Speech Prosody 2020
25-28 May 2020, Tokyo, Japan

# Temporal variability in four Alemannic dialects and its influence on the respective varieties of Swiss Standard German

*Urban B. Zihlmann*

Phonetics Laboratory, University of Zurich, Switzerland
urban.zihlmann@uzh.ch

## Abstract

Languages, dialects, and speakers have been found to show variability in articulation rate and speech rhythm. The present study examines whether temporal variability found between four Alemannic (ALM) dialects is also present in the same speakers' Swiss Standard German (SSG) varieties. The results suggest that ALM interferences are not systematic. Whereas in ALM dialects, Chur speakers show the quickest articulation rate and Brig speakers the slowest, amongst the SSG varieties, Zurich speakers articulated the most quickly and Bern the most slowly. As for further rhythm metrics, the insights of this study provide evidence that consonantal variability is more telling regarding dialect/variety identification than vocalic measures.

**Index Terms**: Temporal variability, Alemannic, Swiss Standard German, dialectology

## 1. Introduction

In German-speaking Switzerland (CH), two varieties are used, one for spoken, and one for written communication, thus creating a "diglossic" situation [1]. Orally, Alemannic (ALM) dialects are used except in Samnaun in the canton of Grisons, where a Bavarian dialect is spoken [2]. The official written language is Swiss Standard German (SSG). However, with the advent of messenger apps that resemble oral communication, ALM has become more popular in informal written communication [3]. Meanwhile, SSG is also spoken, mostly in educational settings, military, clerical, and official political speeches as well as in many broadcasting contexts (e.g., news); moreover, the standard language is used in interactions with German speakers who do not speak ALM, with foreigners and in communication with deaf people for them to be able to lip-read [4]. Native German-speaking Swiss are able to switch instantly, implying that the linguistic situation in CH is not a standard-dialect continuum [5]. Consequently, contact between the two varieties is unavoidable, and thus they influence one another on multiple levels [6]. This study analyses four ALM dialects and their respective SSG varieties in the timing domain.

### 1.1. Alemannic dialects in Switzerland

ALM dialects in CH belong to three subvariants: Low ALM, limited to the city of Basel, High ALM, spoken to the north of the Alps, and Highest ALM, spoken in the Alpine region [7]. Given that High and Highest ALM are the most frequently spoken varieties, this article will focus on them.

By overlaying phonetic, morphosyntactic, and lexical isoglosses, ALM dialects can dialectologically be divided into northern/southern and eastern/western halves [8]. E.g., while northern dialects display diphthongs in hiatus positions, like in [ˈɒb̥əˌg̊hei̯ə] 'to fall down', southern dialects have a monophthong and produce the same word like [ˈɑb̥əˌg̊hiːə]. The arguably best-known east-west distinction is the realisation of /e/ in the word *Bett* 'bed'. In the west, it is pronounced [ˈb̥ɛt], and in the east [ˈb̥et]. If the east-west and north-south differences are averaged, four quadrants emerge. For this study, a dialect of a densely populated city from each quadrant was analysed, i.e. Bern (BE) in the northwest, Zurich (ZH) in the northeast, Brig (VS; as the city is located in the canton of Valais) in the southwest, and Chur (GR; as the city is located in the canton of Grisons) in the southeast.

### 1.2. Articulation rate

Several studies have found evidence for ALM dialects to yield different articulation rates (AR), i.e. the number of syllables uttered in a given time frame without pauses [9]. Specifically, the results of these studies suggest that BE speakers enunciate more slowly (5.0 syllables/second = syl/s) than speakers from ZH and VS (5.8 syl/s), with speakers from GR lying somewhere in between [10, 11, 12]. [11] explain their results with BE speakers engaging in phrase-final lengthening, as vowels in prepausal position tend to be extended way more than in other dialects, leading to a lower average AR.

However, AR is not as stable as it might appear. Multiple factors have been proven to influence it, amongst which are mood, speaking style, sex, utterance length, and age [12]. As for sex, women have been found to articulate more slowly, and thus, more carefully [13] in many Western languages, including British [14] and American English [15], Swedish [16], Dutch [17], German Standard German (GSG) [18], and even ALM [12, 19]. It has been linked to women being more concerned about social status, which they linguistically convey by employing fewer non-standard and more prestigious features [20]. Regarding utterance length, [21] and [22] have found Dutch speakers to engage in "anticipatory shortening", meaning that the longer the utterance, the shorter the average syllable duration tends to be. Regarding age, [12] has found a slight effect in that older speakers articulated marginally more quickly even though this is diametrically opposed to what earlier studies had found [22, 23] and might be due to an artefact of the data-gathering process as the dataset was heavily skewed towards young people [12]. However, as the present study does not include two age groups, no conclusions about age can be drawn. Moreover, AR differences have also been found within varieties of a given language (see [24] for New Zealand and American English; [25] for Dutch and Flemish; [26] for French, Belgian and Swiss French. However, many studies report rather strong intra-speaker variability, implying that findings on AR must be taken with a grain of salt.

Regarding SSG, only a few studies on AR have been conducted so far. [27] states without providing empirical evidence that SSG is slower than GSG, which he attributes to the lack of training. To collect empirical data, [28] thus measured the AR

 10.21437/SpeechProsody.2020-127

of SSG and compared it to GSG and Austrian Standard German (ASG). She found that SSG shows indeed the lowest AR (5.29 syl/s), followed by ASG (5.39 syl/s), and GSG (5.52 syl/s). While the SSG-GSG and ASG-GSG differences were statistically significant, the SSG-ASG ones were not.

In conclusion, while AR is far from being entirely predictable due to a myriad of correlating factors, certain tendencies for languages as well as for varieties of languages can be identified. Amongst ALM, a tendency for BE to articulate slowly seems to exist. The same can also be said about SSG when it is compared to other standard German varieties.

### 1.3. Rhythmic variables

Several studies have looked at the intonational and segmental characteristics of ALM dialects (cf. [11, 29, 30]. The results suggest that region-specific segmental variation exists in terms of segmental temporal features in that western dialects (primarily BE and VS) typically have shorter phrases than eastern ones (ZH and GR), and each dialect has its own lengthening pattern, be it phrase-initially, phrase-finally, or in the penultimate syllable. BE, e.g., shows a slow AR with phrase-final lengthening, whereas ZH shows more phrase-initial lengthening on top of phrase-final lengthening. Furthermore, VS was found to produce the most regular syllable pattern. Continuing the research on intonation, [31, 32] have analysed rhythmic variability in ALM dialects with regards to the east-west and the north-south divide. They focussed on eight cities, two in each quadrant (amongst which the four cities of this study), and found that vocalic-interval variability is an important feature to discriminate ALM dialects (V = vowel, C = consonant), especially *varcoV*, but also *ΔV* and *nPVI-V* (for an explanation of all abbreviations see Table 1). Northern dialects show less vocalic variability, which [31, 32] explain with Alpine dialects tending to use non-reduced vowels word-finally. Moreover, BE was found to have a statistically significantly lower *%V* value. Furthermore, [31] report that rhythmic differences are noticeable especially between the Alpine and Midland regions, i.e. the north-south divide. Lastly, they report that consonantal variability is less important to distinguish ALM dialects as only *rPVI-C* and *nPVI-C* showed statistically significant interdialectal differences. In a later study, the same group of researchers assessed whether the sentence material used influences the outcome of rhythm studies, which they confirmed [32]. This means that to compare rhythmic qualities of dialects, similar sentences with a comparable syllable structure must be used.

But insights regarding rhythmic discrimination of dialects in languages other than ALM can also be taken into consideration. [33] have found that *%V* and *VarcoV* are the most adequate variables to hold British dialects apart based on rhythm. Furthermore, [34] demonstrated that *nPVI-V* is a suitable measurement to highlight interdialectal rhythmic differences in British dialects. They concluded that rhythmic measurements for vowels are more telling than consonantal ones. [35] analysed nine Italian dialects and argues that *ΔC* is a valid measurement to assess syllable complexity, and *nPVI-V* reflects vowel reduction quite accurately.

### 1.4. Research questions

Two questions guided the current study. (1) Do ALM dialects differ in their temporal properties? And (2) do the dialect-specific SSG varieties differ in their rhythmic properties? Based on the literature review, I hypothesise that (1) BE shows the slowest AR, while ZH and VS will show the fastest AR. Furthermore, I hypothesise that (2) BE will also show the slowest AR in SSG compared to the other dialects' SSG varieties.

## 2. Data and method

### 2.1. Speakers, wordlists, and data collection

32 speakers (50% male/female; aged 17–32; mean = 22.5; standard deviation (SD) = 3.42) were recruited, 8 each from BE, GR, VS, and ZH. At least one of each speaker's parents had to have grown up with the same dialect. Also, to reduce possible dialect-contact interference, the speakers had to be residing in the city they had been raised in, and they were not allowed to have spent more than one year outside said city. Should one parent have grown up in another German-speaking country, the participants were excluded. The speakers received a reimbursement of CHF 15.– per 30 min of testing.

For the recordings in each ALM dialect and SSG, disyllabic words were chosen whose first nucleus is either /a i u/ and whose second onset is either /p b̥ t d̥ k ɡ̊ l n s z̥/ in the four possible vowel-consonant sequences VC, VCː, VːC, and VːCː. This resulted in a corpus of approximately 12,000 words. ALM words were represented in [36]'s spelling system. The words were originally designed for analysing VC quantity and reused for the purpose of this study.

To familiarise the participants with the recording situation, first an interview was conducted eliciting biographical and sociolinguistic information about the speakers. Subsequently, the recording of the target words took place, consisting of six blocks that alternated three times between randomised ALM and SSG word lists. To prevent list-reading effects, generic carrier phrases were used. ALM carrier phrases has 6 syllables, with 8 (for BE), 9 (for GR and VS), and 10 segments (for ZH). The SSG carrier phrase had 12 segments and 7 syllables. For the recordings *SpeechRecorder* (version 3.28.0) [37] was used. Whenever possible, the participants were recorded in a sound-attenuated booth at the University of Zurich with the interface *USBPre® 2* by *Sound Devices* and the microphone *NT2-A* by *RØDE* (at 16-bit/44.1 kHz in mono, stored as .WAV). If it was not possible to record in Zurich, it was done in the field with the same interface model and the microphone *Opus 54.16/3* by *BeyerDynamic* (at 16-bit/44.1 kHz in mono, stored as .WAV).

The audio files were automatically segmented with an *R* script (courtesy of Markus Jochim MA, LMU Munich) using the *Munich AUtomatic Segmentation (MAUS) System* [38, 39] with the language setting *General Swiss German* for ALM, and *Standard German* for SSG. The processed sentences were then uploaded to the *EMU Speech Database Management System* [40], where each sentence was corrected by hand by the author (78%) and, for reasons of time, by two additional researchers (22%). For the entire corpus, corrections were made to the boundaries at the beginning and at the end of the utterance to measure utterance length, and to the complete target word. Additionally, to assess further rhythmic variables, a subset containing all sentences with the bilabial plosives of all eight varieties was created, where all boundaries (including the carrier phrases) were manually corrected with *Praat* [41].

### 2.2. Analyses and statistics

For the entire dataset, AR was calculated in syl/s. Furthermore, the rhythmic variables shown (and described) in Table 1 were measured in the subset. C intervals (or V intervals, respectively) refer to the duration of 1≤ consecutive C (or V) between two V (or C). Intervals may stretch over syllable, word, or sentence boundaries. Generally, high values reflect high variability in all measurements except *%V* and *rateSeg* [42].

Table 1: *Description of temporal variables measured.*

| Variable | Description |
|---|---|
| *rateSeg* | Rate of segments without pauses |
| *meanV/C* | Mean durations of V/C intervals |
| *%V* | Share of vocalic speech in % [43] |
| *ΔV/C* | SD of V/C intervals [43] |
| *varcoV/C* | Rate-normalised SD of V/C intervals [35, 44] |
| *rPVI-V/C* | Mean differences between consecutive V/C intervals [45] |
| *nPVI-V/C* | Rate-normalised mean differences between consecutive V/C intervals: V/C reduction [45] |

Utterance duration was measured in *R* [46] with a script courtesy of Markus Jochim, LMU Munich. The rhythmic variables were measured in *Praat* [43] with two scripts: the first script, *CVtier Creator* by Volker Dellwo and Sandra Schwab, indicated which of the segments were vocalic, and which were consonantal. Subsequently, the resulting annotation was analysed with the second script, *Duration Analyzer*, created by Volker Dellwo. The statistical analysis was also done in *R* [46] with help of Sandra Schwab and included linear mixed-effects models (LMM) with the packages *lme4* [47] with *dialect*, *variety*, or *sex* as fixed factors, and either one of the rhythmic variables as dependent variable (DepVar). *Speaker* and *target word* were included as random factors. Visual control of the residual plots did not result in any noticeable deviations from homoscedasticity or normality. The Tukey method was applied for the post-hoc pairwise comparisons conducted with *lsmeans* [48].

# 3. Results

Table 2 shows mean ARs and utterance duration (based on about 12,000 sentences). The ARs of either variety are rather similar. In fact, the fastest and slowest mean ALM values are merely 0.19 syl/s apart. The SSG are similar: the fastest and slowest mean values are separated by 0.27 syl/s. For either variety, BE shows the biggest and GR the smallest SD.

Table 2: *Mean AR (in syl/s), SD (in parentheses), and utterance duration (in ms) by dialect and variety.*

| Variety | BE | GR | VS | ZH |
|---|---|---|---|---|
| ALM | 4.71 (.75)<br>1274 ms | 4.75 (.40)<br>1262 ms | 4.56 (.53)<br>1317 ms | 4.66 (.59)<br>1287 ms |
| SSG | 5.06 (.76)<br>1384 ms | 5.16 (.38)<br>1363 ms | 5.09 (.43)<br>1375 ms | 5.33 (.38)<br>1313 ms |

All LLMs were run for ALM and SSG independently. LMMs with *AR* as dependent variable, *dialect* as fixed factor, and *speaker* and *target word* as random factors do not result in any statistically significant differences in ALM and SSG. Neither do LMMs with the same design but *dialect* and *sex* as fixed factors (with interaction term).

Table 3 shows the statistically significant results for the ALM temporal variables. For the sake of readability, values >1.0 are reported with one decimal place, values <1.0 with two. Only BE and ZH differ statistically significantly from one another in *rateSeg* (*p*=.014) and *meanV* (*p*=.017). As for *meanC*, statistically significant differences were found between eastern and western dialects (each *p*<.001). In the remaining consonantal variables, VS differs from all other dialects for *varcoC* (each *p*<.001), *ΔC* (BE: *p*<.001, GR/ZH: each *p*=.007), and *nPVI-C* (BE: *p*=.025, GR: *p*=.019, ZH: *p*<.001). Regarding the vocalic parameters, GR differs statistically significantly from VS (*p*=.003) and ZH (*p*=.005) for *%V*, from BE (*p*=.016) and VS (*p*=.011) for *rPVI-V*, and from all dialects (each *p*<.001) for *nPVI-V*. Furthermore, ZH also frequently displays statistically significant differences with other dialects, especially for *varcoV* (BE: *p*=.015, GR: *p*=.002, VS: *p*=.073), even though VS-ZH lies between *p*=.050–.100.

Table 3: *Results of the temporal variables in ALM.*

| DepVar | BE | GR | VS | ZH | *F*-/*p*-value |
|---|---|---|---|---|---|
| *rateSeg* | 10.9 | 11.9 | 11.4 | 12.6 | $F_{3,32.0}$=4.2, *p*=.012 |
| *meanV* | 0.12 | 0.10 | 0.11 | 0.09 | $F_{3,31.7}$=4.5, *p*=.010 |
| *meanC* | 0.15 | 0.12 | 0.15 | 0.12 | $F_{3,31.4}$=4.2, *p*<.001 |
| *%V* | 43.7 | 46.5 | 40.8 | 41.0 | $F_{3,31.3}$=7.0, *p*<.001 |
| *ΔC* | 0.07 | 0.06 | 0.04 | 0.06 | $F_{3,31.6}$=12.7, *p*<.001 |
| *varcoV* | 0.53 | 0.51 | 0.55 | 0.63 | $F_{3,31.5}$=6.7, *p*=.001 |
| *varcoC* | 0.47 | 0.51 | 0.29 | 0.48 | $F_{3,31.2}$=35.8, *p*<.001 |
| *rPVI-V* | 84.5 | 56.6 | 85.6 | 73.6 | $F_{3,31.7}$=5.3, *p*=.005 |
| *nPVI-V* | 73.8 | 56.3 | 74.6 | 75.7 | $F_{3,31.6}$=12.7, *p*<.001 |
| *nPVI-C* | 47.1 | 47.5 | 36.4 | 52.5 | $F_{3,31.1}$=8.0, *p*<.001 |

Table 4 shows the statistically significantly different results for the SSG temporal variables.

Table 4: *Results of the temporal variables in SSG.*

| DepVar | BE | GR | VS | ZH | *F*-/*p*-value |
|---|---|---|---|---|---|
| *meanC* | 0.14 | 0.13 | 0.13 | 0.12 | $F_{3,31.3}$=4.1, *p*=.015 |
| *ΔC* | 0.07 | 0.06 | 0.06 | 0.05 | $F_{3,31.6}$=7.0, *p*=.001 |
| *varcoC* | 0.53 | 0.47 | 0.46 | 0.43 | $F_{3,31.7}$=4.8, *p*=.007 |
| *rPVI-C* | 88.7 | 72.4 | 75.4 | 59.7 | $F_{3,31.5}$=5.8, *p*=.003 |
| *nPVI-C* | 61.2 | 53.7 | 56.5 | 48.9 | $F_{3,31.9}$=4.2, *p*=.013 |

Statistically significant differences were only found for C parameters. Clear distinctions exist mainly between BE and ZH, namely for *meanC* (*p*=.012), *ΔC* (*p*<.001), *varcoC* (*p*=.008), *rPVI-C* (*p*=.002), and *nPVI-C* (*p*=.013). Differences between *p*=.050–.100 have been found for BE-VS for *varcoC* (*p*=.088) and *ΔC* (*p*=.061), and for BE-GR for *ΔC* (*p*=.067).

# 4. Discussion

Some methodological remarks for the discussion must be made. Firstly, it has been shown that sentence material does have an influence on the outcome [32], so precautions for comparability were made. In the rather extensive corpus of this study, ALM carrier phrases consisted of dialect-specific allomorphs, whereas the SSG ones were identical. Nevertheless, it cannot be ruled out that this study's insights would have been different if other sentences had been used. Secondly, all subjects were young and educated; it might be the case that older speakers or individuals with a lower level of education behave differently, especially when reading SSG sentences. Thirdly, despite briefing the participants to read as naturally as possible, it is still read speech, which might differ from actual spontaneous speech.

## 4.1. Articulation and segment rate

Contrary to what [19] have found for ALM (and [15–18] for other Western languages), the results that entail more dialects and are based on more and longer utterances than [19]'s data do not show any sex differences. Considering the results obtained in the context of this study, [20]'s claim that women are more worried about social status and thus speak more carefully cannot be supported, at least for read speech.

A rather surprising finding is that BE does not articulate the most slowly due to phrase-final-lengthening phenomena. Rather, VS shows the lowest AR in ALM. Furthermore, ZH is not the fastest dialect but comes third behind BE, which is not congruent with [11–13, 19]. GR, which typically shows more intermediate values, has indeed the highest AR. Reasons for this cannot be found on the syllable level, as all sentences consisted of the same number of syllables. Rather, it must be attributable to interspeaker variation as evident by the large SD. As a matter of fact, already [19] mention that to generalise, the insights of studies on speaking rate conducted so far must be verified by analysing an extensive corpus consisting of both controlled sentences as well as spontaneous speech.

As regards ALM *rateSeg*, a difference can be identified between ZH and BE in that ZH shows the quickest and BE the slowest one, which is indeed compatible with [11]. In fact, BE-ZH are the only ones to show statistically significant differences in *rateSeg*, thus within the data set of this study, only BE and ZH could be told apart by means of analysing *rateSeg*. Therefore, an east-west divide regarding *rateSeg* as reported by [11, 29, 30] cannot be confirmed with the results of this study.

The SSG AR results, in contrast, are compatible with the literature. BE is the slowest variety, followed by VS and GR, while ZH speakers articulate the most quickly. The SSG results might be more telling about AR than the ALM ones as all speakers produced identical sentences. However, as it has been the case for ALM, BE's SD is approximately double the amount of GR, VS, and ZH, which again suggests that much interspeaker variation is present amongst BE speakers. Nevertheless, it appears that the average SSG ARs correspond to the dialect-specific stereotypes.

### 4.2. Rhythmic variables

For ALM, two major findings can be provided by this study. First, the dialect that shows the most different consonantal rhythm metrics within the dialects is VS. Evidence for this claim can be found when looking at *varcoC*, *ΔC*, and *nPVI-C*. Only VS differs statistically significantly from the other dialects. The strongest differences can be identified in *varcoC*, meaning that the consonantal segments of VS show the greatest duration variability even after normalisation. Furthermore, regarding *meanC*, an east-west divide is reinforced in that western dialects show longer mean consonantal durations than eastern ones. The rounded values for *meanC* even are identical for the two eastern and the two western dialects, respectively.

Second, GR's vowels show the most divergent vocalic rhythmic measurements. First and foremost, this applies to *nPVI-V*, where GR behaves statistically significantly differently than the other three dialects. When the non-normalised values are taken into consideration, i.e. *rPVI-V*, no statistical significance between GR and ZH exists, however. In terms of *%V*, it appears that GR is especially vocalic compared to the other dialects, even though the GR-BE comparison was not statistically significant. As for *varcoV*, ZH contrasts with the other dialects in that its speakers show a statistically significantly higher normalised SD than BE and GR, while the difference with VS lies just outside the 95% confidence level. Nevertheless, a tendency towards a bigger SD compared to VS exists. Regarding *meanV*, only BE-ZH was found to differ statistically significantly, which is most likely due to *rateSeg* being higher for ZH; thus, their mean vowel duration is shorter accordingly.

For SSG, one of the major findings is that differences between the varieties manifest themselves primarily on the consonantal level. These differences, however, are mostly due to the BE-ZH disparities. One reason responsible for this can be found in BE having a slower AR. Consequently, the non-normalised values will inevitably be higher for BE. Exactly the opposite is true for ZH. As the ZH SSG variety is articulated the most quickly, non-normalised values will be smaller. None of the vocalic rhythmic variables showed statistically significant differences between SSG varieties. Thus, regarding vowels, a speaker's dialect does not interfere with SSG rhythm. Instead, the strongest measurement to distinguish SSG varieties from one another turned out to be *ΔC*, which again appears to be related to the slower AR.

### 4.3. Contrast between articulation rate and segment rate

An issue to be discussed is the counterintuitive result of BE having a comparatively quick AR, yet a slow segment rate. This can be explained by the fact that the BE carrier phrase consisted of the same number of syllables as the other dialects' ones (i.e. 6), yet of 8 segments rather than 9 (GR, VS) or 10 (ZH). This implies that while BE speakers generally spoke quickly, the fact that their syllable structure was less complex, they took longer on average to produce the segments.

## 5. Conclusions

This study has analysed temporal variability amongst four ALM dialects and their respective SSG varieties. Considering ALM, GR showed the fastest syllable rate, followed by BE and ZH, whereas VS showed the slowest one. This was surprising as BE has been reported to articulate slowly by several studies. When *rateSeg* was considered, however, BE did indeed show the slowest rate, suggesting that even though they might speak quickly, they are perceived as slow given the comparatively greater syllable duration. On the other hand, ZH showed the fastest *rateSeg*, while GR and VS were in between. Moreover, VS shows the most divergent C properties, while GR shows the most divergent V properties.

In SSG, speakers did behave as expected regarding AR in that BE articulated the most slowly, and ZH the fastest. These two SSG varieties also differed the most in the dataset. Furthermore, the results suggest that in order to distinguish regional varieties of SSG, consonantal variables are more valuable than vocalic variables. Future research, however, must shed more lights on the issue at hand. For this, more ALM dialects can be included, and further analyses can be done (e.g., Principal Component Analyses, Linear Discriminant Analyses, or Multiple Correspondence Analyses to see which parameters correlate most with the dependent variables). Lastly, the issue of interspeaker variation should also be explored.


## 6. Acknowledgements

This study was funded by the Swiss National Science Foundation (SNSF), grant Nr. 164377.


## 7. References


[1] C. Ferguson, "Diglossia," *Word*, vol. 15, pp. 325–340, 1959.

[2] W. Haas, "Die deutschsprachige Schweiz," in *Die viersprachige Schweiz*, H. Bickel and R. Schläpfer, Eds. Aarau: Sauerländer, 2000, pp. 57–138.

[3] C. Dürscheid, "Von der Mündlichkeit zur Schriftlichkeit? Normen und Variation in der Internetkommunikation," in *informationen zur deutschdidaktik (ide)*, M. Dannerer and U. Esterl, 2011, pp. 93–100.

[4] F. Rash, *The German Language in Switzerland – Multilingualism, Diglossia and Variation*. Bern: Peter Lang, 1998.

[5] I. Hove, "Zur Unterscheidung des Schweizerdeutschen und der (schweizerischen) Standardsprache," in *Sprechen, Schreiben, Hören – Zur Produktion und Perzeption von Dialekt und Standardsprache zu Beginn des 21. Jahrhunderts*, H. Christen and E. Ziegler, Eds. Wien: Praesens, 2008, pp. 63–82.

[6] I. Hove, *Die Aussprache der Standardsprache in der deutschen Schweiz*. Berlin, Boston: Max Niemeyer, 2002.

[7] H. Christen, E. Glaser, and M. Friedli, *Kleiner Sprachatlas der deutschen Schweiz*. Frauenfeld: Huber, 2013.

[8] R. Hotzenköcherle, *Die Sprachlandschaften der deutschen Schweiz.* Aarau: Sauerländer, 1984.

[9] T. Crystal and A. House, "Articulation rate and the duration of syllables and stress groups in connected speech," *Journal of the Acoustical Society of America*, vol. 88, pp. 101–112, 1990.

[10] A. Leemann and B. Siebenhaar, "Intonational and temporal features of Swiss German," *Proceedings of the ICPhS*, Saarbrücken, Germany, pp. 957–960, 2007.

[11] A. Leemann and B. Siebenhaar, "Statistical modeling of F0 and timing of Swiss German dialects," *Proceedings of Speech Prosody*, Chicago, USA, pp. 11–14, 2010.

[12] A. Leemann, "Analyzing dialectal variation in articulation rate using crowdsourced speech data," *Journal of Linguistic Geography*, vol. 2, no. 4, pp. 76–96, 2017.

[13] M. Meyerhoff, "Variation and Gender," in *Handbook of Language, Gender, and Sexuality*, S. Ehrlich, M. Meyerhoff and J. Holmes, Eds. Hoboken: John Wiley & Sons, 2003, pp. 98–118.

[14] S. P. Whiteside, "Temporal-based acoustic-phonetic patterns in read speech: some evidence for speaker gender differences," *Journal of the International Phonetic Association*, vol. 26, pp. 23–40, 1996.

[15] E. Jacewicz, R. A. Fox, C. O'Neill and J. Salmons "Articulation rate across dialect, age, and gender," *Language Variation and Change*, vol. 45, pp. 233–256, 2009.

[16] B. Nordberg, "En undersökning av språket i Eskilstuna," *Sprakvard*, vol. 3, pp. 7–15, 1971.

[17] D. Brouwer and R. van Hout, "Gender-related variation in Amsterdam vernacular," *International Journal of the Sociology of Language*, vol. 94, pp. 99–122, 1992.

[18] A. P. Simpson, "Phonetische Datenbanken des Deutschen in der empirischen Sprachforschung und der phonologischen Theoriebildung," *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel*, vol. 33, 1998.

[19] A. Leemann, M.-J. Kolly and V. Dellwo, "Crowdsourcing regional variation in speaking rate through the iOS app 'Dialäkt Äpp'," *Proceedings of Speech Prosody*, pp. 217–221, 2014.

[20] P. Trudgill, "Sex, covert prestige and linguistic change in the urban British English of Norwich," *Language in Society*, vol. 1, no. 2, pp. 179–195, 1972.

[21] S. Nooteboom, *Production and perception of vowel duration: A study of durational properties of vowels in Dutch*. PhD thesis, Rijksuniversiteit Utrecht, The Netherlands, 1972.

[22] H. Quené, "Multilevel modeling of between-speaker and within-speaker variation in spontaneous speech tempo," *Journal of the Acoustical Society of America*, vol. 123, no. 2, pp. 1104–1113, 2008.

[23] J. Yuan, M. Liberman and C. Cieri, "Towards an integrated understanding of speaking rate in conversation," *Proceedings of Interspeech*, pp. 541–544, 2006.

[24] M. P. Robb, M. A. Maclagan and Y. Chen, "Speaking rates of American and New Zealand varieties of English," *Clinical Linguistics & Phonetics*, vol. 18, no. 1, pp. 1–15, 2004.

[25] J. Verhoeven, G. de Pauw and H. Kloots, "Speech rate in a pluricentric language: A comparison between Dutch in Belgium and the Netherlands," *Language and Speech*, vol. 47, no. 3, pp. 297–308, 2004.

[26] S. Schwab and M. Avanzi, "Regional variation and articulation rate in French," *Journal of Phonetics*, vol. 48, pp. 96–105, 2015.

[27] K. Meyer, *Wie sagt man in der Schweiz? Wörterbuch der schweizerischen Besonderheiten. Duden Taschenbücher 22*. Mannheim/Wien/Zurich: Duden, 1989.

[28] C. Ulbrich, *Phonetische Untersuchungen zur Prosodie der Standardvarietäten des Deutschen in der Bundesrepublik Deutschland, in der Schweiz und in Österreich*. Frankfurt am Main: Peter Lang, 2005.

[29] A. Leemann and L. Zuberbühler, "Declarative sentence intonation patterns in 8 Swiss German dialects," *Proceedings of Interspeech*, pp. 1768–1771, 2010.

[30] A. Leemann, *Swiss German Intonation Patterns.* Amsterdam: John Benhamins, 2012.

[31] A. Leemann, V. Dellwo, M.-J. Kolly and S. Schmid, "Rhythmic variability in Swiss German dialects," *Proceedings of Speech Prosody*, pp. 607–610, 2012.

[32] A. Leemann, V. Dellwo, M.-J. Kolly and S. Schmid, "Disentangling sources of rhythmic variability between dialects," *Proceedings of Speech Prosody*, pp. 693–697, 2014.

[33] L. White and S. L. Mattys, "Calibrating rhythm: First language and second language studies," *Journal of Phonetics*, vol. 35, pp. 501–522, 2007.

[34] E. Ferragne and F. Pellegrino, "Rhythm in Read British English: Interdialect variability," *Proceedings of International Conference on spoken Language Processing*, pp. 1573–1576, 2004.

[35] S. Schmid, "Syllable typology and the rhythm class hypothesis: evidence from Italo-Romance dialects," in *Phonological Typology of Syllable and Word Languages in Theory and Practice*, J. Caro Reina and R. Szczepaniak, Eds. Berlin/New York: Mouton de Gruyter, 2011, pp. 421–454.

[36] E. Dieth, *Schwyzertütschi Dialäktschrift. Leitfaden einer einheitlichen Schreibweise für alle Dialekte. Nach den Beschlüssen der Schriftkommission der Neuen Helvetischen Gesellschaft (Gruppe Zürich)*. Zurich: Orell Füssli, 1938.

[37] C. Draxler and K. Jänsch, "SpeechRecorder - a Universal Platform Independent Multi-Channel Audio Recording Software," *Proceedings of LREC*, pp. 559–562, 2004.

[38] F. Schiel, "Automatic Phonetic Transcription of Non-Prompted Speech," *Proceedings of the ICPhS*, pp. 607–610, 1999.

[39] T. Kisler, U. D. Reichel and F. Schiel, "Multilingual processing of speech via web services," *Computer Speech & Language*, vol. 45, pp. 326–347, 2017.

[40] R. Winkelmann, J. Harrington and K. Jänsch, "EMU-SDMS: Advanced Speech Database Management and Analysis in R," *Computer Speech & Language*, vol. 45, pp. 392–410, 2017.

[41] P. Boersma and D. Weenink, "Praat: doing phonetics by computer," www.praat.org, version 6.1.03, retrieved 01.09.2019.

[42] V. Dellwo, "Choosing the right rate normalization method for measurements of speech rhythm," in *La dimensione temporale del parlato*, S. Schmid, M. Schwarzenbach and D. Studer-Joho, Eds. Torriana: EDK, 2009, pp. 13–32.

[43] F. Ramus, M. Nespor and J. Mehler, "Correlates of linguistic rhythm in speech signal," *Cognition*, vol. 73, pp. 268–292, 1999.

[44] V. Dellwo, "Rhythm and Speech Rate: A Variation Coefficient for deltaC," in *Language and language-processing*, P. Karnowski and P. Szigeti, Eds. Torriana: EDK, 2006, pp. 13–32.

[45] E. Grabe and E. L. Low, "Durational variability in speech and the Rhythm Class Hypothesis," in *Laboratory Phonology*, vol. 7, V. Gussenhoven and N. Warner, Eds. Berlin/New York: Mouton de Gruyter, 2002, pp. 515–545.

[46] R Core Team, "R: A language and environment for statistical computing. R Foundation for Statistical Computing," www.R-project.org, version 3.6.0, retrieved 30.04.2019.

[47] D. M. Bates, M. Mächler, B. Bolker and S. Walker, "Fitting Linear Mixed-Effects Models Using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.

[48] R. V. Lenth, "Least-Squares Means: The R Package lsmeans," *Journal of Statistical Software*, vol. 69, no. 1, pp. 1–33, 2016.