

Research Article

Franka Zebe*

Vowel and consonant quantity in two Swiss German dialects and their corresponding varieties of Standard German: effects of region, age, and tempo

<https://doi.org/10.1515/phon-2022-0017>

Received May 30, 2022; accepted June 12, 2023; published online July 7, 2023

Abstract: The diglossic situation in German-speaking Switzerland entails that both an Alemannic dialect and a Swiss standard variety of German are spoken. One phonological property of both Alemannic and Swiss Standard German (SSG) is contrastive quantity not only in vowels but also in consonants, namely lenis and fortis. This study aims to compare vowel and plosive closure durations as well as articulation rate (AR) between Alemannic and SSG in the varieties spoken in a rural area of the canton of Lucerne (LU) and an urban area of the canton of Zurich (ZH). In addition to the segment durations, an additional measure of vowel-to-vowel + consonant duration ($V/(V + C)$) ratios is calculated in order to account for possible compensation between vowel and closure durations. Stimuli consisted of words containing different vowel-consonant (VC) combinations. The main differences found are longer segment durations in Alemannic compared to SSG, three phonetic vowel categories in Alemannic that differ between LU and ZH, three stable $V/(V + C)$ ratio categories, and three phonetic consonant categories lenis, fortis, and extrafortis in both Alemannic and SSG. Most importantly, younger ZH speakers produced overall shorter closure durations, calling into question a possible reduction of consonant categories due to a contact to German Standard German (GSG).

Keywords: Alemannic; articulation rate; consonant duration; Swiss Standard German; vowel duration

***Corresponding author: Franka Zebe**, Department of Computational Linguistics, Phonetics Laboratory, University of Zurich, Rämistrasse 71, Zurich, 8006, Switzerland, E-mail: franka.zebe@uzh.ch

 Open Access. © 2023 the author(s), published by De Gruyter.  This work is licensed under the Creative Commons Attribution 4.0 International License.

1 Introduction

German-speaking Switzerland constitutes a textbook example of a diglossia (Ferguson 1959), where the low variety is represented by a regional Alemannic dialect and the high variety by Swiss Standard German (SSG). Nevertheless, most studies focus on Alemannic dialects rather than SSG, while almost none compares the two to each other. A phonological property of Alemannic dialects is the presence of quantity contrasts not only in vowels but also in consonants, labelled as ‘lenis’ (i.e., shorter consonants) and ‘fortis’ (i.e., longer consonants) (e.g., Fleischer and Schmid 2006). As Maddieson (1984) states, only 19.6 % of languages have distinctive vowel quantity. It has also been argued by Laver (1996) that contrastive consonant duration is even rarer than contrastive vowel duration. Therefore, it can be assumed that the occurrence of both vowel and consonant quantity contrasts is particularly rare. This study aims to provide additional information to a general typology of vowel and consonant quantities.

Phonotactically, all possible sequences of long and short vowels and consonants are legal in Alemannic as well as in SSG, resulting in four different vowel-consonant combinations (VC combinations): short + short (VC), short + long (VC:), long + short (V:C), and long + long (V:C:). In contrast, other southern German varieties, such as Eastern Central Bavarian (spoken in the area of Vienna, Austria) and Western Central Bavarian (spoken in the area of Munich, Germany), have a complementary length pattern, meaning that only the combinations V:C and VC: are phonotactically legal (Moosmüller and Brandstätter 2014; Seiler 2005). Following the apparent-time method (Bailey et al. 1991; Labov 1994), recent investigations revealed that this is still true for older speakers of Eastern and Western Central Bavarian, while younger speakers behave differently, also producing the previously phonotactically illegal combination V:C:, possibly due to contact to German Standard German (GSG), where all four combinations are legal (Kleber 2017; Moosmüller 2007; Moosmüller and Brandstätter 2014).

This study aims to better understand vowel and consonant (i.e., plosive closure) durations in Alemannic and SSG and how they relate to each other, comparing speakers from an urban area in the canton of ZH to speakers living in a rural area in the canton of LU. Thus, the productions of the four VC combinations in those two regions of German-speaking Switzerland were examined. While the Alemannic dialect of ZH has been studied quite extensively, also with regard to several durational and rhythmic properties (e.g., Fleischer and Schmid 2006; Leemann et al. 2012; Leemann and Siebenhaar 2010; Nocchi and Schmid 2006; Pellegrino et al. 2021; Schmid 2004; Zihlmann 2020a, 2020b, 2021a, 2021b), it is usually not taken into account that its contact to GSG might have an influence on those metrics. Rather, it is

often presented as a prototype of Alemannic in German-speaking Switzerland. This is why in this study more rural areas in the canton of LU are also investigated. In addition, it has been shown that SSG is influenced by Alemannic (Hove 2002; Zihlmann 2020b). While the general temporal patterns in SSG are certainly similar to those in Alemannic, it has not been investigated yet if there are age-related differences in segment durations. This is why this study, for the first time, compares those durational measurements between younger and older speakers.

An additional goal is to investigate articulation rate (AR), measured in terms of mean syllable duration (MSD), which has been shown to differ across regions in German-speaking Switzerland (Leemann 2016; Leemann and Siebenhaar 2010; Zihlmann 2020a), as well as its influence on the four VC combinations to investigate their stability. Consequently, vowel and consonant durations were investigated not only in normal but also in fast speech tempo. In addition to the segment durations by themselves, the vowel-to-vowel + consonant duration ratio ($V/(V + C)$ ratio), also known as Proportional Vowel Duration (PVD) (Kohler 1979; Zihlmann 2020b), was chosen as a further measure. This metric is an important perceptual cue for both voicing (Kohler 1979) and the vowel length contrast (Kleber 2017). It calculates the percentage of the vowel within a $V + C$ sequence. Because some compensation between vowel and consonant durations is expected and the $V/(V + C)$ ratio takes this into account, it is a helpful addition to the segment durations analyzed in this study.

This study aims, first and foremost, to compare the durational measurements in Alemannic to those in SSG. The second main goal of the study is to investigate how stable they are in each group of speakers (LU and ZH, younger and older speakers) and across tempo.

1.1 Diglossia in German-speaking Switzerland

The dialects spoken in German-speaking Switzerland comprise Low Alemannic, which is limited to the city of Basel, High Alemannic, spoken in the north, and Highest Alemannic spoken in the south (Arquint et al. 1982; Christen 2019). The two dialects investigated in this study belong to the High Alemannic group and will simply be referred to as Alemannic in this paper.

It is typical in German-speaking Switzerland to speak Alemannic in most situations in everyday life. SSG is learned in school and spoken only in specific formal situations, e.g., parliament sessions or formal news reports. For written communication, mainly SSG is used with some exceptions, e.g., written commercials, which may also be written in dialect. However, the diglossic configuration has been changing since 1980, when it has become increasingly common to also communicate in written Alemannic, e.g., in personal letters, emails, and, more recently, in online

chats or on social media (Siebenhaar 2006). While it has been shown that one cannot infer a specific Alemannic dialect when analyzing vowel and consonant durations in SSG, recent findings revealed that speakers use the same durational categories in SSG as in Alemannic, suggesting that, in general, SSG is influenced by Alemannic rather than the other way around (Zihlmann 2020b). The predominant use of Alemannic in everyday life and the fact that it is used increasingly in the written language implies that the diglossic situation in German-speaking Switzerland is relatively stable. Even so, the impact of GSG has not been investigated so far. This is why this study focuses on an urban area with high contact to GSG and a rural area, as well as two age groups.

1.2 Articulation rate, vowel durations, and consonant durations

1.2.1 Articulation rate

Taking a look at AR, among the factors by which it can be influenced are age and region. There is broad consensus that speakers of older age speak with a slower AR compared to younger speakers (Jacewicz et al. 2009; Quené 2008; Schwab and Avanzi 2015). Results from Jacewicz et al. (2009), who investigated AR in terms of syllables per second in different regions of the USA, additionally revealed that age-related differences in AR are also dependent on the region itself, as older speakers turned out to have a significantly slower AR only in one of the two investigated regions, although there was still a trend in the second region (Jacewicz et al. 2009). Quené (2008) conducted a similar study comparing the MSD of Dutch speakers from The Netherlands to those from Flanders in spontaneous speech. He also confirmed significant differences between younger and older speakers with older speakers having a slower AR (Quené 2008). Schwab and Avanzi (2015) investigated AR, measured in terms of MSD, in French-speaking Switzerland, Belgium, and France focusing on both read and spontaneous speech. Similarly to the findings by Jacewicz et al. (2009), their results showed that the effect of age is also dependent on the region itself (Schwab and Avanzi 2015). The aforementioned studies also revealed general regional differences in AR (Jacewicz et al. 2009; Quené 2008; Schwab and Avanzi 2015).

Regarding German-speaking Switzerland, Leemann (2016) conducted a crowdsourcing study using the Swiss App called “Dialäkt-Äpp” (Kolly and Leemann 2013), in which the general public can record words in their variety of Alemannic and give information about their exact location. He measured the duration between the two vowel onsets in a set of isolated disyllabic words (Leemann 2016). Results revealed that speakers from the eastern area generally have a faster AR than those from the western area of German-speaking Switzerland (Leemann 2016). Indeed, AR in

Alemannic seems to decrease gradually from east to west (Leemann 2016). Zihlmann (2020a) confirms that speakers from ZH have a faster AR (measuring syllables per second) than those from Berne for SSG. Surprisingly, he found the opposite results for Alemannic with speakers from Berne speaking faster than those from ZH (Zihlmann 2020a). The results also showed that the AR is generally slower in Alemannic than in SSG, although that comparison was not focus of the study (Zihlmann 2020a). This study aims to replicate the findings of the aforementioned studies, using MSD as a measure.

Thus, the first hypothesis (H1) of this study is (a) older speakers should have a slower AR, i.e., longer MSD, than younger speakers, (b) LU speakers should have a slower AR than ZH speakers in SSG (while it is not quite clear for Alemannic), and (c) Alemannic speech should be slower than SSG speech.

One relevant issue regarding AR is whether the listener can perceive the differences in tempo or not. According to Quené (2007), the just noticeable difference (JND) for AR lies at 5 %. This is why the data in this study was not only analyzed in terms of significance, but also in terms of the JND to reach conclusions on whether possible differences are perceivable or not.

1.2.2 Lenis and fortis obstruents

As mentioned above, Alemannic dialects display a phonological quantity contrast for vowels (Schmid 2004) as well as for obstruent consonants (Fleischer and Schmid 2006; Kraehenmann 2001, 2003; Nocchi and Schmid 2006). In addition, both lenis and fortis obstruents are voiceless and in plosives, there is no difference between the two categories with regard to Voice Onset Time (VOT); instead, closure duration has been proven to be the relevant acoustic correlate for the phonological lenis versus fortis contrast (Dieth 1950; Enstrom and Spörri-Bütler 1981; Fleischer and Schmid 2006; Fulop 1994; Kraehenmann 2001; Ladd and Schmid 2018; Willi 1996). An exception to this general pattern comes from loanwords from GSG or from English, some of which are pronounced with aspirated plosives (Fleischer and Schmid 2006; Ladd and Schmid 2018; Schifferle 2010).

Now, it has been shown that ‘voiced’ and ‘voiceless’ plosives have an effect of F0 on the vowel following the consonant in several languages such as American English (Hanson 2009), French and Italian (Kirby and Ladd 2015, 2016), and German (Kirby et al. 2020). This effect indicates that F0 is higher in the vowel onsets following voiceless than in those following voiced consonants, likely due to articulatory gestures and higher air pressure (House and Fairbanks 1953; Kingston and Diehl 1994; Kohler 1985; Lehiste and Peterson 1961). This pattern has also been observed for ZH German lenis and fortis plosives as well (Ladd and Schmid 2018). Nevertheless, it is no controversy that closure duration constitutes the primary phonetic correlate that

aids to distinguish between lenis and fortis obstruents in Alemannic dialects (Kraehenmann 2001; Kraehenmann and Lahiri 2008; Nocchi and Schmid 2006; Willi 1996).

At this point, a remark on terminology is in order. Several reasons have led to the decision to choose the lenis versus fortis terminology as opposed to the terms ‘singleton’ and ‘geminate’ for this study. First, in Alemannic, there are cases where the contrast between lenis and fortis is neutralized, e.g., in obstruent clusters (Dieth 1950; Moulton 1986; Würth 2020). The outcome of this neutralization is often referred to as ‘half fortis’ (Dieth 1950; Moulton 1986), whereas no specific term exists within the singleton versus geminate terminology. Although neutralization is not a part of this study, it makes sense in general to stick with the lenis versus fortis terminology in languages for which cases like these have already been taken into account.

The second reason to use the terms lenis and fortis is that, based on closure duration, three types of plosives can be distinguished in some Swiss German varieties, namely lenis, fortis, and extrafortis (Schmid 2019; Zihlmann 2020b). The term ‘extrafortis’ refers to obstruents that are significantly longer in duration than fortis ones and occur after short vowels, while fortis obstruents are produced following long vowels (Zihlmann 2020b). The three-way contrast in closure duration was confirmed for the Alemannic dialects spoken in Berne and ZH (Zihlmann 2020b), according to the phonotactic context (VC combinations): Whereas the V:C: combination yielded the typical durations of fortis consonants, the duration of the consonants in the VC: combination is significantly longer, which is why it may reasonably be called ‘extrafortis’. As these three categories are not found in all dialectal areas, Zihlmann (2020b) argues that they are two phonological categories with some regions producing three phonetic categories.

Note that Zihlmann (2020b) is the first study that investigated consonant duration not only in Alemannic dialects, but also in four different varieties of SSG (ZH, Berne, Chur, and Brig), where he found the same tripartite pattern (lenis, fortis, and extrafortis). The same three-way distinction in closure duration was most recently also confirmed for both Alemannic and SSG productions by LU speakers (Zebe 2022). Ham (2001) also found a three-way distinction in closure duration produced by speakers of the Bernese dialect, suggesting the terms ‘lenis’, ‘fortis’, and ‘geminate’, but that terminology could lead to further confusion, given that ‘geminate’ has also been used as a synonym of ‘fortis’ (Kraehenmann 2001; Würth 2020).

Ultimately, it seems more advisable to use the terms ‘lenis’ versus ‘fortis’ (and possibly ‘half fortis’ and ‘extrafortis’) for both Alemannic and SSG to avoid further terminological inconsistencies with respect to consonant duration.

1.2.3 Vowel and consonant durations

Even though the Alemannic and SSG varieties represent an excellent foundation for it, combined research on vowel and consonant quantity is rare, even more so including both Alemannic and SSG. Highly relevant to this approach is the aforementioned study from Zihlmann (2020b), who investigated vowel and consonant durations in the Alemannic dialects as well as the respective SSG varieties of four regions in German-speaking Switzerland, namely ZH, Berne, Chur, and Brig. In particular, Zihlmann (2020b) analyzed vowel and consonant durations as well as $V/(V + C)$ ratios. He found that, while the quantitative measurements are stable in general, the main difference between Berne (in the west) and ZH (in the east) were the consonant durations in both Alemannic and SSG, with speakers from Berne producing longer normalized consonant durations than those from ZH (Zihlmann 2020a). This leads to the second hypothesis (H2) that LU speakers are expected to produce longer consonant durations, as they are also situated in the west of ZH.

Regarding vowel durations, the results from the available research are not consistent. In previous studies, speakers from Berne in some cases produced longer normalized vowel durations compared to ZH, while in other cases they showed the opposite pattern (Zihlmann 2020a, 2020b). Therefore, it is unclear what to expect in terms of vowel duration in the current study. The stability among the varieties investigated by Zihlmann (2020b) turned out to be particularly striking in SSG, where there is no indication that speakers from different regions produce significantly distinct segment durations.

Regarding $V/(V + C)$, Zihlmann (2020b) concluded on three broad categories for both Alemannic and SSG, the first one being $V:C$ with the highest vocalic proportion, the second one consisting of both $V:C$ and VC with an intermediate vocalic proportion, and the third one being VC with the smallest vocalic proportion. It is, therefore, highly likely that in the current study the results will look similar, leading to the third hypothesis (H3), according to which (a) no significant differences in $V/(V + C)$ ratios between regions are expected and (b) three categories of $V + C$ sequences are expected.

Summarizing the description of segmental durations in Alemannic, it must be pointed out that – from a phonological point of view – vowel quantity and consonant quantity are both distinctive and independent from each other, given the existence of numerous minimal pairs which are based solely on the contrast between short and long vowels or between short and long consonants (cf. Fleischer and Schmid 2006), e.g., Zurich German $/z̥iḃə/$ (VC) ‘seven’ versus $/z̥i:ḃə/$ (V:C) ‘to sieve’, $/z̥itə/$ ‘costum’ (VC:) versus $/z̥i:tə/$ ‘page’ (V:C:), $/lɑḃə/$ (VC) ‘store’ versus $/lɔtə/$ (VC:) ‘lath’, $/hu:ḃə/$ (V:C) ‘bonnet’ versus $/hu:pə/$ (V:C:) ‘horn of a vehicle’. Only the allophonic *extrafortis*

category is phonotactically predictable (e.g., /'lɔtə/ 'lath' pronounced as [lɔt:ə]) in some dialects; nevertheless, the legal combination of a short vowel and a short consonant in /laḑə/ (pronounced as [laḑə]) shows that stressed syllables need not to be heavy but may as well be light. Therefore, quantity constitutes the only relevant phonological primitive and segment durations cannot be derived from higher prosodic constraints of syllable weight.

2 Methodology

Fourty speakers in total from the cantons of LU and ZH were recorded in both Alemannic and SSG in two tempo conditions, i.e. normal and fast speech tempo.

2.1 Speakers

Of the 40 speakers recorded, 20 were from LU, 10 speakers (5 female) were younger with ages between 25 and 32 years (mean = 28.7, SD = 1.94) and 10 speakers (5 female) were older with ages between 47 and 64 (mean = 58.0, SD = 6.30) at the time of the first recording. The other 20 speakers were from ZH, 10 speakers (5 female) in the younger group with ages between 18 and 28 (mean = 23.0, SD = 2.98) and 10 speakers (5 female) in the older group with ages between 58 and 69 (mean = 64.1, SD = 3.75) at the time of the first recording. All of the speakers grew up in either LU or ZH and at least one but in most cases both of their parents also come from the same canton. With the exception of four speakers in LU and five speakers in ZH (one of them still going to secondary school) the speakers either finished university or were still studying at university at the time of the recordings.

2.2 Stimuli

2.2.1 Alemannic

A detailed overview of all stimuli can be found in the Appendix of this paper. Stimuli consisted of 34 disyllabic target words for the speakers from LU and 31 for the speakers from ZH with the stress always being on first syllable. All target words were part of a series of three or four words (e.g. Side (VC) 'silk', miide (V:C) 'to avoid', Siite (V:C:) 'side', Sitte (VC:) 'manners'). The nucleus of the first syllable contained one of the vowels /a i u/ (short or long) followed by one of the plosives /b d ɡ p t k/. The target

words were embedded in a carrier sentence each with a total of six to eight syllables (e.g. Das isch *Side* vo China. ‘This is silk from China.’).

2.2.2 SSG

Stimuli for speakers from LU and ZH were the same for the SSG productions. They consisted of 23 disyllabic target words with the stress always being on the first syllable. Again, all target words were part of a series of three or four words (e.g. Tube (V:C) ‘tube’, Lupe (V:C:) ‘magnifying glass’, Suppe (VC:) ‘soup’). The target words contained the same vowels and consonants as in the Alemannic stimuli. They were embedded in a carrier sentence each with a total of seven syllables (e.g. Er will die *Tube* nehmen. ‘He wants to take the tube.’).

2.3 Procedure

Whenever possible, the participants were recorded in a soundproof booth at the Phonetics Laboratory of the University of Zurich using a personal computer with the interface USBPre[®] 2 (Sound Devices) and the microphone NT2-A (RØDE). Otherwise, they were recorded in a quiet room using a laptop computer with the same interface and the microphone Opus 54.16/3 (BeyerDynamic). For the recording software, SpeechRecorder 3.8.0 (Draxler and Jänsch 2004) was used, except for the interviews (cf. below), which were recorded using Audacity (Audacity Team 2017). All recordings had a sample rate of 16 bit/44.1 kHz and were saved as .wav-files. Prior to the first recording session, participants signed a declaration of consent. They had two appointments, the first one lasting about 75 min, the second one about 105 min. The participants received a reimbursement of 15CHF per 30 min, resulting in a payment of 60CHF for most participants.

At the first appointment, an interview of about 10 min was conducted to gather information about the sociolinguistic background of the participants and for them to get used to the recording situation. After the interview, the training phase began. Participants were instructed to read three sentences with different target words both in normal and in fast speech tempo. These recordings served as the basis for establishing each speaker’s time limit for each sentence during the rest of the experiment, which was implemented using a time bar. The average duration for normal and fast speech tempo was calculated separately for each speaker. This value plus 400 ms served as the basis for the time bar. The speakers read four further test sentences in each speech tempo first silently, and when the time bar appeared, aloud. Participants were instructed to repeat the sentences before the time indicated by the bar was over. During the appearance of the time bar, the written sentences were not

presented anymore. An additional purpose of the training phase was for the participants to get used to the Swiss German spelling system by Dieth (1986), which was used for the Alemannic production task. After the training, the actual experiment took place. Each sentence was repeated five times in both tempo conditions, resulting in 340 recordings for each participant from LU and 310 recordings for each participant from ZH. All stimuli were recorded in one session and presented in semi-random order, with two of the same stimuli never appearing twice in a row. Participants were instructed to start the sentences from the beginning in case they made a mistake.

At the second appointment, the same participants' productions of SSG were recorded. With the first six training sentences (three for each tempo), again, the durations for the time bars were calculated for each speaker. After this, the participants read four further training sentences in normal and fast speech tempo, first silently, then aloud with the time limit given by the time bar. For the SSG stimuli, the sentences were written in the official standard German orthography. After the training, the experiment began. Each sentence was repeated five times for each tempo, resulting in 230 recordings for each participant from both LU and ZH. All stimuli were recorded in one session and presented in semi-random order, with two of the same stimuli never appearing twice in a row. Participants were instructed to start the sentences from the beginning in case they made a mistake. After the production experiment, the speakers participated in a perception experiment also focusing on segment durations, which is not part of this study. Lastly, the conductor of the experiment filled out the main information about the speakers, i.e., name, speaker ID, place of residence, date of birth, and date of recording, in a form, before further preparing the data for the analysis.

2.4 Data preparation

After the recordings were saved, they were automatically segmented using WebMAUS (Schiel 1999), selecting in the language annotation settings the option German Dieth (CH) for the Alemannic recordings and German (DE) for the SSG recordings. The phonetic segments were manually adjusted with the EMU-webApp (Winkelmann and Raess 2014), using the following procedure: The beginning and end segments of each sentence were corrected if necessary. For the sentences beginning with a plosive, the release was used as the beginning of the first segment. If speakers paused within a sentence, the pause segment was removed and its duration was subtracted from the duration of the sentence. If there was more than one longer pause, the recording was not included in the analysis. Each of the phonetic segments of the target words was precisely adjusted. The plosives were

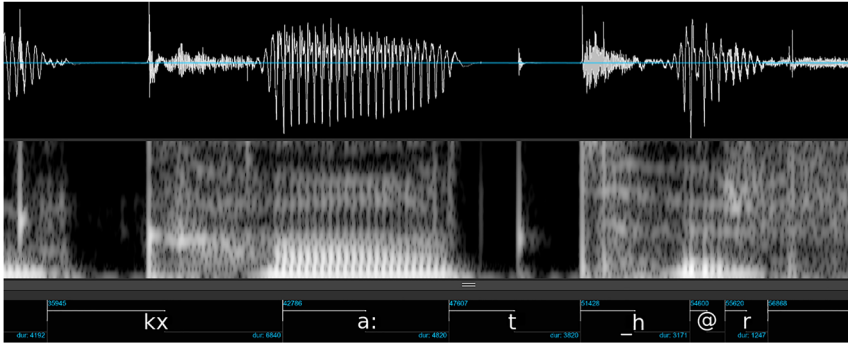


Figure 1: Segmentation of the target word *Kater* (tomcat), pronounced as [ˈkxɔːtəɾ], spoken by LU_0001, a young female speaker from LU, in SSG.

segmented into two phases: closure, indicated by /b̥ d̥ ɡ̊/ for lenis and /p t k/ for fortis, and release (VOT), indicated by _h (which was not included in the analyses). As VOT is known to not significantly differ between lenis and fortis plosives, this study focuses on closure duration only. An example for the segmentation can be seen in Figure 1.

If participants produced a target word in a different way than expected, i.e., accidentally produced another word or pronounced the word incorrectly, the recording was excluded from the analysis. Further recordings were excluded if participants made mistakes that resulted in an alteration of the number of syllables, the recording was cut at the beginning or end, or the quality of the recordings that were not conducted at the University of Zurich was not sufficient. Unfortunately, for some speakers from LU, there were quality issues that resulted in a relatively high number of exclusions. The rest of the aforementioned exclusion criteria only included a small number of recordings. Ultimately, from the total of 13,000 of the Alemannic recordings (6,800 for LU, 6,200 for ZH), 11,182 were used for the analysis (5,514 for LU, 5,668 for ZH). From the total of 9,200 of the SSG recordings (4,600 each from LU and ZH), 8,431 were used for the analysis (4,310 for LU and 4,121 for ZH).

2.5 Measurements

The duration of each sentence, the duration of each target word, and the word-medial vowel and closure durations were measured using the emuR package (Winkelmann et al. 2021) in RStudio (version 2022.07.2; R version 4.0.2; RStudio Team 2020), which provides a direct link between the EMU-webApp and R. In this study, only the measurements of the closure duration of the plosives were included, as VOT

is assumed not to differ between lenis and fortis plosives. The sentences were measured in milliseconds, and afterwards, the AR was calculated in terms of MSD by dividing the duration of the sentences by the number of syllables of each sentence (pauses had already been excluded). To obtain the relative vowel and closure durations, their absolute duration was divided by the word duration. This normalization procedure is appropriate considering the similar phonotactic and prosodic makeup of the words used for the stimuli, i.e., disyllabic trochees. Lastly, $V/(V + C)$ ratios were calculated by dividing the duration of the vowel by the duration of the whole $V + C$ sequence.

2.6 Statistical analyses

Statistical analyses were performed in RStudio (RStudio Team 2020). Linear mixed-effects models were fitted for all the analyses using the *lme4* (version 1.1-29) and the *lmerTest* (version 3.1-3) packages (Bates et al. 2015; Kuznetsova et al. 2017). For the model regarding AR, the dependent variable was *MSD*. Fixed effects were *variety* (Alemannic vs. SSG), *tempo* (normal vs. fast speech tempo), *region* (LU vs. ZH), and *age* (younger vs. older), including two-way interactions. Random intercepts were added for *speaker* and *word*.

For the models regarding the durations, either *relative vowel duration*, *relative closure duration* or *V/(V + C) ratio* were defined as the dependent variable. Fixed effects were *variety*, *VC combination*, *region*, *age*, and *tempo*, including two-way interactions. Random intercepts were added for *speaker* as well as *word* nested within *VC combination*.

A Type II ANOVA was calculated for each model using the R package *car* (version 3.1-0; Fox and Weisberg 2018), which yielded the Chi-square and *p* values reported in the Results section of this paper. For the interactions that turned out to be significant, pairwise comparisons using Tukey's tests were calculated using the R package *emmeans* (version 1.7.2; Lenth 2021). Additionally, Tukey's tests were calculated in case of a significant effect of *VC combination* in order to compare the durations of each combination to each other.

The .csv files as well as the R script for the analyses will be available under osf.io/y5c7d.

3 Results

In this section, the results from the statistical models and the pairwise comparisons are presented for AR, relative vowel and closure durations, and $V/(V + C)$ ratios.

Table 1a: Mean MSD and standard deviations (in parentheses) in ms and increase/decrease in % (%I/D) in Alemannic.

Tempo	LU		%I/D	ZH		%I/D
	Young	Old		Young	Old	
Normal		204.14 (40.54)			219.12 (39.04)	7.34
	199.07 (42.94)	209.55 (37.07)	5.27	210.56 (37.38)	227.46 (38.82)	8.02
Fast		175.11 (43.73)			178.45 (31.01)	1.64
	171.05 (38.44)	179.38 (29.75)	4.87	179.45 (31.72)	177.47 (30.27)	−1.10

Table 1b: Mean MSD and standard deviations (in parentheses) in ms and increase/decrease in % (%I/D) in SSG.

Tempo	LU		%I/D	ZH		%I/D
	Young	Old		Young	Old	
Normal		207.18 (31.71)			210.57 (28.90)	1.91
	192.73 (20.95)	220.42 (34.04)	14.37	200.08 (27.25)	221.47 (26.43)	5.70
Fast		173.47 (27.09)			171.17 (22.76)	−1.33
	162.48 (21.88)	183.52 (27.51)	12.95	165.79 (23.31)	176.34 (20.96)	6.30

3.1 Articulation rate

The means and standard deviations of the MSD per group are shown in Tables 1a (Alemannic) and 1b (SSG).

The linear mixed-effects model revealed significant main effects of *age* and *tempo* on MSD, as can be seen in Table 2. As expected, older speakers had an overall longer MSD compared to younger speakers, which can also be seen in Figure 2.

Pairwise comparisons of the significant interactions revealed that LU and ZH speakers did not differ significantly from each other when comparing them in dialect and standard speech separately as well as in fast and normal tempo separately, although a strong trend occurred: Similarly to the findings by Zihlmann (2020a), speakers from the western area of LU did not have a slower AR than those from ZH. Instead, the opposite pattern occurred not only in Alemannic, in accordance with the findings in Zihlmann (2020a), but also in SSG, which is even more surprising. This difference might be due to the fact that read speech, which was also investigated by Zihlmann (2020a), was analyzed, while previous research mainly focused on spontaneous speech (Leemann and Siebenhaar 2010). Another surprising result was that

Table 2: Statistical ANOVA output of the linear mixed-effects model for MSD.

	Chisq	Df	Pr (<Chisq)
<i>Variety</i>	1.0027	1	0.31665
<i>Region</i>	0.4017	1	0.52619
<i>Age</i>	5.3897	1	0.02026*
<i>Tempo</i>	14,373.2223	1	<2e-16***
<i>Variety:region</i>	166.7418	1	<2e-16***
<i>Variety:age</i>	24.5996	1	<2e-16***
<i>Variety:tempo</i>	2.421	1	0.11972
<i>Region:age</i>	0.0866	1	0.76855
<i>Region:tempo</i>	228.9082	1	<2e-16***
<i>Age:tempo</i>	287.4623	1	<2e-16***

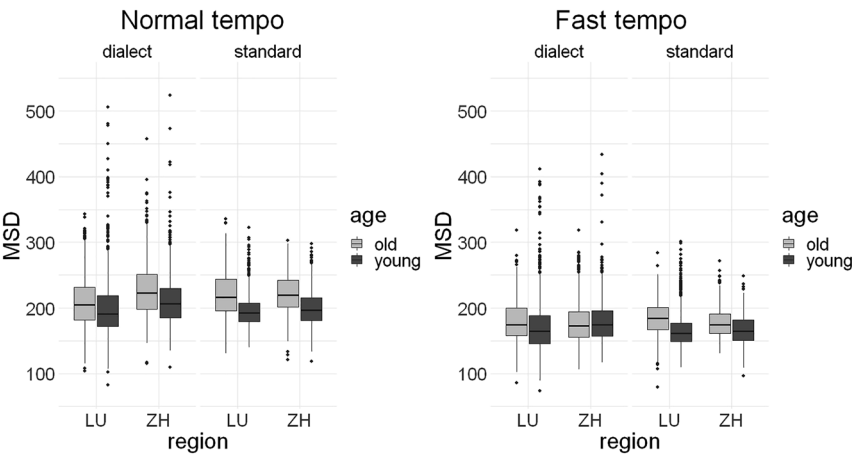


Figure 2: MSD in ms (y axis) in normal (left) and fast (right) speech tempo in both dialect and standard speech of LU and ZH speakers (x axis); younger speakers in dark gray, older speakers in light gray.

the age difference turned out to be significant only in SSG ($z = 3.266$, $p = 0.0011$), probably due the large amount of variability in Alemannic for the younger speakers, which can be seen in Figure 2. Comparing Alemannic to SSG, ZH speakers had a significantly higher MSD in Alemannic than in SSG ($z = 7.191$, $p < 0.0001$), while it was the other way around for LU speakers ($z = -7.051$, $p < 0.0001$).

As proposed by Quené (2007), the JND, i.e., the perceptual threshold, for AR lies at 5 %. The percentage values of increase (positive numbers) and decrease (negative numbers) are therefore also shown in Tables 1a and 1b. Despite not all being

significant, most of the regional and age-related differences are above the JND, particularly in the normal speech condition. The difference between Alemannic and SSG did not reach the 5 % in any case. To conclude, the age-related differences in AR were most stable and confirm H1a, while the regional differences yielded surprising results, leading to a rejection of H1b. The differences between Alemannic and SSG are significant but likely not perceivable. Therefore, when taking the JND into account, H1c ultimately must be rejected.

3.2 Relative vowel durations

The means and standard deviations of the relative vowel durations per group are shown in Tables 3a (Alemannic) and 3b (SSG).

Table 3a: Mean relative vowel durations and standard deviations (in parentheses) in Alemannic.

Tempo	Category	LU		ZH	
		Young	Old	Young	Old
Normal	VC:	0.193 (0.059)	0.210 (0.060)	0.194 (0.052)	0.197 (0.053)
	VC	0.267 (0.089)	0.267 (0.083)	0.206 (0.051)	0.216 (0.060)
	V:C:	0.299 (0.072)	0.320 (0.066)	0.291 (0.058)	0.321 (0.069)
	V:C	0.392 (0.086)	0.431 (0.068)	0.388 (0.066)	0.451 (0.070)
Fast	VC:	0.201 (0.059)	0.219 (0.058)	0.206 (0.053)	0.218 (0.057)
	VC	0.264 (0.092)	0.265 (0.078)	0.216 (0.052)	0.221 (0.065)
	V:C:	0.290 (0.071)	0.316 (0.066)	0.296 (0.059)	0.316 (0.072)
	V:C	0.380 (0.084)	0.418 (0.068)	0.386 (0.070)	0.428 (0.076)

Table 3b: Mean relative vowel durations and standard deviations (in parentheses) in SSG.

Tempo	Category	LU		ZH	
		Young	Old	Young	Old
Normal	VC:	0.154 (0.041)	0.164 (0.037)	0.157 (0.039)	0.159 (0.040)
	VC	0.217 (0.084)	0.217 (0.090)	0.211 (0.068)	0.208 (0.074)
	V:C:	0.263 (0.048)	0.283 (0.050)	0.275 (0.052)	0.296 (0.056)
	V:C	0.327 (0.077)	0.343 (0.079)	0.309 (0.074)	0.345 (0.080)
Fast	VC:	0.162 (0.043)	0.174 (0.041)	0.173 (0.043)	0.172 (0.045)
	VC	0.222 (0.078)	0.223 (0.091)	0.224 (0.064)	0.221 (0.070)
	V:C:	0.254 (0.050)	0.274 (0.049)	0.277 (0.058)	0.303 (0.057)
	V:C	0.326 (0.082)	0.338 (0.084)	0.327 (0.080)	0.346 (0.082)

Table 4: Statistical ANOVA output of the linear mixed-effects model for relative vowel durations.

	Chisq	Df	Pr (>Chisq)
<i>Variety</i>	107.3764	1	<2.2e-16***
<i>VC combination</i>	135.694	3	<2.2e-16***
<i>Region</i>	0.0001	1	0.992977
<i>Age</i>	12.5015	1	0.0004066***
<i>Tempo</i>	9.2076	1	0.0024102**
<i>Variety:VC combination</i>	66.0066	3	3.055e-14***
<i>Variety:region</i>	104.095	1	<2.2e-16***
<i>Variety:age</i>	37.5208	1	9.045e-10***
<i>Variety:tempo</i>	19.2723	1	1.133e-05***
<i>VC combination:region</i>	69.7562	3	4.814e-15***
<i>VC combination:age</i>	377.556	3	<2.2e-16***
<i>VC combination:tempo</i>	13.8473	3	<2.2e-16***
<i>Region:age</i>	0.0076	1	0.930673
<i>Region:tempo</i>	29.6663	1	5.132e-08***
<i>Age:tempo</i>	3.7364	1	0.053238

The statistical model revealed significant main effects of *variety*, *VC combination*, *age*, and *tempo*, also shown in Table 4. Relative vowel durations were overall longer in dialect compared to standard speech. Regarding age, older speakers produced longer vowel durations compared to younger speakers. Comparing the two tempo conditions to each other, the relative vowel duration turned out to have higher values in fast compared to normal speech.

The pairwise comparisons for the significant interaction revealed the following: There was no significant regional difference in either dialect or standard and in either normal or fast tempo for none of the *VC combinations*. An age-related difference was found for both Alemannic ($z = 3.893$, $p = 0.0001$) and SSG ($z = 2.488$, $p = 0.0129$), with older speakers producing longer vowels, but only significantly so in the *VC combinations* V:C ($z = 6.197$, $p < 0.0001$) and V:C: ($z = 4.075$, $p < 0.0001$). The combination V:C was significantly longer in normal compared to fast speech ($z = -4.827$, $p < 0.0001$), while VC ($z = 3.727$, $p = 0.0002$) and VC: ($z = 1.169$, $p < 0.0001$) were longer in fast compared to normal speech. This explains the direction of the main effect of tempo and is also not surprising, as longer segments are known to be shortened with acceleration in AR a lot more in comparison to short segments (e.g. Arvaniti 1999; Klatt 1973).

To have a better overview on how many vowel categories there are and on possible differences between the factors of interest, pairwise comparisons were made between the consecutive *VC combinations* from shortest to longest. This is

shown in Tables 5a (Alemannic) and 5b (SSG) and is also visible in Figures 3a (normal tempo) and 3b (fast tempo).

The results show that there are three phonetic categories of relative vowel durations in Alemannic, which are rather stable across tempo, at least for the ZH

Table 5a: *z* Ratios and *p* values of the pairwise comparison (Tukey test) of the relative vowel durations in Alemannic.

Tempo		VC: – VC		VC – V:C:		V:C: – V:C	
		<i>z</i> Ratio	<i>p</i> Value	<i>z</i> Ratio	<i>p</i> Value	<i>z</i> Ratio	<i>p</i> Value
Normal	LU						
	Young	–2.788	0.0272*	–1.19	0.6331	–3.823	0.0008**
	Old	–2.839	0.0234*	–1.515	0.4284	–4.494	<0.0001**
	ZH						
Fast	Young	–0.611	0.9286	–3.52	0.0024*	–4.324	0.0001**
	Old	–0.894	0.8082	–4.395	0.0001**	–5.758	<0.0001**
	LU						
	Young	–2.485	0.0623	–0.878	0.8163	–3.743	0.001*
	Old	–2.172	0.131	–1.608	0.3738	–4.114	0.0002**
	ZH						
	Young	–0.5	0.9592	–3.17	0.0083*	–3.803	0.0008**
	Old	–0.268	0.9932	–9.254	<0.0001**	–4.864	<0.0001**

Table 5b: *z* Ratios and *p* values of the pairwise comparison (Tukey test) of the relative vowel durations in SSG.

Tempo		VC: – VC		VC – V:C:		V:C: – V:C	
		<i>z</i> Ratio	<i>p</i> Value	<i>z</i> Ratio	<i>p</i> Value	<i>z</i> Ratio	<i>p</i> Value
Normal	LU						
	Young	–1.266	0.585	–1.889	0.2327	–2.083	0.1588
	Old	–0.965	0.7696	–2.444	0.069	–1.93	0.2154
	ZH						
Fast	Young	–1.553	0.4057	–2.429	0.0717	–1.24	0.6012
	Old	–1.372	0.5173	–3.193	0.0077*	–1.831	0.2586
	LU						
	Young	–1.446	0.4704	–1.265	0.585	–2.262	0.107
	Old	–1.138	0.666	–1.814	0.2668	–1.993	0.1905
	ZH						
	Young	–1.419	0.4871	–1.89	0.2325	–1.489	0.444
	Old	–1.219	0.6145	–2.874	0.0211*	–1.317	0.5521

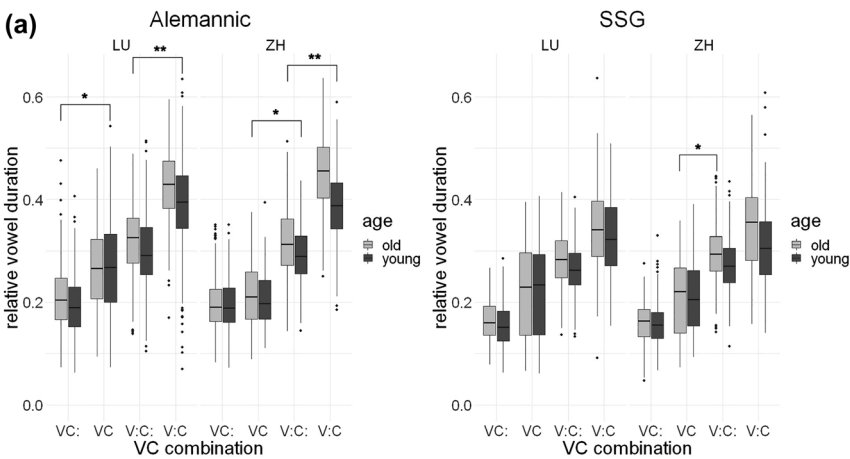


Figure 3a: Relative vowel durations (y axis) in normal speech tempo in Alemannic (left) and SSG (right) for all four VC combinations (x axis); younger speakers in dark gray, older speakers in light gray.

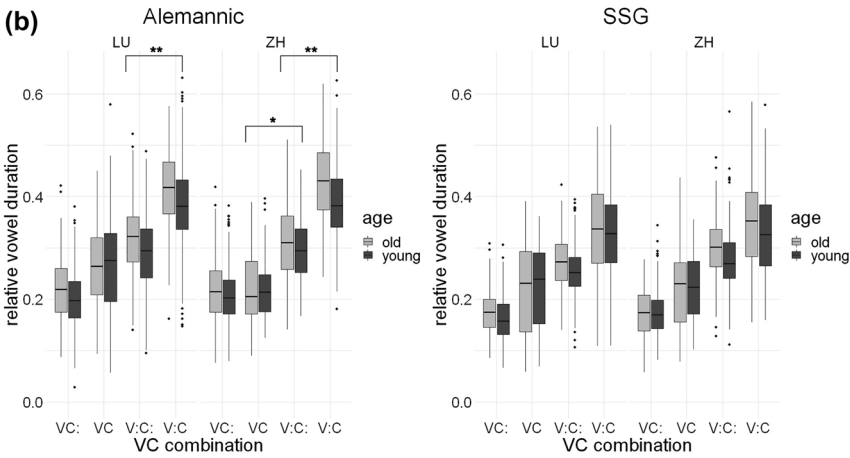


Figure 3b: Relative vowel durations (y axis) in fast speech tempo in Alemannic (left) and SSG (right) for all four VC combinations (x axis); younger speakers in dark gray, older speakers in light gray.

speakers. Still, they differ from each other between regions: While LU speakers differentiate so much between VC: and VC that VC and V:C do not differ significantly from each other, ZH speakers produce the short vowels in VC: and VC almost identically. The categories for ZH speakers are consistent with the evidence from Zihlmann (2020b), who found the same for Bernese and ZH speakers and invoked the

possibility for more than two (i.e., long and short) vowel categories. The findings of the present study seem to confirm the existence of three categories, at least for Alemannic. Still, the categories for LU speakers seem to be unique and are one of the major differences between the two regions in Alemannic. The categories are much more similar in SSG. In addition, it is apparent that the vowel durations, specifically of the combination V:C, are shorter in standard speech, leading them to be closer together and to a reduction to only two categories, i.e., short and long vowels. Although this is what the analysis revealed, it is clear that there are still quite large differences between *VC combinations*. The main question here remains if these differences are perceivable.

3.3 Relative closure durations

The means and standard deviations of the relative closure durations per group are shown in Tables 6a (Alemannic) and 6b (SSG).

Table 7 shows the model output, which revealed main effects of *variety*, *VC combination*, *region*, *age*, and *tempo*. Similarly to the results on the relative vowel durations, closure durations were significantly longer in dialect compared to standard productions. Generally, LU speakers produced significantly longer closure durations compared to ZH speakers, which confirms H2, while older speakers produced longer durations than younger speakers. As for tempo, the normal tempo condition resulted in longer relative closure durations compared to the fast tempo condition.

Moving on to the pairwise comparisons, the difference between fast and normal speech is significant for all combinations: While the relative closure durations are

Table 6a: Mean relative closure durations and standard deviations (in parentheses) in Alemannic.

Tempo	Category	LU		ZH	
		Young	Old	Young	Old
Normal	V:C	0.151 (0.042)	0.150 (0.044)	0.143 (0.036)	0.141 (0.036)
	VC	0.192 (0.065)	0.184 (0.065)	0.154 (0.049)	0.164 (0.055)
	V:C:	0.266 (0.058)	0.274 (0.055)	0.248 (0.051)	0.269 (0.041)
	VC:	0.356 (0.078)	0.364 (0.082)	0.348 (0.073)	0.378 (0.065)
Fast	V:C	0.528 (0.043)	0.152 (0.041)	0.147 (0.038)	0.148 (0.040)
	VC	0.195 (0.063)	0.193 (0.073)	0.154 (0.049)	0.167 (0.051)
	V:C:	0.262 (0.059)	0.263 (0.056)	0.239 (0.046)	0.262 (0.044)
	VC:	0.346 (0.081)	0.343 (0.080)	0.334 (0.071)	0.350 (0.063)

Table 6b: Mean relative closure durations and standard deviations (in parentheses) in SSG.

Tempo	Category	LU		ZH	
		Young	Old	Young	Old
Normal	V:C	0.129 (0.034)	0.127 (0.033)	0.119 (0.031)	0.122 (0.032)
	VC	0.170 (0.061)	0.172 (0.078)	0.146 (0.058)	0.168 (0.065)
	V:C:	0.234 (0.060)	0.241 (0.059)	0.201 (0.040)	0.227 (0.039)
	VC:	0.300 (0.058)	0.311 (0.066)	0.270 (0.050)	0.303 (0.057)
Fast	V:C	0.131 (0.036)	0.130 (0.034)	0.118 (0.041)	0.127 (0.033)
	VC	0.168 (0.060)	0.168 (0.064)	0.150 (0.059)	0.174 (0.064)
	V:C:	0.238 (0.060)	0.241 (0.056)	0.204 (0.043)	0.226 (0.039)
	VC:	0.287 (0.061)	0.305 (0.065)	0.264 (0.054)	0.298 (0.057)

Table 7: Statistical ANOVA output of the linear mixed-effects model for relative closure durations.

	Chisq	Df	Pr (>Chisq)
<i>Variety</i>	701.3359	1	<2.2e-16***
<i>VC combination</i>	299.6935	3	<2.2e-16***
<i>Region</i>	8.7334	1	0.0031244**
<i>Age</i>	4.6181	1	0.0316359*
<i>Tempo</i>	39.0896	1	4.048e-10***
<i>Variety:VC combination</i>	13.4204	3	<2.2e-16***
<i>Variety:region</i>	36.0178	1	1.955e-09***
<i>Variety:age</i>	11.3375	1	0.0007596***
<i>Variety:tempo</i>	15.5133	1	8.193e-05***
<i>VC combination:region</i>	23.7374	3	2.834e-05***
<i>VC combination:age</i>	125.388	3	<2.2e-16***
<i>VC combination:tempo</i>	148.2991	3	<2.2e-16***
<i>Region:age</i>	2.8578	1	0.09093
<i>Region:tempo</i>	0.1485	1	0.699966
<i>Age:tempo</i>	0.6735	1	0.411822

longer in fast speech for V:C ($z = 3.060$, $p = 0.0022$) and VC ($z = 2.373$, $p = 0.0176$), they are longer in normal speech for V:C: ($z = -3.009$, 0.0026) and VC: ($z = -12.472$, $p < 0.0001$). This could be expected as longer segments are shortened more in fast speech than short segments (Arvaniti 1999; Klatt 1973). When comparing younger to older speakers, older speakers produce longer closures only in SSG ($z = 2.672$, $p = 0.0075$). Looking at Tables 6a and 6b and at Figures 4a and 4b, it is apparent that this is probably due to the productions from the younger ZH speakers. There is a

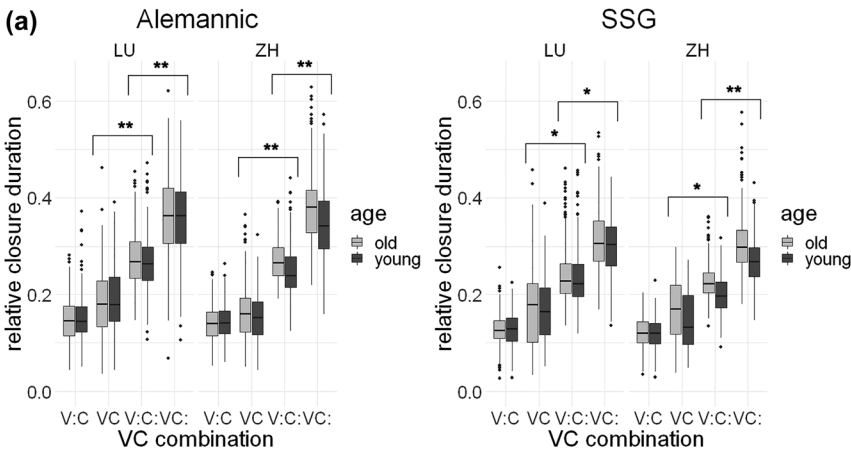


Figure 4a: Relative closure durations (y axis) in normal speech tempo in Alemannic (left) and SSG (right) for all four VC combinations (x axis); younger speakers in dark gray, older speakers in light gray.

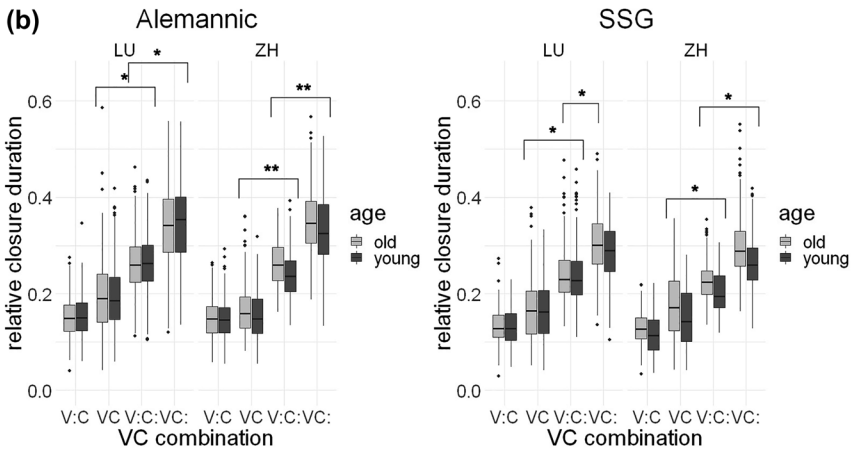


Figure 4b: Relative closure durations (y axis) in fast speech tempo in Alemannic (left) and SSG (right) for all four VC combinations (x axis); younger speakers in dark gray, older speakers in light gray.

much larger age-related difference for ZH speakers than for LU speaker in both Alemannic and SSG. Furthermore, the productions from the older ZH speakers look much more similar to those from the LU speakers. An explanation for this could be that younger speakers from the urban area of ZH, have increasing contact to GSG. While GSG speakers' productions of word-medial fortis consonants are also longer than those of lenis consonants (Jessen 1998), this difference is much smaller than for

Swiss-German speakers. Instead, in GSG, aspiration is the main cue to differentiate between lenis and fortis (Jessen 1998).

Similarly to the vowel durations, pairwise comparisons for the *VC combinations* were calculated, which are shown in Tables 8a (normal tempo) and 8b (fast tempo).

Table 8a: *z* Ratios and *p* values of the pairwise comparison (Tukey test) of the relative closure durations in Alemannic.

Tempo		V:C – VC		VC – V:C:		V:C: – VC:	
		<i>z</i> Ratio	<i>p</i> Value	<i>z</i> Ratio	<i>p</i> Value	<i>z</i> Ratio	<i>p</i> Value
Normal	LU						
	Young	–2.05	0.1697	–3.936	0.0005**	–4.563	<0.0001**
	Old	–2.207	0.1214	–4.24	0.0001**	–4.333	0.0001**
	ZH						
Fast	Young	–0.706	0.8948	–4.906	<0.0001**	–5.449	<0.0001**
	Old	–1.396	0.5017	–5.472	<0.0001**	–5.839	<0.0001**
	LU						
	Young	–2.116	0.1479	–3.391	0.0039*	–3.905	0.0005**
	Old	–2.56	0.0512	–3.029	0.0131*	–3.555	0.0021*
	ZH						
	Young	–0.528	0.9523	–4.388	0.0001**	–5.182	<0.0001**
	Old	–1.118	0.6783	–4.99	<0.0001**	–4.617	<0.0001**

Table 8b: *z* Ratios and *p* values of the pairwise comparison (Tukey test) of the relative closure durations in SSG.

Tempo		V:C – VC		VC – V:C:		V:C: – VC:	
		<i>z</i> Ratio	<i>p</i> Value	<i>z</i> Ratio	<i>p</i> Value	<i>z</i> Ratio	<i>p</i> Value
Normal	LU						
	Young	–1.557	0.4035	–2.777	0.0281*	–3.02	0.0135*
	Old	–1.83	0.259	–2.906	0.0192*	–3.208	0.0073*
	ZH						
Fast	Young	–1.142	0.6637	–2.773	0.0284*	–3.743	0.001*
	Old	–2.109	0.1502	–2.983	0.0151*	–4.151	0.0002**
	LU						
	Young	–1.659	0.3456	–3.073	0.0114*	–2.504	0.0593
	Old	–1.779	0.2836	–3.19	0.0078*	–3.153	0.0088*
	ZH						
	Young	–1.301	0.5623	–2.583	0.0481*	–3.255	0.0062*
	Old	–2.022	0.1797	–2.584	0.0481*	–3.801	0.0008**

Firstly, there are clearly three phonetic categories in Alemannic, i.e., lenis, fortis, and extrafortis, which are stable across tempo. With the exception of the LU speakers in fast speech tempo (where V:C: and VC: still are close to being significantly different from each other), those three categories are also found in SSG. This also means that, although younger ZH speakers behave differently from the other groups, they still maintain these three categories at this point.

As stated above, some compensation between vowel and closure duration is expected. Thus, although some significant differences in vowel and consonant categories have been revealed at this point, the entire V + C sequence should also be taken into account. That's why the additional measure of V/(V + C) ratios was selected for this study and its results presented in the following section.

3.4 V/(V + C) ratios

The means and standard deviations of the V/(V + C) ratios per group are shown in Tables 9a (Alemannic) and 9b (SSG).

As can be seen in Table 10, the model for the V/(V + C) ratios revealed significant main effects for *variety*, *VC combination*, and *tempo*. In general, the ratios in dialect speech have higher values compared to standard speech. As for tempo, the values of the ratios are higher in fast speech tempo.

Looking at the significant interactions, the tempo difference turned out to be significant for V:C (higher ratios in normal speech) ($z = -3.892, p = 0.0001$) and VC: (higher ratios in fast speech) ($z = 13.656, p < 0.0001$), confirming, again, that longer segments are shortened more than shorter ones with an acceleration of speech tempo. Furthermore, it was revealed that the higher values of the ratios in Alemannic

Table 9a: Mean V/(V + C) ratios and standard deviations (in parentheses) in Alemannic.

Tempo	Combination	LU		ZH	
		Young	Old	Young	Old
Normal	VC:	0.351 (0.085)	0.367 (0.092)	0.358 (0.080)	0.341 (0.070)
	V:C:	0.526 (0.098)	0.537 (0.086)	0.540 (0.071)	0.540 (0.069)
	VC	0.579 (0.086)	0.593 (0.086)	0.575 (0.085)	0.570 (0.080)
	V:C	0.716 (0.086)	0.742 (0.071)	0.729 (0.070)	0.760 (0.063)
Fast	VC:	0.368 (0.090)	0.393 (0.091)	0.383 (0.082)	0.383 (0.076)
	V:C:	0.523 (0.102)	0.545 (0.089)	0.552 (0.071)	0.542 (0.080)
	VC	0.569 (0.096)	0.582 (0.091)	0.587 (0.086)	0.568 (0.087)
	V:C	0.709 (0.084)	0.733 (0.066)	0.721 (0.071)	0.740 (0.073)

Table 9b: Mean V/(V + C) ratios and standard deviations (in parentheses) in SSG.

Tempo	Combination	LU		ZH	
		Young	Old	Young	Old
Normal	VC:	0.340 (0.071)	0.348 (0.062)	0.367 (0.063)	0.342 (0.064)
	V:C:	0.530 (0.091)	0.542 (0.084)	0.577 (0.062)	0.564 (0.065)
	VC	0.556 (0.092)	0.559 (0.096)	0.595 (0.078)	0.556 (0.084)
	V:C	0.713 (0.074)	0.726 (0.073)	0.719 (0.069)	0.733 (0.075)
Fast	VC:	0.362 (0.080)	0.365 (0.072)	0.396 (0.072)	0.364 (0.076)
	V:C:	0.519 (0.088)	0.533 (0.082)	0.575 (0.069)	0.571 (0.067)
	VC	0.565 (0.097)	0.563 (0.096)	0.605 (0.087)	0.561 (0.092)
	V:C	0.708 (0.078)	0.716 (0.077)	0.734 (0.081)	0.726 (0.075)

Table 10: Statistical ANOVA output of the linear mixed-effects model for V/(V + C) ratios.

	Chisq	Df	Pr (>Chisq)
<i>Variety</i>	39.9897	1	2.553e-10***
<i>VC combination</i>	1033.026	3	<2.2e-16***
<i>Region</i>	3.3227	1	0.068328
<i>Age</i>	0.3778	1	0.538762
<i>Tempo</i>	33.2214	1	8.224e-09***
<i>Variety:VC combination</i>	22.7385	3	4.578e-05***
<i>Variety:region</i>	86.4261	1	<2.2e-16***
<i>Variety:age</i>	23.6955	1	1.128e-06***
<i>Variety:tempo</i>	0.1347	1	0.713597
<i>VC combination:region</i>	36.4881	3	5.904e-08***
<i>VC combination:age</i>	99.7632	3	5.904e-08***
<i>VC combination:tempo</i>	168.9555	3	<2.2e-16***
<i>Region:age</i>	1.8448	1	0.174394
<i>Region:tempo</i>	11.6868	1	0.0006294***
<i>Age:tempo</i>	0.9925	1	0.319133

compared to SSG are only significant for the combinations VC ($z = 3.097$, $p = 0.0020$) and VC: ($z = 7.586$, $p < 0.0001$). Furthermore, there are significant regional differences in SSG, with higher ratios in ZH compared to LU speakers ($z = -2.906$, $p = 0.0037$). As for age, only V:C ($z = 2.118$, $p = 0.0341$) turned out to be significantly different, with older speakers producing higher ratios.

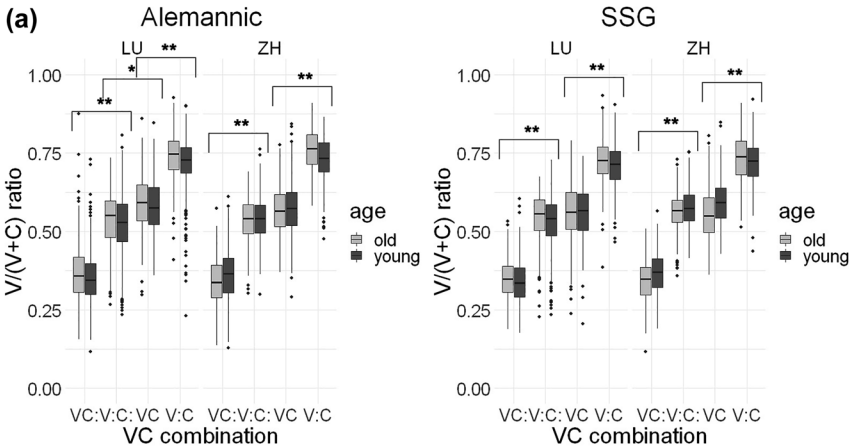


Figure 5a: $V/(V + C)$ ratios (y axis) in normal speech tempo in Alemannic (left) and SSG (right) for all four VC combinations (x axis); younger speakers in dark gray, older speakers in light gray.

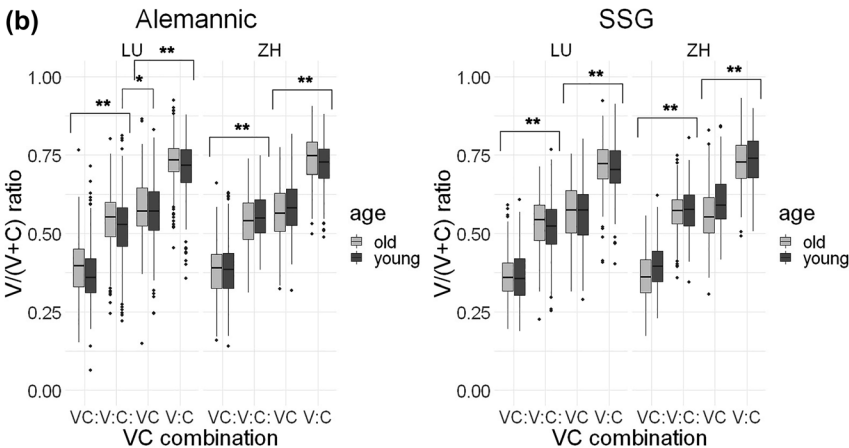


Figure 5b: $V/(V + C)$ ratios (y axis) in fast speech tempo in Alemannic (left) and SSG (right) for all four VC combinations (x axis); younger speakers in dark gray, older speakers in light gray.

Overall, the $V/(V + C)$ ratios are relatively similar to each other when comparing varieties, age groups and tempo conditions, which is also shown in Figures 5a (normal tempo) and 5b (fast tempo).

The most obvious difference seems to be that LU speakers produce four categories of VC combinations, which can also be seen in Tables 11a and 11b. The

Table 11a: z Ratios and p values of the pairwise comparison (Tukey test) of the V/(V + C) ratios in Alemannic.

		VC: – V:C:		V:C: – VC		VC – V:C	
Tempo		z Ratio	p Value	z Ratio	p Value	z Ratio	p Value
Normal	LU						
	Young	–1.813	<0.0001**	–3.422	0.0035*	–8.755	<0.0001**
	Old	–1.725	<0.0001**	–3.451	0.0031*	–9.424	<0.0001**
	ZH						
Fast	Young	–8.922	<0.0001**	–1.773	0.2862	–8.015	<0.0001**
	Old	–9.758	<0.0001**	–1.432	0.4793	–9.896	<0.0001**
	LU						
	Young	–8.52	<0.0001**	–2.648	0.0404*	–8.101	<0.0001**
	Old	–8.469	<0.0001**	–1.841	0.2542	–8.813	<0.0001**
	ZH						
	Young	–7.814	<0.0001**	–1.58	0.3901	–6.611	<0.0001**
	Old	–7.349	<0.0001**	–1.216	0.6169	–8.405	<0.0001**

Table 11b: z Ratios and p values of the pairwise comparison (Tukey test) of the V/(V + C) ratios in SSG.

		VC: – V:C:		V:C: – VC		VC – V:C	
Tempo		z Ratio	p Value	z Ratio	p Value	z Ratio	p Value
Normal	LU						
	Young	–8.539	<0.0001**	–0.24	0.9951	–7.264	<0.0001**
	Old	–8.691	<0.0001**	0.143	0.999	–7.729	<0.0001**
	ZH						
Fast	Young	–11.628	<0.0001**	–0.502	0.9585	–6.656	<0.0001**
	Old	–12.201	<0.0001**	0.719	0.8897	–9.461	<0.0001**
	LU						
	Young	–6.883	<0.0001**	–1.263	0.5864	–6.229	<0.0001**
	Old	–7.428	<0.0001**	–0.605	0.9307	–6.684	<0.0001**
	ZH						
	Young	–8.286	<0.0001**	–0.908	0.8007	–5.482	<0.0001**
	Old	–9.441	<0.0001**	0.699	0.8973	–7.266	<0.0001**

difference between V:C: and VC is not entirely stable, as it is not maintained in fast tempo (for older speakers) and in SSG. It is not clear if V:C: and VC can be considered two separate categories because they are also much closer together than VC: to V:C: or VC to V:C. At the same time, ZH speakers clearly produce three categories of V + C

sequences in both Alemannic and SSG. For now, it seems more reasonable to assume three categories in both LU and ZH, in accordance with the analysis proposed by Zihlmann (2020b). All in all, these findings confirm H3a, stating no regional differences, and H3b, expecting three categories of $V/(V + C)$ ratios.

The reported results will be further discussed in the next section.

4 Discussion and conclusion

The main goal of this study was to compare AR, vowel and closure durations as well as $V/(V + C)$ ratios between Alemannic and SSG, focusing on speakers from rural areas of LU and from urban areas of ZH. Furthermore, the stability of these measures across age and tempo was investigated.

4.1 Discussion of the results

Based on the results of this study, no strong conclusions can be formed on the differences in AR between Alemannic and SSG or between LU and ZH. Even the age-related differences, which turned out to be the most stable, are not as strong as expected, at least in Alemannic, where there is a lot of variability in the group of younger speakers. Ultimately, AR alone does not seem to be a reliable measure to distinguish between Alemannic and SSG or between LU and ZH. At least in the case of this study, it is rather a measure to control for when analyzing segment durations that are dependent on speech tempo.

The most prominent difference between Alemannic and SSG turned out to be that both vowel and closure durations are significantly longer in dialect compared to standard speech. Regarding the vowel durations, LU and ZH speakers clearly differ in terms of their vowel categories in Alemannic: While the three categories for LU speakers consisted of (1) VC:, (2) VC and V:C:, and (3) V:C, those for ZH speakers consisted of (1) VC: and VC, (2) V:C:, and (3) V:C. There seems to be a specific vowel pattern used by the speakers from ZH, which they only produce in Alemannic.

Note that the relative closure durations are the only measure with a main effect of region, which was due to LU speakers producing longer closures than ZH speakers. As stated above, this was most probably due to the shorter closure durations produced by younger speakers from ZH. This is possibly the most important finding in this study, as it might be due to a contact to GSG, which, in turn, could lead to a sound change in progress, reducing the phonetic consonant categories to lenis and fortis. It is particularly striking that younger speakers from ZH do not only produce shorter closure durations in SSG, but also in Alemannic. As for now, the distinction between

lenis, fortis, and extrafortis is still found in all groups, but the trend found here might develop into consonant productions much closer to those found in GSG in those regions that have high contact to GSG. This result offers a new perspective on Alemannic and SSG, as it has been assumed, until now, that SSG is only influenced by Alemannic (Hove 2002; Zihlmann 2020b). The next step here would be to investigate not only closure duration, but also VOT, which is known to be the primary cue to distinguish between lenis and fortis plosives in GSG. If younger speakers from ZH produce shorter closure durations and longer VOTs, there might be a change due to a trading relation between the two, meaning the more they aspirate, the shorter the closure becomes. A similar trend has been found for German speakers of the Bavarian and Saxon dialect, where VOT becomes increasingly important for younger speakers (Kleber 2018). At this point, it is speculative if the situation is similar for ZH speakers, but it is worth further investigation in both production and perception.

All in all, although some remarkable differences – mainly the difference in segment durations between Alemannic and SSG, the different vowel categories in LU compared to ZH speakers in Alemannic, and the shorter closure durations produced by younger ZH speakers – were found when looking at the segment durations on their own, the $V/(V + C)$ ratios are remarkably stable.

4.2 Remarks/limitations

There are some limitations due to practicability and time reasons that need to be mentioned, since they might have influenced the results to a certain amount.

Firstly, the number of speakers might not have been enough to detect either regional or generational differences in speaking behaviors. Yet, this limitation should be partly compensated by the high number of tokens per speaker. For future research, it is worth considering more participants and a lower number of tokens per speaker, considering that the inevitable manual adjustment of the segmentation process is a time-consuming task.

The speaker groups could also have been more homogenous in terms of age. The mean ages of the older speaker groups differed six years from each other between LU and ZH. Not only were the older speakers from LU younger than those from ZH, but they also differed more in age within the group. Furthermore, the younger LU group was, on average, almost six years older than the younger ZH group, meaning that the age difference between the younger and older speakers was somewhat smaller in LU than in ZH. Despite that, there was still a mean difference of nearly 30 years between the age groups in LU, which certainly can be sufficient for generational comparisons. The mean age difference between the groups from ZH was about 40 years, which is even more suitable for age-dependent comparisons.

Furthermore, some of the speakers tended to produce the fast sentences as fast as they possibly could, instead of speaking in a fast but natural speech tempo. It remains unclear how this may have influenced the results. Since it was the case for speakers from both regions, the groups are still comparable to each other.

Concerning the stimuli, there are three main remarks. First, it would have been optimal for the dialect data to have the same target words for both regions to make a better comparison. This, however, would result in a smaller number of stimuli, as both lexical and phonetic properties differ between the two regions. Therefore, the comparisons between regions concern groups of VC combinations rather than words. Second, the SSG stimuli did not have a category for bilabial plosives + [i], or one for velar plosives + [u]. This could have influenced the results. Since no minimal pairs could be found in these categories of SSG, the only solution would have been to remove these categories from the Alemannic stimuli as well and this, in turn, would have led to a further reduction in stimulus number. It ultimately was decided that the higher stimulus number was more important in this case. Third, the usage of the spelling system by Dieth (1986) for the Alemannic recordings but the standard German orthography for the SSG recordings might have influenced some of the participants' productions. Yet, this effect should be minor because the speakers read the sentences silently before speaking out loud and did not have the text in front of their eyes while speaking. Additionally, all unexpectedly pronounced tokens were excluded from the dataset.

To summarize, for future studies, a higher number of speakers per group is considered most important. If possible, a more homogeneous group in terms of age is also preferable; yet, an age difference of 30 years for one regional group and of 40 years for the other one is also acceptable considering the difficulties of participant recruiting.

4.3 Outlook

To conclude, the main findings of this study show that the most prominent difference between Alemannic and SSG lie in the longer segment durations in dialect compared to standard speech. Furthermore, there are three phonetic vowel categories in Alemannic that differ between LU and ZH. Regarding SSG, perceptual data would be needed to obtain the number of vowel categories, but there are at least two. The three V/(V + C) ratio categories and the phonetic consonant categories lenis, fortis, and extrafortis are stable across tempo, age, and region in both Alemannic and SSG. Most importantly, younger ZH speakers produced overall shorter closure durations, even if they still conserve the three consonant categories. This should be investigated

further by focusing on both the closure duration and the VOT of plosive consonants in ZH German in both production and perception.

Acknowledgments: First, I am indebted to Urban Zihlmann, Patrick Côte and Sabrina von Rotz for collecting the data. Urban Zihlmann and Melissa Bruno also helped with the manual corrections of annotated segment boundaries. Furthermore, I am grateful to Sandra Schwab for her valuable support with the statistical analysis. Lastly, I would like to thank Stephan Schmid and Felicitas Kleber for suggestions and feedback.

Statement of ethics: This research is in accordance with the principles of the Helsinki Declaration on human research participants' rights and is at all points adhered to the ethics procedures as established at the Faculty of Arts and Social Science of the University of Zurich. All participants gave informed consent by signing an appropriate declaration of consent before the experiment was conducted.

Funding sources: This study was funded by the Swiss National Science Foundation (SNSF), grant nr. 164377 and 190005.

Appendix

A1: Alemannic stimuli for LU speakers.

Vowel	Place of articulation (plosive)	VC combination	Target word	Carrier sentence	Translation
a	Bilabial	VC	Rabbi ['raβi]	Är esch mòu Rabbi gsee.	He once was a rabbi.
		V:C	Raben ['ra:βə]	Ech ha zwee Raabe gsee.	I saw two ravens.
		V:C:	Rappen ['rap:ə]	Si hèt zää Rappe zaaut.	She payed ten (Swiss) cents.
	Alveolar	VC:	–	–	–
		VC	–	–	–
		V:C	Kader ['kxɑ:ðər]	Si hèt s is Kaader gschafft.	She made it into the cadre.
		V:C:	Kater ['kxɑ:tər]	Är hèt e Kaater deheime.	He has a tomcat at home.
		VC:	Cutter ['kat:ər]	Ech bruuche ne Cutter zom Schniide.	I need a cutter to cut.
	Velar	VC	mag' ['maŋə]	Ech glòub ech mage nò.	I think I can still do it.

(continued)

Vowel	Place of articulation (plosive)	VC combination	Target word	Carrier sentence	Translation
i		V:C	Magen [ˈmaːgə]	Är hëtt de Maage gschpöört.	He felt the stomach.
		V:C	«tschaagge» [ˈtʃa:kə]	Si meint ech tschaagge hôt.	She means I dawdle today.
		VC:	Zagge [ˈʦak:ə]	Är hëd e Zagge ab.	He is a bit crazy.
	Bilabial	VC	Biber [ˈbi:bər]	Es hëtt e Biber im Zoo.	There is a beaver in the zoo.
		V:C	treiben [ˈtri:bə]	Si hëtt doch ,triibe‘ gseit.	But she said ‘drive’.
		V:C	Gripen [ˈɡri:pə(n)]	Si hënd de Griipen abgleent.	They voted against the Gripen (military aircraft).
	Alverolar	VC:	tippen [ˈtip:ə]	Si mönd das tippe a de Kasse.	They have to register a tip at the cash desk.
		VC	in der [ˈi dər]	Si esch no i de Schuëu.	She is still at school.
		V:C	Seide [ˈzi:də]	Si chòufed Siide vo China.	They buys silk from China.
	Velar	V:C	Seite [ˈzi:tə]	Das gòòt bes Siite nūünzä.	This goes till page nineteen.
		VC:	Sitte [ˈzit:ə]	Di guete Sitte send verbii.	The good manners are over.
		VC	Igel [ˈiɡu]	Mer hënd en Igu deheime.	We have a hedgehog at home.
		V:C	beigen [ˈbi:ɡə]	Ech tue chli biige hôt Òòbe.	I pile up a bit tonight.
		V:C	Miiggu [ˈmi:ku]	Ech ha de Miiggu gsee.	I saw Miiggu (nick-name for Emil).
		VC:	ticken [ˈtik:ə]	Ech ghööre s Ticke vo de Uhr.	I hear the ticking of the clock.
	Bilabial	VC	Rubel [ˈrubəl]	Si mönd met Rubel zaale.	They have to pay with roubles.
		V:C	Haube [ˈhu:bə]	Si hëtt e Huube gnäät.	She sewed a bonnet.
		V:C	hupen [ˈhu:pə]	Ech tue chli huupe bem Faare.	I honk a bit while I drive.
u		VC:	super [ˈʒup:ər]	Das esch scho super gsee.	This was great indeed.
	Alveolar	VC	Pudding [ˈpuːdɪŋ]	Ech ha gëschter Pudding ggässe.	Yesterday, I ate pudding.

(continued)

Vowel	Place of articulation (plosive)	VC combination	Target word	Carrier sentence	Translation
Velar		V:C	Bude ['bʊ:ɖə]	Är hëtt e Buude ghaa.	He had a stall.
		V:C:	Pute ['pu:tə]	Ech ha lieber Puute as Schtruuss.	I prefer turkey to ostrich.
		VC:	Butter ['bʊt:ər]	Ech muesch no Butter chòufe.	I still have to buy butter.
		VC	Sugus ['zʊɣʊs]	Ech ha nes Sugus im Muu.	I have a ‘Sugus’ (Swiss candy) in my mouth.
		V:C	saugen ['zʊ:ɣə]	Du muesch chli suuge zom Trenke.	You have to suck a bit to drink.
		V:C:	guugge ['ɣʊ:kə]	Ech tue chli guugge hõt zòòbe.	I play a wind instrument tonight.
		VC:	gugguus ['ɣʊk:u:s]	Ech sètt nò gug-guus sääge.	I should say hello.

A2: Alemannic stimuli for ZH speakers.

Vowel	Place of articulation (plosive)	VC combination	Target word	Carrier sentence	Translation
a	Bilabial	VC	Rabbi ['rɑɸi]	Èr isch mal Rabbi gsii.	He once was a rabbi.
		V:C	Raben ['rɑ:ɸə]	Ich ha zwee Raabe gsee.	I saw two ravens.
		V:C:	Rappen ['rɑp:ə]	Si hätt zää Rappe zalt.	She payed ten (Swiss) cents.
	Alveolar	VC:	–	–	–
		VC	–	–	–
		V:C	Kader ['kxɑ:ɖər]	Si hätt s is Kaader gschafft.	She made it into the cadre.
		V:C:	Kater ['kxɑ:tər]	Èr hätt en Kaater dehei.	He has a tomcat at home.
		VC:	–	–	I need a cutter to cut.

(continued)

Vowel	Place of articulation (plosive)	VC combination	Target word	Carrier sentence	Translation
i	Velar	VC	Cutter [ˈkat:ər]	Ich bruuche en Cutter zum Schniide.	–
		V:C	–	–	–
		V:C	hagen [ˈhø:ɡə]	De Puur findet haage wichtig.	The farmer finds fencing in important.
		V:C:	Haken [ˈhø:kə]	Ich han s an Haagge ghänkt.	I hung it on the hook
	Bilabial	VC:	Backe [ˈbøk:ə]	Si hätt sich i d Bagge bbisse.	She bit herself in the cheek.
		VC	Sieben (7) [ˈzi:bə]	Was git dänn sibe maal nüün?	What equals seven times nine?
		V:C	sieben [ˈzi:bə]	Muesch s Määl na siibe zerscht.	You have to sieve the flour first.
		V:C:	Gripen [ˈɡri:pə(n)]	Si händ de Griipen abgleent.	They voted against the Gripen (military aircraft).
	Alverolar	VC:	Sippe [ˈzi:p:ə]	Di ganzi Sippe chunnt.	The whole clan comes.
		VC	Seide [ˈzi:də]	Das isch Side vo China.	This is silk from China.
		V:C	meiden [ˈmi:də]	Ich tuen en miide bim Ässe.	I avoid him while eating.
		V:C:	Seite [ˈzi:tə]	Das gaat bis Siite nüünzää.	This goes till page nineteen.
	Velar	VC:	Sitte [ˈzi:tə]	Di guete Sitte sind verbii.	The good manners are over.
		VC	Tiger [ˈti:ɡər]	Ich wett en Tiger sträichle.	I want to stroke the tiger.
		V:C	beigen [ˈbi:ɡə]	Ich tue biige hüt zaabig.	I pile up tonight.
		V:C:	–	–	–
	Bilabial	VC:	Tigger [ˈtik:ər]	Ich tuen en Tigger i d Milch.	I put the milk watcher in the milk.
VC		Stube [ˈtu:bə]	Mir sind i de Stube gsii.	We were in the living room.	
V:C		Haube [ˈhu:bə]	Si hätt e Huube gnèët.	She sewed a bonnet.	
V:C:		hupen [ˈhu:pə]	Ich tue chli huupe bim Faare.	I honk a bit while I drive.	

(continued)

Vowel	Place of articulation (plosive)	VC combination	Target word	Carrier sentence	Translation
	Alveolar	VC:	Suppe [ˈzʊp:ə]	Ich wett e Suppe choche.	I want to cook a soup.
		VC	Pudding [ˈpuːdɪŋ]	Ich han geschter Pudding ggässe.	Yesterday, I ate pudding.
		V:C	Bude [ˈbʊ:ɖə]	Ich bin i de Buude gsii.	I was in the stall.
		V:C:	Pute [ˈpu:tə]	Ich han lieber Puute als Schtruuss.	I prefer turkey to ostrich.
	Velar	VC:	Butter [ˈbʊt:ər]	Ich mues na But-ter chaufe.	I still have to buy butter.
		VC	Sugus [ˈzʊɡʊs]	Ich han es Sugus im Muul.	I have a ‘Sugus’ (Swiss candy) in my mouth.
		V:C	saugen [ˈzʊ:ɡə]	Du muesch chli suuge zum Trinke.	You have to suck a bit to drink.
		V:C:	–	–	–
		VC:	Gugger [ˈɡʊɡ:ər]	Ich han en Gug-ger ghöört.	I heard a cuckoo.

A3: SSG stimuli.

Vowel	Place of articulation (plosive)	VC combination	Target word	Carrier sentence	Translation
a	Bilabial	VC	Rabbi [ˈrabi]	Er wollte Rabbi werden.	He wanted to become a rabbi.
		V:C	Rabe [ˈra:bə]	Ich soll doch Rabe lesen.	I should read ‘raven’.
		V:C:	–	–	–
	Alveolar	VC:	Rappe [ˈrapə]	Ich soll doch Rappe lesen.	I should read ‘(Swiss) cents’.
		VC	–	–	–
		V:C	Kader [ˈka:ɖə]	Er hat das Kader besetzt.	He occupied the cadre.
		V:C:	Kater [ˈka:tə]	Er will den Kater füttern.	He wants to feed the tomcat.

(continued)

Vowel	Place of articulation (plosive)	VC combination	Target word	Carrier sentence	Translation
i	Velar	VC:	Cutter [ˈka:tə]	Er muss den Cutter kaufen.	He must buy the cutter.
		VC	–	–	–
		V:C	Hagen [ˈha:ɡən]	Er muss auf Hagen warten.	He must wait for Hagen (first name).
		V:C:	Haken [ˈha:kən]	Sie muss noch Haken kaufen.	She still has to buy hooks.
		VC:	hacken [ˈhakən]	Er hat noch hacken müssen.	He still had to chop.
	Bilabial	VC	–	–	–
		V:C	–	–	–
		V:C:	–	–	–
		VC:	–	–	–
	Alveolar	VC	Widder [ˈviðə]	Er wollte Widder streicheln.	He wanted to stroke rams.
		V:C	wieder [ˈvi:də]	Ich wollte wieder sagen.	I wanted to say 'again'.
		V:C:	Bieter [ˈbi:tə]	Er will doch Bieter werden.	He wants to become a bidder.
		VC:	bitter [ˈbi:tə]	Das hat doch bitter geschmeckt.	That tasted bitter.
	Velar	VC	Tigger [ˈti ɡə]	Sie hat dich Tigger genannt.	She called you 'Tigger'.
		V:C	Tiger [ˈti:ɡə]	Sie will den Tiger füttern.	She wants to feed the tiger.
		V:C:	–	–	–
		VC:	Ticker [ˈti:kə]	Ich soll den Ticker nehmen.	I should take the ticker.
u	Bilabial	VC	–	–	–
		V:C	Tube [ˈtu:bə]	Er will die Tube nehmen.	He wants to take the tube.
		V:C:	Lupe [ˈlu:pə]	Er kann die Lupe brauchen.	He can make use of the magnifying glass.
		VC:	Suppe [ˈzupə]	Ich will die Suppe kochen.	I want to cook the soup.
	Alveolar	VC	Pudding [ˈpuɖɪŋ]	Sie wollte Pudding kochen.	She wanted to cook pudding.
		V:C	Puder [ˈpu:də]	Sie muss noch Puder kaufen.	She still needs to buy powder.
		V:C:	Pute [ˈpu:tə]	Er muss die Pute kaufen.	He must buy the turkey.

(continued)

Vowel	Place of articulation (plosive)	VC combination	Target word	Carrier sentence	Translation
Velar		VC:	Butter ['but e]	Ich muss noch Butter kaufen.	I still have to buy butter.
		VC	–	–	–
		V:C	–	–	–
		V:C:	–	–	–
		VC:	–	–	–

References

Arquint, Jachen Curdin, Iso Camartin & Robert Schläpfer. 1982. *Die Viersprachige Schweiz* [Quadrilingual Switzerland]. Zürich & Köln: Benziger.

Arvaniti, Amalia. 1999. Effects of speaking rate on the timing of single and geminate sonorants. In *Proceedings of the 14th International Congress of Phonetics Sciences*, vol. 1, 595–598. Berkeley: University of California.

Audacity Team. 2017. *Audacity(R): Free audio editor and recorder [Computer program]*. Audacity Team.

Bailey, Guy, Tom Wikle, Jan Tillery & Lori Sand. 1991. The apparent time construct. *Language Variation and Change* 3(3). 241–264.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.

Christen, Helen. 2019. Alemannisch in der Schweiz [Alemannic in Switzerland]. In Joachim Herrgen & Jürgen Erich Schmidt (eds.), *Deutsch*, vol. 4, 246–279. Berlin & Boston: De Gruyter Mouton.

Dieth, Eugen. 1950. *Vademekum der Phonetik. Phonetische Grundlagen für das wissenschaftliche und praktische Studium der Sprachen*. Unter Mitwirkung von Rudolf Brunner [vademeccum of phonetic and phonological basics for the scientific and practical study of languages, in collaboration with Rudolf Brunner]. Bern: Francke.

Dieth, Eugen. 1986. *Schwyzertütschi Dialäktschrift* [Swiss German dialect spelling system], 2nd edn. Aarau: Sauerländer.

Draxler, Christoph & Klaus Jänsch. 2004. SpeechRecorder-a universal platform independent multi-channel audio recording Software. In *Proceedings of the IV. International conference on language resources and evaluation*, 559–562. Lisbon, Portugal: Citeseer.

Enstrom, Daly & Sonja Spörri-Büttler. 1981. A voice onset time analysis of initial Swiss German stops. *Folia Phoniatica* 33. 137–150.

Ferguson, Charles A. 1959. Diglossia. *Word* 15(2). 325–340.

Fleischer, Jürg & Stephan Schmid. 2006. Zurich German. *Journal of the International Phonetic Association* 36(2). 243–253.

Fox, John & Sanford Weisberg. 2018. *An R companion to applied regression*, 2nd edn. Thousand Oaks: Sage.

Fulop, Sean A. 1994. Acoustic correlates of the fortis/lenis contrast in Swiss German plosives. *Calgary Working Papers in Linguistics* 16. 55–63.

- Ham, William Hallett. 2001. *Phonetic and phonological aspects of geminate timing*, vol. 15. New York: Routledge.
- Hanson, Helen M. 2009. Effects of obstruent consonants on fundamental frequency at vowel onset in English. *The Journal of the Acoustical Society of America* 125(1). 425–441.
- House, Arthur S. & Grant Fairbanks. 1953. The influence of consonant environment upon the secondary acoustic characteristics of vowels. *The Journal of the Acoustical Society of America* 25. 105–113.
- Hove, Ingrid. 2002. *Die Aussprache der Standardsprache in der deutschen Schweiz* [The pronunciation of standard speech in German-speaking Switzerland]. Berlin & Boston: Max Niemeyer.
- Jacewicz, Ewa, Robert A. Fox, Caitlin O'Neill & Joseph Salmons. 2009. Articulation rate across dialect, age, and gender. *Language Variation and Change* 21(2). 233–256.
- Jessen, Michael. 1998. *Phonetics and phonology of tense and lax obstruents in German*. Amsterdam: John Benjamins Publishing.
- Kingston, John & Randy L. Diehl. 1994. Phonetic knowledge. *Language* 70. 419–454.
- Kirby, James & D. Robert Ladd. 2015. Stop voicing and F0 perturbations: Evidence from French and Italian. In *Proceedings of the 18th international congress of phonetic sciences*. Glasgow, Scotland.
- Kirby, James & D. Robert Ladd. 2016. Effects of obstruent voicing on vowel F0: Evidence from 'true voicing' languages. *The Journal of the Acoustical Society of America* 140(4). 2400–2411.
- Kirby, James, Felicitas Kleber, Jessica Siddins & Jonathan Harrington. 2020. Effects of prosodic prominence on obstruent-intrinsic F0 and VOT in German. In *Proceedings of the 10th international conference on speech prosody 202*, 210–214. Tokyo, Japan.
- Klatt, Dennis H. 1973. Interaction between two factors that influence vowel duration. *The Journal of the Acoustical Society of America* 54(4). 1102–1104.
- Kleber, Felicitas. 2017. Complementary length in vowel–consonant sequences: Acoustic and perceptual evidence for a sound change in progress in Bavarian German. *Journal of the International Phonetic Association* 50(1). 1–22.
- Kleber, Felicitas. 2018. VOT or quantity: What matters more for the voicing contrast in German regional varieties? Results from apparent-time analyses. *Journal of Phonetics* 71. 468–486.
- Kohler, Klaus. 1979. Dimensions in the perception of fortis and lenis plosives. *Phonetica* 36. 332–343.
- Kohler, Klaus. 1985. F0 in the perception of lenis and fortis plosives. *The Journal of the Acoustical Society of America* 78. 21–32.
- Kolly, Marie-José & Adrian Leemann. 2013. *Dialäkt Äpp: Dialektologie Vermitteln – Dialekte Ermitteln* [Dialect app: Convey dialectology – establish dialects]. Zürich, Switzerland.
- Kraehenmann, Astrid. 2001. Swiss German stops: Geminate all over the word. *Phonology* 18(1). 109–145.
- Kraehenmann, Astrid. 2003. *Quantity and prosodic asymmetries in Alemannic: Synchronic and diachronic perspectives*. Berlin & New York: Mouton de Gruyter.
- Kraehenmann, Astrid & Aditi Lahiri. 2008. Duration differences in the articulation and acoustics of Swiss German word-initial geminate and Singleton stops. *The Journal of the Acoustical Society of America* 123(6). 4446–4455.
- Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen. 2017. LmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13). 1–26.
- Labov, William. 1994. *Principles of language change: Internal factors*. Oxford: Blackwell.
- Ladd, D. Robert & Stephan Schmid. 2018. Obstruent voicing effects on F0, but without voicing: Phonetic correlates of Swiss German lenis, fortis, and aspirated stops. *Journal of Phonetics* 71. 229–248.
- Laver, John. 1996. *Principles of phonetics*. Cambridge: Cambridge University Press.
- Leemann, Adrian. 2016. Analyzing geospatial variation in articulation rate using crowdsourced speech data. *Journal of Linguistic Geography* 4(2). 76–96.

- Leemann, Adrian & Beat Siebenhaar. 2010. Statistical modeling of F0 and timing of Swiss German dialects. In *Proceedings of the speech prosody 2010-fifth international conference*. Chicago, USA.
- Leemann, Adrian, Volker Dellwo, Marie-José Kolly & Stephan Schmid. 2012. Rhythmic variability in Swiss German dialects. In *Proceedings of the speech prosody 2012*, 607–661. Shanghai, China.
- Lehiste, Ilse & Gordon E. Peterson. 1961. Some basic considerations in the analysis of intonation. *The Journal of the Acoustical Society of America* 33. 419–425.
- Lenth, Russel V. 2021. *Emmeans: Estimated marginal means, aka least-squares means*. Available at: <https://CRAN.R-project.org/package=emmeans> (accessed 20 October 2021).
- Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press.
- Moosmüller, Sylvia. 2007. On some timing aspects of the Viennese dialect. *The Phonetician* 95. 19–27.
- Moosmüller, Sylvia & Julia Brandstätter. 2014. Phonotactic information in the temporal organization of Standard Austrian German and the Viennese dialect. *Theoretical and Empirical Approaches to Phonotactics and Morphonotactics* 46. 84–95.
- Moulton, William G. 1986. Sandhi in Swiss German dialects. In Henning Andersen (ed.), *Sandhi phenomena in the languages of Europe*, 385–392. Berlin, New York: De Gruyter Mouton.
- Nocchi, Nadia & Stephan Schmid. 2006. Labiodentale Konsonanten Im Schweizerdeutschen [Labiodental consonants in Swiss German]. In *Raumstrukturen im Alemannischen*, 25–35. Feldkirch (Neugebauer): Klausmann H. (edn).
- Pellegrino, Elisa, Lei He & Volker Dellwo. 2021. Age-related rhythmic variations: The role of syllable intensity variability. *TRANEL-Travaux Neuchâtelois de Linguistique* 74. 167–185.
- Quené, Hugo. 2007. On the just noticeable difference for tempo in speech. *Journal of Phonetics* 35(3). 353–362.
- Quené, Hugo. 2008. Multilevel modeling of between-speaker and within-speaker variation in spontaneous speech tempo. *The Journal of the Acoustical Society of America* 123(2). 1104–1113.
- RStudio Team. 2020. *RStudio: Integrated development for R*. Boston: PBC.
- Schiel, Florian. 1999. Automatic phonetic transcription of non-prompted speech. In *International Congress of Phonetic Sciences*, 607–661. San Francisco, USA.
- Schifferle, Hans-Peter. 2010. Zunehmende Behauchung. Aspirierte Plosive im modernen Schweizerdeutsch [Increasing aspiration. Aspirated plosives in modern Swiss German]. In *Alemannische Dialektologie: Wege in Die Zukunft*, 43–55. Stuttgart: Steiner.
- Schmid, Stephan. 2004. Zur Vokalquantität in der Mundart der Stadt Zürich [Vowel quantity in the dialect of the city Zurich]. *Linguistik Online* 2. 93–116.
- Schmid, Stephan. 2019. Wie viele Typen von Plosiven gibt es in Schweizerdeutschen Dialekten [How many types of plosives are there in Swiss German dialects]. Presented at Workshop in Gedenken an Sylvia Moosmüller, April 8, Vienna, Austria.
- Schwab, Sandra & Mathieu Avanzi. 2015. Regional variation and articulation rate in French. *The Impact of Stylistic Diversity on Phonetic and Phonological Evidence and Modeling* 48. 96–105.
- Seiler, Guido. 2005. On the development of the Bavarian quantity system. *Interdisciplinary Journal for Germanic Linguistics and Semiotic Analysis* 10(1). 103–129.
- Siebenhaar, Beat. 2006. Code choice and code-switching in Swiss-German internet relay chat rooms. *Journal of Sociolinguistics* 10(4). 481–506.
- Willi, Urs. 1996. *Die segmentale Dauer als phonetischer Parameter von 'fortis' und 'lenis' bei Plosiven im Zürichdeutschen: Eine akustische und perzeptorische Untersuchung* [The segmental duration as phonetic parameter of 'fortis' and 'lenis' in plosives in Zurich German]. Stuttgart: Franz Steiner Verlag.
- Winkelmann, Raphael & Georg Raess. 2014. Introducing a web application for labeling, visualizing speech and correcting derived speech signals. In *Proceedings of the 9th International conference on language resources and evaluation*, 4129–4133. Reykjavic, Iceland.

- Winkelmann, Raphael, Klaus Jaensch, Steve Cassidy & Jonathan Harrington. 2021. *EmuR: Main package of the EMU speech database management system*.
- Würth, Kathrin. 2020. *Consonant quantity and positional Neutralisation–Heusler’s Law and Winteler’s Law in Zurich German*. Zurich: University of Zurich dissertation.
- Zebe, Franka. 2022. Durational consonant categories in Alemannic and Swiss Standard German across tempo and age. In *Proceedings of the 11th international conference on speech prosody 2022*. Lisbon, Portugal.
- Zihlmann, Urban. 2020a. Temporal variability in four Alemannic dialects and its influence on the respective varieties of Swiss Standard German. In *Proceedings of the 10th international conference on speech prosody*, 620–624. Tokyo, Japan.
- Zihlmann, Urban. 2020b. Vowel and consonant length in four Alemannic dialects and their influence on the respective varieties of Swiss Standard German. *Wiener Linguistische Gazette* 86. 1–46.
- Zihlmann, Urban. 2021a. Investigating speaker individuality in the Swiss Standard German of four Alemannic dialect regions: Consonant quantity, vowel quality, and temporal variables. *Loquens* 7(1). 1–13.
- Zihlmann, Urban. 2021b. Vowel quality in four Alemannic dialects and its influence on the respective varieties of Swiss Standard German. *Journal of the International Phonetic Association* 53. 1–28.