



Deconfounding the Effects of Competition and Attrition on Dialect Across the Lifespan

Karen V. Beaman, R. Harald Baayen, & Michael Ramscar
Eberhard Karls Universität Tübingen

Corpora for Language and Aging Research (CLARe 4)
University of Helsinki, Finland
February 27 – March 1, 2019

Beaman, Baayen, and Ramscar – CLARe4 Helsinki – February 2019

Page 1

A considerable body of research shows that dialects are receding across the globe, and nowhere is this more evident than in Europe.

There are also widespread assumptions that, as individuals age, their mental **capabilities** “decline”, and as a consequence, they lose aspects of their language.

However, growing evidence from cognitive studies on aging and language usage indicates that, rather than lose linguistic forms, speakers actually gain extensive quantities of new lexical material over the course of their lifespan.

As people grow older, their knowledge naturally expands:

- they experience new things (e.g., in schools, on the job),
- they face various new life events (e.g., graduation, marriage, childbirth),
- they tackle new challenges (e.g., baking, mountain climbing).

As a result of these undertakings, they encounter new and original words which they add to their vocabulary to describe these experiences. Some linguists see language development as a process in which speakers obtain greater awareness of the standard language over their lifespan, gained through their increasing participation in various educational, commercial, and public institutions.

So the question we asked our ourselves: what if dialect is not really receding, rather it just appears so, because the standard language is expanding?

----- 1 min

Hypotheses



- 1) rather than lose dialect, speakers gain a massive amount of new lexical knowledge that is not spoken about in the dialect, which exerts a cumulative and competitive influence on their vocabularies and cognitive processing abilities; and
- 2) speakers are more likely to retain dialect forms when frequencies are high and words are drawn from *early* experiences, and to lose dialect forms when frequencies are low and words are more relevant to *later* life experiences.

So, we put forth two hypotheses:

[CLICK] (1) rather than lose dialect, speakers actually gain a massive amount of new lexical knowledge that is not spoken about in the dialect, which exerts a cumulative and competitive influence on their vocabularies and cognitive processing abilities;

[CLICK] (2) speakers are more likely to retain dialect forms when frequencies are high and words are drawn from early experiences, and to lose dialect forms when frequencies are low and words are more relevant to later life experiences.

Spoiler alert: as we will show, our results provide proof for the first hypothesis, but we were completely wrong about the second one!

----- 1 min

Swabian

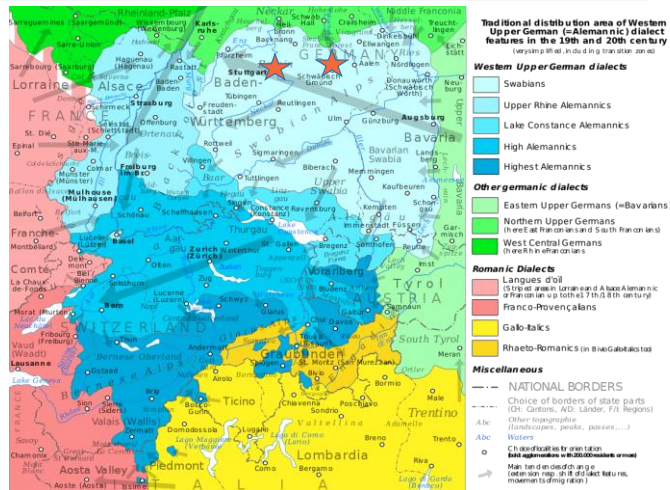


Swabian or Schwäbisch

is a High German dialect, belonging to the Alemannic family, spoken by just over 800,000 people.

Two communities:

- Stuttgart area
- Schwäbisch Gmünd



Beaman, Baayen, and Ramscar – CLARe4 Helsinki – February 2019

Page 3

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.

[CLICK] Two communities have been selected for this research:

- the large international city of Stuttgart and its surrounding suburbs
- the semi-rural, mid-sized town of Schwäbisch Gmünd and the surrounding rural villages.

----- 30 secs

Two Speech Communities



Beaman, Baayen, and Ramscar – CLARe4 Helsinki – February 2019

Page 4

Stuttgart is the heart of Swabia. It is a large urban area with over one million inhabitants and is home to many well-known global firms, such as Daimler-Mercedes-Benz, Porsche, Bosch, and Siemens.

[CLICK] Schwäbisch Gmünd lies 100 kilometers east of Stuttgart. With 60,000 inhabitants, it is a typical mid-sized German town, surrounded by small rural villages with 77% of the land dedicated to woodland and agriculture.

----- 30 secs

Some Swabian Features



Palatalization of coda -st

machst ~ *machsch* 'do/make'
gehst ~ *gehsch* 'go'
darfst ~ *darfsch* 'may'
nächst ~ *nächsch* 'next'
letzt ~ *letscht* 'last'
meistens ~ *meischstens* 'most'

Front Rounded Vowels

möglich ~ *meeglich* 'possible'
schön ~ *schee* 'pretty'
Bäume ~ *Baim* 'trees'
Freund ~ *Fraind* 'friend'
Küche ~ *Kiche* 'kitchen'
müde ~ *mide* 'tired'

Diphthong Shift

kein ~ *kôi* 'none'
gleich ~ *glôi* 'same'
allein ~ *allôi* 'alone'
daheim ~ *dahôim* 'at home'
weiß ~ *wôiẞ* 'I know'
nein ~ *nôi* 'no'

Irregular Verb Formation

gehen ~ *gange* 'go'
verstehe ~ *verstâh* 'understand'
stehen ~ *stande* 'stand'
wollen ~ *welle* 'want'
haben ~ *hen han khet* 'have'
tun ~ *doe* 'do/make'

I've identified over 30 linguistic variables that I'm investigating in Swabian: phonological, morpho-syntactic, lexical.

To give you a little taste, here four of the most productive and salient ones:

[CLICK] Palatalization of /st/ in syllable-coda position: *machst* and *gehst* are pronounced as *machsch* and *gehsch*.

[CLICK] There are a number of front rounded vowels that are unrounded in Swabian: *möglich* is *meeglich*, *Bäume* is *Baim*, *Küche* is *Kiche*.

[CLICK] Shifting of the /ai/ diphthong: words like *kein*, *allein*, *daheim* are pronounced as *kôi*, *allôi*, *dahôim*.

[CLICK] A number of irregular verbs: *gange* for *gehen*, *stande* for *stehen*, and *welle* for *wollen*.

----- 1 min

Swabian: Loved or Loathed



wenn i Urschwâbe hör, also die mǎ gar ned versteht, des denkt mǎ immer, des isch e Fremdsprache ja, ... muss mǎ halt manchmal de Kopf schüttle, aber so find i des ... kôî schlimme Sprach ... 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-82)

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 reject this village solidarity'

(Helmut-17)

Attitudes toward Swabian vary: it is either loved or loathed. It is highly stigmatized by some and adored by others, as these two quotations show:

[CLICK] Bertha in 1982, said: 'if I hear really old-Swabian, that you can't even understand, then you always think, that's a foreign language, yeah, ... sometimes you have to shake your head, but I don't think it's a bad language ... I think a dialect is never bad.'

[CLICK] Helmut in 2017, said: 'nowadays my children are actually ashamed of Swabian, well they associate Swabian with something they don't like.... they reject this village solidarity'

You'll notice a large number of dialect features in Bertha's comment and the complete absence in Helmut's; Bertha is one of the speakers who has changed her dialect the least over the years, and Helmut is one of those who's change the most. He's radio moderator for the local station and he says his kids laugh when he speaks Swabian.

----- 30 secs

Methods



- **Sociolinguistic Interviews**

- Labovian-style, casual interview questions, ca. one hour
- Native Swabian-speaking interviewers, “friend-of-friend”
- Same interview instrument and same topics discussed in 1982 and 2017

- **Transcription/Annotation**

- Completed in ELAN, native German speakers, Swabian orthography
- Words tagged as:
 - Standard, e.g., *habe* ‘have’
 - Vernacular, e.g., *hab* ‘have’
 - Swabian, e.g., *han* ‘have’

} Dialect

[CLICK] The methods used in this study consist of semi-structured sociolinguistic interviews, conducted by native Swabian speakers with me in attendance in the role of a friend-of-a-friend. To increase compatibility across years, the same survey instrument was used in both 1982 and 2017, following the same structure and covering the same topics.

[CLICK] The initial transcriptions were completed in ELAN by native German speakers, following a well-documented set of transcription guidelines and using a standard orthography specifically adapted for Swabian.

From 40 hours of interviews, over 160,000 words were extracted and tagged as either Swabian-specific, general Vernacular or Standard German.

For example, with the verb *haben* ‘to have’,
--*habe* is identified as the Standard form,
--*hab* as the Vernacular variant (with the reduction of the final ‘e’),
--*han* as the Swabian variant (an irregular verb in the dialect).

Because the aim of this investigation is to look at overall dialect usage, we grouped the Vernacular and Swabian-specific forms together [CLICK] (henceforth called, “dialect” forms) to contrast them with the standard German forms.

----- 1 min

Corpus: Panel Study

20 Panel Speakers:
– 1982 & 2017

2 Communities:
– 7 from Stuttgart
– 13 from Gmünd

2 Genders:
– 11 men
– 9 women

2 Education levels:
– 14 with Abitur
– 6 without Abitur

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

The corpus consists of 20 panel speakers, recorded first in 1982 and then re-recorded 35 years later in 2017.

Seven speakers are from Stuttgart and 13 from Schwäbisch Gmünd

11 are men and 9 are women.

14 of the 20 speakers were students in 1982 who completed their *Abitur* ‘German high school diploma / college preparatory exam’.

Most speakers are of the same age group (18-25 in 1982 and 53-60 in 2017) and socioeconomic status (middle class).

Four speakers were in their late 50's in 1982, and hence their late 80's in 2017.

All speakers were coded for SOI – Swabian Orientation Index – in each year.

----- 30 secs

Swabian Orientation Index (SOI)



Swabian Allegiance:

- 1-1. Self-Declared Swabian: Are you a 'real' Swabian?
5=definitely, 4=maybe, 3=don't know, 2=not really, 1=no
- 1-2. Non-Swabian Friends: Do you have friends who are NOT Swabian?
5=no, 4=a few, 3=don't know, 2=many, 1=a lot
- 1-3. Swabian Ridicule: Do they laugh at how you speak?
5=always, 4=sometimes, 3=don't know, 2=not really, 1=not at all
- 1-4. Accommodation: Do you change how you speak?
5=not at all, 4=a little, 3=don't know, 2=a lot, 1=always

Swabian Cultural Competence:

- 3-1. Swabian Knowledge: Are there different Swabian dialects?
5=considerable, 4=some, 3=don't know, 2=not much, 1=none
- 3-2. Swabian Specialties: Do you know how to make Spätzle? Maultaschen?
5=of course, 4=somewhat, 3=don't know, 2=not well, 1=not at all
- 3-3. Swabian People & Jokes: Do you know [various well-known Swabians]?
5=of course, 4=somewhat, 3=don't know, 2=not well, 1=not at all
- 3-4. Swabian Activities: Do you participate in Hocketse & local activities?
5=always, 4=some, 3=don't know, 2=not much, 1=never

Swabian Language Attitudes:

- 2-1. Opinion of Swabian Language: What do you think of the Swabian language?
5=super, 4=good, 3=don't know, 2=not good, 1=awful
- 2-2. Job Prospects for Swabians: Is it difficult to find a job when you speak Swabian?
5=great, 4=good, 3=no impact/don't know, 2=maybe some, 1=very difficult
- 2-3. Swabians Speaking German: Is it odd when a Swabian speaks standard German?
5=very odd/awful, 4=funny, 3=don't know, 2=good, 1=great
- 2-4. Non-Swabians Speaking Swabian: Is it odd when a non-Swabian speaks Swabian?
5=very odd/awful, 4=funny, 3=don't know, 2=good, 1=great

Swabian Language Usage:

- 4-1. Parents Speak Swabian: Do your parent speak Swabian?
5=both, 3=one, 1=neither
- 4-2. Swabian with Friends & Family: Do you speak Swabian with ...?
5=considerable, 4=some, 3=don't know, 2=not much, 1=none
- 4-3. Swabian with Neighbors: Do you speak Swabian with ...?
5=considerable, 4=some, 3=don't know, 2=not much, 1=none
- 4-4. Swabian with Others: Do you speak Swabian with ...?
5=considerable, 4=some, 3=don't know, 2=not much, 1=none

Modelled after Hoffman and Walker's ethnic identity index, the Swabian Orientation Index (SOI) is derived from speakers' answers to 16 questions posed in the interview, covering:

[CLICK] (1) **their allegiance and feelings** about being Swabian,

[CLICK] (2) **their attitudes** towards the Swabian language,

[CLICK] (3) **their knowledge of Swabian culture**, people and icons, and

[CLICK] (4) **their self-reported answers** to whether they speak Swabian or standard German with family, friends, neighbors, and others.

The 16 questions were evaluated on a five-point scale and averaged to create an overall score for each speaker in each year, from one for the lowest to five for the highest Swabian orientation

----- 30 secs

Types and Tokens



‘THE CAT IS ON THE MAT’

- WORD TYPE – a unique letter string
- WORD TOKEN – a specific instance of a WORD TYPE
- TEXT LENGTH is measured by the number of WORD TOKENS
- VOCABULARY SIZE is measured by the number of WORD TYPES

DATASET	1982	2017
TYPES	17,707	17,134
TOKENS	72,560	90,414

TOKENS	1982	2017
DIALECT	22,401	20,795
STANDARD	50,149	69,619

Beaman, Baayen, and Ramscar – CLARe4 Helsinki – February 2019

Page 10

Before diving into our analysis, for a lexical frequency analysis, it's important to make a distinction between TYPES and TOKENS.

[CLICK] For example, in the sentence, 'the cat is on the mat,' there are 6 TOKENS and 5 TYPES.

[CLICK] WORD TYPE refers to any unique letter string, delineated by spaces or punctuation marks in the transcript.

[CLICK] WORD TOKEN refers to a specific instance of a WORD TYPE that occurs or reoccurs in the transcript.

[CLICK] TEXT LENGTH is measured by number of WORD TOKENS,

[CLICK] VOCABULARY SIZE is measured by the number of WORD TYPES.

[CLICK] This table shows the number of TYPES and TOKENS by recording year for the 20 speakers in our Swabian corpus.

[CLICK] And this table shows the breakdown of TOKENS between dialect and standard. It is interesting to note that the standard words are more than **double** the dialect words in 1982 and more than **triple** in 2017, providing a first indication that, rather than **lose** dialect words, speakers actually **gain** a large number of standard language words over the course of their lifespan.

----- 1 min

Lexical Productivity



- Challenges in lexical productivity analysis:
 - VOCABULARY SIZE increases with TEXT LENGTH
 - intrinsic order in aggregate data could skew the results
- VOCABULARY GROWTH CURVE is calculated by counting the number of TOKENS within equally spaced measurement points throughout the text and graphing the corresponding count of each WORD TYPE.
- RANDOMISATION WINDOW performs Monte Carlo-like permutations on subsections of the text at predefined measurement points

[CLICK] A major challenge in conducting quantitative analyses of lexical productivity is dealing with texts of differing lengths. Naturally the longer the transcript, the more TYPES and TOKENS we'll have. The goal is to make sure we compare VOCABULARY SIZE for the same number of TOKENS.

A second challenge is to avoid any intrinsic order in the aggregate data which could skew the results, such as the loquacious and erudite speakers versus the more reticent speakers.

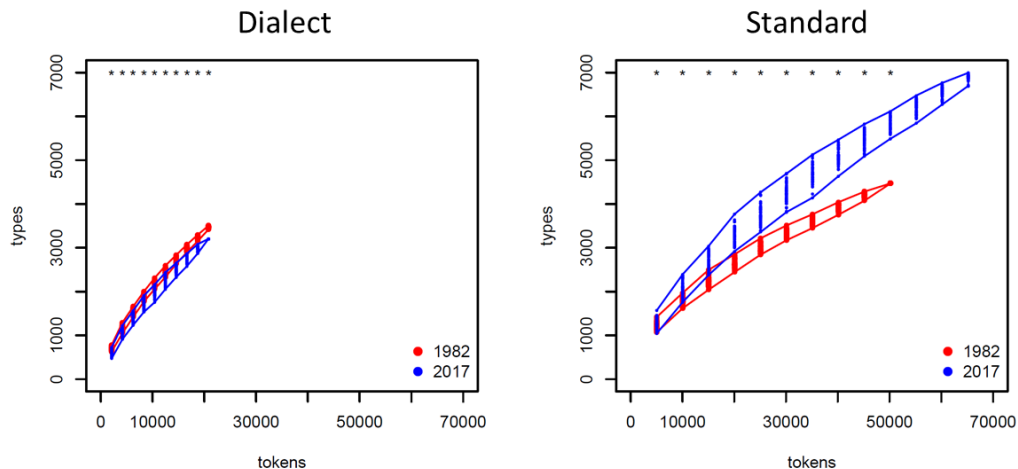
[CLICK] To work around the first problem, we calculate a VOCABULARY GROWTH CURVE by counting the number of TOKENS within equally spaced measurement points throughout the text and graphing the corresponding count of each WORD TYPE. This gives us a curve that depicts the rate at which the vocabulary increases.

[CLICK] To deal with the second problem, we use a PARTIAL RANDOMISATION technique. Rather than randomise the full transcript, as that would disrupt the discursive structure of the text, we permute the order of the speakers, and this gives us a distribution of vocabulary size at different text lengths, which shows patterns in the variability across samples.

Let's look at an example to illustrate this.

----- 1 min

Vocabulary Growth Curves



Beaman, Baayen, and Ramscar – CLARE4 Helsinki – February 2019

Page 12

This plot depicts the **dialect** vocabulary growth curves for our 20 panel speakers over the 35-year timeframe: 1982 is shown in red and 2017 in blue. TOKENS are shown on the horizontal axis and WORD TYPES on the vertical axis.

The results of the RANDOMISATION process are displayed as vertical bars made up of dots representing the mean values for the individual permutations. The outer boundary of each vocabulary growth curve is shown as a POLYGON that connects the minimum and maximum vocabulary sizes generated by the randomization process.

The asterisks at the top signify that there is a significant difference between the measured intervals.

Looking at this plot, it is quickly obvious from the overlapping red and blue polygons that there has been little change in speakers' **dialect** vocabulary over the 35-years.

[CLICK] In contrast, here is corresponding **standard** vocabulary growth curve for the 20 speakers. The large blue polygon shows that speakers have considerably enriched their **standard** language, adding over 3,000 new words – more than 25% increase.

These findings provide solid support for our hypothesis that, rather than **lose** dialect, in fact speakers **gain** an immense amount of additional lexical knowledge that is not dialect, creating competition between the vocabularies and making it “appear” as if dialect has been “lost”.

These results replicate many other studies that show vocabulary size increases with age. Keuleers and colleagues claim that “age is by far the most important variable in predicting vocabulary size.... every day lived represents an opportunity for acquisition of vocabulary and that existing vocabulary is not forgotten.”

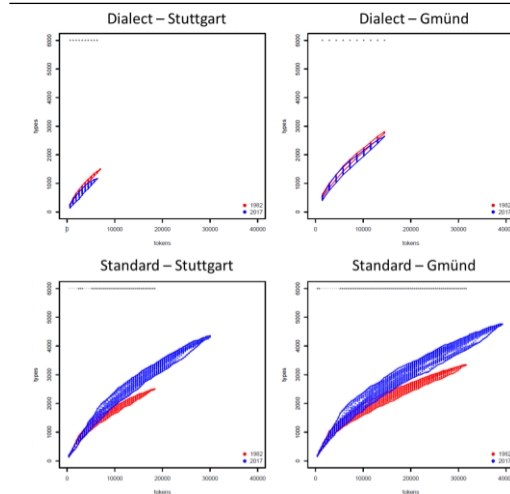
It appears that, for our Swabian speakers, the wisdom gained through added experience is manifested in the standard language rather than in dialect.

It is also interesting to note that the dialect vocabularies in 1982 and 2017 (on the left) are quite similar, which can be observed in how the polygons overlap for most of the trajectory. The two vocabularies only begin to disassociate about three quarters into the curve and are not completely disassociated until the last interval. Yet, for the standard vocabulary (on the right), the two trajectories disassociate much sooner, almost from the beginning, signifying that the standard language vocabularies in 1982 and 2017 are considerably more dissimilar. You have only to think of the internet explosion, that has occurred since 1982, to appreciate the vocabulary differences between these two time points.

This leads us to the premise that the domains and contexts in which dialect is spoken have changed little over the years, whereas the domains in which the standard language is encountered are vast and multifarious.

----- 2 mins 15 secs

Community and Education (1 of 2)



Beaman, Baayen, and Ramscar – CLARE4 Helsinki – February 2019

Page 13

[CLICK] These plots show the dialect vocabulary growth curves by community, Stuttgart on the left and Schwäbisch Gmünd on the right.

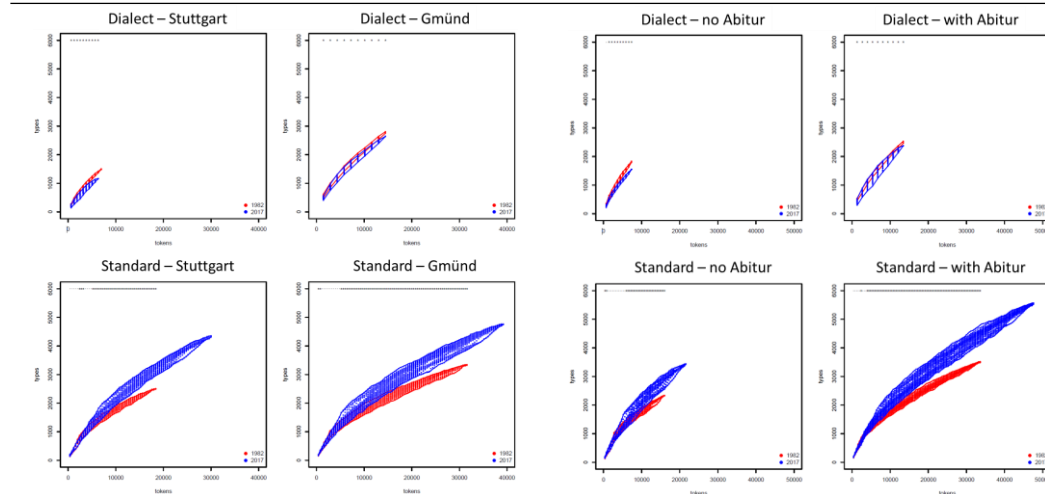
First off, we note that people from Schwäbisch Gmünd are much more talkative than those from Stuttgart, in that they produce more tokens and more word types. Based on our ethnographic observations of the speakers in these communities, we know that people from Gmünd place a high value on their dialect, which is strengthened in the social setting via intense and frequent communication with friends and family. They manifest a strong orientation to Swabia, and dialect provides a conduit for indexing their identity and bonding with the people around them. In the urban centre of Stuttgart, social connections are weaker and looser; hence, communication tends to be briefer and to the point.

[CLICK] These plots show the corresponding standard vocabulary curves for each community, which again confirms the fact that speakers have substantially enlarged their standard language vocabulary over the years.

We also note that speakers from Stuttgart have gained even more standard words than those from Gmünd, almost doubling their standard vocabulary size, which we attribute to the fact that urban life typically comprises more diverse experiences than are found in smaller, semi-rural towns, as well as to the large non-Swabian population: over half of Stuttgart's inhabitants have at least one foreign born parent.

-----1 min

Community and Education (2 of 2)



Beaman, Baayen, and Ramscar – CLARE4 Helsinki – February 2019

Page 14

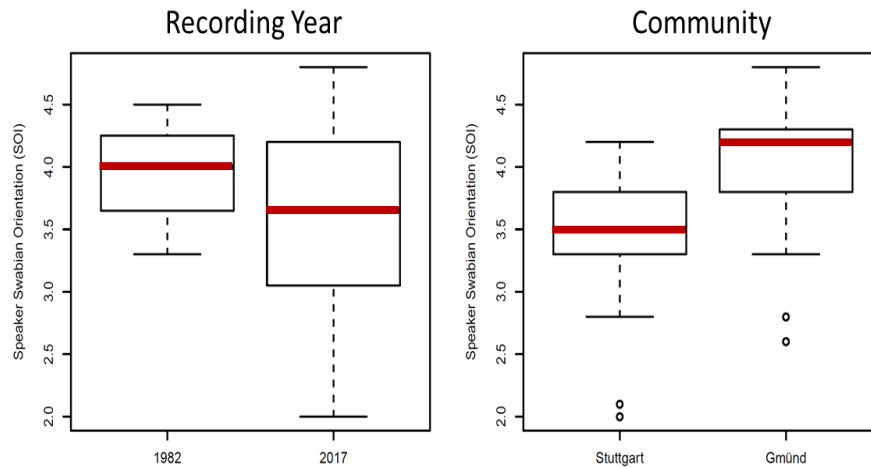
These plots on the right show the speakers' vocabulary growth rate by level of education: those with and without an *Abitur*.

From the top panels, there is little change in the use of **dialect** based on educational attainment: both groups of speakers have retained most of their dialect over the years. However, from the bottom panels, we see growth in the **standard** language for both groups of speakers, particularly those with an *Abitur*.

These results can certainly be attributed to the fact that the standard language is reinforced in school, and, indeed, many studies have confirmed the association between a loss of dialect forms and higher levels of education. Increased **standard** language vocabulary clearly reflects the contact that the more educated group has to the standard language register.

----- 45 secs

Swabian Orientation Index (SOI)



Beaman, Baayen, and Ramscar – CLARe4 Helsinki – February 2019

Page 15

We now turn to Swabian orientation. These plots depict the changing prominence of Swabian over the 35 years: the left panel shows orientation by year and the right panel shows orientation by community.

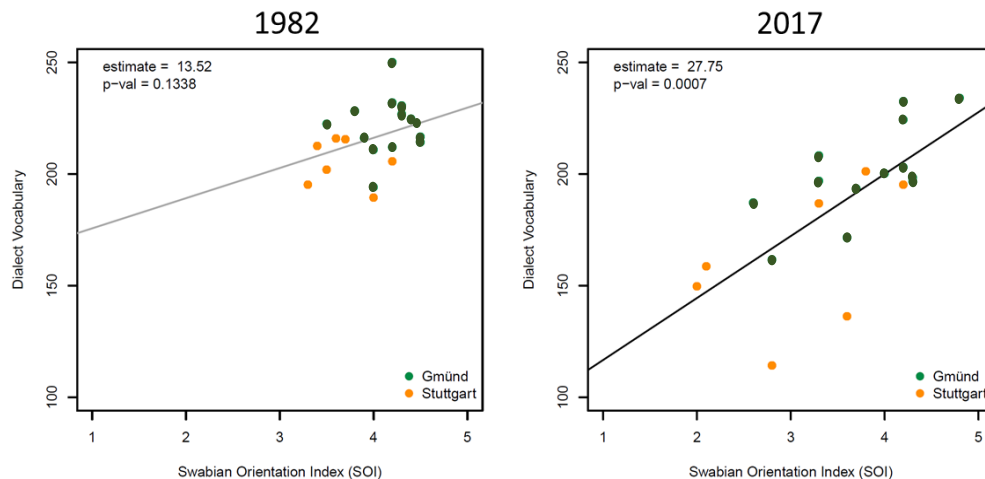
[CLICK] With a mean of 4.0 in 1982, Swabian orientation played a very powerful role. [CLICK] However, by 2017, orientation scores for these same speakers has fallen to an average of 3.6 and with a much broader spread.

We see similar skewing in orientation scores by community, with [CLICK] Stuttgart showing lower overall scores and than [CLICK] Schwäbisch Gmünd.

[CLICK] These plots make it evident that the notion of Swabian identity has changed dramatically over the years, especially for Stuttgart.

----- 45 secs

Dialect Vocabulary and Orientation



Beaman, Baayen, and Ramscar – CLARE4 Helsinki – February 2019

Page 16

We now look at the effect of Swabian orientation on our individual speakers and their propensity to use dialect.

These plots show the mean **dialect** vocabulary size on the vertical axis and Swabian orientation on the horizontal, 1982 is on the left and 2017 on the right. The Stuttgart speakers are denoted by orange dots and the Gmünd speakers by green dots.

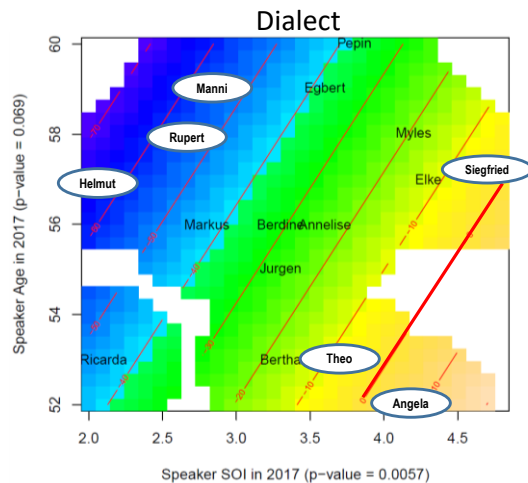
Our first observation is the [CLICK] tight clustering of speakers in the upper right corner in 1982 versus the [CLICK] more dispersed placement of speakers in 2017. These speakers' similar patterns of dialect usage imply that Stuttgart and Schwäbisch Gmünd were more homo-GEN-eous in 1982 than they have become in 2017. By 2017, for some speakers, Swabian orientation has declined concomitant with their dialect usage, particularly for the Stuttgart speakers.

Still, we see a number of speakers, those from Schwäbisch Gmünd, who have retained their high Swabian orientation and dialect vocabulary. The trend is clear: the higher the Swabian orientation score, the larger the dialect vocabulary; and conversely, the lower the speakers' orientation, the smaller the dialect vocabulary.

This leads us to question: who are the speakers who have changed their vocabulary the most, and what are the reasons behind this change?

----- 1 min

Individual Patterns of Change



Beaman, Baayen, and Ramscar – CLARE4 Helsinki – February 2019

Page 17

Individual patterns of linguistic change have been shown to complement and enhance insights gained from community change. So we now take a deeper dive into the individual speakers and the change across their lifespans.

[CLICK] Using generalised additive mixed models, this graph helps to visualize the differences in **dialect** vocabulary change for our speakers. Speaker age in 2017 is shown on the vertical axis and speaker orientation in 2017 on the horizontal axis. The contour lines delineate vocabulary change, [CLICK] with the zero line demarcating no change. Higher values are shown in shades of yellow and smaller values in shades of blue.

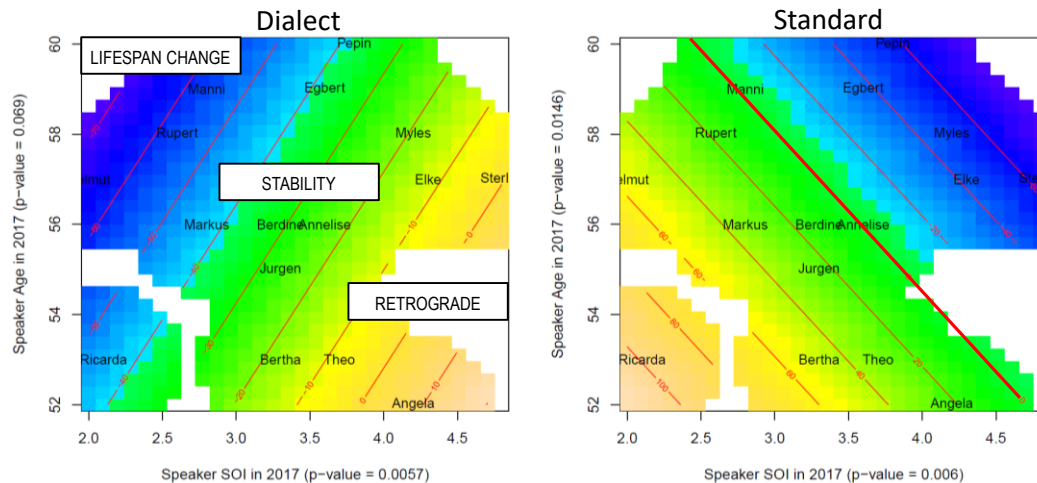
In the lower right corner is [CLICK] Angela, who has actually gained dialect words over the years. We see [CLICK] Siegfried on the cusp, along with Theo in the yellow zone. These three speakers have high orientation scores and have retained most of their dialect over the years.

At the far left, we observe [CLICK] Helmut, along with Rupert and Manni, in the blue zone, who have lost the most dialect vocabulary. These three “businessmen”, because of their work, have extensive contact with speakers outside of Swabia and show the lowest Swabian orientation scores.

In contrast to the composite diagram we saw earlier, this individual view shows that Swabian vocabulary richness has diminished over the 35 years for some speakers and unmistakably establishes the high correlation that Swabian orientation has on an individual’s vocabulary.

----- 1 min 15 secs

Individual Patterns of Change



Beaman, Baayen, and Ramscar – CLARE4 Helsinki – February 2019

Page 18

The right panel portrays speakers' **standard** vocabulary change over the time period, and we immediately see the inverse.

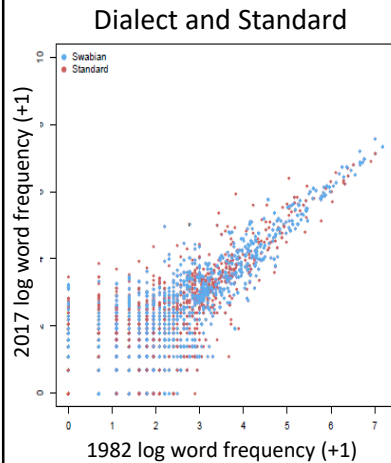
[CLICK] The zero-contour line, showing no vocabulary change, is now in the middle. We describe this as a “very shallow sea” since all of the speakers are quite similar. All speakers have gained standard vocabulary over the years. The results of our linear regression model again confirm that both orientation and age are significant predictors of speakers' **standard** vocabulary gain. As Swabian orientation scores decrease, speakers' standard vocabulary increases; as people age, their vocabulary expands, and this expansion comes in the form of the standard language.

The colors depicting **dialect** change visually reveal the three classic patterns of individual change as identified by Sankoff. [CLICK] The blue zone denotes **LIFESPAN CHANGE**, that is, those speakers moving in the direction of the overall community change by speaking less dialect and more standard; [CLICK] the green zone marks **SPEAKER STABILITY**, those individuals resisting change and continuing to use a similar amount of dialect; and, [CLICK] the yellow zone portrays **RETROGRADE CHANGE**, speakers moving in the opposite direction of the general community change and speaking more dialect today than they did in 1982.

These results underscore the criticality of incorporating individual lifespan analyses into general trend studies to tease out forces that are otherwise hidden in community-wide averages.

----- 1 min 15 secs

Frequency Effects



Beaman, Baayen, and Ramscar – CLARe4 Helsinki – February 2019

Page 19

We now turn to our second hypothesis. Recall that our expectation was that high frequency words would more likely to be retained in the dialect and that low frequency words would more likely to be lost.

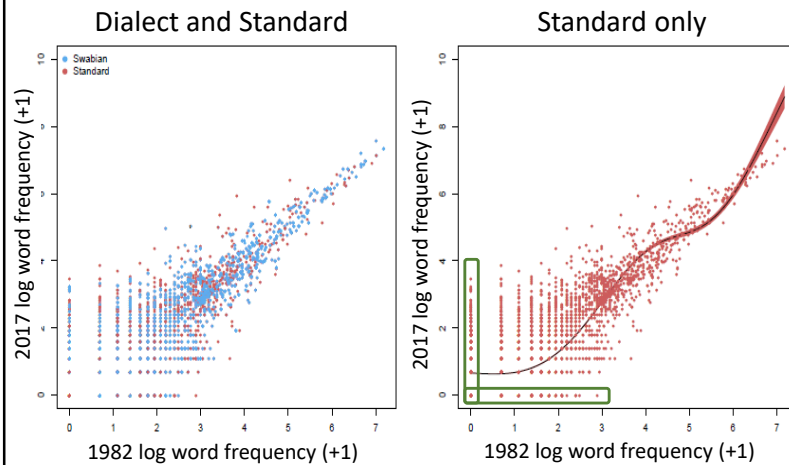
[CLICK] We tested this by comparing word frequencies between 1982 and 2017.

This plot depicts word frequency for **dialect** in blue and **standard** in red. The horizontal axis shows word frequency in 1982 and the vertical axis word frequency in 2017.

Due to large number of overlapping points, the trend is not immediately obvious.

----- 30 secs

Frequency Effects



Beaman, Baayen, and Ramscar – CLARe4 Helsinki – February 2019

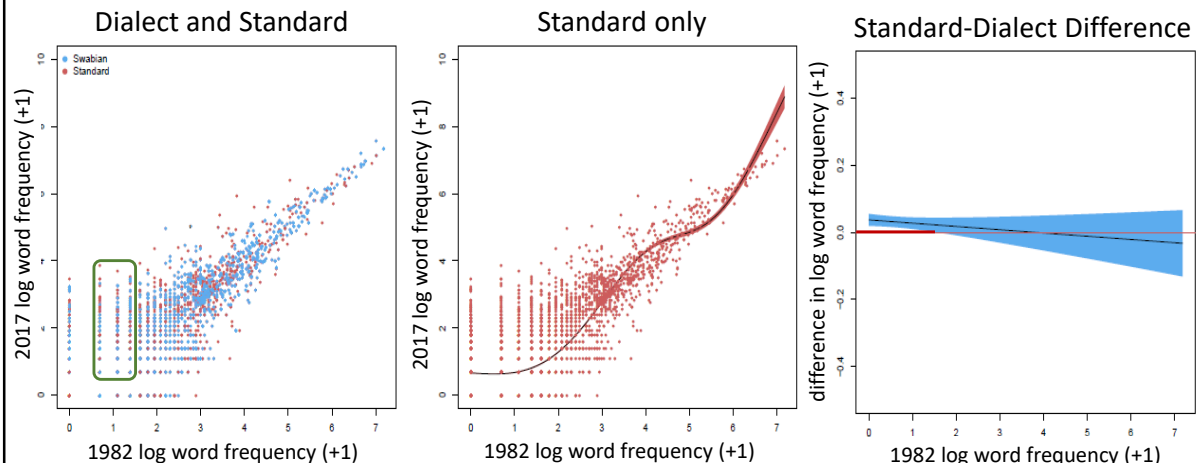
Page 20

So in this panel we pull out the **standard** word frequencies (red points) to illustrate the main trend.

As we would expect, we see -- in the lower left corner -- that there are [CLICK] a number of **standard** words used in 1982 that were not used in 2017, and [CLICK] there are also standard words used in 2017 that were not used in 1982. Just think, were no cell phones or Twitter accounts back in 1982.

----- 15 secs

Frequency Effects



Beaman, Baayen, and Ramscar – CLARe4 Helsinki – February 2019

Page 21

The right panel shows the difference between the dialect and standard word frequencies. The only significant difference is in the low frequency **dialect** words, where we see, [CLICK] signified by the red line marking the zero confidence interval, that it is the low frequency **dialect** words in 1982 that are being used slightly more in 2017. As we move to higher frequency words, there is no significant difference, signified by the blue shaded area.

[CLICK] It's in this area where we see the low frequency words from 1982 being used slightly more in 2017.

Hence, contrary to our hypothesis -- and to standard lore -- that high frequency words would more likely be more retained and low frequency words would more likely die out -- we have found the exact the opposite! In fact, it's the low frequency that have become more frequent.

It appears that there is a large domain in which Swabian is simply not used because it is not appropriate, and this domain has grown even larger over the years. As people age, rather than "forget" their dialect, they develop greater fluency and expertise with the standard language and this creates a cumulative and competitive effect on dialect usage.

This finding supports the idea that Swabian vocabulary is not being lost. However, it is even more far-reaching than we initially thought, and is contrary to many studies that show pervasive dialect levelling, particularly with low frequency words.

----- 45 secs

In Summary



- Vocabulary size increases with age and experience
- Later life experiences come in the form of the standard language
- Swabian is not in decline, rather restricted to specific domains of use
- Swabian Orientations influence levels of dialect usage
- Low frequency dialect words have become slightly more frequent
- Age of acquisition suggests early acquired words are more accessible

In summary:

[CLICK] We've clearly seen that **vocabulary increases with age and experience**; and, at least for our Swabian panel speakers, these

[CLICK] **later life experiences come in the form of the standard language** which doesn't necessarily entail a concomitant loss of dialect. As one of our shrewd speakers commented: "there are things you can say in Swabian that you can't say in standard German, and there are things you can say in standard German that you can't say in Swabian." Imagine trying to have a scientific discussion with a colleague in Swabian. In fact, even standard German is falling out of favour in this domain, where increasingly English is the language of choice.

[CLICK] We've seen that **Swabian is not in decline** for all speakers, rather its usage is restricted to specific domains, such as making "Spätzle" or drinking a "Virtelle" with friends in a "Kneipe". The usage of Swabian has become proportionally marginalized due to the massive influx of standard language words over the years.

[CLICK] The notion of **Swabian Orientation is extremely powerful** in the influence it plays on speakers' propensity to speak more dialect or more standard.

[CLICK] Finally, contrary to what we expected, **low frequency dialect words have become more frequent**, which implies that

[CLICK] an **age of acquisition effect** may be at play, suggesting that words acquired earlier are more accessible than those learned later in life.

Returning to the main thesis of our paper, we need to be careful not to confound the effects of competition and attrition on dialect usage and dive deeper into the individual differences between speakers -- their personal orientation and the 'dialect identity' they wish to convey -- to better understand changes in lexical productivity across the lifespan.

----- 1 min 30 secs

Thank you!



CORRESPONDING AUTHOR:

Karen V. Beaman
Queen Mary, University of London
Eberhard Karls University of Tübingen
karenbeamanvslx@gmail.com

Thank you!

References



- Auer, P. 2018. "Dialect Change in Europe -- Leveling and Convergence." Pp. 159–76 in *The Handbook of Dialectology*, edited by C. Boberg, J. Nerbonne, and D. Watt. Oxford: Wiley-Blackwell.
- Baayen, R. H.. 2001. *Word Frequency Distributions*. Kluwer Academic Publishers.
- Giles, H., D. J. Taylor, and R. Bourhis. 1973. "Towards a Theory of Interpersonal Accommodation through Language: Some Canadian Data." *Language in Society* 2(2):177–92.
- Hoffman, M. F. and J. A. Walker. 2010. "Ethnolects and the City: Ethnic Orientation and Linguistic Variation in Toronto English." *Language Variation and Change* 22:37–67.
- Keuleers, E., M. Stevens, P. Mandera, and M. Brysbaert. 2015. "Word Knowledge in the Crowd: Measuring Vocabulary Size and Word Prevalence in a Massive Online Experiment." *The Quarterly Journal of Experimental Psychology* 68(8):1665–92.
- Ramscar, M., P. Hendrix, B. Love, and R. H. Baayen. 2013. "Learning Is Not Decline: The Mental Lexicon as a Window into Cognition across the Lifespan." *The Mental Lexicon* 8(3):450–81.
- Ramscar, M., P. Hendrix, C. Shaoul, P. Milin, and R. H. Baayen. 2014. "The Myth of Cognitive Decline: Non-Linear Dynamics of Lifelong Learning." *Topics in Cognitive Science* 6(1):5–42.
- Sankoff, Gillian. 2018. "Language Change Across the Lifespan." *Annual Review of Linguistics* 4:297–316.
- Trudgill, Peter. 1986. *Dialects in Contact*. Oxford: Blackwell Publishing.
- Wagner, S. E. and I. Buchstaller. 2017. *Panel Studies in Variation and Change*. Routledge.
- Wieling, M., J. Nerbonne, and R. H. Baayen. 2011. "Quantitative Social Dialectology: Explaining Linguistic Variation Geographically and Socially." *PLoS ONE* 6(9):1–14.