuttgart: Franz Steiner Verlag.

Keuleers, Emmanuel, Michael Michaël Stevens, Pawel Mandera, and Marc Brysbaert. 2015. "Word Knowledge in the Crowd: Measuring Vocabulary Size and Word Prevalence in a Massive Online Experiment." *The Quarterly Journal of Experimental Psychology* 68(8): 1665–92.

Köpke, Barbara, and Monika Schmid. 2004. "Language Attrition: The next Phase." In *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues*, eds. Monika S. Schmid, Barbara Köpke, Merel Keijzer, and Lina Weilemar. John Benjamins, 1–43.

Labov, William. 1963. "The Social Motivation of a Sound Change." *Word* 19(3): 273–309.

———. 1964. "Stages in the Acquisition of Standard English." In *Social Dialects and Language Learning*, ed. Roger W. Shuy. Champaign, IL: National Council of Teachers of English, 77–104.

———. 1966. The Social Stratification of English in New York City *The Social Stratification of English in New York City*. 2nd ed. Cambridge: Cambridge University Press.

———. 1984. *Field Methods of the Project on Linguistic Change and Variation*.

McCabe, David P. et al. 2010. "The Relationship between Working Memory Capacity and Executive Functioning: Evidence for a Common Executive Attention Construct." *Neuropsychology* 24(2): 222–43.

Milroy, Lesley. 1987. *Language and Social Networks*. 2nd Ed. Oxford: Basil-Blackwell.

Nikolov, Plamen, and Alan Adelman. 2019. "Do Pension Benefits Accelerate Cognitive Decline? Evidence from Rural China." *IZA Institute of Labor Economics. Discussion Paper Series. IZA DP No. 12524* August: 1–49.

Nordberg, Bengt, and Eva Sundgren. 1998. *On Observing Real-Time Language Change: A Swedish Case Study*. SoLiD nr 10 (=FUMS Rapport nr 190). Uppsala University, Uppsala.

Park, Denise C. et al. 2002. "Models of Visuospatial and Verbal Memory across the Adult Life Span." *Psychology and Aging* 17(2): 299–320.

Pedersen, Inge Lise. 2005. "Processes of Standardisation in Scandinavia." *Dialect Change: Convergence and Divergence in European Languages*: 171–95.

Ramscar, Michael et al. 2014. "The Myth of Cognitive Decline: Non-Linear Dynamics of Lifelong Learning." *Topics in Cognitive Science* 6(1): 5–42.

Ramscar, Michael, Peter Hendrix, Bradley Love, and R. Harald Baayen. 2013. "Learning Is Not Decline: The Mental Lexicon as a Window into Cognition across the Lifespan." *The Mental Lexicon* 8(3): 450–81.

Ramscar, Michael, Ching Chu Sun, Peter Hendrix, and R. Harald Baayen. 2017. "The Mismeasurement of Mind: Life-Span Changes in Paired-Associate-Learning Scores Reflect the 'Cost' of Learning, Not Cognitive Decline." *Psychological Science* 28(8): 1171–79.

Sankoff, David, and Suzanne Laberge. 1978. "The Linguistic Market and the Statistical Explanation of Variability." In *Linguistic Variation: Models and Methods*, ed. David Sankoff. New York: Academic Press, 239–50.

Sankoff, Gillian. 2006. "Age: Apparent Time and Real Time." *Encyclopedia of Language and Linguistics*. (1): 110–16.

Sankoff, Gillian, and Hélène Blondeau. 2013. "Instability of the [r] ~ [R] Alternation in Montreal French: An Exploration of Stylistic Conditioning of a Sound Change in Progress." In *Rhotics: New Data and Perspectives*, eds. Lorenzo Spreafico and Alessandro Vietti. Bozen-Bolzano: Free University of Bozen-Bolzano Press, 249–65.

Schilling-Estes, Natalie. 2004. "Constructing Ethnicity in Interaction." *Journal of Sociolinguistics* 8(2): 163–95.

Schilling-Estes, Natalie, and Walt Wolfram. 1999. "Alternative Models of Dialect Death: Dissipation vs. Concentration." *Language* 75(3): 486–521.

Schmidt, Jürgen Erich. 2011. "Formation of and Change in Regiolects and (Regional) Dialects in German." *Taal en Tongval* 63(1): 143–73.

Shors, Tracey J., Megan L. Anderson, Daniel M. Curlik II, and Miriam S. Nokia. 2012. "Use It

or Lose It: How Neurogenesis Keeps the Brain Fit for Learning." *Behavioral Brain Research* 227(2): 450–58.

Smith, Jennifer, and Mercedes Durham. 2012. "Bidialectalism or Dialect Death? Explaining Generational Change in the Shetland Islands, Scotland." *American Speech* 87(1): 57–88.

Streck, Tobias, and Peter Auer. 2012. "Das Raumbildene Signal in Der Spontansprache: Dialektometrische Untersuchungen Zum Allemannischen in Deutschland." *Zeitschrift für Dialektologie und Linguistik* 79(2): 149–88.

Trudgill, Peter. 1986. *Dialects in Contact*. Oxford: Blackwell Publishing.

———. 1992. "Dialect Contact, Dialectology and Sociolinguistics." In *Sociolinguistics Today: International Perspectrves,* eds. Kingsley Bolton and Helen Kwok. London: Routledge, 71–79.

Wagner, Suzanne Evans. 2012. "Age Grading in Sociolinguistic Theory." *Linguistics and Language Compass* 6(6): 371–82.

Wagner, Suzanne Evans, and Isabelle Buchstaller. 2017. *Panel Studies of Variation and Change*. eds. Suzanne Evans Wagner and Isabelle Buchstaller. New York: Routledge.

Wagner, Suzanne Evans, and Sali A. Tagliamonte. 2017. "What Makes A Panel Study Work? Research and Participant in Real Time." In *Panel Studies of Variation and Change,* eds. Suzanne Evans Wagner and Isabelle Buchstaller. New York: Routledge, 213–32.

Wieling, Martijn, Simonetta Montemagni, John Nerbonne, and R. Harald Baayen. 2014. "Lexical Differences between Tuscan Dialects and Standard Italian: Accounting for Geographic and Socio-Demographic Variation Using Generalized Additive Mixed Modeling." *Language* 90(3): 669–92.

Wieling, Martijn, John Nerbonne, and R. Harald Baayen. 2011. "Quantitative Social Dialectology: Explaining Linguistic Variation Geographically and Socially." *PLoS ONE* 6(9): 1–14.

Wittenburg, Peter et al. 2006. "ELAN: A Professional Framework for Multimodality Research." *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive*: 1556–59.

Xue, Baowen et al. 2018. "Effect of Retirement on Cognitive Function: The Whitehall II Cohort Study." *European Journal of Epidemiology* 33(10): 989–1001.

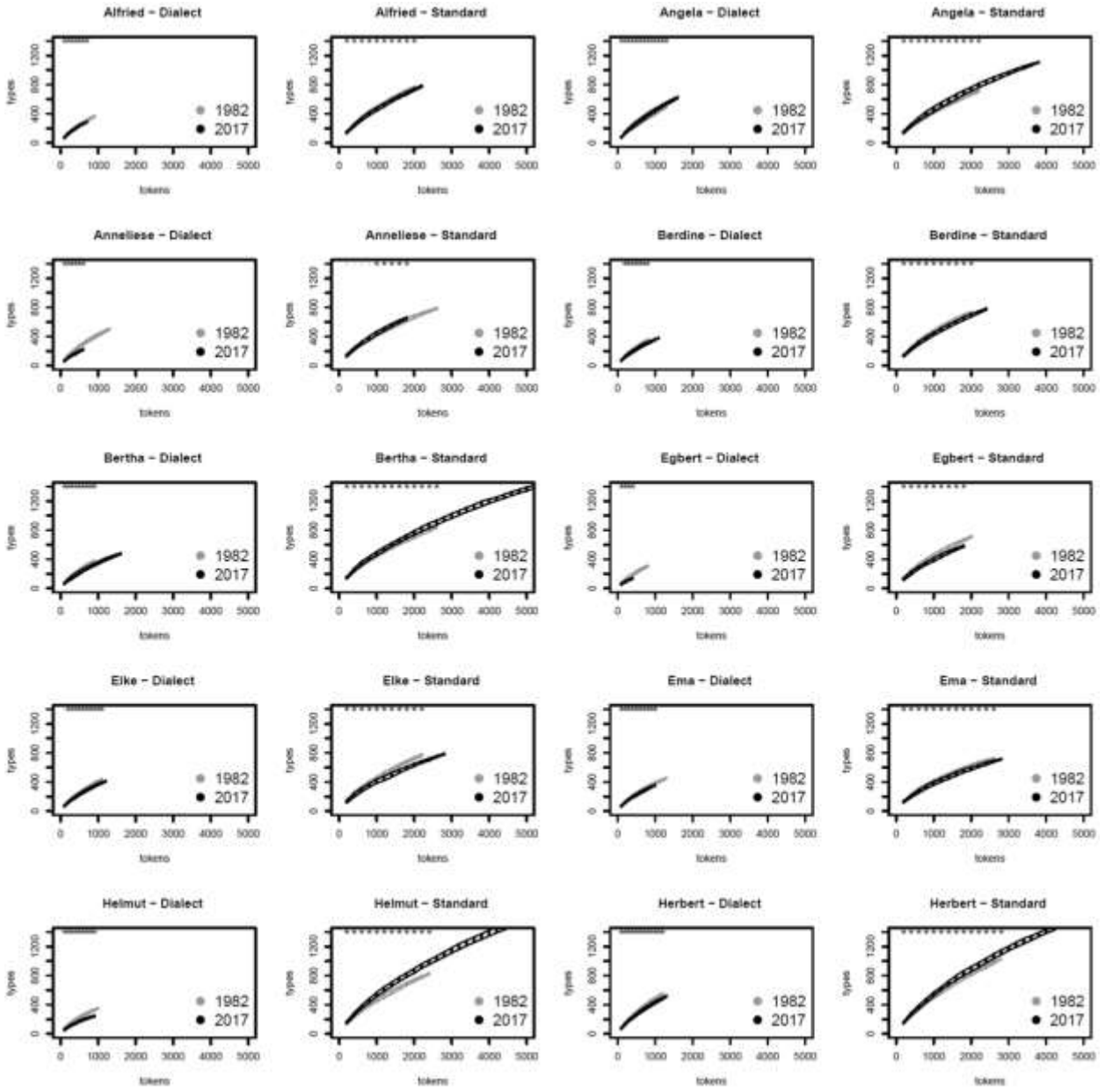

Figure 9. Vocabulary Growth by Speaker

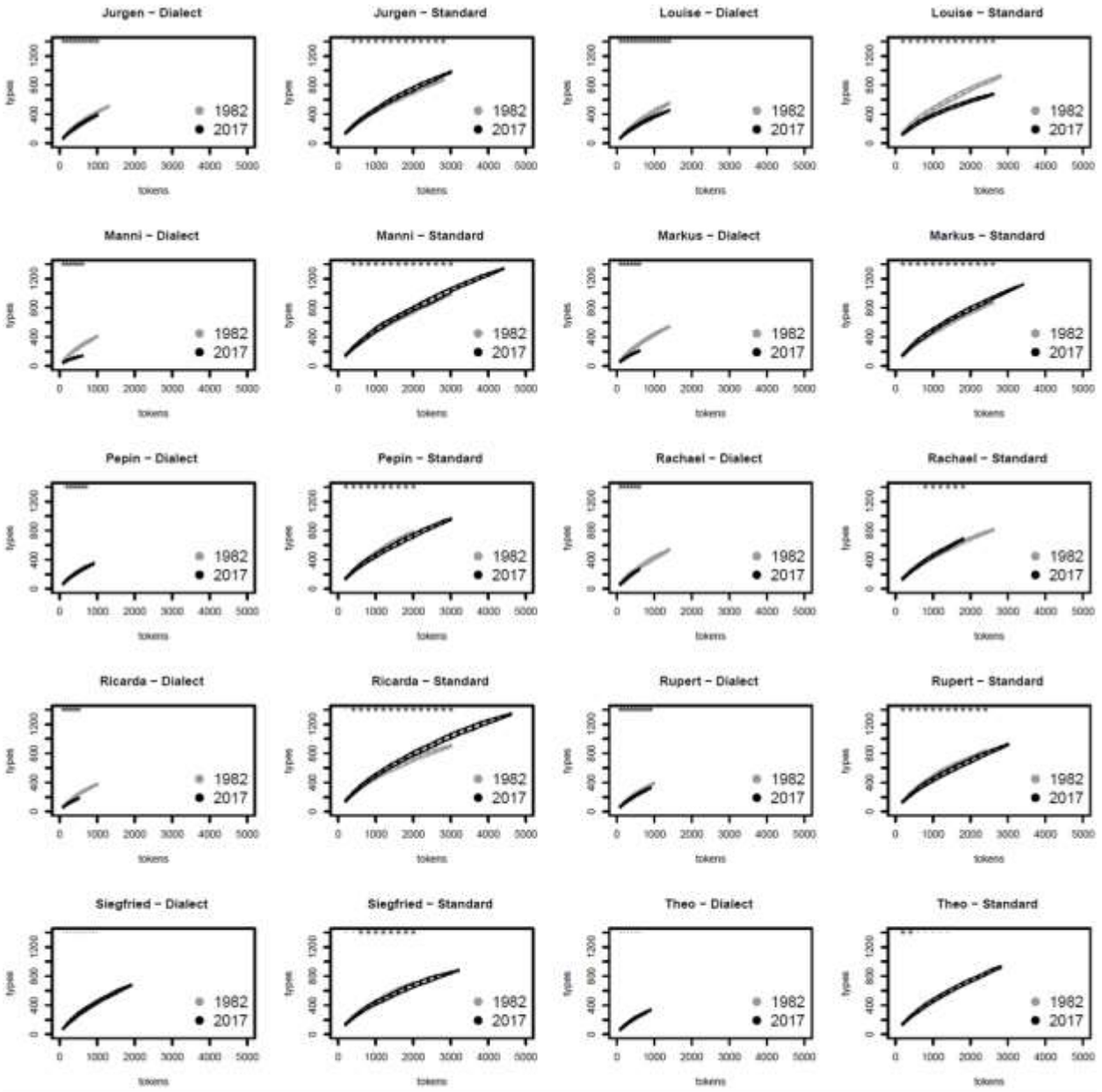

Figure 9. Vocabulary Growth by Speaker (continued)

Swabian Allegiance:

1-1. Self-Declared Swabian: Are you a 'real' Swabian?
5=definitely, 4=maybe, 3=don't know, 2=not really, 1=no

1-2. Non-Swabian Friends: Do you have friends who are NOT Swabian?
5=no, 4=a few, 3=don't know, 2=many, 1=a lot

1-3. Swabian Ridicule: If yes, do they laugh at how you speak?
5=always, 4=sometimes, 3=don't know, 2=not really, 1=not at all

1-4. Accommodation: If yes, do you change how you speak?
5=not at all, 4=a little, 3=don't know, 2=a lot, 1=always

Swabian Language Attitudes:

2-1. Opinion of Swabian Language:
5=super, 4=good, 3=don't know, 2=bad, 1=awful

2-2. Job Prospects for Swabian Speakers:
5=no impact, 4=good, 3=don't know, 2=maybe some, 1=very difficult

2-3. Opinion of Swabians Speaking Standard German:
5=very odd/awful, 4=funny, 3=don't know, 2=good, 1=great

2-4. Opinion of Non-Swabians Speaking Swabian:
5=very odd/awful, 4=funny, 3=don't know, 2=good, 1=great

Swabian Cultural Competence:

3-1. Swabian Knowledge, e.g., different types of Swabian
5=considerable, 4=some, 3=don't know, 2=not much, 1=none

3-2. Swabian Specialties, e.g., *Spätzle*, *Maultaschen*, *Moscht*
5=of course, 4=somewhat, 3=don't know, 2=not well, 1=not at all

3-3. Swabian People and Jokes, e.g., comedians, tongue-twisters, jokes
5=of course, 4=somewhat, 3=don't know, 2=not well, 1=not at all

3-4. Swabian Activities, e.g., *Hocketse* and local festivals
5=always, 4=some, 3=don't know, 2=not much, 1=never

Swabian Language Usage:

4-1. Swabian Usage by Parents
5=both parents speak Swabian, 3=only one parent speaks Swabian, 1=neither parent speaks Swabian

4-2. Swabian Usage with Family and Friends
one point for each; half point if both Swabian and standard German are used:
1=parents, 1=siblings, 1=relatives, 1=partner, 1=friends

4-3. Swabian Usage with Neighbours
two points for each; one point if both Swabian and standard German are used
2=neighbors who are older, 2=neighbors who are younger

4-4. Swabian with Others
one point for each; half point if both Swabian and standard German are used:
1=acquaintances, 1=people in bus/train, 1=teachers, 1=colleagues, 1=boss

Table 5. Swabian Orientation Index (SOI) Questions

DIALECT		1982			2017			Vocab
Pseudonym	Community	Tokens	Types	Hapax	Tokens	Types	Hapax	Change
Alfried	Gmünd	917	214	149	717	203	138	-11
Angela	Gmünd	1313	217	158	1692	233	173	16
Anneliese	Gmünd	1366	223	164	676	172	114	-51
Berdine	Gmünd	833	216	160	1101	197	134	-19
Bertha	Stuttgart	932	216	156	1648	187	130	-29
Egbert	Stuttgart	897	189	132	471	136	79	-53
Elke	Gmünd	1149	212	154	1215	197	136	-15
Ema	Stuttgart	1325	206	145	1022	195	135	-10
Helmut	Stuttgart	1005	195	131	998	150	96	-46
Herbert	Gmünd	1294	250	190	1394	224	165	-26
Jurgen	Gmünd	1317	228	163	1095	208	149	-20
Louise	Gmünd	1467	231	169	1406	200	138	-30
Manni	Stuttgart	1026	216	154	620	114	60	-101
Markus	Gmünd	1449	227	166	691	162	102	-65
Pepin	Stuttgart	750	213	155	965	201	140	-11
Rachael	Gmünd	1496	225	161	668	198	149	-26
Ricarda	Stuttgart	1079	202	143	589	159	96	-43
Rupert	Gmünd	1012	211	149	929	187	131	-24
Siegfried	Gmünd	1098	232	169	1982	234	168	2
Theo	Gmünd	676	194	134	916	193	134	-1

STANDARD		1982			2017			Vocab
Pseudonym	Community	Tokens	Types	Hapax	Tokens	Types	Hapax	Change
Alfried	Gmünd	2138	715	463	2376	683	447	-33
Angela	Gmünd	2382	629	399	3852	692	450	62
Anneliese	Gmünd	2687	630	385	1937	649	414	18
Berdine	Gmünd	2042	689	437	2598	647	420	-42
Bertha	Stuttgart	2684	670	418	6061	718	474	48
Egbert	Stuttgart	2091	663	418	1967	582	370	-82
Elke	Gmünd	2368	676	446	2896	594	371	-82
Ema	Stuttgart	2698	586	355	2913	548	335	-38
Helmut	Stuttgart	2457	685	445	6930	809	553	124
Herbert	Gmünd	2800	769	514	5246	827	579	58
Jurgen	Gmünd	2820	659	417	3107	704	468	45
Louise	Gmünd	2958	697	450	2735	548	317	-150
Manni	Stuttgart	3165	708	470	4434	741	504	33
Markus	Gmünd	2726	693	444	3402	737	491	44
Pepin	Stuttgart	2181	725	491	3193	683	457	-42
Rachael	Gmünd	2721	645	405	1922	678	450	33
Ricarda	Stuttgart	3198	672	404	4757	742	488	71
Rupert	Gmünd	2516	692	448	3116	654	442	-38
Siegfried	Gmünd	2076	663	419	3235	624	383	-39
Theo	Gmünd	1451	523	355	2942	525	356	1

Vernacular	22,401	4,317	3,104	20,795	3,752	2,567	-566
Standard	50,159	13,390	8,584	69,619	13,382	8,770	-7
TOTAL	72,560	17,707	11,688	90,414	17,134	11,337	-573

Table 6. Individual Speaker Statistics

Code	Description	IPA	Orthography	Example	Level	Family
AIS1	MHG /i:/ Diphthong Shift	[ɔɪ] ~ [aɪ]	êi ~ ei	wêiß	PHO	SWG
AIS2	MHG /ei/ Diphthong Shift	[əɪ] ~ [aɪ]	ôi ~ ei	kôin	PHO	SWG
ANN	Nasalisation /an/	[ã] ~ [a]	ã ~ a	mã	PHO	SWG
DAS	Definite Neuter Article	[dɛs] ~ [das]	des ~ das	des	PHO	REG
EDP	Verbal Plural Inflection	[ət] ~ [ən]	et ~ en	machet	MOR	ALM
FRV1	Unrounded Front Vowel /ö/	[ɛ] ~ [ø]	ee ~ ö	meeglich	PHO	SWG
FRV2	Unrounded Diphthong /eu/	[aɪ] ~ [ɔɪ]	ai ~ äu	Baim	PHO	SWG
FRV3	Unrounded Diphthong /ü/	[iə] ~ [y]	ii ~ ü	Kiche	PHO	SWG
FRV4	Diphthongisation /u/	[uə] ~ [u]	ue ~ u	muess	PHO	SWG
IRV1	Irregular Verb 'gehen'	[gangə] ~ [ge:ən]	gange ~ gehe	gange	MOR	SWG
IRV2	Irregular Verb 'stehen'	[ʃtandə] ~ [ʃte:ən]	stande ~ stehe	stande	MOR	SWG
IRV3	Irregular Verb 'haben'	[hɛn] ~ [ha:bən]	hen ~ haben	hen	MOR	SWG
IRV4	Irregular Verb 'wollen'	[vɛlə] ~ [vo:lən]	wolle ~ wollen	wolle	MOR	SWG
IRV5	Irregular Verb 'tun'	[doə] ~ [tu:n]	doe ~ tun	doe	MOR	SWG
LEO	Long /e:/ Opening	[æ] ~ [e:]	ää ~ e	Läährer	PHO	REG
LXS	Swabian Lexical Item	NA	dåhanne ~ da hinten	dåhanne	LEX	SWG
NEG	Negative Marker	[nedə]/[ed] ~ [nɪçt]	nedde/et ~ nicht	nedde/ette	PHO	REG
PVB	Periphrastic Subjunctive	NA	dääd ~ würde	dääd beeinflusse	MOR	SWG
SAF1	Swabian Affix '-le'	NA	-le ~ -chen/-lein	Mädle	MOR	ALM
SAF1B	Swabian Affix '-le'	NA	bissle ~ bisschen	bissle	MOR	ALM
SAF2	Swabian Affix 'ver-'	NA	ver- ~ er-/zer-	verzähle	MOR	SWG
SAF3	Swabian Affix 'na-'	NA	nå- ~ hin-	någange	MOR	SWG
SAF4	Swabian Affix 'sau-'	NA	sau- ~ sehr	saukalt	MOR	SWG
SAF5	Swabian Affix 'ge-'	NA	θ ~ ge-	hin[ge]kriegt	MOR	REG
SFV	Stop-Fricative Variation	[ɪç] ~ [ɪk]	ich ~ ig	wenig	PHO	REG
STP6	Palatalisation coda /st/ - 6 Verbs	[ʃt] ~ [st]	sch ~ st	kannsch	PHO	ALM
STPI	Palatalisation coda /st/ - 'ist'	[ʃt] ~ [st]	sch ~ st	isch	PHO	ALM
STPO	Palatalisation coda /st/ - non-verbs	[ʃt] ~ [st]	sch ~ st	nächschte	PHO	ALM
STPV	Palatalisation coda /st/ - other verbs	[ʃt] ~ [st]	sch ~ st	machsch	PHO	ALM
ULO1	Lowering /u/ before n - prefix	[o] ~ [u]	oo ~ un	oomeeglich	PHO	SWG
ULO2	Lowering /u/ before n - other forms	[o] ~ [u]	o ~ u	gesond	PHO	SWG
ULO3	Lowering /u/ - other forms	[o] ~ [u]	o ~ u	dorch	PHO	SWG

Table 7. Swabian Linguistic Features Under Investigation