The Cognitive Coherence of Sociolects Across the Lifespan: A Panel Study of Swabian German

Karen V. Beaman, Fabian Tomaschek, and Konstantin Sering

Abstract

In a thought-provoking article, Guy (2013:63) claims that "lectal coherence … [implies] that variables are correlated; if they are not, the cognitive and social reality of the 'sociolect' is problematic." Considerable linguistic research has established that, for structural reasons, variables are correlated; however, structural correlation does not imply sociolectal coherence. Thus the question arises: do multiple variables cluster or correlate, loosely, tightly, or not at all, based on social factors, such as age, gender, community belonging, salience, prestige/stigma, or other external factors (Guy and Hinskens 2016)? This paper explores the unresolved question of cognitive sociolectal coherence by investigating a panel of 20 speakers of Swabian, a dialect spoken in southwestern Germany, across a 35-year lifespan. The corpus consists of Labovian-style sociolinguistic interviews, and the data comprise 20 phonological and morphosyntactic linguistic features, coded for a binary distinction between the dialect and standard variant. In concurrence with Guy (2013), the findings support the notion that sociolectal coherence is more multi-dimensional than previously believed. However, clear patterns have emerged, indicating that sociolectal coherence may lie in more cognitive constructs such as the type, status, and stigmatisation of the variable itself.

Keywords: language variation and change, linguistic coherence, covariation, panel studies, lifespan change, dialects, German, Swabian.

1. Introduction

Since Guy's (2013) thought-provoking article investigating the cognitive coherence of sociolects in Brazilian Portuguese, considerable debate has ensued as to whether the notion of *covariation* of multiple linguistic features across different levels of the grammar and within specific social groups constitutes *sociolectal coherence*.[1] With some studies uncovering some level of covariation and others finding little or none, the verdict is still out on the viability of covariation as a meaningful heuristic for determining sociolectal coherence. In fact, Guy's (2013) own research found that "some sociolectal cohesion does exist, but it may be weaker and more multidimensional, than is commonly assumed" (Guy 2013:63). While some research has explored the concept of sociolectal coherence (e.g., Meyerhoff and Klaere 2017, Newlin-Lukowicz 2016; Oushiro 2016, Oushiro and Guy 2015; Tamminga 2019; Woo, Gadanidis, and Nagy, this volume), Guy laments the dearth of research into whether linguistic features really do cluster or co-occur within individuals and communities and across multiple levels of the grammar (Guy 2013:64).

To address Guy's lament, this research takes up three questions: (1) do linguistic features cluster, correlate or co-occur within a speech variety, and if so, to what extent; (2) are some kinds of language varieties (e.g., local dialects) more coherent than others; and (3) in what ways do the characteristic variables associated with a dialect or speech community covary? (Guy and Hinskens 2016:4)? Accordingly, this study aims to explore more broadly the concept of sociolectal coherence with a large number of linguistic features, in two speech communities, and

[1] This paper is dedicated to Greg Guy whose exceptional insight and ceaseless inspiration has encouraged many scholars, in particular the first author and principal investigator of this research, to never stop searching for solutions to sociolinguists' greatest quandaries. The authors would like to thank Peter Auer, Jenny Cheshire, and Naomi Nagy for their review and feedback on earlier drafts of this work. Of course, any deficiencies remaining are our own.

across a 35-year lifespan of 20 dialect speakers. We first provide a brief review of the findings from previous studies that have looked at covariation across the grammar. Next, the data and methods for the current study are explained, covering the data collection and preparation process and describing the linguistic features and the external predictors considered. The analysis and results are organised into three areas: (1) measures of sociolectal coherence through correlation analyses of 20 dialect features, (2) drivers of sociolectal coherence explored through linear regression modelling, and (3) isolation of the linguistic features exerting the greatest impact on sociolectal coherence through principal components analysis. We conclude with a review of the major empirical findings from this research and some thoughts on the theoretical role of coherence in studies of language variation and change.

2. Background

This research takes as its starting point, Guy and Hinskens' (2016:1) definition of coherence: "the extent that linguistic features systematically covary, they can be characterized as displaying coherence." It is generally accepted that linguistic features tend to cluster – that is, bound together by patterns of correlation – for structural reasons (e.g., vocalic chain shifts (Labov 1966), morphosyntactic priming (Bock and Griffin 2000), structural or parametric relationships (Guy 2013); however, what is not well established is whether variables cluster and covary for social reasons and whether these clusters form distinct sociolects. The conceptual construct of a sociolect is analogous to a language, dialect, or ethnolect, including as well class/status-based varieties and styles/registers. Guy (2013) defines a sociolect as a "cluster of variables" that identify a specific social group, arguing that "reified social varieties will necessarily encompass multiple sociolinguistic variables" (Guy 2013:64). If variables do not cluster within sociolects, Guy claims, then “the cognitive and social reality of the ‘sociolect’ is problematic” (Guy 2013:64).

The concept of sociolectal coherence has recently received some attention in the literature, yet little research conclusively supports its existence or role in a theory of language variation and change (e.g., Guy and Hinskens 2016; Meyerhoff and Klaere 2017; Newlin-Lukowicz 2016; Oushiro 2016; Oushiro and Guy 2015; Tamminga 2019; Woo, Gadanidis, and Nagy, this volume.). In his investigation of four Brazilian Portuguese variables, Guy (2013) discovered some evidence for sociolectal coherence, particularly in women’s tendency to use higher status variables, however, men showed no such coherence, raising the question whether non-standard phonology is more indexical of masculinity, a constraint that overpowers coherence. In their review of six studies from Guy & Hinskens’ (2016) *Lingua* issue dedicated to the topic of coherence, Woo, Gadanidis, & Nagy (this volume) observed that little over half of the linguistic variables investigated demonstrated coherence, while their own findings on heritage Cantonese spoken in Toronto shown even less covariation: only six out of 21 variable pairs were significantly correlated ($p<.05$) (and only two under the Spearman test). Oushiro & Guy (2015) also found little covariation in their investigation of six features of Brazilian Portuguese and concluded that coherence may be better explained through “structural similarities” and phonic salience (Naro 1981) than through social groupings.

Studies in sociolectal coherence have employed different methods in analysing whether linguistic features covary within specific language varieties. One of the earliest such studies was carried out by Horvath and Sankoff (1987) who investigated variation in four vowels in Sydney Australia using principal components analysis (PCA), a linear clustering method that determines similarities between groups based solely on linguistic criteria. More recently, Meyerhoff and

Klaere (2017) used constrained correspondence analysis (CCA), which incorporates researcher designated “must-link” constraints (e.g., “village membership”) to guide the clustering algorithms (a semi-supervised method in contrast to the unsupervised PCA method). Guy (2013), Oushiro & Guy (2015), Newlin-Lukowicz (2016), Tamminga (2019), and Woo et al. (this volume) used multivariate analyses to obtain factor weights (i.e., speakers’ tendency to use an innovative or nonstandard variant, derived through speaker random effects or residuals), and performed cross-correlations using either Pearson’s *r* or Spearman’s *rho*, depending on whether the data were normally or non-normally distributed. Drawing on the learnings from these studies, we continue the quest to understand sociolectal coherence by exploring covariation with a larger set of variables, in two speech communities, and across two points in time under the premise: “the more variables we model at once, the more sociolinguistically informative our models will be” (Meyerhoff and Klaere 2017:42).

3. Data and Methods

This section describes the corpus, the data collection and preparation process, and the methodologies employed in this investigation.

3.1 Data Collection and Preparation

3.1.1. Swabian Dialect

Swabian or *Schwäbisch* is a high Alemannic dialect spoken in the southwestern German state of Baden-Wurttemberg (see Figure 1). *Ethnologue* reports there are 820,000 speakers of Swabian or about one percent of the German population (Ethnologue 2019).

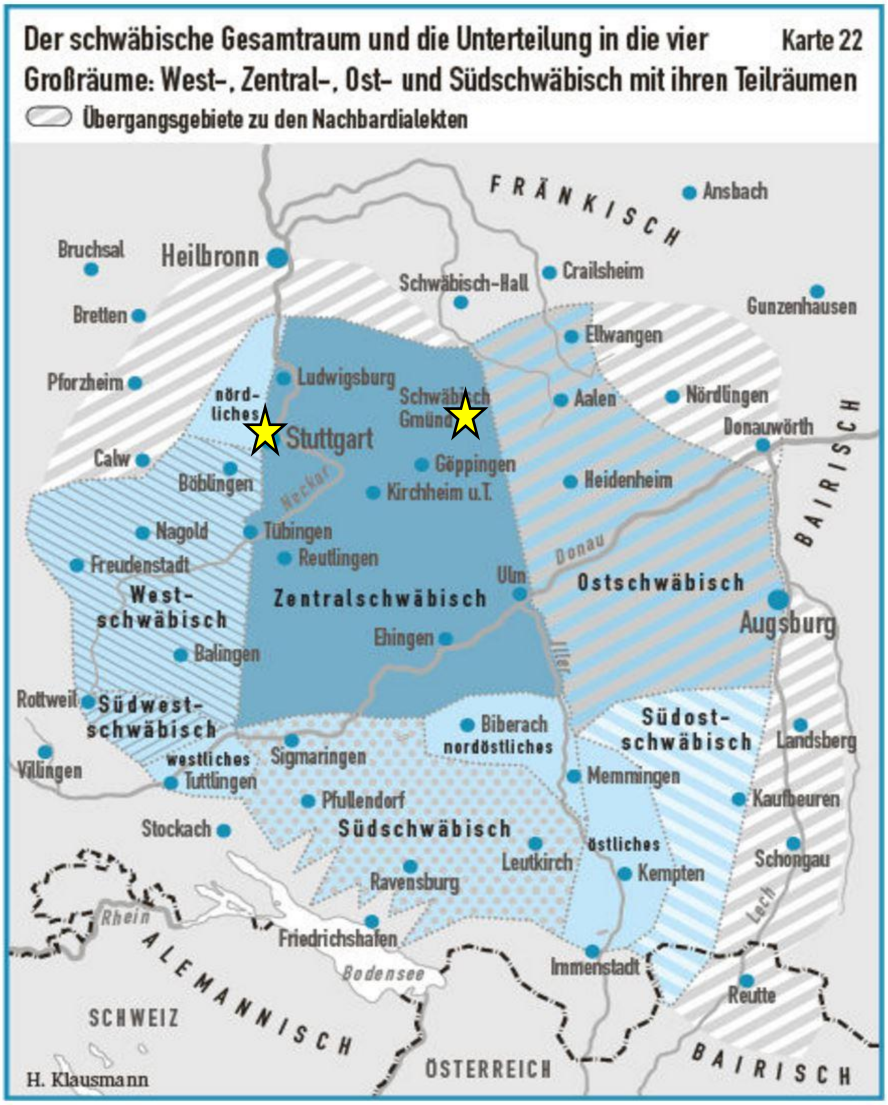

Figure 1. Map of the Swabian Dialect Area (Vogt 1977).

Attitudes toward the Swabian language vary widely: it is loved by some and loathed by others, as the following quotes from some of the speakers in the study demonstrate:

(1) Markus 2017:
so sind ja Dialekte sicher au anderweitig weniger geworden bis ausgestorben, weil se einfach e bissel -ähm- Zeichen waren fe weniger Bildung, oder se waren en irgend e Stigma
'so yeah dialects have certainly become otherwise less [common] almost extinct, because they were simply a little -ähm- sign for lower education, or they were somehow a stigma'

(2) Rachael 2017:
Schwäbisch des heert ja jääder also, und nâ han i immer e bissle Minderwertigkeitskomplex, aber des isch e Blödsinn, und des kommt bloß bei --- bei dr Sprââch,
'Swabian yeah you hear everywhere, and then I always have a little inferiority complex, but that's nonsense, and it comes just from --- from the language'

(3) Helmut 2017:
ist dieses Wechselspiel ganz deutlich geworden, auf der einen Seite dieses Gefühl in der Öffentlichkeit, du darfst nicht Mundart sprechen, weil du gleich dann nicht ernst genommen wirst, auf der anderen Seite halt man merkt, dass da einfach eine Sehnsucht danach ist.
'this interplay has become really obvious, on one side this feeling in public, you shouldn't speak dialect, because you'll immediately not be taken seriously, on the other side, you notice that there is simply a longing for it'

3.1.2. Swabian Corpus

The data for this study comprise speakers from two communities in central Swabia: the large urban centre of Stuttgart and its surrounding suburban towns and the mid-sized, semi-rural town of Schwäbisch Gmünd and its surrounding rural villages. Stuttgart is the heart of Swabia and is home to over one million inhabitants. Schwäbisch Gmünd lies 100 kilometres east of Stuttgart, with just over 60,000 inhabitants, 77% of its land is woodland and agriculture.

Twenty speakers born and raised in Swabia, first interviewed in 1982 and then reinterviewed in 2017-2018, provide the opportunity to examine changes in sociolectal coherence across the lifespan. Table 1 presents a summary of the corpus by standard sociodemographic groups. The 20 panel speakers are all middle class, most are of the same age group (in their 20's in 1982 and their 50's in 2017) and most have completed their *Abitur* 'German college preparatory exam.' They comprise two social networks – seven speakers from Stuttgart and thirteen from Schwäbisch Gmünd – and were close-knit groups in 1982, all friends and family of the interviewers. Over the 35-year timespan both networks have grown more open and dispersed: some members have moved away, others have grown apart, and many of the friends have lost contact with one another. This changing social structure and the urban-semi-rural divide provide the justification for our treatment of the two communities and two recording periods as distinct *sociolects* (Guy 2013).

Swabian Panel Study		Schwäbisch Gmünd		Stuttgart		TOTAL
		Hi Edu	Lo Ed	Hi Ed	Lo Ed	
1982	Men	0	1	0	0	**1**
31-60 years	Women	0	2	0	1	**3**
1982	Men	6	0	4	0	**10**
18-30 years	Women	3	1	1	1	**6**
2017	Men	0	1	0	0	**1**
61-90 years	Women	0	2	0	1	**3**
2017	Men	6	0	4	0	**10**
31-60 years	Women	3	1	1	1	**6**
TOTALS		**18**	**8**	**10**	**4**	**40**
		26		**14**		

Table 1. Corpus of 20 Swabian panel speakers, in Stuttgart and Schwäbisch Gmünd, recorded in 1982 and in 2017.

3.1.3. Sociolinguistic Interviews

The data were collected via standard Labovian-style sociolinguistic interviews (Labov 1984), conducted by native Swabian speakers with the primary investigator in the role of a friend-of-friend (Milroy 1987). The interviews covered questions about the speaker's childhood, favourite games played, hobbies and interests, as well as questions about the Swabian culture and language and the speaker's participation in local activities and festivals. The same questionnaire was used in 1982 and in 2017, and the interviewers were matched for age and gender with the goal of creating similar interview situations between the two recording periods. Interviews lasted approximately an hour, providing a total of 43 hours of recorded speech.

3.1.4. Transcription and Coding

The interviews were transcribed in ELAN (Wittenburg et al. 2006) by native German speakers, linguistics students at the University of Tübingen, following a strict set of guidelines and using a standard orthography specifically developed for Swabian. All transcripts were proofed four times by the principal investigator to ensure standards were followed, to neutralise any transcriber bias, and verify that all token were correctly identified and annotated. Text grids were automatically extracted from ELAN, and words containing variables under investigation were automatically tagged based on a match in a bespoke Swabian-German Lexicon (SGL). SGL was built from the sociolinguistic interviews from all lexemes with at least one of 45 linguistic features. SGL currently comprises over 10,000 Swabian variants along with their standard German counterparts, English translations, part of speech (initially built from the standard German POS-tagger (Toutanova et al. 2003) and trained for Swabian), and word frequency counts for word stem, lemma, standard variant, and Swabian variant. False starts and repetitions were excluded from the analysis.

3.2. Linguistic Features

Ten phonological and ten morphosyntactic linguistic features were selected for the current analysis based on their productivity and currency in the Swabian dialect (i.e., sufficient number of tokens for analysis in each recording period) (see Table 2). All were coded for a binary distinction between the dialect variant and the standard German variant, as well as for the linguistic level, variable type, saliency, stigmatisation, and status, i.e., change in frequency of usage between 1982 and 2017 (each of these is described in detail below). The selected linguistic features have been extensively researched in the literature (e.g., Ammon and Loewer 1977; Auer, Baumann, and Schwarz 2011; DiWA 2001; Elspaß and Möller 2003; Frey 1975; Russ 1990;

Schwarz 2015; Spiekermann 2008), however, little analysis has been conducted in the sociolinguistic variationist paradigm and none has looked at covariation of linguistic features in Swabian. A short description of each linguistic feature follows.

Linguistic Features	SWG ~ STD	Type	Salient	Stigma	Change
PHONOLOGICAL VARIABLES:					
AIS1 - Diphthong Shift /ai/ (MHG /i:/)	[əɪ] ~ [aɪ]	SWG	low	low	-9.5%
AIS2 - Diphthong Shift /ai/ (MHG /ei/)	[ɔɪ] ~ [aɪ]	SWG	high	high	-16.2%
ANN - Nasalisation /an/	[ã] ~ [an]	SWG	low	high	-18.0%
FRV1 - Unrounded Front Vowel /ö/	[ɛ] ~ [ø]	SWG	low	low	-21.2%
FRV2 - Unrounded Diphthong /eu/	[aɪ] ~ [ɔɪ]	SWG	low	low	-20.6%
FRV3 - Unrounded Front Vowel /ü/	[iə] ~ [y]	SWG	low	low	-14.3%
FRV4 - Diphthongisation /u/ (MHG /uo/)	[uə] ~ [u]	SWG	low	low	-13.7%
LEO - Long /e:/ Lowering	[ɛː] ~ [eː]	REG	low	low	-14.4%
SFV - Stop-Fricative Variation /-ig/	[ɪk] ~ [ɪç]	REG	low	low	0.4%
STPA - Palatalisation syllable coda /-st/	[ʃt] ~ [st]	ALM	high	low	-16.6%
MORPHOSYNTACTIC VARIABLES:					
DAS - Definite Neuter Article 'des'	[dɛs] ~ [das]	REG	high	low	-0.8%
EDP - Plural Verb Inflection /-ed/	[əd] ~ [ən]	SWG	high	low	-31.5%
IRV1 - Irregular Verb 'gehen'	[gangə] ~ [ge:ən]	SWG	high	high	-31.0%
IRV2 - Irregular Verb 'stehen'	[ʃtandə] ~ [ʃte:ən]	SWG	high	high	-20.4%
IRV3 - Irregular Verb 'haben'	[hɛn] ~ [ha:bən]	SWG	low	low	-30.0%
NEG - Negative Marker 'et/net/nette'	[nedə]/[ed] ~ [nɪçt]	REG	high	low	-24.2%
PVB - Periphrastic Subjunctive	[dædə] ~ [vʏʁdə]	SWG	low	low	-16.0%
SAF1 - Diminutive Suffix '-le'	[lə] ~ [çən/laɪ̯n]	ALM	high	low	-14.4%
SAF3 - Swabian Prefix 'nâ-'	[nɔ] ~ [hɪn]	SWG	low	high	-23.6%
SAF5 - Past Participle Marker 'ge-'	[θ] ~ [gə]	REG	low	low	-9.7%

Table 2. Linguistic Features under Investigation. Type = Swabian or Regional; Saliency = high or low; Stigma = high or low; Change calculated as the difference in dialect frequency between 1982 and 2017.

1. AIS1 – Shifting (i.e., lowering) of diphthongs of MHG /iː/ origin is common in Swabian, hence a word such as *Teig* [taɪk] ‘dough’ is realised as *Dêig* [dəɪg] in Swabian.[2]
2. AIS2 – Shifting (i.e., rounding) of diphthongs of MHG /ei/ origin is a stereotypical feature of Swabian, hence a standard German form such as *klein* [klaɪn] ‘small’ is realised as *glôi* [glɔɪ] in Swabian. The Swabian variant is highly stigmatised, considered indicative of Swabian spoken by farmers and in the countryside.
3. ANN – Nasalisation of /a/ before /n/ is a traditional feature of Swabian, hence words such as *man kann* [man kan] ‘one can’ are realised as *mã kã* [mɑ̃ kɑ̃] in Swabian. The Swabian variant is also highly stigmatised.
4. FRV1 – Unrounding and lengthening of the standard German front rounded vowel /ø/ is common in Swabian, hence a word such as *möglich* [møːklɪç] ‘possible’ is realised as *meeglich* [mɛːglɪç] in Swabian.
5. FRV2 – Unrounding of the diphthong /ɔɪ/ is a traditional Swabian feature, hence a word such as *Freund* [fʁɔɪnt] ‘friend’ is realised as *Fraind* [fʁaɪnd] in Swabian.[3]
6. FRV3 – Unrounding of the front vowel /y/ is typical in Swabian, so that standard German forms such as *Küche* [kyːçə] ‘kitchen’ and *Gmünd* [gyːmunt] are realised as *Kiiche* [kɪːçə] and *Gmiind* [gɪːmund] in Swabian.
7. FRV4 – Diphthongisation of /u/ to /uə/ is common in Swabian, hence words such as *muss* [mus] ‘must’ and *gut* [gut] ‘good’ are realised as [muəs] and [guəd] in Swabian.

---

[2] Voicing of plosives (‘p’, ‘t’, and ‘k’ to ‘b’, ‘d’, and ‘g’) is a typical feature of Swabian and other southwestern dialects, but is not evaluated in the current investigation.
[3] Reduction of coda schwa and other word endings is typical in Swabian, but is not investigated in this study.

8. LEO – Lowering of long /e:/ is common in forms such as *lesen* [leːzn] 'read' and *Lehrer* [leːʁɐ] 'teacher' which are pronounced as *lääs* [læːs] and *Läährer* [læːʁɐ] in southern German varieties.
9. SFV – Variation between a stop /k/ in Swabian and a fricative /ç/ in standard German is typical in the suffix -ig, used to convert nouns and verbs into adjectives, hence the standard German form *richtig* [ʁɪçtɪç] 'correct' is realised as [ʁɪçtɪk] in southern German varieties.
10. STPA – Palatalisation of /-st/ to /-ʃt/ in syllable-coda position is a highly productive feature of Swabian and the Alemannic dialects. It is common across the lexicon, but particularly in the second person singular verb formation, hence words such as *machst* [mɑxst] 'you do/make' are realised as *machscht* [maxʃt] in Alemannic.
11. DAS – The use of *des* [dɛs] 'the' for the definite neuter article versus the standard German realisation *das* [das] is widespread throughout southern Germany and increasing.
12. EDP – The standard German present tense plural verb inflexion '-en' is realised as '-ed' in Swabian, hence a standard form such as *sie finden* 'they find' is realised as *sie finded* in Swabian.
13. IRV1 – The verb *gehen* 'go' has an irregular conjugation in Swabian, hence forms such as *ich gehe* 'I'm going' are realised as *i gang*. The Swabian variant is highly salient and highly stigmatised.
14. IRV2 – The verb *stehen* 'stand' has an irregular conjugation in Swabian, hence forms such as *ich stehe* 'I'm going' are realised as *i stand*. The Swabian variant is also highly salient and stigmatised.

15. IRV3 – The verb *haben* 'have' has an irregular conjugation in Swabian, for example, the past participle is typically realised as *ghet* [ɡhɛːt] or [khɛːt] in Swabian versus *gehabt* [ɡəhaːpt] 'had' in standard German.
16. NEG – Use of the negative marker *et/net* or *edde/nedde* versus the standard German variant *nicht* is common in Swabian and other southern German varieties.
17. PVB – The past subjunctive (*Konjunktiv II*) in Swabian is a periphrastic construction formed with *tun* 'do/make' rather than *werden* 'to become', hence *es dääd beeinflusse* 'it would influence' varies with the standard German construction *es würde beeinflussen.*
18. SAF1A – The diminutive suffix -*le* is highly productive feature in the Alemannic dialects, varying with the standard German suffix -*chen* (or the older suffix -*lein)*, creating forms such as *bissle* 'a little' versus standard *bisschen* and *Mädle* 'little girl' versus *Mädchen.*
19. SAF3 – In phrasal verbs, Swabian typically uses the prefix *nââ* [nɔ] instead of *hin* [hɪn], which translates to many different prepositions in English, such as 'in', 'to', 'at', 'away', 'about', 'down', 'there', etc., creating constructions such as *nââkriegt* 'carry off' versus standard German *hinkriegt.* This is feature is characteristic of Swabian from the countryside and hence highly stigmatised.
20. SAF5 – Dropping of past participle prefix -*ge* is a common feature in Swabian, hence forms such as *hen baut* 'have built' vary with the standard German *haben gebaut*.

For these 20 variables in the 40 interviews, a total of 65,155 tokens were extracted, 27,252 from 1982 and 37,903 from 2017, with an average of over 1,300 tokens per speaker in 1982 and over 1,800 tokens per speaker in 2017. Table 3 provides the breakdown on the number of tokens by feature, community and year.

Linguistic Features	Schwäbisch Gmünd		Stuttgart	
	1982	2017	1982	2017
PHONOLOGICAL VARIABLES:				
AIS1 – Diphthong Shift /ai/ (MHG /i:/)	1825	2294	968	1508
AIS1 – Diphthong Shift /ai/ (MHG /ei/)	1615	1759	869	1408
ANN – Nasalisation /an/	1688	1623	819	1177
FRV1 – Unrounded Front Vowel /ö/	535	437	204	297
FRV2 – Unrounded Diphthong /eu/	503	520	218	353
FRV3 – Unrounded Front Vowel /ü/	903	1212	416	759
FRV4 – Diphthongisation /u/ (MHG /uo/)	1128	1223	584	837
LEO – Long /e:/ Lowering	850	1147	491	1109
SFV – Stop-Fricative Variation /-ig/	419	548	184	403
STP – Palatalisation syllable coda /-st/	2843	3356	1553	2390
MORHOSYNTACTIC VARIABLES:				
DAS – Definite Neuter Article 'des'	1428	1940	787	1551
EDP – Plural Verb Inflection '-ed'	696	1201	495	1387
IRV1 – Irregular Verb 'gehen'	203	254	70	134
IRV2 – Irregular Verb 'stehen'	134	131	62	73
IRV3 – Irregular Verb 'haben'	721	1031	378	905
NEG – Negative Marker 'et/net/nette'	968	1133	473	848
PVB – Periphrastic Subjunctive	105	159	58	49
SAF1 – Diminutive Suffix '-le'	374	324	168	203
SAF3 – Swabian Prefix 'nââ'	92	82	44	62
SAF5 – Past-participle Marker 'ge-'	918	1190	463	886
TOTAL	**17,948**	**21,564**	**9,304**	**16,339**

Table 3. Token Counts by Linguistic Feature, Speech Community, and Year.

Based on work by Wolfram and others (Van Hofwegen and Wolfram 2010; Oetting and McDonald 2002), we use a Dialect Density Index (DDI) to characterise the concentration of dialect variants within a sociolect. DDI is a token-based frequency measure that represents the total dialect variants as a percent of the total linguistic features under investigation. For the 20 phonological and morphosyntactic variables evaluated in this study, the overall DDI in 1982 was 55.8% (27,252), which has declined to 38.4% (37,903) in 2017, revealing considerable dialect levelling over the 35-year timespan.

3.3. Predictors and Predictions

Seven external predictors are considered in this investigation: two speaker factors and five cognitive linguistic factors. All predictors were coded for a binary distinction. Each is explained in detail below, along with our prediction as to expected results of the covariation analysis.

1. **Recording Year** – coded as ‘1982’ or ‘2017’.
    - Prediction: speakers in 2017 will show higher levels of coherence since Swabian is undergoing considerable change through dialect levelling, hence, convergence to the standard language acts as a stablising and “consistency factor” (Woo, Gadanidis, and Nagy, this volume).
2. **Speaker Community** – coded as ‘Stuttgart’ or ‘Schwäbisch Gmünd’.
    - Prediction: speakers from Stuttgart will show greater levels of coherence since the urban regiolect of Stuttgart serves as a prestige factor with respect to semi-rural local lect of Schwäbisch Gmünd (Britain 2016; Trudgill 1986).
3. **Variable Level** – coded as ‘phonological’ or ‘morphosyntactic’.
    - Prediction: morphosyntactic variables will show greater coherence than phonological ones, as grammatical variables, with lower levels of frequency, tend to be more linguistically and pragmatically governed than socially constrained (Cheshire 1999; Scherre and Naro 1992); moreover, in moving toward more

standardised varieties, speakers first avoid non-standard syntactic structures which are more directly associated with lower levels of education (Cheshire 2003; Sharma 2005).

4. **Variable Type** – coded as 'Swabian' or 'regional', based on dialectology studies of Swabian and other German varieties (AdA 2011; DiWA 2001; Frey 1975; Russ 1990).
    - Prediction: Swabian-specific variables will show a greater level of coherence than more widely used regional variables, driven by speakers' desire to mark group membership through their variable usage (Beaman 2018; Nycz 2016; Woo, Gadanidis, and Nagy, this volume).
5. **Variable Status** – coded as 'stable' or 'changing' based on a change in the overall frequencies of the dialect variant between 1982 and 2017 (see Table 2); frequency changes less than 10% are considered stable.
    - Prediction: variables that are currently undergoing change will exhibit greater levels of coherence as a result of dialect levelling and increasing consistency with the standard language (Woo, Gadanidis, and Nagy, this volume) and "shared social motivation for the change" (Tamminga 2019).
6. **Variable Salience** – coded as 'high' or 'low'; based on speakers' overt mention of the variable during the sociolinguistic interview ("awareness", "noticeability", "markedness", "perceptually and cognitively prominent" (Kerswill and Williams 2002:81) (Rácz 2013)).
    - Prediction: low salient variables will show higher levels of coherence since forms functioning below the level of awareness may be more automatic and uniform, while higher salient forms will show greater variation as they are more readily available for stylisation and identity construction (Erker 2017; Levon and

Buchstaller 2015; Levon and Fox 2014; Montgomery and Moore 2018; Nycz 2016; Oushiro and Guy 2015; Trudgill 1986).

7. **Variable Stigmatisation** – coded as 'high' or 'low', based on speaker' overt positive or negative comments about specific variables during the interviews.
    - Prediction: variables that are highly stigmatised will show higher levels of coherence because, because speakers may actively try to avoid them and converge toward the standard language (Trudgill 1986).
8. **Interaction Effects** – as with most sociolinguistic features, many of these constraints are expected to interact which will be evaluated through mixed effects linear regression modelling.
    - Prediction: variables with high levels of saliency and stigma will likely be correlated with variables that are undergoing change based on Trudgill's (1986) observation that "speakers modify those features of their own varieties of which they are most aware" (Trudgill 1986:11).

3.4. Measures of Coherence

This paper follows previous sociolinguistic studies that use cross-correlations to determine coherence across multiple linguistic features (see Section 2). The theoretical basis for this approach is that groups of speakers who are consistent in their frequency of dialect usage will exhibit high covariation across variables. However, if variable usage is *ad hoc* or simply unrelated, then speakers will exhibit a random mix of high and low correlation indices (Guy 2013). In short, more coherent lects will show greater levels of covariation.

There is no single "best" measure or generally agreed-upon set of conventions for measuring covariation in a dataset: "all models are wrong but some are useful" (attributed to Box

1976). We tested a number of methods cited in the literature for assessing covariation in a dataset and decided on the following four measures as most indicative of the relative differences in the levels of covariation between different sets of predictors. Following Box (1976:792) "we cannot know that any statistical technique we develop is useful unless we use it."

1. **Significant Pairs** – the percentage of variable pairs (within a given study) that show significant correlations; larger values show a greater number of correlated variables (Guy 2013; Oushiro 2016; Woo, Gadanidis, and Nagy, this volume).
2. **Correlation Mean** – the mean of the values[4] in a correlation matrix using Spearman's *rho* because of the non-normal distribution of our data; higher values imply higher levels of correlation with values close to 0 signifying non-correlation (*cor* and *mean* functions in the *base* R package, version 3.5.3).
3. **Principal Components** – a linear clustering method that identifies the underlying patterns in a dataset by reducing the dimensionality with the aim of explaining the greatest proportion of variation; the values of the first three principal components are provided for comparison (*prcomp* function in the R packages *stats*, version 3.5.3) (Horvath and Sankoff 1987).
4. **Steiger chi-square** – the sum of the squared Fisher transformed correlations distributed as chi square, shown to be particularly effective for controlling Type I error rates, a potential issue in large correlation matrices with small sample sizes; higher chi square (X^2) scores indicate greater collinearity (Steiger 1980) (*cortest.normal* function in the R package *psych*, version 1.8.12).

---

[4] Since mean values are affected by both positive and negative correlations, they more appropriately reflect the level of coherence when correlations are working in opposite directions.

4. Analysis and Results

The analysis and results are organised into three subsections: (1) measurement of the levels coherence in different Swabian sociolects drawn from correlation matrices of dialect frequencies; (2) exploration into the drivers of coherence in Swabian using linear regression analyses to examine the interaction effects between the correlation coefficients of the 20 linguistic features and seven external factors; and, (3) identification of the linguistic features "pulling the most weight" in the most coherent lects and what this may suggest about the relationship between cognitive sociolectal coherence and language change.

4.1. Measuring Coherence

Although correlation matrices are the foundation of all traditional multivariate analyses, there are few techniques for visually depicting the patterns of underlying relationships between linguistic features, particularly for large numbers of variables. To address this challenge, we use the construct of a *correlogram* to graphically portray the correlation values, including (1) the sign (i.e., whether variables are positively or negatively correlated), (2) the scale (i.e., how strongly or weakly the variables are correlated), and (3) the relationships between variables (i.e., by clustering "similar" variables together using principal components) (Friendly 2002; Murdoch and Chow 1996).

To address the first question of this research -- do linguistic features correlate and to what extent -- we created three sets of correlation matrices (recording year, speech community, and linguistic level) using the frequencies of dialect variants for each linguistic feature. Figure 2 presents the resulting correlation matrices for all 20 linguistic features and both communities by recording year. The bottom left triangle of each plot reports the correlation coefficient for each pair of variables, and the top right triangle graphically displays the significant levels (***

p<.001, ** p<.01, * p<.05), with positive correlations represented in darker shades of blue and negative correlations in darker shades of red. The correlation matrices were rendered with the *corrplot.mixed* function in the R package *corrplot*, version 0.84, and significance tests for each pair of variables were conducted with the *cor.mtest* function using Spearman's rank correlation coefficient.

In comparing the correlograms for the two time periods, 2017 appears to be "more coherent" than 1982, as visualised by a greater number of stronger pairwise comparisons (i.e., darker shaded boxes). To test whether the patterns in the two years are statistically different, we use the Steiger X^2 test (*cortest.normal* function in R *psych* package, version 1.8.12), and our visual assumption is confirmed: the two matrices are indeed significantly different from one another ($X^2 = 243.2$, $p < 0.0055$).

Although the number of significant variable pairs is slightly lower in 2017 (55%, 105 out of 190 pairs) than in 1982 (61%, 115 out of 190), the correlation mean is higher (.358 for 1982 and .529 for 2017). This difference brings to light one core distinction in how the current study views sociolectal coherence: while the number of significant pairwise correlations indicates "how many" variables covary, the strength of the correlations between the variables indicates "how tightly" they covary and hence how difficult it may be to uncouple them. As a result of dialect levelling over the 35-year timespan of this study, dialect density is lower in 2017 (DDI=38%) than in 1982 (DDI=56%). The lower dialect density and higher correlation mean in 2017 provides support for our first prediction (from Section 3.3) that the standard language acts as a stablising and "consistency factor" that boosts sociolinguistic coherence (cf. Woo, Gadanidis, and Nagy, this volume).

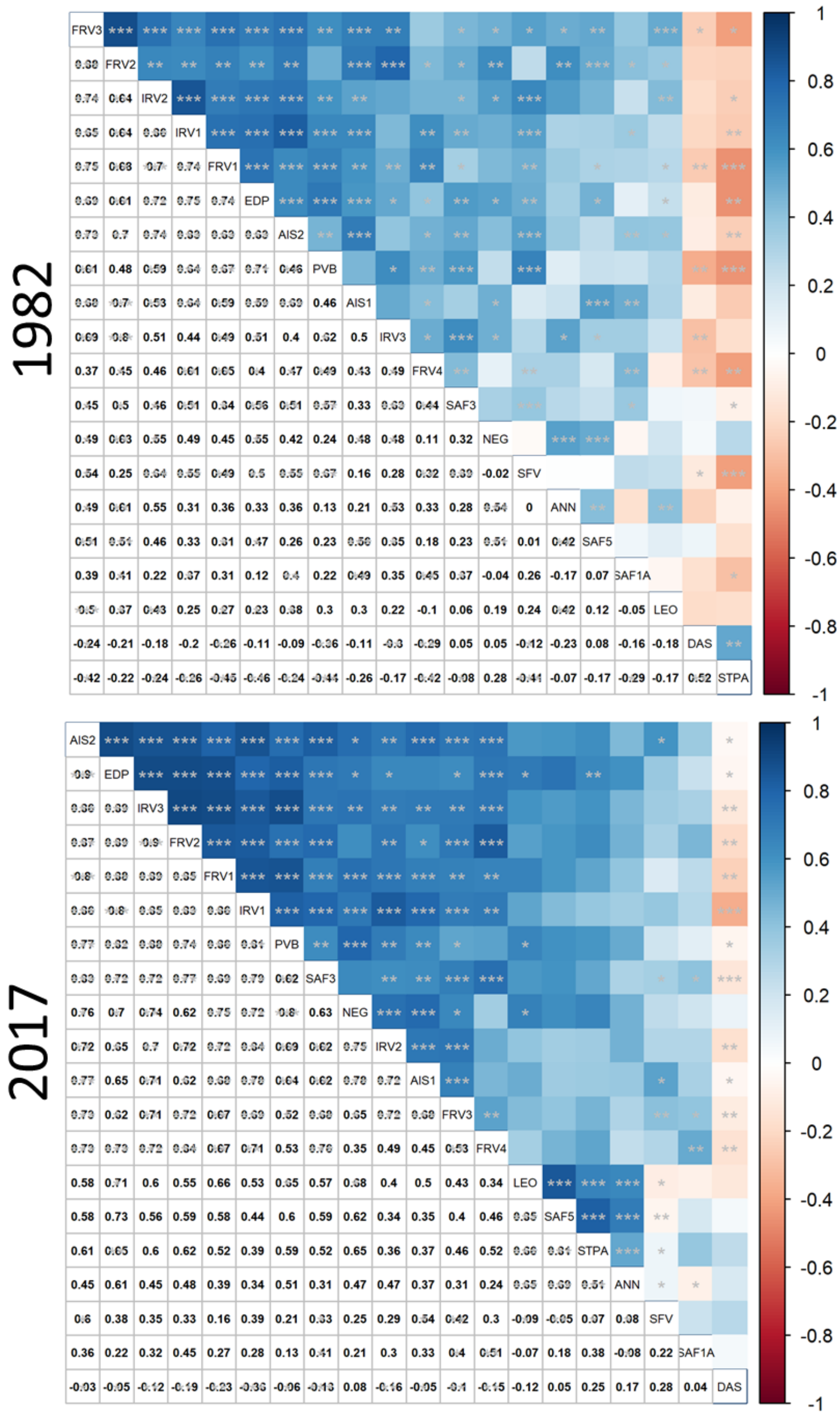

Figure 2. Covariation Analysis for two Recording Years: 1982 and 2017.
Significance levels: *** p<.001, ** p<.01, * p<.05; positive correlations shown in darker shades of blue and negative correlations in darker shades of red; DDI: 1982 = 55.8% (26,475); 2017 = 38.4% (37,111); Significant pairs: 1982 = 60.5% (115/190); 2017 = 55.3% (105/190); Mean: 1982 = .358; 2017 = .529; PC123: 1982 = .693; 2017 = .787; Steiger X2 test of identity of the two matrices = 243.2, df = 190, p <. 0.0055**.**

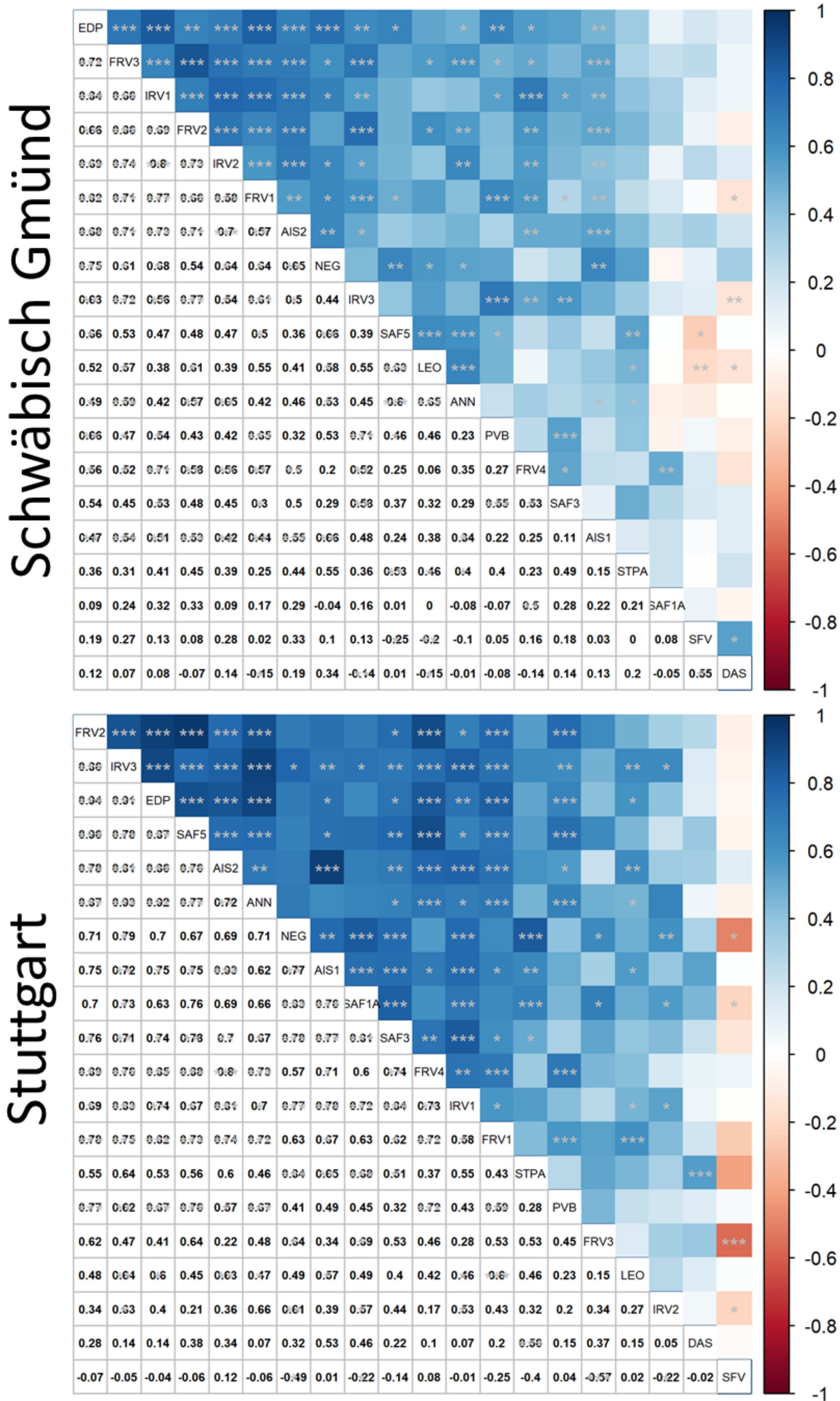

Figure 3. Covariation Analysis for two Communities Schwäbisch Gmünd and Stuttgart. Significance levels: *** p<.001, ** p<.01, * p<.05; positive correlations shown in darker shades of blue and negative correlations in darker shades of red; DDI: Gmünd = 54.3% (38,331); Stuttgart = 33.7% (24,884); Significant pairs: Gmünd = 45.3% (86/190); Stuttgart = 48.4% (92/190); Mean: Gmünd = .426; Stuttgart = .540; PC123: Gmünd = .721; Stuttgart = .783; Steiger X^2 test of identity of the two matrices = 280.19, df = 190, p <. 0.001.

Next, we compare the correlation matrices for the two communities (see Figure 3). Although the number of significant variables pairs is largely the same between the two communities (45% and 48%), Stuttgart shows lower dialect density (34%) and a higher correlation mean (.540) than Schwäbisch Gmünd (DDI=54%, mean=.426), a difference confirmed with the Steiger X^2 test ($X^2 = 280.19$, $p < 0.001$). Thus, we can confirm our second prediction that the more standardised dialect in the large urban centre of Stuttgart brings uniformity and prestige that enhances sociolinguistic coherence.

Further subdivision of the covariation analysis by community and recording year (see Table 4, plots not shown) reveals that the highest dialect density and lowest correlation mean is Schwäbisch Gmünd in 1982 (DDI=62%, mean=.171), while the lowest dialect density and highest correlation mean is in Stuttgart in 2017 (DDI=25%, mean=.488). In fact, Table 4 shows that for all measures of covariation, there is a distinct progression of decreasing dialect density and increasing levels of coherence across the four sociolects: from Schwäbisch Gmünd in 1982 to Schwäbisch Gmünd in 2017 to Stuttgart in 1982 and finally to Stuttgart in 2017. These results provide insight into the second question of this research: are some language varieties more coherent than others. Based on the measures of covariation defined in this study, the large urban centre of Stuttgart in 2017 is the leader in coherence. The sociolect of Schwäbisch Gmünd in 2017 resembles that of Stuttgart in 1982, implying that the smaller, semi-rural community is following the lead of the larger, more prestigious community and suggesting that change emanates from urban centres to surrounding towns and villages following the Cascade/Gravity models (Labov 2003; Trudgill 1974).

Correlation Metrics	Schwäbisch Gmünd		Stuttgart	
	1982	2017	1982	2017
Dialect Density Index (DDI)	62.6%	45.9%	43.0%	24.5%
Percent of Significant Pairs	35.3%	48.9%	48.4%	51.6%
Correlation Mean	.171	.387	.329	.488
Proportion of Variance (PC1-PC3)	.540	.760	.825	.842
Steiger X^2 Test of Significance	245.92	752.09	502.34	576.34
Steiger X^2 Significance Level	**	***	***	***

Table 4. Correlation Measures by Community and Recording Year.
Significance levels: *** p<.001, ** p<.01, * p<.05.

We now turn to the differences in coherence based on linguistic level. Figure 4 shows the correlation matrices for the phonological variables (top panel) and the morphosyntactic variables (bottom panel) covering both communities and both time periods. While dialect density is notably higher for the morphosyntactic variables (59%) than for the phonological ones (35%), the correlation means are quite similar (.551 and .536). The Steiger X^2 test of identity shows no significant difference between the two matrices ($X^2 = 25.49$, $p < 0.99$), leading us to conclude that coherence is not constrained by linguistic level and disconfirming our third prediction that the morphosyntactic variables would exhibit greater levels of coherence than the phonological ones.

This leads us to question what factors are responsible for driving higher or lower levels of coherence, a topic we address in Section 4.2. Another confounding effect when looking at aggregate levels of coherence across speech communities concerns the weight each individual linguistic feature carries in driving levels of coherence, a matter we consider in Section 4.3.

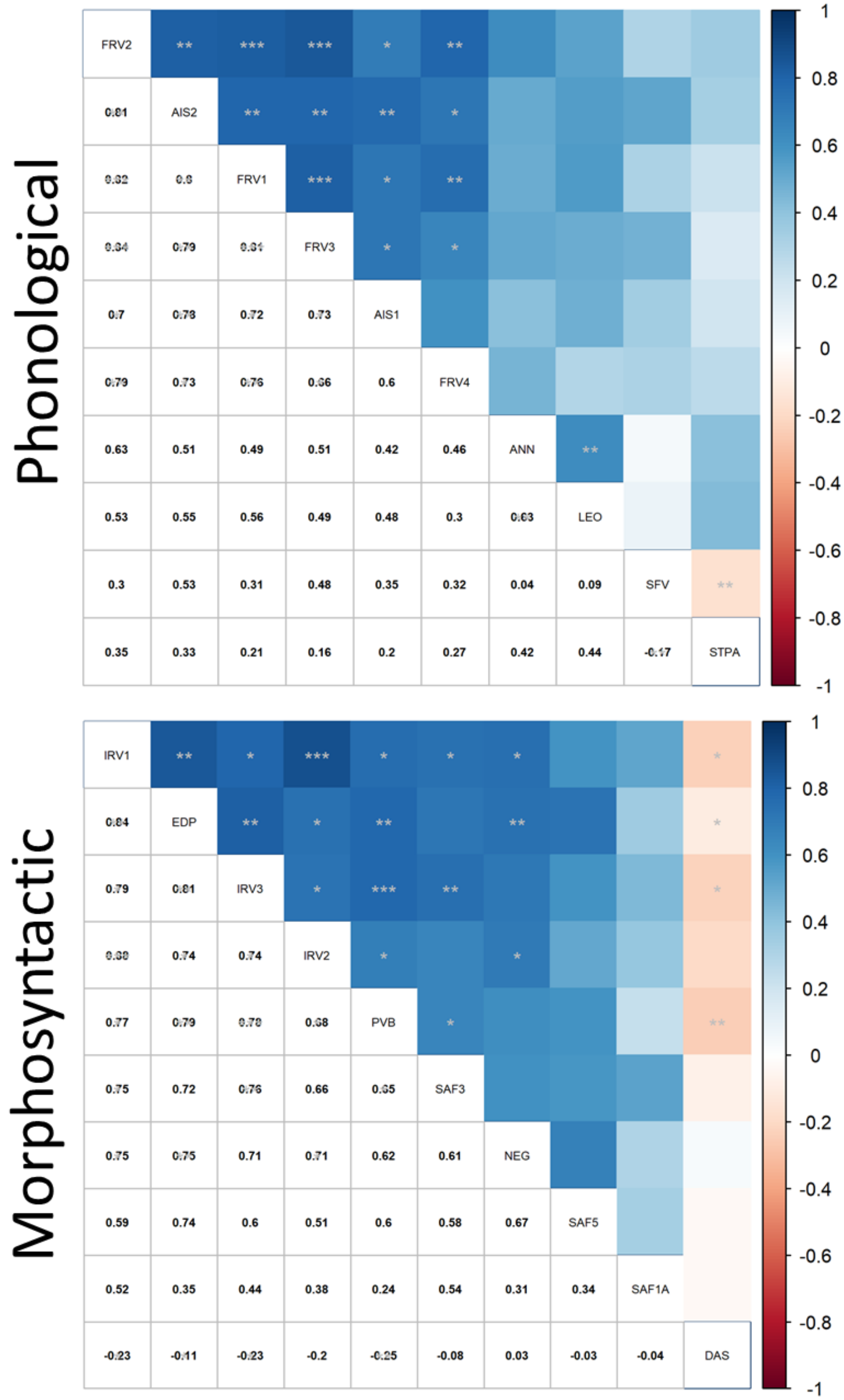

Figure 4. Covariation Analysis for two levels: phonological and morphosyntactic. Significance levels: *** p<.001, ** p<.01, * p<.05; positive correlations shown in darker shades of blue and negative correlations in darker shades of red; DDI: phonological = 35.3% (41,137); morphosyntactic = 58.9% (22,078); Significant pairs: phono = 35.6% (16/45); morphosyn = 44.4% (20/45); Mean: phono = .551; morphosyn = .536; PC123: phono = .828; morphosyn = .834; Steiger X^2 test of identity of the two matrices = 25.49, df = 45, p <. 0.99.

4.2. Driving Coherence

To investigate the factors driving coherence in Swabian, multivariate analyses were conducted with the correlation coefficients for the 20 linguistic features (20 * 20 for two recording years for a total of 760 correlation coefficients) and the seven external predictors described in Section 3.3. Table 5 presents the results of three linear regression models (*lm* function in R package *stats*, version 3.5.3): variable status (top panel), stigmatisation (middle panel), and variable type (bottom panel). Linguistic level and variable salience showed no significant effects, refuting predictions #3 and #6 (see Section 3.3), and thus were removed from the modelling effort. Figure 5 provides a visualisation of the results of the linear regression models by plotting one point (i.e., one correlation coefficient) for each pair of variables in each recording year (*visreg* function in R package *visreg*, version 2.5-1). Higher values indicate positive correlations and smaller values negative correlations with coefficient values of zero indicating non-correlation.

All plots in Figure 5 substantiate the first and second predictions in this study (see Section 3.3) that Stuttgart and the later recording year (2017) would show higher levels of coherence than Schwäbisch Gmünd and the earlier recording year (1982) as a result of the dialect levelling process that has been taking place in Swabia. The regression analyses confirm that the standard language acts as a “consistency factor” driving linguistic change (Woo, Gadanidis, and Nagy, this volume) and that change diffuses from urban centres to outlying communities in a cascading or gravity-like manner (Labov 2003; Trudgill 1974).

Coefficients	Estimate	Std.Error	t-value	p-value	Sig.Lvl
VARIABLE STATUS: (Multiple R^2 = .164; Adjusted R^2 = .157; p < .001)					
(Intercept)	.14548	.03133	4.644	4.37e-06	***
Community-Stuttgart	.34042	.04403	7.732	5.88e-14	***
Year-2017	.30439	.04403	6.914	1.45e-11	***
Status-Stable	-.18318	.18318	-2.506	.0125	*
Stuttgart : 2017 : Changing	-.27248	.06226	-4.376	1.47e-05	***
VARIABLE STIGMA: (Multiple R^2 = .204; Adjusted R^2 = .192; p < .001)					
(Intercept)	.10081	.03388	2.975	.003084	**
Community-Stuttgart	.20844	.04791	4.350	1.68e-05	***
Year-2017	.18288	.04791	3.817	.000154	***
Stigma-High	.06725	.11490	.585	.558636	
Stuttgart : 2017	.08170	.06776	1.206	.228548	
Stuttgart : High Stigma	.27858	.16249	1.715	.087121	.
Year-2017 : High Stigma	.22811	.16249	1.404	.161041	
Stuttgart : 2017 : High	-.82139	.22979	-3.574	.000389	***
VARIABLE TYPE: (Multiple R^2 = .442; Adjusted R^2 = .432 p < .001)					
(Intercept)	.06140	.07038	.872	.383537	
Community-Stuttgart	-.10728	.09954	-1.078	.281773	
Year-2017	.01217	.09954	.122	.902747	
Type-Swabian	.09018	.07596	1.187	.235849	
Stuttgart : Year-2017	.48874	.14077	3.472	.000571	***
Stuttgart : Type-Swabian	.65209	.10743	6.070	2.89e-09	***
Year-2017 : Type-Swabian	.40735	.10743	3.792	.000172	***
2017 : Stuttgart : Swabian	-1.02952	.15193	-6.776	4.23e-11	***

Table 5. Multivariate Analyses of Factors Influencing Coherence in Swabian.
Significance levels: *** p<.001, ** p<.01, * p<.05

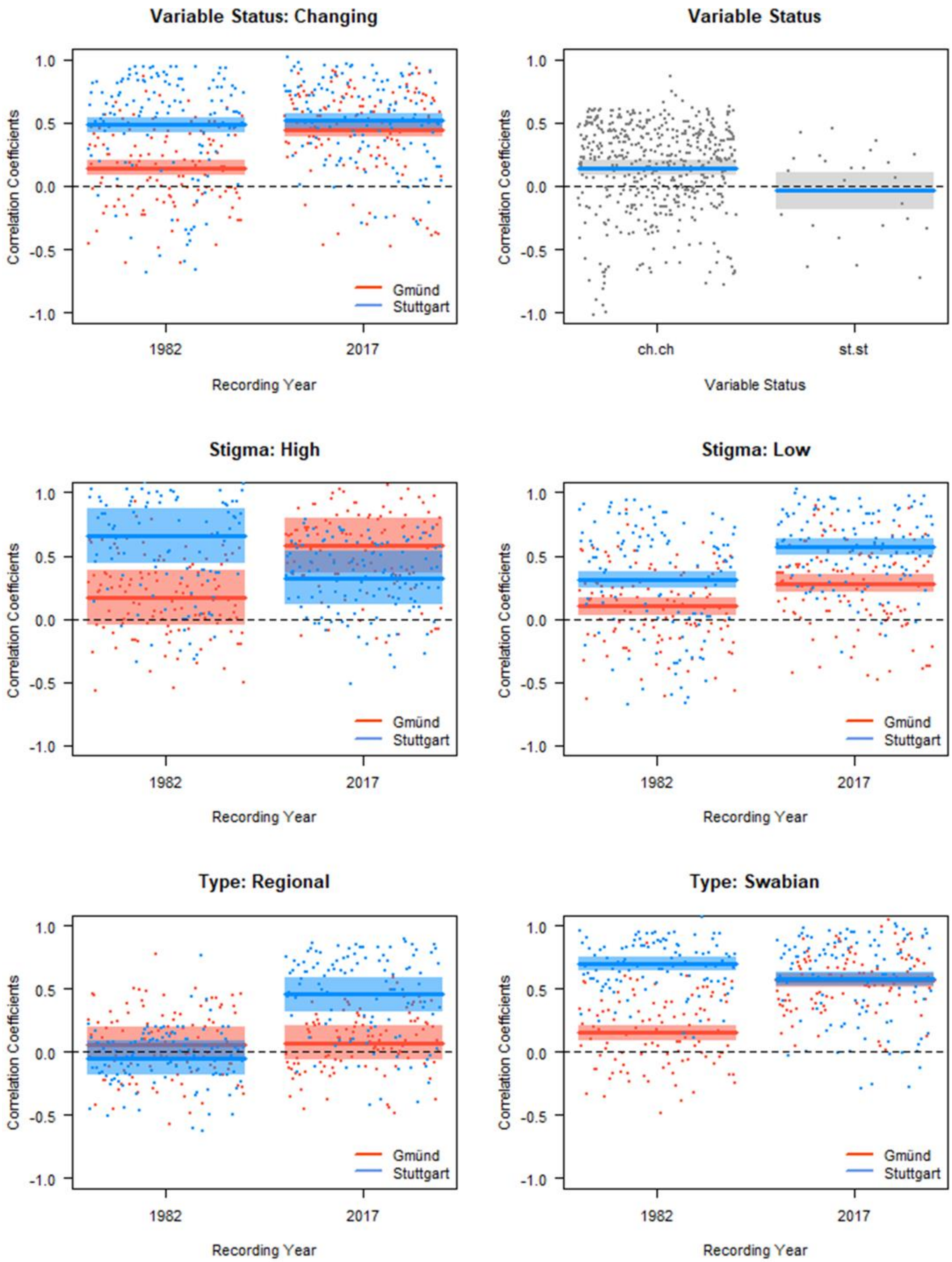

Figure 5. Drivers of Coherence. Multivariate analysis for correlation coefficients for 20 linguistic features by status (changing versus stable, stigmatisation (high versus low), and type (regional versus Swabian).

The top panels in Figure 5 and Table 5 reveal that the changing variables (i.e., those that have decreased by more than 10% over the 35 years, see Table 2) exhibit higher levels of coherence than the stable variables, providing support for our fifth prediction (see Section 3.3), and this change is highly significant for Schwäbisch Gmünd. In fact, the visualisation shows Stuttgart's level of covariation for the changing variables across the two periods has remained the same, while Schwäbisch Gmünd has evolved to become more like Stuttgart. This finding supports both prediction #2, that Stuttgart as the urban centre of Swabia exerts a standardising effect on local dialect varieties (e.g., Auer 2005; Trudgill 1986), and prediction #5, that variables undergoing convergence to the standard will show higher rates of covariation because the change likely reduces complexity in the grammar due to fewer number of variants (again corroborating Woo, Gadanidis, and Nagy's expectation (this volume)).

The middle panels of Figure 5 and Table 5 show the results of the modelling effort based on level of stigmatisation (see Table 2 for how the variables were classified). Confirming prediction #7, the more highly stigmatised variables show higher levels of covariation in 1982, a factor that appears to have become less conclusive in 2017 in Stuttgart than it was in 1982. We surmise that this reversal is due to the fact that many of the high stigma variables have simply died out of the Stuttgart sociolect by 2017, a point we will return to in Section 4.3.

The lower panels of Figure 5 and Table 5 present the results of the modelling effort based on the type of linguistic feature (i.e., Swabian or regional). While there was no difference in the level of covariation with the regional features between Stuttgart and Schwäbisch Gmünd in 1982 (lower left panel), a significant difference has emerged in 2017 with Stuttgart showing a higher level of covariation as a result of greater movement to the standard language. For the Swabian features (lower right panel), Stuttgart shows an even greater degree of covariation in 1982, and by 2017 Schwäbisch Gmünd is following suit, providing support for our fourth prediction that Swabian features cohere more tightly as a result of their common sociohistorical background than regional features which have spread from other areas in Germany.

In sum, the main drivers of covariation in the Swabian situation are recording year and community, and these factors interact with variable status (changing or stable), variable type (Swabian or regional), and level of stigmatisation (high or low). Table 6 provides a summary of the factor predictions from a linear regression model with the three-way interaction effects (*predict* function in the R package *stats*, version 3.5.3). Most noteworthy is that none of these factors had any effect on the coherence of the sociolect in Schwäbisch Gmünd in 1982, yet all have become strong predictors of coherence in the sociolect of Stuttgart in 2017, with the sociolect of Schwäbisch Gmünd in 2017 representing an intermediate variety (cf., Auer's 'intermediate forms' (Auer 1997)). Our data reveal that a decrease in dialect density is concomitant with an increase in coherence (as measured by levels of covariation), again bringing support for the claim that the standard language brings conformity and unity to a sociolect (Woo et al., this volume), at least for the 20 linguistic features analysed in this study.

Predictors	Schwäbisch Gmünd		Stuttgart	
	1982	2017	1982	2017
Status: Changing	.161	.585	.615	.616
Stable	.330	-.043	.280	.365
Stigma: Low	.181	.517	.583	.641
High	.152	.601	.612	.468
Type: Regional	.123	.196	.074	.604
Swabian	.188	.638	.740	.596

Table 6. Drivers of Coherence – Predicted values from multivariate analyses. predictions over .60 are considered to be more likely (shaded in grey).

4.3. Identifying Coherence

We now turn to the third and final question this research seeks to address: in what ways do the characteristic linguistic features of a sociolect covary? To isolate the individual variables pulling the most "coherence weight", we use Principal Components Analysis (PCA) to identify the proportion of variation explained by each linguistic feature. Figure 6 plots the first two principal components for the 20 linguistic features, visually depicting the four sociolects (outlined by the multi-coloured trapezoids) and the variable loadings (arrows) (*autoplot* function in *ggplot2* R package, version 3.2.1). The angles of the arrows indicate the degree of correlation between the linguistic feature and the principal component (i.e., PC1 or PC2), and their length signifies how much of the variation is explained by that feature: specifically, the more horizontal or vertical the arrow, the stronger the correlation with PC1 or PC2 respectively, while the longer the arrow, the greater the explained variance.

The first, most apparent observation in Figure 6 is the vast change in the sociolects between 1982 and 2017. Both Stuttgart and Schwäbisch Gmünd in 1982 have evolved from small, close-knit, homogeneous communities in 1982 (the small and narrow trapezoids) to broad and more heterogeneous communities in 2017 (the long and wide trapezoids). The dialect levelling and standard language convergence occurring in Swabia over the 35-year timespan of this study have affected the panel speakers in different ways: some have moved completely to the standard language (e.g., S015 (Ricarda), S016 (Manni), and S036 (Helmut) found in the far left corner of Figure 6), while others have remained staunch dialect speakers (e.g., S013 (Louise), S021 (Siegfried), and S026 (Berdine) found on the far right side of Figure 6). Our prior research has shown that individuals' personal orientation to the Swabian language and culture is a strong influence in their choice to use dialect features or not (Beaman 2018, 2020; Beaman and Tomaschek 2021). But which features do speakers choose?

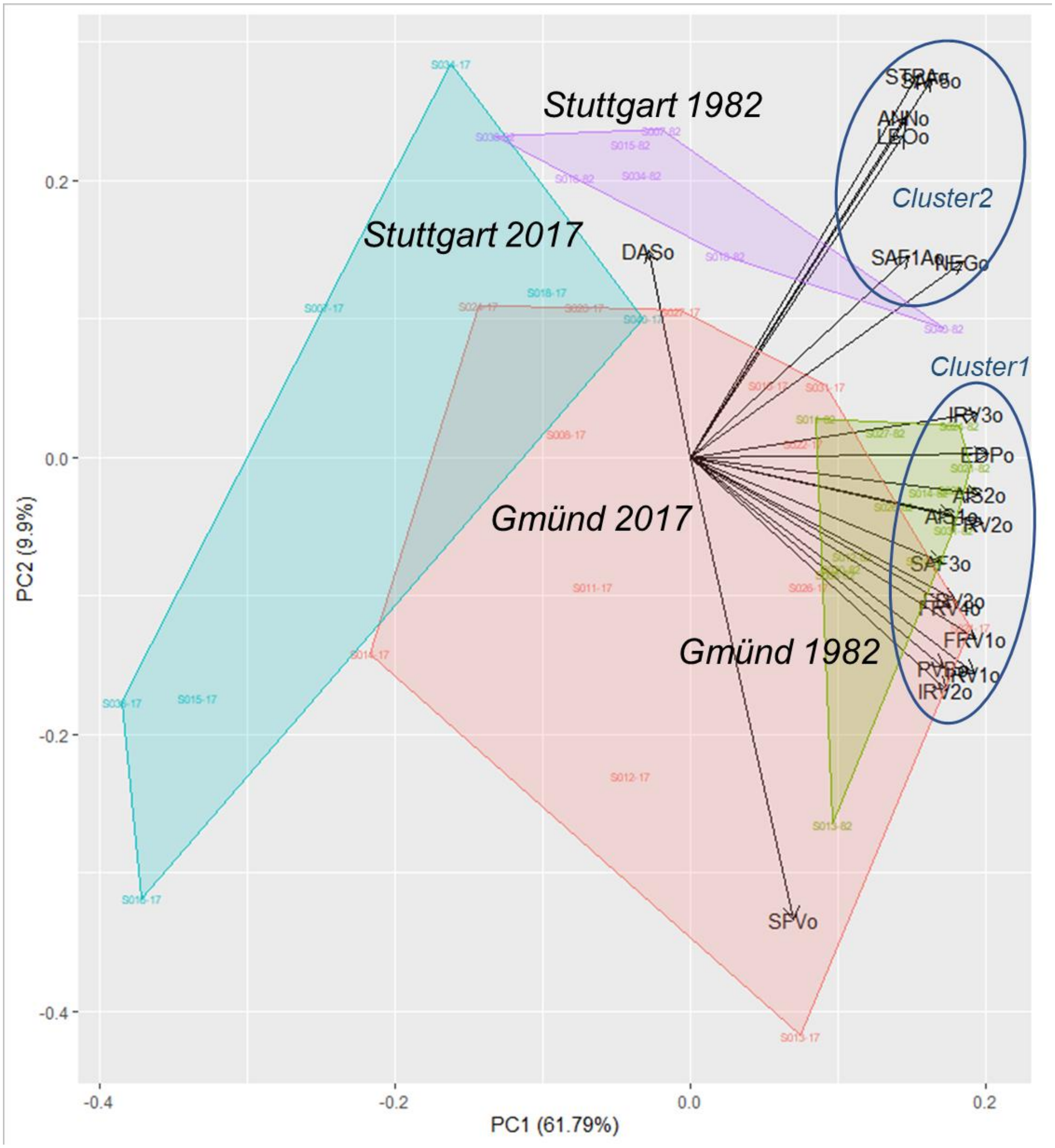

Figure 6. Principal Components Analysis (PCA) for both speech communities and recording years, showing loadings and clustering for the 20 linguistic features; the angle of the arrows indicates the degree of correlation while the length depicts the amount of explained variance.

In searching for the characteristic linguistic features that play a greater role in sociolectal coherence, Figure 6 exposes two important clusters. Cluster 1 is primarily defined by the first principal component: four front rounded vowels (FRV1, FRV2, FRV3, FRV4), two /ai/ diphthongs (AIS1, AIS2), three irregular verb stems (IRV1, IRV2, IRV3), and three morphosyntactic constructions (EDP, SAF3, PVB). These are the traditional, conservative features that are unique to the Swabian dialect (see Table 2); thus, we name this cluster: Traditional Swabian (TS).

The second large cluster of linguistic features lies between PC1 and PC2 and consists of features that are prevalent throughout southern Germany, including Bavaria to the east, Baden to the west, and Switzerland to the south: palatalisation of coda /-st/ (STPA), diminutive suffix '-le' (SAF1A), dropping of past participle 'ge-' (SAF5), negative markers (NEG), and lowering of /e:/ (LEO). Since these dialect features are common in many southern German varieties, we name this cluster: Southern Regional (SR). There is, however, one anomaly in the SR cluster: nasalisation of /an/ (ANN), which is a traditional Swabian feature (Griffen 1992). With nasalisation there appears to be both a local and an individual aspect at play: some speakers and some localities in Swabia are simply more nasal than others. In the current corpus, Schwäbisch Gmünd exhibits more nasalisation than Stuttgart (51% versus 38%), but, in both communities, nasalisation is on the decline (54% in 1982 versus 39% in 2017), with Schwäbisch Gmünd 1982 showing the highest level of nasalisation (60%) and Stuttgart 2017 the lowest (34%).

This leaves two remaining linguistic features to be explained. The definite neuter article (DAS) and stop-fricative variation (SFV) are both strongly aligned along the second principal component, one positive and one negative. While these features are also regional, they are spreading into Swabian from northern German varieties (Spiekermann 2008:186), rather than from other southern varieties, hence, we name this group: Northern Influence (NI). Further support for the difference in these two features can be seen in Figures 2 through 4, which show the weakest (some negative) correlations for SFV and DAS with respect to the other features.

Linguistic Features	Cluster	Schwäbisch Gmünd		Stuttgart	
		1982	2017	1982	2017
FRV2 – Unrounded Diphthong /eu/	TS	.854	.852	.460	1.000
IRV3 – Irregular Verb ‘haben’	TS	.002	.783	.797	.930
EDP – Plural Verb Inflection ‘-ed’	TS	.600	.842	.797	.797
PVB – Periphrastic Subjunctive	TS	.015	.901	.611	.791
FRV1 – Unrounded Front Vowel /ö/	TS	.218	.617	.797	.617
AIS1 – Diphthong Shift /ai/ (MHG /i:/)	TS	.479	.140	.675	.617
FRV4 – Diphthongisation /u/ (MHG /uo/)	TS	.198	.558	.797	.562
AIS1 – Diphthong Shift /ai/ (MHG /ei/)	TS	.229	.744	.735	.510
FRV3 – Unrounded Front Vowel /ü/	TS	.842	.550	.184	.287
SAF3 – Swabian Prefix ‘nââ’	TS	.025	.568	.375	.064
IRV1 – Irregular Verb ‘gehen’	TS	.274	.842	.862	.062
IRV2 – Irregular Verb ‘stehen’	TS	.581	.568	.531	.021
SAF5 – Past-participle Marker ‘ge-’	SR	.399	.234	.327	.797
LEO – Long /e:/ Lowering	SR	.725	.198	.032	.735
ANN – Nasalisation /an/	SR	.218	.239	.460	.675
SAF1A – Diminutive Suffix ‘-le’	SR	.136	.137	.154	.617
STPA – Palatalisation syllable coda /-st/	SR	.005	.435	.011	.562
NEG – Negative Marker ‘et/net/nette’	SR	.617	.560	.862	.413
DAS – Definite Neuter Article ‘des’	NI	.542	.002	.287	.460
SFV – Stop-Fricative Variation /-ig/	NI	.070	.008	.103	.062

Table 7. Proportion of Explained Variances (Spearman’s *r*-squared). Correlation of frequency of dialect variant with the first principal component; proportions over .60 are considered to be more “coherent” (shaded in grey).

Table 7 reports the proportions of explained variances (from Figure 6) for the first principal component for each of the 20 linguistic features for the two communities and the two time periods, sorted by cluster and then Stuttgart 2017 (the most standard of the four sociolects). Proportions over .60 are considered to be more "coherent" and are shaded in grey. Table 7 reveals the strongest coherence with the Traditional Swabian features, followed by the Southern Regional features, and finally the Northern Influence features, in which there little to no covariation. These findings provide strong support for the concept of covariation as a measure of coherence and underscores the importance of considering the sociohistorical linguistic background of the features being investigated.

5. Conclusion

Our aim with this study has been twofold: to examine the theoretical underpinnings of covariation as a measure of sociolectal coherence and to extend empirical analyses of coherence to another language variety and to a broader set of linguistic features. In so doing, we have probed three key questions: First, do variables covary within a sociolect? Our results show that they indeed do, however, coherence is relative, so the crucial question becomes: to what extent do variables covary? This leads to our second key question, are some language varieties more coherent than others? In our analysis of four Swabian sociolects separated by place and time, we found greater coherence in varieties closer to the standard language signalling that 'consistency' and 'prestige' promote less variation (Woo et al., this volume). Finally, in exploring our third question, which linguistic features pull the most "coherence weight", we have seen that features cluster for reasons which can be unearthed in the sociohistorical context. The Swabian dialect is a conservative linguistic variety with a long tradition evoking sundry and opposing images such as "inventive", "hard-working" and "thrifty" but also "backward", "uneducated" and "stingy".

Hence, it is not surprising that the conservative Swabian features would respond differently than the innovative ones entering the dialect from other varieties of German.

We investigated seven external factors and found five to be strong predictors of sociolectal coherence: community membership (Stuttgart), recording year (2017), variable type (Swabian), variable stage (changing), and level of stigma (low). Linguistic level and perceptual salience showed no significant effect on covariation measures. We uncovered significant interaction among the Swabian-specific features that are undergoing change (i.e., higher levels of attrition) and have a low level of stigma (hence, below the level of consciousness). Perhaps, as Guy (2013, 2014) has suggested, the more highly stigmatised features, which are obviously more perceptually salient, are reserved for identity formation, style indexicalities, and stance-taking (e.g., Swabian nasalisation). Guy's approach resolves the disconnect between the concept of a speech community, which promotes collective coherence, and the notion of speaker agency, which advocates individual choice. Our findings provide support for Guy's proposal that the sets of variables that covary belong to the community, while those that do not belong to the individual.

In the end, coherence is a matter of relativity and degree, not absoluteness (Tomaschek, Hendrix, and Baayen 2018). There has been a change over time in the relative level of coherence for the four sociolects, with both Schwäbisch Gmünd and Stuttgart becoming "more coherent" in 2017 than they were in 1982 as the dialect has levelled to the standard language. This leads us to speculate that, in the ensuing years, as Swabian continues to converge toward the standard language, the sociolects will become even more coherent, as more of the changing variables stabilise. Yet, as we have known since the Roman scholar Marcus Terentius Varro (116-27 BC) observed, *consuetudo loquendi est in motu,* 'the vernacular is always in motion' (see Taylor

1975). As some variables stabilise, others begin to change, keeping sociolinguists constantly on the move in search of the cognitive coherence of sociolects.

6. References

AdA. 2011. “Atlas Zur Deutschen Alltagssprache (AdA).” *Website from Universität Salzburg and Université Liège*.

Ammon, Ulrich and Uwe Loewer. 1977. *Schwäbisch. Dialekt/Hochsprache Kontrastiv. Sprachhefte Für Den Deutschunterricht 4*. Düsseldorf: Schwann.

Auer, Peter. 1997. “Co-Occurrence Restrictions between Linguistic Variables: A Case for Social Dialectology, Phonological Theory and Variation Studies.” Pp. 69–100 in *Variation, Change, and Phonological Theory*. Vol. 4, edited by F. Hinskens, R. van Hout, and W. L. Wetzels. Amsterdam: John Benjamins.

Auer, Peter. 2005. “Europe’s Sociolinguistic Unity, or: A Typology of European Dialect/Standard Constellations.” Pp. 7–42 in *Perspectives on Variation: Sociolinguistic, Historical, Comparative*, edited by N. Delbecque, J. van der Auwera, and D. Geeraerts. Berlin: Mouton de Gruyter.

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.

Beaman, Karen V. 2018. “Mobility and Dialect Identity as Factors in Language Change: Findings from a 35-Year Panel and Trend Study in Swabia.” in *Urban Language Research Conference*. University of Graz, Austria, November 1-3, 2018.

Beaman, Karen V. 2020. “Coherence in Real- and Apparent-Time: A Sociolinguistic Study of Swabian. Unpublished PhD Thesis.” Queen Mary University of London.

Beaman, Karen V. and Fabian Tomaschek. 2021. “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*, edited by K. V. Beaman and I. Buchstaller. New York: Routledge.

Bock, Kathryn and Zenzi M. Griffin. 2000. “The Persistence of Structural Priming: Transient Activation or Implicit Learning?” *Journal of Experimental Psychology* 129(2):177–92.

Box, George E. P. 1976. “Science and Statistics.” *Journal of the American Statistical Association* 71(356):791–99.

Britain, David. 2016. “Sedentarism and Nomadism in the Sociolinguistics of Dialect.” Pp. 217–41 in *Sociolinguistics: Theoretical Debates*, edited by N. Coupland. Cambridge: Cambridge University Press.

Cheshire, Jenny. 1999. “Taming the Vernacular: Some Repercussions for the Study of Syntactic Variation and Spoken Grammar.” *Cuadernos de Filogía Inglesa* 8:59–80.

Cheshire, Jenny. 2003. “Social Dimensions of Syntactic Variation: The Case of ‘when’ Clauses.” Pp. 245–61 in *Social Dialectology: In Honour of Peter Trudgill*, edited by J. Cheshire and D. Britain. Amsterdam: John Benjamins.

DiWA. 2001. “Digitaler Wenker-Atlas (DiWA). Ausgabe von Georg Wenkers ‘Sprachatlas Des Deutschen Reichs’. 1888–1923.”

Elspaß, Stephan and Robert Möller. 2003. “Atlas Zur Deutschen Alltagssprache (AdA).” *Liége*

Université and Universität Salzburg.
Erker, Daniel. 2017. "The Limits of Named Language Varieties and the Role of Social Salience in Dialectal Contact: The Case of Spanish in the United States." *Language and Linguistics Compass* 11(1):1–20.
Ethnologue. 2019. "Swabian" edited by D. M. Eberhard, G. F. Simons, and C. D. Fennig. *Ethnologue: Languages of the World. Twenty-Second Edition. SIL International*.
Frey, Eberhard. 1975. *Stuttgarter Schwäbisch: Laut- Und Formenlehre Eines Stuttgarter Idiolekts*. Marburg: N. G. Elwert Verlag.
Friendly, Michael. 2002. "Corrgrams: Exploratory Displays for Correlation Matrices." *American Statistician* 56(4):316–24.
Griffen, Toby D. 1992. "Nasalization in Swabian: Studies on the Nasalization Process." *Working Paper*.
Guy, Gregory R. 2013. "The Cognitive Coherence of Sociolects: How Do Speakers Handle Multiple Sociolinguistic Variables?" *Journal of Pragmatics* 52:63–71.
Guy, Gregory R. and Frans Hinskens. 2016. "Linguistic Coherence: Systems, Repertoires and Speech Communities." *Lingua* 172–173:1–9.
Van Hofwegen, Janneke and Walt Wolfram. 2010. "Coming of Age in African American English: A Longitudinal Study." *Journal of Sociolinguistics* 14(4):427–55.
Horvath, Barbara M. and David Sankoff. 1987. "Delimiting the Sydney Speech Community." *Language in Society* 16(2):179–204.
Kerswill, Paul and Ann Williams. 2002. "'Salience' as an Explanatory Factor in Language Change: Evidence from Dialect Levelling in Urban England." Pp. 81–110 in *Language Change: The Interplay of Internal, External and Extra-Linguistic Factors*, edited by M. C. Jones and E. Esch. Berlin: Walter de Gruyter.
Labov, William. 1966. "The Linguistic Variable as a Structural Unit." *Washington Linguistics Review* 3:4–22.
Labov, William. 1984. *Field Methods of the Project on Linguistic Change and Variation*.
Labov, William. 2003. "Pursuing the Cascade Model." Pp. 9–22 in *Social Dialectology: In Honour of Peter Trudgill*, edited by D. Britain and J. Cheshire. Amsterdam: John Benjamins.
Levon, Erez and Isabelle Buchstaller. 2015. "Perception, Cognition, and Linguistic Structure: The Effect of Linguistic Modularity and Cognitive Style on Sociolinguistic Processing." *Language Variation and Change* 27(3):319–48.
Levon, Erez and Sue Fox. 2014. *Social Salience and the Sociolinguistic Monitor: A Case Study of ING and TH-Fronting in Britain*. Vol. 42.
Meyerhoff, Miriam and Steffen Klaere. 2017. "A Case for Clustering Speakers and Linguistic Variables: Big Issues with Smaller Samples in Language Variation." Pp. 23–46 in *Language Variation - European Perspectives VI*, edited by I. Buchstaller and B. Siebenhaar. John Benjamins.
Meyerhoff, Miriam and James A. Walker. 2006. "The Persistence of Grammatical Constraints: 'Urban Sojourners' from Bequia." *University of Pennsylvania Working Papers in Linguistics* 12(2):131–43.
Meyerhoff, Miriam and James A. Walker. 2013. "An Existential Problem: The Sociolinguistic Monitor and Variation in Existential Constructions on Bequia (St. Vincent and the Grenadines)." *Language in Society* 42(4):407–28.
Milroy, Lesley. 1987. *Language and Social Networks*. 2nd Ed. Oxford: Basil-Blackwell.

Montgomery, Chris and Emma Moore. 2018. “Evaluating S(c)Illy Voices: The Effects of Salience, Stereotypes, and Co-Present Language Variables on Real-Time Reactions to Regional Speech.” *Language* 94(3):629–61.
Murdoch, D. J. and E. E. Chow. 1996. “A Graphical Display of Large Correlation Matrices.” *The American Statistician* 50(2):178–80.
Naro, Anthony J. 1981. “The Social and Structural Dimensions of a Syntactic Change.” *Language* 57(1):63–98.
Newlin-Lukowicz, Luiza. 2016. “Co-Occurrence of Sociolinguistic Variables and the Construction of Ethnic Identities.” *Lingua* 172–173:100–115.
Nycz, Jennifer. 2016. “Awareness and Acquisition of New Dialect Features.” Pp. 62–79 in *Awareness and Control in Sociolinguistic Research*, edited by A. Babel. Cambridge: Cambridge University Press.
Oetting, Janna B. and Janet L. McDonald. 2002. “Methods for Characterizing Participants’ Nonmainstream Dialect Use in Child Language Research.” *Journal of Speech, Language, and Hearing Research* 45(3):505–18.
Oushiro, Livia. 2016. “Social and Structural Constraints in Lectal Cohesion.” *Lingua* 172–173:116–30.
Oushiro, Livia and Gregory R. Guy. 2015. “The Effect of Salience on Co-Variation in Brazilian Portuguese.” *University of Pennsylvania Working Papers in Linguistics (Selected Papers from NWAV43)* 21(2):156–66.
Rácz, Péter. 2013. *Salience in Sociolinguistics: A Quantitative Approach*. Berlin: De Gruyter Mouton.
Russ, Charles V. J. 1990. *The Dialects of Modern German: A Linguistic Survey*. Oxford: Routledge.
Scherre, Maria Marta Pereira and Anthony J. Naro. 1992. “The Serial Effect on Internal and External Variables.” *Language Variation and Change* 4(1):1–13.
Schwarz, Christian. 2015. *Phonologischer Dialektwandel in Den Alemannischen Basisdialekten Südwestdeutschlands Im 20. Jahrhundert: Eine Empirische Untersuchung Zum Vokalismus*. Stuttgart: Franz Steiner Verlag.
Sharma, Devyani. 2005. “Dialect Stabilization and Speaker Awareness in Non-Native Varieties of English.” *Journal of Sociolinguistics* 9(2):194–224.
Spiekermann, Helmut. 2008. *Sprache in Baden-Württemberg. Merkmale Des Regionalen Standards*. Tübingen: Max Niemeyer Verlag.
Steiger, James H. 1980. “Testing Pattern Hypotheses on Correlation Matrices: Alternative Statistics and Some Empirical Results.” *Multivariate Behavioral Research* 15(3):335–52.
Tamminga, Meredith. 2019. “Interspeaker Covariation in Philadelphia Vowel Changes.” *Language Variation and Change* 31(2):119–33.
Taylor, Daniel J. 1975. *Declinatio: A Study of the Linguistic Theory of Marcus Terentius Varro (Studies in the History of Linguistics, Volume 2)*. Amsterdam: John Benjamins.
Tomaschek, Fabian, Peter Hendrix, and R. Harald Baayen. 2018. “Strategies for Addressing Collinearity in Multivariate Linguistic Data.” *Journal of Phonetics* 71:249–67.
Toutanova, Kristina, Dan Klein, Christopher D. Manning, and Yoram Singer. 2003. “Feature-Rich Part-of-Speech Tagging with a Cyclic Dependency Network.” *Proceedings of HLT-NAACL* 252–59.
Trudgill, Peter. 1974. “Linguistic Change and Diffusion: Description and Explanation in Sociolinguistic Dialect Geography.” *Language in Society* 2:215–46.

Trudgill, Peter. 1986. *Dialects in Contact*. Oxford: Blackwell Publishing.
Wittenburg, Peter, Hennie Brugman, Albert Russel, Alex Klassmann, and Han Sloetjes. 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.
Woo, Joyce, Tim Gadanidis, and Naomi Nagy. n.d. “SUBMITTED. Co-Variation in Toronto’s Heritage Cantonese: Primarily a Phonological Process.”