Paper submitted to Proceedings of 10th International Conference on Language Variation in Europe (ICLaVE10) held on June 26-28, 2019 at the Fryske Akademy in Leeuwarden, Netherlands. Nanna Haug Hilton, Remco Knooihuizen, and Hans Van de Velde (editors). Series on Studies in Language Variation. John Benjamins Publishing House.

Exploring an approach for modelling lectal coherence

by Karen V. Beaman
Queen Mary University of London and Eberhard Karls Universität Tübingen

Abstract

This paper presents an exploratory approach for modelling and measuring the concept of lectal coherence – the logical unity of idiolects, dialects, sociolects, regiolects, etc. – and how coherence can shape variation and foster or constrain language change. Twelve phonological and morpho-syntactic features of Swabian, a dialect spoken in southwestern Germany, exemplify differences in lectal coherence across two communities (Stuttgart and Schwäbisch Gmünd) and two points in time (1982 and 2017). Following the traditional quantitative variationist approach pioneered by Labov (1963), coupled with Guttman-like (1944) implicational scaling, and drawing on concepts from the order and lattice theory of mathematics (Partee, Ter Meulen, and Wall 1993, Ch. 11), the proposed model brings together three views of coherence – covariation, implicational scaling, and lattice theory – to demonstrate a holistic approach to the theory of linguistic coherence and its influence on language change. The research question this investigation explores is: does lectal coherence enable or inhibit linguistic change? The hypothesis to be tested through this study is that more coherent lects are less vulnerable to change and convergence while less coherent lects are more susceptible.

Keywords: sociolinguistics, language variation, language change, linguistic coherence, dialects, quantitative models, longitudinal studies, lifespan change, German.

Introduction

Fifty years ago, Weinreich, Labov, and Herzog (1968:188) observed that "idiolects do not provide the basis for self-contained or internally consistent grammars," rather it is the grammar of the speech community, governed by social factors, which reflects regularity and coherence and where linguistic change occurs. Hence, one approach toward explaining the regularity of linguistic variation and orderly heterogeneity is the notion of coherence. According to Guy & Hinskens (2016), the concept of orderly heterogeneity implies that "speech communities are sociolinguistically coherent .... [meaning that] the community should collectively behave in parallel: variants (or rates of use of variants) that index a given style, status, or a social characteristic should co-occur" (Guy & Hinskens 2016:2). These authors claim that "to the extent that linguistic variables systematically co-vary [i.e., exhibit similar frequencies and distributions], they can be characterized as displaying coherence" (Guy & Hinskens 2016:1).

Co-variation is one method for determining coherence; however, another approach utilises Guttman (1944) "scalogram analysis" to identify the underlying, orderly structure of the variation revealing implicational-like patterns (Bickerton 1973; DeCamp 1972; Fasold 1970; Greenberg 1963; Rickford 2002). A recent variation analysis using implicational scaling techniques is Ghyselen & Van Keymeulen's (2016) study of the Belgian dialect of *tussentaal*. These researchers found that, as a result of dialect loss, destandardisation, and demotisation, the dialect-standard constellation in Flanders has transformed from a diglossic into a largely diaglossic repertoire. They argue that *tussentaal* "is not just a random idiolectal mix of dialect features, but that it is structured by implicational principles shared across the speech community" (Ghyselen & Van Keymeulen 2016:15). In fact, "speakers do not randomly mix dialect features when speaking *tussentaal*; clear patterns were found whereby the presence of one dialect feature automatically implies the presence of other features" (Ghyselen & Van Keymeulen 2016:14).

Auer's (1997) concept of “co-occurrence restrictions” advocates a similar method for categorising repertoires and partitioning them “along continua of standard-dialect realizations” (Auer 1997:95). He maintains that tight, bi-directional co-occurrence restrictions (i.e., strong coherence) dichotomise lects while loose, uni-directional ones (i.e., weak coherence) promote greater variation which can stimulate language change (cf. Auer’s 'intermediate forms'). Remarking on the role of social factors, Auer adds, “it seems that given the appropriate social backing, any co-occurrence may be turned upside down” (Auer 1997:95). The overall concern with linear scaling, whether bi-directional or uni-directional, is in its strictness and inability to account for inherent linguistic variation or the influence of social factors. Hence, the challenge for the current study in characterising linguistic coherence is to generalise the concept of an implicational structure to one more representative and inclusive of the myriad factors influencing orderly heterogeneity.

The aim of this paper to explore a method for modelling and measuring linguistic coherence across varieties. A major assumption underlying this research is that greater lectal coherence implies that changes in one variant trigger changes in another variant such that multiple related variables co-occur within a unified variety. The overall hypothesis of this study is that more coherent lects are more resistant to change, while less coherent lects are more vulnerable to change, paralleling Milroy's (1987) findings that the most closed social networks are the least resistant to change. To test this hypothesis, a new methodological construct based on variable frequency analysis, implicational scaling, and lattice theory from abstract algebra is explored, called the Lectal Lattice.[1]

[1] The author wishes to thank James T. Garrett for suggesting the lattice concept to depict lectal coherence and for developing the initial R script to portray it. Of course, any deficiencies remaining are entirely my own.

Data and Methods

This section describes the data and methods employed in this investigation, covering the speech communities, data collection and preparation, the dependent linguistic variables, and the extra-linguistic predictors.

Speech Communities

This research investigates the use of Swabian or *Schwäbisch*, a High German dialect belonging to the Alemannic family, which is spoken by just over 800,000 people or one percent of the German population (see Figure 1). Two communities were selected for this research: the large international city of Stuttgart and its surrounding suburbs and the mid-sized town of Schwäbisch Gmünd and the neighbouring rural villages. Stuttgart, a large urban area with over one million inhabitants, is the heart of Swabia and home to many well-known global firms, including Daimler-Mercedes-Benz, Porsche, Bosch, and Siemens. With 60,000 inhabitants, Schwäbisch Gmünd lies 100 kilometres east of Stuttgart. A typical mid-sized German town, *Gmiind*, as it is called by the locals, is surrounded by small rural villages with 77% of the land dedicated to woodland and agriculture.

Attitudes toward Swabian vary: it is either loved or loathed. It is highly stigmatised by some and adored by others, as these two quotations from native Swabians demonstrate:

> *(1)*
> *wenn i Urschwâbe hör, also die mã gar ned versteht, des denkt mã immer, des isch e Fremdsprache ja, … muss mã halt manchmal de Kopf schüttle, aber so find i des … kôi schlimme Sprâch … i find e Dialekt isch nie schlecht*
> ‘if I hear really old-Swabian, that you can‘t even understand, then you always think, that’s a foreign language, yeah, … sometimes you just have to shake your head, but I don‘t think it‘s a bad language … I think a dialect is never bad.’ (Bertha 1982)

(2)
meine Kinder schämen sich sogar heutzutage Schwäbisch, also die verbinden Schwäbisch mit irgendwas, was sie nicht möchten.… dieser dörfliche Zusammenhalt stoßen die eher ab.
'nowadays my children are actually ashamed of Swabian, well they associate Swabian with something they don't like…. they are more likely to reject this village solidarity' (Helmut 2017)

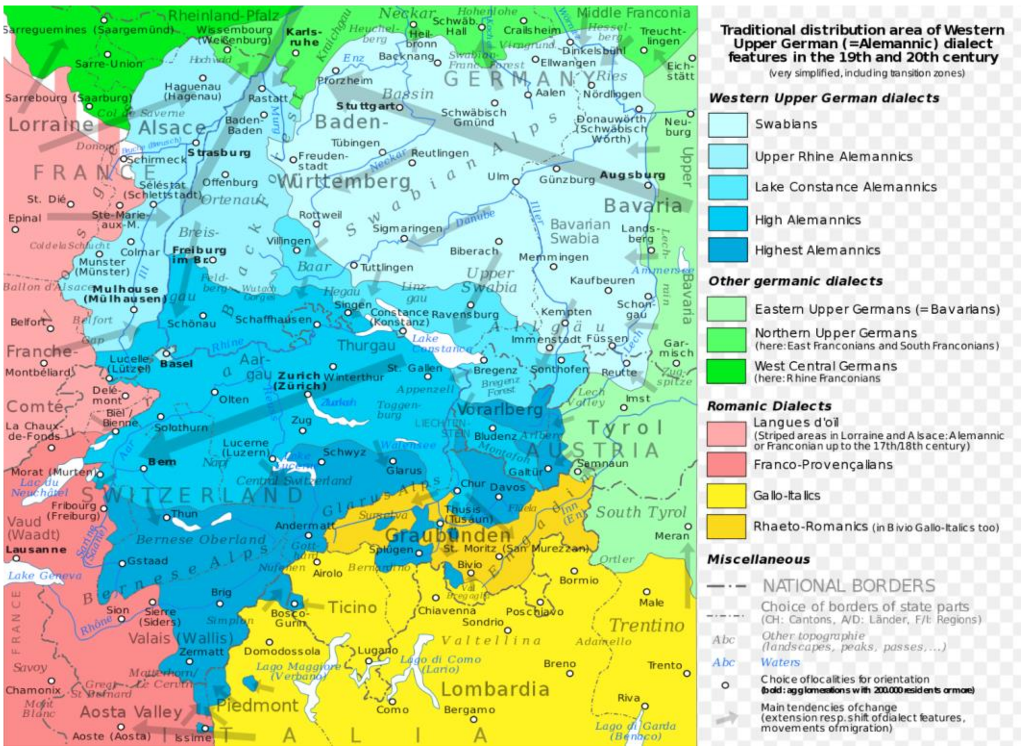

Figure 1. *Map of the Swabian and Alemannic Dialect families.*

Swabian Corpus

This data for this study are drawn from a real-time Panel Study of 20 native Swabian speakers, first recorded 1982 and then re-interviewed between 2017-2018: 13 speakers were from Schwäbisch Gmünd and 7 speakers from Stuttgart; 11 were women and 9 men; 16 speakers were between the ages of 18 and 25 in 1982 and between 53 and 60 in 2017, and four speakers were between 48 and 53 in 1982 and 83 and 88 in 2017; 14 of the 20 speakers completed their *Abitur*, the German college preparatory exam, and all were of a similar socio-economic status, quasi upper-middle class.

The data collection methods followed the Labovian sociolinguistic paradigm, consisting of semi-structured sociolinguistic interviews, conducted by native Swabian speakers with the primary investigator in attendance in the role of friend-of-a-friend (Milroy and Milroy 1985). To increase compatibility across years, the same survey instrument and interviewing techniques were used in both 1982 and 2017, covering questions about the speaker's childhood, games, friends, hobbies, local festivals and activities, and attitudes toward the Swabian language.

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 the Swabian dialect forms. All transcripts were verified by the principal investigator to ensure that standards were followed and to neutralise any potential transcriber bias. The dataset consists of 40 interviews (20 from 1982 and 20 from 2017), comprising 43 recorded hours for a total of 162,964 words, 72,550 in 1982 and 90,414 in 2017 (the interviews were slightly longer in 2017 and in Schwäbisch Gmünd).

Linguistic Variables

The dependent variables investigated in this study are 12 Swabian dialect features – six phonological and six morphosyntactic – all highly representative of the rich palette of features available to the Swabian speaker (see Figure 2). All variables were coded for a binary distinction between the dialect variant and the standard German variant. A brief description of each variable follows.

Code	Name	Swabian ~ Standard	Examples (Swabian Orthography)
	PHONOLOGICAL VARIABLES:		
AIS1	MHG /i:/ Diphthong Shift	[əɪ] ~ [aɪ]	*da dued mã in den Zylinder obe der Dêig nei* 'then you put [it] into the cylinder above the dough'
ANN	Nasalisation /an/	[ã] ~ [an]	*mã kã es mit em normale [Mehl] mache* 'one can make it with a normal [flour]'
FRV1	Unrounded Front Vowel /ö/	[ɛ] ~ [ø]	*so guet wie meeglich probier es* 'as good as possible [I] try it'
FRV2	Unrounded Diphthong /eu/	[aɪ] ~ [ɔɪ]	*bin gern auf Baim gestiege* '[I] liked to climb trees'
LEO	Long /e:/ Opening	[ɛː] ~ [eː]	*dâ e baar Jâhr lääbe* 'then live a few years there'
STP	Palatalisation coda /st/	[ʃt] ~ [st]	*da darfsch ja bloß hundertdreißig fahre in Italien* 'then you're only allowed to drive 130 in Italy'
	MORPHOSYNTACTIC VARIABLES:		
EDP	Plural Verb Inflection	[əd] ~ [ən]	*die finded es wichtich* 'they find it important'
IRV1	Irregular Verb 'gehen'	[gangə] ~ [ge:ən]	*willsch du an Telefon gange* 'do you want to answer the telephone'
IRV3	Irregular Verb 'haben'	[hɛn] ~ [ha:bən]	*mr hen e aldes Haus khet* 'we have had an old house'
PVB	Periphrastic Subjunctive	'dääd' ~ 'würde'	*es dääd beeinflusse* 'it should influence'
SAF1	Swabian Affix '-le'	'-le'~'-chen'/'-lein'	*dass er en Mädle mâg un se ihn mâg* 'that he likes a girl and she likes him'
SAF5	Swabian Affix 'ge-'	[θ] ~ [gə]	*un hen hier e Haus [ge]baut* 'and they have built a house here'

Figure 2. *Swabian linguistic variables under investigation.*

1. **Rounding of diphthong of MHG /i:/ origin** (AIS1) is a stereotypical feature of Swabian, hence standard German forms such as *klein* [klaɪn] 'small' and *allein* [alaɪn] 'alone' are realised as *glôi* [glɔɪ] and *allôi* [alɔɪ] in Swabian.
2. **Nasalisation of /a/ before /n/** (ANN) is a traditional feature of Swabian, hence words such as *man kann* [man kan] 'one can' and *Anfang* [anfaŋ] 'offer' are realised as *mã kã* [mã kã] and *Ãfang* [ãfɑŋ] in Swabian.
3. **Unrounding of the front vowel /ö/** (FRV1) is typical in Swabian, so that standard German words such as *möglich* [møːklɪç] 'possible' and *schön* [ʃøːn] 'beautiful' are realised as *meeglich* [mɛːglɪç] and *schee* [ʃɛː] in Swabian.
4. **Unrounding of the front vowel /ü/** (FRV3) 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].

5. **Opening of long /e:/** (LEO) is a regional dialect feature, so that standard German forms such as *lesen* ‘read’ [leːzn] and *Lehrer* ‘teacher’ [leːʁɐ] are pronounced as *lääse* [læːs] and *Läährer* [læːʁɐ] in Swabian.
6. **Palatalisation of /st/ in syllable-coda position** (STP) is a highly productive feature of Swabian and the Alemannic dialects. It is common in the second person singular verb formation, hence words such as *machst* [mɑxst] ‘you do/make’ and *nächst* [nɛːçst] ‘next’ are realised as *machscht* [maxʃ] and *nächsht* [nɛːçʃ] in Swabian.
7. **Present tense plural verb inflexion** (EDP) *-en* in standard German is realised as *-ed* in Swabian, so that standard forms such as *sie finden* ‘they find’ and *sie gehen* ‘they go’ are realised as *sie finded* and *sie ganged* in Swabian.
8. **Verb *gehen* ‘go’** (IRV1) has an irregular conjugation in Swabian, hence forms such as *ich gehe* ‘I go’ and *weitergehen* ‘continue’ are realised as *i gang* and *wêitergâht.*
9. **Verb *haben* ‘have’** (IRV3) has an irregular conjugation in Swabian, for example, the past participial has different realisations depending on voicing, *ghet, ghed, khet*, or *khed* in Swabian versus *gehabt* ‘had’ in standard German.
10. **Periphrastic subjunctive *tun* ‘do/make’** (PVB) is typical in Swabian, so forms such as *er dääd lache* ‘he would laugh’ and *es dääd beeinflusse* ‘it would influence’ vary with the standard German periphrastic subjunctive using *werden* ‘to become’, *er würde lachen* and *es würde beeinflussen.*
11. **Diminutive suffix *-le*** (SAF1) is highly productive and varies with the standard German suffix *-chen* (or the older affix *-lein).* Hence, forms like *Mädle* ‘little girl’, *Tellerle* ‘little plate’, and *Unterschiedle* ‘small difference’ vary with standard German forms *Mädchen, Tellerlein,* and *klein Unterschied*.

12. **Dropping of past participle prefix *-ge*** (SAF5) is a common feature in Swabian, hence forms such as *hen kriegt* 'have received' and *isch umzoge* 'has moved' vary with the standard German forms *haben gekriegt* and *ist umgezogen*.

Strict adherence to the *principle of accountability* was ensured through the use of a bespoke Swabian-German Lexicon (SGL), compiled from all words in the 40 transcripts which contained a token (Swabian or standard German) of one of the 12 Swabian features under investigation. False starts and repetitions were excluded. In total, 50,875 tokens for the 12 linguistic variables were extracted, 21,714 from 1982 and 29,161 from 2017, with an average of over 1,000 tokens per speaker in 1982 and over 1,400 tokens per speaker in 2017.

Extra-linguistic Predictors

Due to space limitations, only two extra-linguistic factors are discussed in the current study: (1) two recording years (1982 and 2017) and (2) two communities (Stuttgart and Schwäbisch Gmünd). Additional social factors influencing the Swabian dialect situation have been reported on elsewhere (Beaman 2020).

Analysis and Results

The analysis and results begin with an overview of the changing dialect situation in Swabia with respect to the 12 linguistic variables under investigation. Next, the Lectal Lattice is described and its construction explained, followed by an examination into the linguistic coherence in the two communities across the two recording periods. Finally, the differences and potential advantages of the Lectal Lattice over other linear models and graphical representations are discussed.

Linguistic Variables

Figure 3 shows the results of a generalised linear regression model (*glmer* function in the package *lme4* version 1.1-21) on the change in use of the 12 linguistic variables across the two recording periods. The phonological variables are on the left and the morphosyntactic ones on the right, sorted by decreasing probability of occurrence in 1982. As is quickly apparent, all variables indicate highly significant attrition across the two recording periods. We also see, with the exception of the two Swabian affixes ('-*le*' and '*ge*-'), that the morphosyntactic variables have receded significantly more than the phonological ones. Further investigation into the diminutive suffix (-*le*) suggests that this variable may be more lexical than morphological, while dropping of the past participle prefix (*ge*-) may be more a case of phonological reduction than a true morphological distinction.

Phonological Variables

Variable	Year	n	lodds	prob	diff	sig
STP	1982	4761	1.0209	73.5%	-14.8%	***
st~ʃt	2017	5716	0.3531	58.7%		
ANN	1982	2717	-0.3574	41.2%	-16.6%	***
an~ã	2017	3027	-1.1245	24.5%		
FRV3	1982	1747	-0.7085	33.0%	-15.6%	***
ɑɪ~ɔɪ	2017	2692	-1.5589	17.4%		
LEO	1982	1827	-0.7873	31.3%	-10.9%	***
e~æ	2017	3291	-1.3648	20.4%		
FRV1	1982	1365	-1.0740	25.5%	-13.1%	***
ø~e	2017	1401	-1.9615	12.3%		
AIS1	1982	3914	-1.5848	17.0%	-9.2%	***
aɪ~ɔɪ	2017	4975	-2.4723	7.8%		

Morphosyntactic Variables

Variable	Year	n	lodds	prob	diff	sig
EDP	1982	628	3.3772	96.7%	-37.3%	***
ən~əd	2017	954	0.3800	59.4%		
PVB	1982	122	0.7723	68.4%	-31.1%	***
'tun'	2017	181	-0.5178	37.3%		
IRV1	1982	266	0.7516	68.0%	-51.4%	***
gaŋə	2017	418	-1.6163	16.6%		
IRV3	1982	1022	0.2948	57.3%	-35.5%	***
hɛn	2017	1843	-1.2758	21.8%		
SAF1	1982	1707	-1.1095	24.8%	-12.9%	***
-lə	2017	2277	-1.9970	12.0%		
SAF5	1982	1638	-1.2181	22.8%	-11.1%	***
gə~ɸ	2017	2386	-2.0182	11.7%		

Figure 3. *Swabian linguistic variable predictions by recording year and variable type, generated by the R predict function based on the results of a linear regression model with glmer.*

Dialect Change in Swabia

Figure 4 depicts the changing dialect situation in Swabia by analysing the 12 linguistic variables using Principal Components Analysis (PCA) (*prcomp* function in package *stats*, version 3.5.3), a type of linear modelling which has been used by many sociolinguists as a heuristic for grouping speakers based solely on their linguistic behaviour (Horvath and Sankoff 1987). PCA reduces the dimensionality of multivariate data to a small set of derived factors (i.e., principal components), each representing a summarisation of the linguistic features that co-occur with high frequency. In Figure 4, PC1 for the phonological variables is plotted on the horizontal axis and PC1 for the morphosyntactic variables on the vertical axis. The upper right corner approximates 100% usage of all dialect variants, while the lower left corner verges toward 100% usage of the standard German variants. The plus signs mark each speaker's dialect usage in 1982, and the dots denote each speaker's usage in 2017. With two exceptions, all speakers have experienced dialect attrition as can be seen by the overall trend with the plus signs (1982) at the top of the graph (indicating greater dialect usage) and the dots (2017) at the bottom of the graph (revealing greater standard usage).

The dotted ellipse at the top of Figure 4 (drawn at two standard deviations from the mean of the group) encircles the speakers from Schwäbisch Gmünd in 1982. The small, compact nature of this ellipse indicates that there was considerable homogeneity among the speakers in 1982 with regards to the use of these dialect variants. The dashed ellipse in the middle of Figure 4 encircles all speakers in 1982, signalling a stronger tendency toward the use of standard variants when the speakers from the urban centre of Stuttgart are incorporated into the model. Finally, the longest ellipse encircles all speakers in 2017, highlighting two key findings: (1) the Swabian dialect has moved closer to the standard language in 2017 than it was in 1982, as seen by the placement of the pluses (in the upper right) and the dots (in the lower left), and (2) there is noticeably greater diversity in dialect and standard usage in 2017 than there was 1982 (as demonstrated by the size of the 2017 ellipse). Drawing on ethnographic observations, in 1982 both communities exhibited many, dense, multiplex social relationships, whereas by 2017, community ties have weakened and social connections become considerably more dispersed, particularly in Stuttgart. In fact, many of the Stuttgart speakers, who were all close friends in 1982, had completely lost contact with one another by 2017, requiring considerable espionage on the part of the principal investigator to locate these individuals to re-interview them.

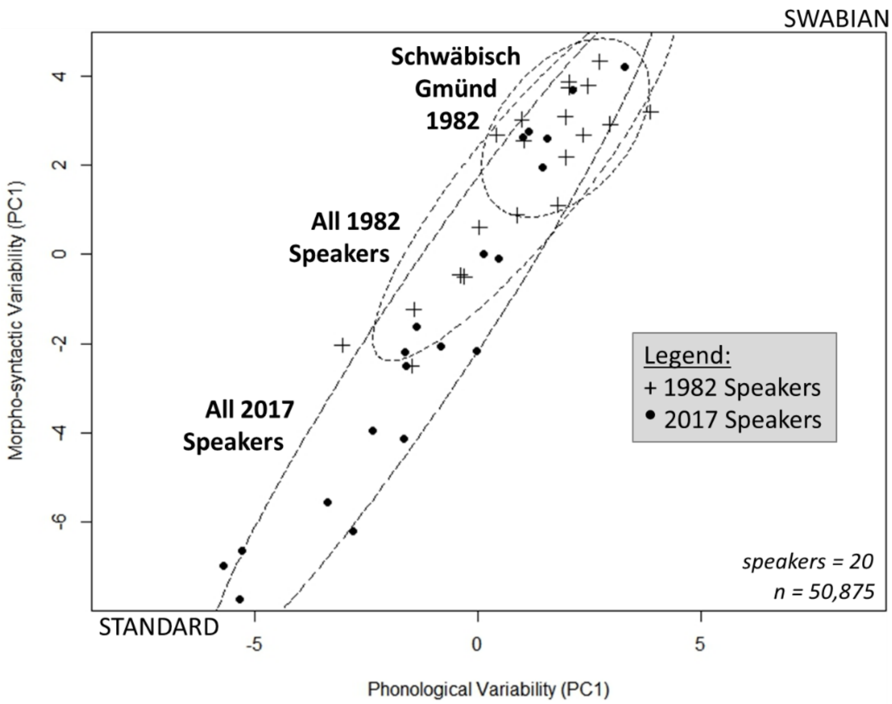

***Figure 4**. Change in Swabian dialect usage for 12 linguistic variables for two communities over a 35-year time span (1982 to 2017).*

Lectal Coherence

The theoretical question this paper seeks to address is to what extent do varying lects reflect coherence. By measuring the level of coherence in a given lect (e.g., the 1982 Schwäbisch Gmünd speakers) we can compare it with other lects (e.g., the 2017 Stuttgart speakers) and thereby examine the impact that coherence has on language variation and change. To address this challenge, this paper proposes a LATTICE, a concept drawn from both the order theory of mathematics and universal algebra (Partee et al. 1993). Linguists have used LATTICES in phonology, syntax, semantics, neurolinguistics and computational linguistics, but not yet in sociolinguistics or variation studies. A LATTICE is an abstract structure that uses binary relations to examine the hierarchical or implicational relationships within a given set of elements. A LATTICE generalises the data from a straight line (such as x implies y implies z) to a multi-dimensional picture, depicted by a Hasse diagram, as illustrated in Figure 5.

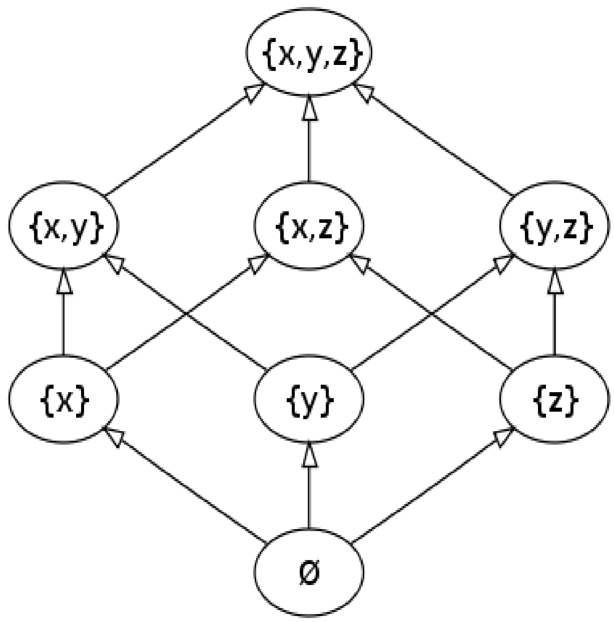

***Figure 5**. Lattice demonstrating sets and subsets as visualised by a Hasse diagram.*

A LATTICE consists of PARTIALLY ORDERED SETS, called POSETS, in which every two elements have a least upper bound, called a JOIN, and a greatest lower bound, called a MEET. The relationship between the variables is one of inclusion. For any two elements, you can move up the LATTICE to find an element that includes both (the JOIN) or step down the LATTICE to find an element that includes both (the MEET). LATTICES exhibit the principle of DUALITY, which means that they function equally in both directions – top-down or bottom-up. Thus, in turning a LATTICE upside down, the MEETS become JOINS, and the JOINS become MEETS.

Pairwise Comparisons

The first step in developing a Lectal Lattice is to create post hoc PAIRWISE COMPARISONS for each speakers' set of linguistic variables, arranged in two-by-two contingency tables. Figure 6 illustrates an example POSET for speaker Angela in 1982. The 12 linguistic variables (AIS1 through STP) generate a POSET of 144 pairs of variables. Using the Suissa & Shuster Exact test with the Holm-Bonferroni method (Holm 1979), each pair of variables is tested to determine whether there is a significant difference in frequency of usage. When a statistical difference is found (specifically, when the variable in the row is lower than the variable in the column maintaining the implicational order) the pair is assigned a 1, otherwise a 0 is assigned. In a POSET every pair of variables need not be related significantly for the pattern to be valid, allowing for uncertainties or inadequacies or unknowns in the dataset, which of course is common with sociolinguistic data.

Angela 1982

	AIS1	ANN	EDP	FRV1	FRV3	IRV1	IRV3	LEO	PVB	SAF1	SAF5	STP
AIS1	--	0	1	0	0	1	0	0	0	1	0	1
ANN	0	--	1	1	1	1	0	0	0	1	0	1
EDP	0	0	--	0	0	0	0	0	0	0	0	0
FRV1	0	0	1	--	0	0	0	0	0	1	0	0
FRV3	0	0	1	0	--	0	0	0	0	1	0	1
IRV1	0	0	0	0	0	--	0	0	0	0	0	0
IRV3	0	0	1	0	0	0	--	0	0	1	0	1
LEO	0	0	1	0	0	1	0	--	0	1	0	1
PVB	0	0	0	0	0	0	0	0	--	0	0	0
SAF1	0	0	0	0	0	0	0	0	0	--	0	0
SAF5	0	0	1	1	1	1	0	0	0	1	--	1
STP	0	0	0	0	0	0	0	0	0	0	0	--

→

Angela 1982 SORTED Rank = 28

	ANN	SAF5	AIS1	LEO	IRV3	FRV3	FRV1	PVB	IRV1	STP	EDP	SAF1
ANN	--	0	0	0	0	1	1	0	1	1	1	1
SAF5	0	--	0	0	0	1	1	0	1	1	1	1
AIS1	0	0	--	0	0	0	0	0	1	1	1	1
LEO	0	0	0	--	0	0	0	0	1	1	1	1
IRV3	0	0	0	0	--	0	0	0	0	1	1	1
FRV3	0	0	0	0	0	--	0	0	0	1	1	1
FRV1	0	0	0	0	0	0	--	0	0	0	1	1
PVB	0	0	0	0	0	0	0	--	0	0	0	0
IRV1	0	0	0	0	0	0	0	0	--	0	0	0
STP	0	0	0	0	0	0	0	0	0	--	0	0
EDP	0	0	0	0	0	0	0	0	0	0	--	0
SAF1	0	0	0	0	0	0	0	0	0	0	0	--

Figure 6*. One speaker's* POSET *illustrating pairwise comparisons for 12 linguistic variables. 0 = non-significant pair and 1 = significant pair based on Suissa & Shuster (1985) Exact test (p<.05) using the Holm-Bonferroni method.*

The speakers' POSETS are then sorted first by significant pairs and then according to the frequency of the dialect variant, generating a new sorted POSET as exemplified on the right in Figure 6. The sorted POSETS are RANKED by summing the significant pairwise comparisons. In Figure 6, Angela in 1982 has a RANK of 28 because there are 28 significant pairwise comparisons in the 12 linguistic variables under investigation. RANK allows us to calculate the DISTANCE between two different lects (two idiolects in this example), a value that denotes the number of pairs that would have to change for the two lects to be identical.

Angela 1982 SORTED Rank = 28

	ANN	SAF5	AIS1	LEO	IRV3	FRV3	FRV1	PVB	IRV1	STP	EDP	SAF1
ANN	--	0	0	0	0	1	1	0	1	1	1	1
SAF5	0	--	0	0	0	1	1	0	1	1	1	1
AIS1	0	0	--	0	0	0	0	0	1	1	1	1
LEO	0	0	0	--	0	0	0	0	1	1	1	1
IRV3	0	0	0	0	--	0	0	0	0	1	1	1
FRV3	0	0	0	0	0	--	0	0	0	1	1	1
FRV1	0	0	0	0	0	0	--	0	0	0	1	1
PVB	0	0	0	0	0	0	0	--	0	0	0	0
IRV1	0	0	0	0	0	0	0	0	--	0	0	0
STP	0	0	0	0	0	0	0	0	0	--	0	0
EDP	0	0	0	0	0	0	0	0	0	0	--	0
SAF1	0	0	0	0	0	0	0	0	0	0	0	--

V

Markus 1982 SORTED Rank = 28

	SAF5	ANN	LEO	AIS1	FRV1	FRV3	IRV3	PVB	IRV1	EDP	SAF1	STP
SAF5	--	0	0	1	1	1	1	0	1	1	1	1
ANN	0	--	0	0	0	0	0	0	1	1	1	1
LEO	0	0	--	0	0	0	0	0	1	1	1	1
AIS1	0	0	0	--	0	0	0	0	0	1	1	1
FRV1	0	0	0	0	--	0	0	0	0	1	1	1
FRV3	0	0	0	0	0	--	0	0	0	1	1	1
IRV3	0	0	0	0	0	0	--	0	0	1	1	1
PVB	0	0	0	0	0	0	0	--	0	0	0	0
IRV1	0	0	0	0	0	0	0	0	--	0	0	0
EDP	0	0	0	0	0	0	0	0	0	--	0	0
SAF1	0	0	0	0	0	0	0	0	0	0	--	0
STP	0	0	0	0	0	0	0	0	0	0	0	--

=

Node 101 1982 JOINED POSET Rank = 31

	SAF5	ANN	LEO	AIS1	IRV3	FRV1	FRV3	PVB	IRV1	EDP	SAF1	STP
SAF5	--	0	0	1	1	1	1	0	1	1	1	1
ANN	0	--	0	0	0	1	1	0	1	1	1	1
LEO	0	0	--	0	0	0	0	0	1	1	1	1
AIS1	0	0	0	--	0	0	0	0	1	1	1	1
IRV3	0	0	0	0	--	0	0	0	0	1	1	1
FRV1	0	0	0	0	0	--	0	0	0	1	1	1
FRV3	0	0	0	0	0	0	--	0	0	1	1	1
PVB	0	0	0	0	0	0	0	--	0	0	0	0
IRV1	0	0	0	0	0	0	0	0	--	0	0	0
EDP	0	0	0	0	0	0	0	0	0	--	0	0
SAF1	0	0	0	0	0	0	0	0	0	0	--	0
STP	0	0	0	0	0	0	0	0	0	0	0	--

***Figure 7**. Joining POSETS with nearest neighbours based on DISTANCE to create NODES in the lattice. 0 = non-significant pair and 1 = significant pair based on Suissa & Shuster (1985) Exact test (p<.05) using the Holm-Bonferroni method. Dark grey cells highlight pairs joined in creating the new POSET.*

In the next step, neighbouring POSETS, i.e., those that are most similar, are mathematically JOINED; specifically, all neighbours lying at the same minimum DISTANCE are joined one by one. Figure 7 demonstrates Angela's POSET being joined with Markus' POSET, to create a new POSET which will become NODE 101 in the LATTICE. In this example, there are seven joined pairs, indicated in dark grey, which is the mathematical DISTANCE between Angela and Markus' lects. To build the LATTICE, all POSETS are connected with their nearest neighbours and joined into new POSETS. It's POSETS within POSETS – or "turtles all the way down" – and up, of course, to maintain the DUALITY of the lattice.

The Lectal Lattice

Figure 8 presents a preliminary Lectal Lattice for the 20 Swabian speakers in 1982. The vertical axis represents the RANK, and the horizontal axis represents the FILE, i.e., the left-to-right

right line-up of the individual lects based on the first principal component. The LATTICE was created with standard R functions, including *plot*, *points*, *lines* and *text*. It is technically a SUB-LATTICE because it does not display all of the points in the LATTICE, rather only the significant ones. It is also a SEMI-LATTICE because it depicts only the JOINS or the upward trajectory and not the MEETS or the downward trajectory. This SUB-SEMI-LATTICE greatly simplifies the visualisation by eliminating redundant and irrelevant information.

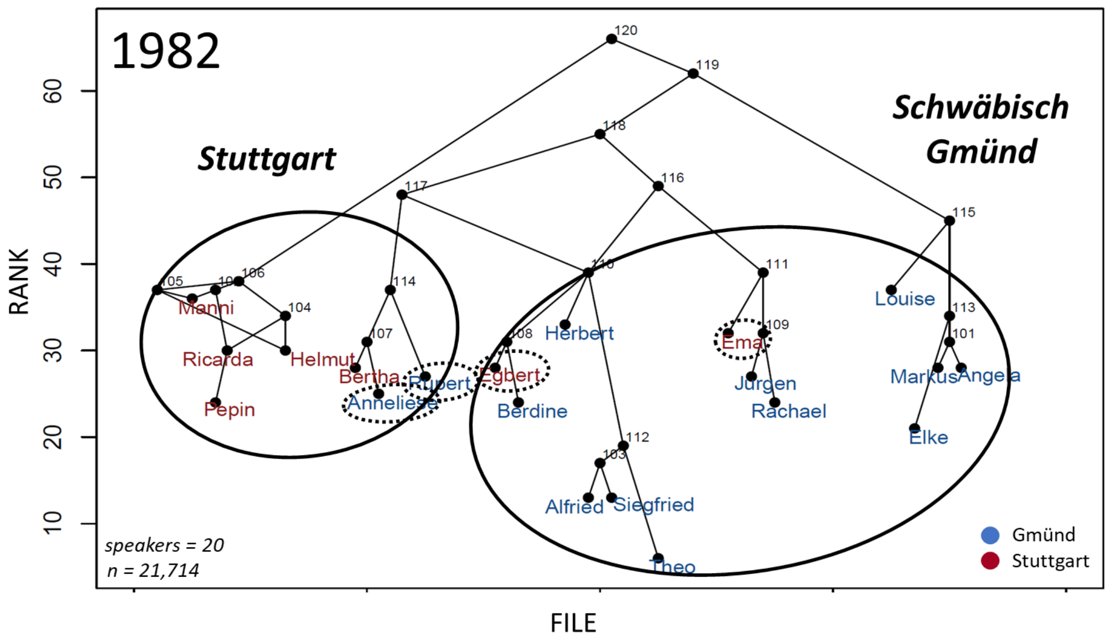

Figure 8. *Lectal Lattice for 20 Panel speakers in Stuttgart and Schwäbisch Gmünd in 1982.*

Each point in the Lectal Lattice represents a lect, either a single idiolect or a group of lects that have been joined, such as a dialect, a sociolect, a regiolect, or even a particular style, stance, or register. The points for each speaker's idiolect form the foundation of the lattice, which are labelled in Figure 8 with the speakers' pseudonyms. On the far right in Figure 8, NODE 101 from Figure 7 is visible showing the JOIN of Angela and Markus' lects.

From this picture, we can easily see that the speakers fall into two fairly distinct groups, the speakers from Stuttgart on the left and the speakers from Schwäbisch Gmünd on the right, with only a few exceptions or outliers, which can be explained. Rupert and Anneliese from Schwäbisch Gmünd are grouped with the more standardised supra-regional lect of Stuttgart. These two speakers are a few years older and have a higher level of education than the other speakers from Schwäbisch Gmünd (both are studying to be doctors, one a PhD, the other an MD), which may explain why they use more standardised forms than their cohorts. Ema from Stuttgart is grouped with the more conservative dialect speakers in Schwäbisch Gmünd. She is one of the oldest speakers, hence her dialect usage reflects a greater number of traditional Swabian forms.

Turning to 2017, Figure 9 presents the preliminary Lectal Lattice for the same 20 panel speakers 35 years later. On the right side of the graphic, we see some preservation of the conservative Schwäbisch Gmünd lect; however, on the left, we now see a very different picture. Over the 35-year timespan of this study, some speakers from Schwäbisch Gmünd have "fused" with speakers from Stuttgart, moving in the direction of greater standardisation, a supralocal lect or regiolect, a lect that is not necessarily geographically situated, rather one that is linguistically closer to the standard language. This finding provides support for the supposition that the Swabian dialect is undergoing levelling, changing from a geographical or horizontal variety to a sociolectal or vertical variety, as a result of the extensive social and demographic changes taking place in contemporary German society (Auer 2005).

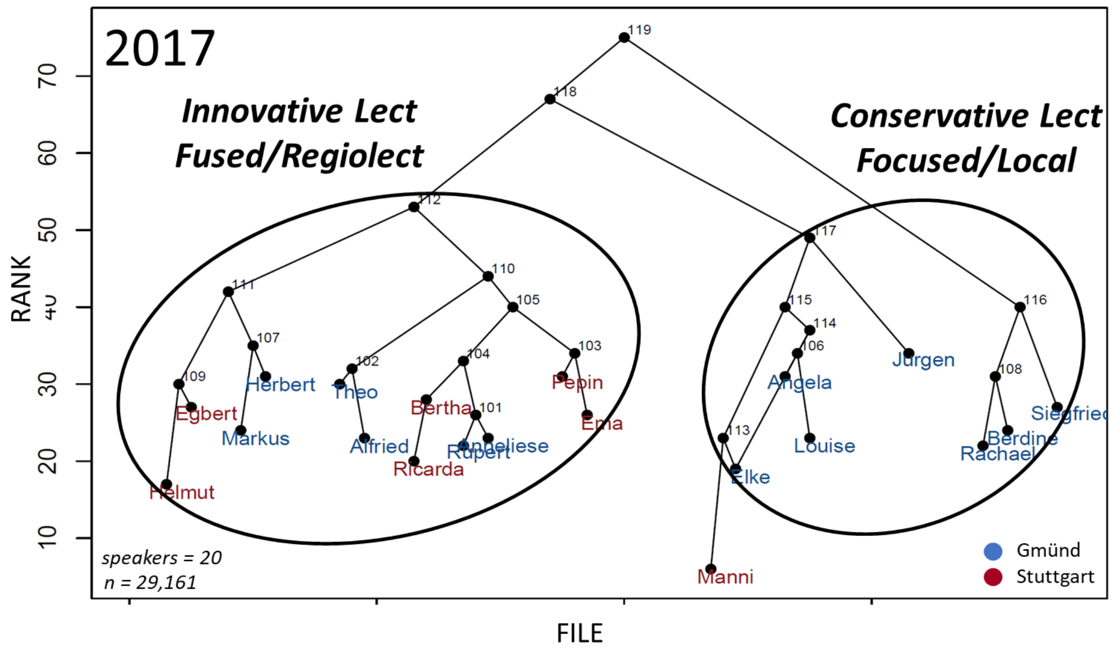

Figure 9. *Lectal Lattice for 20 Panel speakers in Stuttgart and Schwäbisch Gmünd in 2017.*

Still, Figure 9 reveals that more than half of the Schwäbisch Gmünd speakers have retained their conservative dialect features over the years. Considerable research has established that dialect attrition and retention is highly influenced by speakers' notions of local orientation (or 'dialect identity') and interlocutor accommodation (Auer and Hinskens 2005; Hinskens, Auer, and Kerswill 2005; Trudgill 1986), a phenomenon that has been reported on elsewhere (Beaman 2018, 2020). Many of the speakers who have moved away from the conservative lect to the fused regiolect are those who live and work regularly with speakers from other dialect groups, such as Markus and Rupert who both travel extensively across Germany for their work. Speakers who have retained the conservative dialect, such as Angela and Siegfried, maintain the strongest local orientation to Swabia (Baayen, Beaman, and Ramscar 2020; Beaman 2018, 2020; Beaman and Tomaschek 2020).

PCA and the Lattice

To illustrate the differences and advantages of the Lectal Lattice, we turn to a comparison with Principal Component Analysis (PCA). Figure 10 presents the standard PCA results for the 20 Swabian panel speakers, 1982 on the left and 2017 on the right. PC1 is on the horizontal axis and PC2 on the vertical axis. In 1982, PC1 and PC2 together account for 62% of the variation, and in 2017 PC1 and PC2 together account for 82% of the variation.

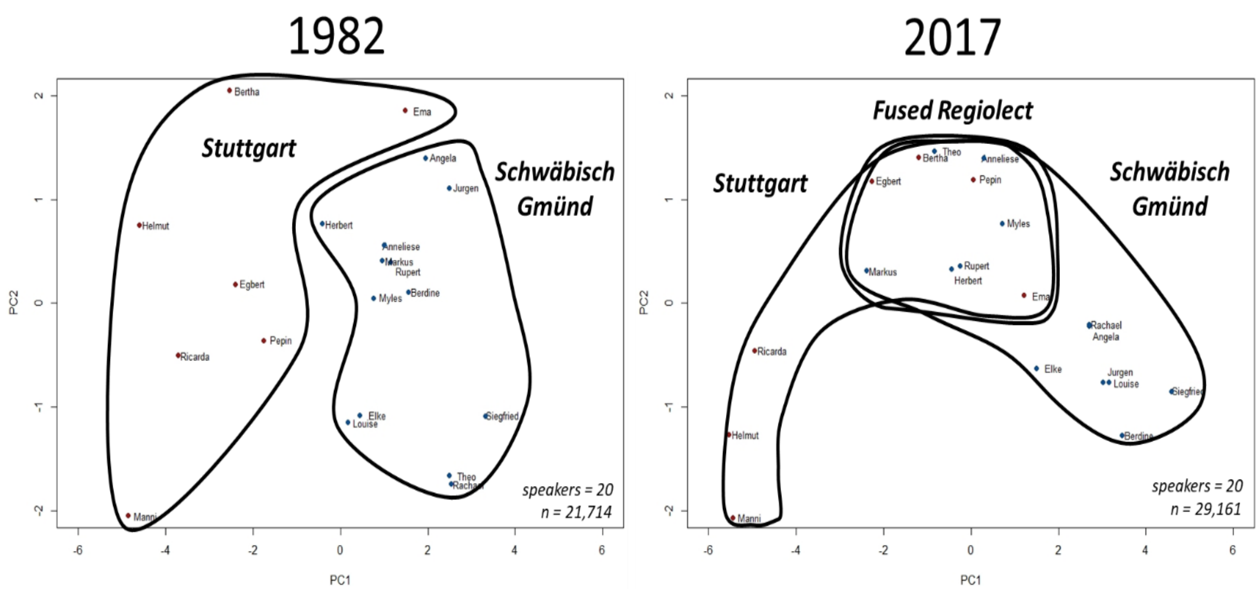

Figure 10. *Principal Components Analysis (PCA) depicting distinct Stuttgart and Schwäbisch Gmünd lects in 1982 and the fusing of lects in 2017.*

The PCA results are quite similar to the Lectal Lattice, albeit with a different graphical representation. In 1982, we again see two very distinct lects, Stuttgart speakers on the left and Schwäbisch Gmünd speakers on the right. The PCA for 2017 also depicts a changing picture of the dialect situation in Swabia. We can still delineate the Stuttgart and Schwäbisch Gmünd groups; however, we see a fusing of the two lects in the middle of the graph. The PCA in Figure 10 corroborates the findings from the Lectal Lattice in Figure 9, both of which reveal greater diversity among the Swabian speakers in 2017 than in 1982.

Both PCA and the Lectal Lattice are linear models that uncover significant groupings of speakers based on linguistic factors alone. One advantage of the Lectal Lattice over PCA is in the graphical display: the hierarchical depiction of related lects exposes the underlying relationships between lects without the researcher having to run multiple PCA models and manually compare and contrast the results. Another key advantage of the Lectal Lattice over PCA is in the calculation of the distance between lects: with principal components, distance is calculated based only on the frequencies of the variables; with the Lectal Lattice, distance is determined based on both the frequencies and the order of the variables. However, the greatest advantage of the Lectal Lattice over PCA and other linear models is the ability to calculate, measure and compare how closely different lects and groups of lects cohere, as explained in the following section.

Implicational Coherence

The motivation behind the Lectal Lattice is the development of a method to evaluate the coherence of lects, that is, how tightly (or loosely) multiple variables co-occur within a given lect. With a quantifiable, objective measure of coherence, we can test the overall hypothesis of this research that more coherent lects are more resistant to change, while less coherent lects are more vulnerable to change. The POSETS in the Lectal Lattice provide a method for quantitatively assessing the level of coherence in any given lect by measuring the number of significant pairwise comparisons that follow the implicational pattern. Quite simply, implicational coherence is calculated by summing the 1's above the diagonal (those following the implicational pattern), subtracting the 1's below the diagonal (those deviating from the pattern), and then dividing by the total number of significant pairs above the diagonal. The following formula describes the calculation for implicational coherence (IC):

$$IC = \frac{\sum_{i=1}^{n} x_i^{\omega} - \sum_{i=1}^{n} x_i^{\beta}}{\sum_{i=1}^{n} x^{i}}$$

To illustrate how the IC formula works, Figure 11 shows the POSETS for the top-most nodes from the Lectal Lattices in 1982 (NODE 120 from Figure 8) and 2017 (NODE 119 from Figure 9). Following the IC formula, in 1982 there are 60 significant pairwise comparisons above the diagonal and 6 below the diagonal, (60 – 6) / 66 = .818, signifying a highly coherent lect in 1982 at 82%. However, the POSET for the top-most node in 2017 reveals double the number of deviants (12 versus 6) and hence an implicational coherence of only 68% (a difference significant at p<.001 based on a chi-square test).

1982

Node 120 IC = 82% Rank = 66

	AIS1	SAF5	ANN	LEO	FRV1	FRV3	PVB	IRV1	IRV3	STP	SAF1	EDP
AIS1	--	1	1	1	1	1	1	1	1	1	1	1
SAF5	1	--	1	1	1	1	1	1	1	1	1	1
ANN	0	1	--	1	1	1	1	1	1	1	1	1
LEO	0	0	1	--	1	1	1	1	1	1	1	1
FRV1	0	0	1	0	--	0	0	1	1	1	1	1
FRV3	0	0	1	0	0	--	0	1	1	1	1	1
PVB	0	0	0	0	0	0	--	0	1	1	1	1
IRV1	0	0	0	0	0	0	0	--	1	1	1	1
IRV3	0	0	0	0	0	0	0	0	--	1	1	1
STP	0	0	0	0	0	0	0	0	1	--	0	1
SAF1	0	0	0	0	0	0	0	0	0	0	--	0
EDP	0	0	0	0	0	0	0	0	0	0	0	--

2017

Node 119 IC = 68% Rank = 75

	SAF5	AIS1	FRV3	FRV1	ANN	LEO	IRV3	IRV1	EDP	PVB	STP	SAF1
SAF5	--	1	1	1	1	1	1	1	1	1	1	1
AIS1	1	--	1	1	1	1	1	1	1	1	1	1
FRV3	0	1	--	1	1	1	1	1	1	1	1	1
FRV1	0	0	1	--	1	1	1	1	1	1	1	1
ANN	0	0	1	1	--	1	1	1	1	1	1	1
LEO	0	0	1	0	1	--	1	1	1	1	1	1
IRV3	0	0	1	0	1	1	--	1	1	0	1	1
IRV1	0	0	0	0	0	0	0	--	1	0	1	1
EDP	0	0	0	0	1	0	0	0	--	0	1	1
PVB	0	0	0	0	0	0	0	0	0	--	1	1
STP	0	0	0	0	0	0	0	0	1	0	--	1
SAF1	0	0	0	0	0	0	0	0	0	0	0	--

***Figure 11**. POSETS with pairwise comparisons for 12 variables for the top-most nodes in the 1982 and 2017 Lectal Lattices, demonstrating implicational coherence percentages.*

Figure 12 presents the Lectal Lattices from 1982 and 2017 with the implicational coherence percentages displayed for each node. It is interesting to note that all NODES in the 2017 lattice indicate lower levels of coherence than the NODES in the 1982 lattice, demonstrating the pervasive breakdown in coherence of the Swabian dialect over the last 35 years.

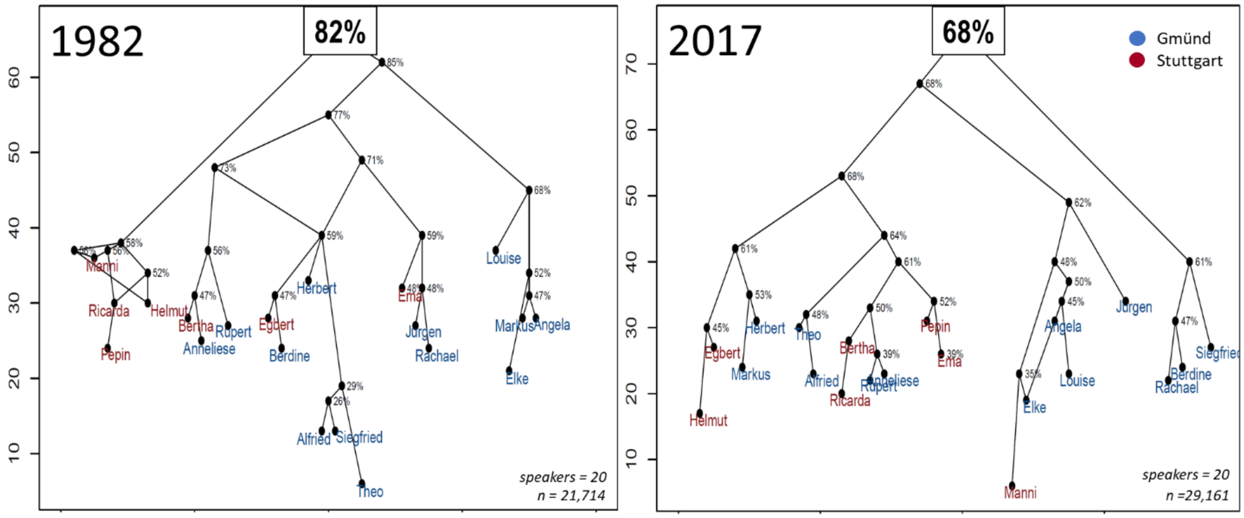

Figure 12. *Lectal Lattices for 20 Swabian Panel speakers in 1982 and in 2017 displaying implicational coherence for each node in the lattice.*

It is worth noting that an implicational scale can be drawn for any POSET using traditional notation to show the patterning of the variables for that lect. Following is the implicational pattern for Swabian spoken in 1982 (Node 120), both the main pattern and the deviant pattern:

$$\mathrm{AIS1} < \mathrm{SAF5} < \mathrm{ANN} < \mathrm{LEO} < \left\{\begin{matrix}\mathrm{FRV1}\\ \mathrm{FRV3}\end{matrix}\right\} < \left\{\begin{matrix}\mathrm{PVB}\\ \mathrm{IRV1}\end{matrix}\right\} < \mathrm{IRV3} < \mathrm{EDP1};$$

$$\mathrm{AIS1} > \mathrm{SAF5} > \left\{\begin{matrix}\mathrm{ANN}\\ \mathrm{FRV1}\\ \mathrm{FRV3}\end{matrix}\right\} > \mathrm{STP}$$

In short, the Lectal Lattice is based on implicational patterns that are derived from the ordering between variables, yet it also considers the frequency of the variables in deriving the order. It is also not as strict as implicational scaling because it allows for variation in the variable pairings by factoring in the effect of deviants rather than ignoring them.

Conclusion

Early in the variationist paradigm, Fasold (1970) argued that the combination of frequency analysis and implicational scaling leads to more revealing insights than either approach on its own (Fasold 1970:562). This paper advances Fasold's claim by presenting a preliminary approach for combining statistically significant frequency differences with implicational scales using a new mathematical construct to measure levels of lectal coherence between lects. The Lectal Lattice offers several benefits over other approaches in identifying unified lects and assessing coherence. First, it provides superior explanatory value over principal components by exposing the significant relationships between variables based on pairwise comparisons. Second, rather than a single linear chain, the Lectal Lattice is multidimensional, rendering a single graph that reveals the logical groupings and hierarchical ordering of similar lects. Third, LATTICE methodology with its variable POSETS proposes an independent statistical method for calculating the level of coherence of different lects. Fourth, the Lectal Lattice suggests a method for testing the hypothesis of this investigation that less coherent lects are more vulnerable to change, potentially providing insight into the actuation of linguistic change. Finally, LATTICE theory supports Weinreich, Labov, and Herzog that coherence or "orderly heterogeneity" is found in the aggregate grammar of the speech community rather than in the individual, reinforcing the widely-held premise that individuals in a community behave in parallel, reflecting regularity and coherence. As Lectal Lattice revealed, the Swabian of 2017 is significantly less coherent than the Swabian of 1982 (see Figure 12), demonstrating the breakdown in dialect usage over the last 35 years and suggesting that modern Swabian remains highly vulnerable to ongoing change and continued levelling with the standard language.

References

Auer, Peter. 1997. "Co-Occurrence Restrictions between Linguistic Variables: A Case for Social Dialectology, Phonological Theory and Variation Studies." *Amsterdam Studies in the Theory and History of Linguistic Science Series* 4:69–100.

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 and Frans Hinskens. 2005. "The Role of Interpersonal Accommodation in a Theory of Language Change." Pp. 335–57 in *Dialect Change: Convergence and Divergence in European Languages*, edited by P. Auer, F. Hinskens, and P. Kerswill. Cambridge: Cambridge University Press.

Baayen, R.Harald, Karen V. Beaman, and Michael Ramscar. 2020. "Deconfounding the Effects of Competition and Attrition on Dialect across the Lifespan: A Panel Study Investigation." 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.

Beaman, Karen V. 2018. "Identity and Place: The Changing Role of Swabian in Modern Germany." in *Conference on Language, Place and Periphery*. University of Copenhagen. January 18-19, 2018.

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)*, edited by A. Ziegler, S. Edler, N. Kleczkowski, and G. Oberdorfer. Amsterdam: John Benjamins Publishing Company.

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*, edited by K. V. Beaman and I. Buchstaller. New York: Routledge.

Bickerton, Derek. 1973. "The Nature of a Creole Continuum." *Language* 49(3):640–69.

DeCamp, David. 1972. "Toward a Generative Analysis of a Post-Creole Speech Continuum." Pp. 349–70 in *Pidginization and Creolization of Languages*, edited by D. Hymes. Cambridge: Cambridge University Press.

Fasold, Ralph W. 1970. "Two Models of Socially Significant Linguistic Variation." *Language* 46(3):551–63.

Ghyselen, Anne-Sophie and Jacques Van Keymeulen. 2016. "Implicational Scales in Colloquial Belgian Dutch." *Dialectologia et Geolinguistica* 24:1–17.

Greenberg, Joseph H. 1963. "Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements." Pp. 73–113 in *Universals of Language*, edited by J. H. Greenberg. London: MIT Press.

Guttman, Louis. 1944. "A Basis for Scaling Qualitative Data." *American Sociological Review*

9(2):139–50.

Guy, Gregory R. and Frans Hinskens. 2016. “Linguistic Coherence: Systems, Repertoires and Speech Communities.” *Lingua* 172–173:1–9.

Hinskens, Frans, Peter Auer, and Paul Kerswill. 2005. *The Study of Dialect Convergence and Divergence: Conceptual and Methodological Considerations*. edited by P. Auer, F. Hinskens, and P. Kerswill. Cambridge: Cambridge University Press.

Holm, Sture. 1979. “A Simple Sequentially Rejective Multiple Test Procedure.” *Scandinavian Journal of Statistics* 6(2):65–70. Retrieved (http://www.jstor.org/stable/10.2307/4615733).

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

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

Milroy, James and Lesley Milroy. 1985. “Linguistic Change, Social Network and Speaker Innovation.” *Journal of Linguistics* 21(2):339–84.

Milroy, Lesley. 1987. *Language and Social Networks*. Second Edition (first published in 1980). Oxford: Basil-Blackwell.

Partee, Barbara H., Alice G. ter Meulen, and Robert E. Wall. 1993. *Mathematical Methods in Linguistics (Vol. 30)*. Dordrecht: Kluwer Academic Publishers.

Rickford, John R. 2002. “Implicational Scales.” Pp. 142–67 in *The Handbook of Language Variation and Change*, edited by J. K. Chambers, P. Trudgill, and N. Schilling-Estes. Blackwell Publishing.

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

Weinreich, Uriel, William Labov, and Marvin I. Herzog. 1968. “Empirical Foundations for a Theory of Language Change.” Pp. 95–188 in *Directions for Historical Linguistics*, edited by W. Lehmann and Y. Malkiel. Austin: University of Texas.

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)* 1556–59.