Loquens 7(1)
Julio 2020, e070
eISSN 2386-2637
https://doi.org/10.3989/loquens.2020.070

Investigating speaker individuality in the Swiss Standard German of four Alemannic dialect regions: Consonant quantity, vowel quality, and temporal variables

Urban Zihlmann

University of Zurich
urban.zihlmann@uzh.ch ORCID: https://orcid.org/0000-0001-5515-7388

Submitted: 23/07/2020; Accepted: 12/12/2020; Published on line: 07/06/2021

Citation / Cómo citar este artículo: Urban Zihlmann (2020). Investigating speaker individuality in the Swiss Standard German of four Alemannic dialect regions: Consonant quantity, vowel quality, and temporal variables. *Loquens, 7*(1), e070. https://doi.org/10.3989/loquens.2020.070

ABSTRACT: While German-speaking Switzerland manifests a considerable amount of dialectal diversity, until the present day the phonetic interrelation of Alemannic (ALM) dialects and spoken Swiss Standard German (SSG) has not been studied with an acoustic phonetic approach on the speaker level.

In this study, out of a pool of 32 speakers (controlled for sex, age, and education level) from 4 dialectologically distinct ALM areas, 16 speakers with 2 dialects were analysed regarding SSG consonant duration (in words whose ALM equivalents may or may not have a geminate), 8 speakers from the city of Bern (BE) were analysed for vowel quality, and 32 speakers were analysed for temporal variables, i.e., articulation rate (*AR*) and vocalic-speech percentage (*%V*).

Results reveal that there is much intradialectal inter- and intraspeaker variation in all three aspects scrutinised, but especially regarding vowel quality of BE SSG mid vowels and temporal variables. As for consonant quantity, while intradialectal interspeaker variation was observed, speakers showed a tendency towards normalised SSG consonant durations that resemble the normalised consonant durations in their ALM dialect. In general, these results suggest that a speaker's dialect background is only one factor amongst many that influence the way in which Swiss Standard German is spoken.

Keywords: speaker individuality; Alemannic dialects; Swiss Standard German; consonant quantity; vowel quality; temporal features.

RESUMEN: *Investigación de la individualidad del locutor en el alemán estándar de Suiza en cuatro regiones dialectales alemánicas: cantidad consonántica, cualidad vocálica y variables temporales.*– Aunque la Suiza de habla alemana cuente con una diversidad dialectal considerable, hasta hoy no se ha estudiado la interrelación fonética entre los dialectos alemánicos (ALM) y el alemán estándar suizo (SSG) con un enfoque acústico a nivel de locutor.

La muestra para este estudio se compone de 32 informantes (controlados por sexo y edad) procedentes de cuatro distintas regiones dialectales ALM. De 16 locutores de 2 dialectos se analiza la duración consonántica en palabras SSG cuyos equivalentes ALM pueden tener o faltar una geminada. De 8 locutores de la ciudad de Berna se analizan distintos timbres vocálicos. Además, para todos los 32 locutores se calculan dos variables temporales, o sea la velocidad de articulación (*AR*) y el portentaje vocálico del habla (*%V*).

Los resultados revelan que existe mucha variación inter- e intraindividual en todos los tres aspectos investigados, pero sobre todo en el timbre vocálico de las vocales medias en BE SSG y en las variables temporales. En relación a la cantidad consonántica, se ha observado cierta variación intradialectal entre varios locutores, pero al mismo tiempo muchos locutores muestran duraciones consonánticas normalizadas SSG que se parecen a las duraciones consonánticas normalizadas en su propio dialecto ALM. En general, estos resultados sugieren que el dialecto alemánico es solo uno entre varios factores que influyen en la pronunciación del alemán estándar suizo.

Palabras clave: individualidad del locutor; dialectos alemánicos; alemán estándar de Suiza; cantidad consonántica; timbre vocálico; variables temporales.

1. INTRODUCTION

Given that German-speaking Switzerland comprises a stable 'diglossia' (Ferguson, 1959) that consists of Alemannic (ALM) dialects typically used in oral communication, and Swiss Standard German (SSG) typically used for written purposes, it is inevitable that the two varieties influence one another on many levels, including the morphosyntactic, the pragmatic, the lexical, and the phonetic one (see Hove, 2002; Ammon et al., 2004; Hove, 2008; Christen et al., 2010; Guntern, 2011). By focusing on the phonetic level, this study explores speaker individuality in spoken SSG while keeping in mind the ALM dialect situation in Switzerland. The corpora used for the analysis consist of data collected for three previous studies that assessed dialect-specific ALM influences on SSG vowel and consonant quantity (Zihlmann, 2020a), SSG vowel quality (Zihlmann, 2021), and SSG temporal features (Zihlmann, 2020b). As in all studies intradialectal interspeaker variation had been observed, the current study will deal with speaker individuality in three features that showed statistically significant interdialectal differences.

The article is structured as follows. First, a dialectological description of German-speaking Switzerland is provided, followed by a summary of the three studies whose corpus is further examined as well as their relevant research background. Subsequently, the methodology and results of the speaker-specific analyses are presented and discussed.

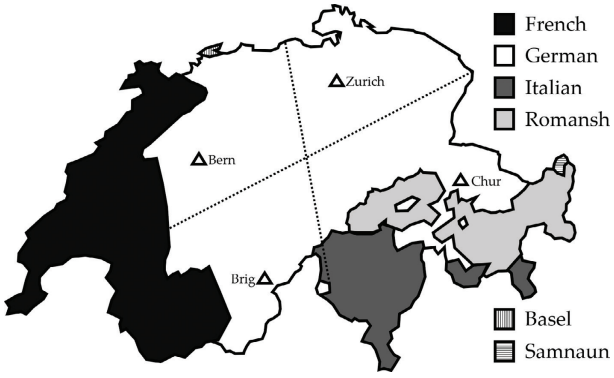
1.1. Alemannic dialects in Switzerland

According to Christen et al. (2013, pp. 28–30), four main dialect areas exist in German-speaking Switzerland, three of which belong to the ALM dialect family. In Basel, Low ALM is spoken, otherwise north of the Alps, High ALM is used. The dialects spoken in the Alpine southern region of German-speaking Switzerland belong to Highest ALM, except in Samnaun, where a Bavarian dialect is used (Haas, 2000, p. 71). As Low ALM and Bavarian make up for only a little part of the dialect diversity in German-speaking Switzerland, they are not part of the dialects scrutinised in this article.

The High and Highest ALM area has four identifiable broad regions (see Figure 1) that originate in the division into a northern and southern part and an eastern and western one (Haas, 2000, p. 67). These divisions are based on averaging different isoglosses that approximately run along the same geographic areas. Although much variation can still be observed within each quadrant, assuming these four regions can help order the vast amount of dialectal diversity. This study focusses on one representative of each of the four regions, i.e., the dialects of Bern (BE), Chur (GR),[1] Brig (VS),[2] and Zurich (ZH).

[1] Chur is located in the Canton of Grisons, whose official abbreviation is 'GR'.
[2] Brig is located in the Canton of Valais, whose official abbreviation is 'VS'.

Figure 1: Languages of Switzerland with an approximation of the four ALM dialect quadrants and the location of the four cities whose dialects were scrutinised.



1.2. ALM interferences in SSG

Before dialectal influences on SSG are discussed, it is important to note that no generally accepted SSG pronunciation norm exists as it is the case for Germany (e.g., Siebs, 1969; Duden, 2015). Rather, there is a lot of articulatory leeway when Swiss people speak SSG. As Hove (2002, p. 6) claims, it is a popular idea in German-speaking Switzerland that Swiss people try to speak German 'correctly' but fail to do so, which is why they employ ALM sounds. From the set of these ALM sounds, the general public seems to have an unwritten agreement over which ones are more and which ones are less acceptable (Guntern, 2012, p. 103), a situation which Hove calls 'language convention' (2002, p. 6). Obviously, it is difficult to predict how SSG is spoken exactly, as this 'language convention' does not categorically exclude the emergence of certain stigmatised dialectal features. There is nevertheless a greater likelihood for some variants to occur, i.e., those that are not stigmatised, and for others not to occur, i.e., those that are stigmatised. This SSG variety based on likelihood of occurrence within the framework of the 'SSG language convention' (Hove, 2002, p. 6) has been referred to as *typical SSG* (Zihlmann, 2020a, p. 8), a term that will also be used throughout this study.

1.2.1. Consonant quantity

Together with the ALM and SSG vowel-quantity systems that will not be discussed here, Zihlmann (2020a) analysed how region-specific ALM and SSG consonant-quantity systems are. Results suggested that while there might be lexical differences between ALM dialects regarding specific words which contain short/long consonants, the basic phonological quantity patterns are shared. Differently put, ZH might use C for the /l/ in *Pille* 'pill' ([ˈpilə]) and VS C: ([ˈpilːə]) but the way in which the two regions contrast short consonants with long consonants does, on average, not show interdialectal differences, neither in ALM nor in SSG. On the lexical level, however, differences between SSG varieties

were observed in that SSG words whose consonantal quantity differed from their ALM equivalents showed a tendency to be produced closer to the ALM quantity patterns, indicating that some speakers show ALM quantity interference in SSG.

This can be illustrated with the SSG word *Bullen* 'cops', which has a geminated /l/ in typical SSG (/ˈbʊlːən/) as well as in the BE and VS ALM equivalents (/ˈbʊlːə/) but a singleton in the GR and ZH ALM equivalents (/ˈbʊlə/). Here, Zihlmann (2020a, p. 28) has found that the average normalised SSG consonant duration of the two dialects whose ALM equivalent did not match SSG quantity tended to be shorter. However, while for the average normalised consonant durations for the /l/ in *Bullen* 'cops', no statistically significant differences were found between the four dialect regions, Zihlmann (2020a, p. 30) reports a statistically significant difference for the average BE and GR SSG Proportionate Vowel Duration[3] (PVD).[4] Nevertheless, Zihlmann (2020a) mentioned without going into further detail that it was very speaker-specific whether or not ALM interferences occurred, i.e., whether or not SSG PVD matched the expected PVD of the speaker's dialect more closely.

Given the statistically significant differences between the average PVD of BE and GR SSG speakers, this study will elaborate on the two dialect regions' speaker-individual quantity patterns of *Bullen* 'cops'.

1.2.2. Vowel quality

Region specificity of ALM and SSG vowel qualities as well as the way in which they interrelate were examined in Zihlmann's (2021) study. As for ALM, it was reported that GR high vowels tended to be pronounced less on the periphery compared to the other dialects. /ɛː œː ɔː/ showed interdialectal variation too in that BE speakers realised them non-stereotypically with qualities between /eː øː oː/ and /ɛː œː ɔː/ as already documented in the Linguistic Atlas of German-speaking Switzerland (SDS, 1962, maps 95, 99, 102). Especially for BE mid vowels, however, much interspeaker variation was observed. Regarding <a>-sounds, BE speakers realised them closest to [ɑː], GR speakers more in the front as [aː], and VS and ZH speakers produced them further back resembling [ɒː], as already found by Christen et al. (2010, pp. 167–168).

When the region-specific average SSG varieties were scrutinised, it was found that vowel quality was mostly adopted from the speakers' respective ALM dialect. Nevertheless, changes were observed for BE and ZH SSG in that either variety's mean <a>-sounds were realised statistically significantly more in the centre than in their respective ALM dialect. Regarding mid vowels, it was again BE SSG that differed from the other dialects' SSG varieties. Its realisation of /eː øː oː/ tended to be lower than elsewhere, while its /ɛː/, orthographically represented by <ä>, showed instances of being produced as either [eː], [ɛː], or [æː] with much interspeaker variation observed. As the results are based on mean values per dialect region, it is unclear whether Zihlmann's (2021) insights are region-specific or speaker-specific, however.

Therefore, to understand the variation better, this study will analyse SSG vowel quality in more detail. As the most amount of variation was reported for the BE SSG realisations of <ä>, the focus will lie on speaker individuality for BE SSG <ä>.

1.2.3. Temporal variability

Between-speaker temporal variability (i.e., speech rate and rhythm) has been studied by several researchers in multiple languages. Asadi et al. (2018, p. 163) identified the share of vocalic speech in percent (*%V*; coined by Ramus et al., 1999), and articulation rate *(AR)* to be the best indicators for interspeaker variation in Persian, whilst the rate-normalised average difference between consecutive vowel intervals (*n-PVI-V*; coined by Grabe & Low, 2002), turned out to be the least useful measurement. Similar findings were reported by Wiget et al. (2010), who found *%V* to be the parameter that yielded the most insightful interspeaker differences amongst English speakers. This was confirmed by Leemann et al. (2014b), who analysed ZH ALM speakers, and by Dellwo et al. (2015), who analysed standard German speakers from Germany and speakers of ZH ALM independently. However, neither one of those studies examined the same speakers in two varieties. This could prove to be insightful as studies have shown individual speech styles to be stable between a speaker's native and foreign language (see de Jong, 2018; de Jong & Mora, 2019).

Zihlmann (2020b) thus analysed region-specific rhythmic variability in ALM and SSG, verifying whether the temporal characteristics of ALM dialects stay stable when the same speakers switch to SSG. The analysis was based on a further examination of the stimuli used for Zihlmann's (2020a) study. Therefore, only isolated sentences could be used for the measurements, which could potentially affect the temporal variables. Nevertheless, it was found that *AR* (in syllables/second; syl/s) and segment rate *(SR)* (in segments/second; seg/s) show different mean results due to some ALM varieties, in Zihlmann's (2020b) case BE ALM, having less complex syllable structures. While in ALM, GR speakers had the fastest *AR*, the fastest *SR* was observed for ZH speakers. Similarly, while VS had the slowest *AR* in ALM, BE speakers had the slowest *SR*. Moreover, VS speakers were observed to have the most consonantal variability, and vowels were most variable in GR ALM. In SSG, BE speakers showed the slowest and ZH speakers the fastest *AR*. All in all, however, much interspeaker

[3] PVD (Kohler, 1979), also called 'VC ratio' (Kleber, 2017) or 'V: V+C ratio quotient' (Ham, 2001), equals V/(V+C) with V and C being vowel and consonant duration in milliseconds. Therefore, PVD is a measurement for showing how much percent of a VC sequence is vocalic.

[4] In this context, Zihlmann (2020a, p. 41) states that these inferences are not dialect-specific per se since they can occur in all ALM dialects with a quantity mismatch to SSG.

variation and little region specificity was found. When SSG varieties differed amongst the regions, these differences were mostly of consonantal nature.

To shed more light on regional SSG differences, this study will examine Zihlmann's (2020b) results on the speaker level. The focus will lie on speaker individuality in *AR* and *%V*, which turned out to be good indicators to for interspeaker variation in several languages, even though they did not show statistically significant interdialectal differences in Zihlmann's (2020b) study.

1.3. ALM dialect identification in SSG

Given the three studies presented have found interdialectal differences, the question arises whether a SSG speaker's ALM dialect origin can be auditorily identified. In fact, empirical evidence from a perception experiment (Guntern, 2011, p. 177) suggests that listeners are indeed able to identify the ALM origin of a SSG speaker. However, the accuracy rate differs depending on the dialect. Specifically, it was 90% correct recognition for VS, 75% for BE, 40% for GR, and 25% for ZH. Guntern claims that cues for this auditory dialect localisation in spoken SSG lie primarily on the level of vowel quality (and consonant quality). More precisely, she explains that listeners relied on the quality of the short vowels /ɪ ʊ ʏ ɔ œ ɛ/, the long vowels /i: u: y: o: ø: e:/, the <a>-sounds, the diphthongs as well as the quality of /r l k/ when successfully judging a SSG speaker's ALM dialect origin (Guntern, 2011, p. 181). However, she adds that those ALM dialects that were identified the most seem to contain phonetic features that are rather rare within German-speaking Switzerland, as it is the case for VS ALM. Thus, if ALM interferences in SSG take place, it is easier for listeners to infer the dialectal origin of a SSG speaker if said interferences are caused by phonetic features that are unique to a given ALM dialect.

1.4. Possible applications

Next to the field of dialectology, the insights of this study could potentially be of interest to forensic caseworkers. Given that dialects are an essential element of a speaker's identity (Rose, 2002, pp. 44–48; Leemann et al., 2018, p. 81), the way in which ALM dialects colour the same speaker's SSG, and the possible stability of certain features could prove beneficial, e.g., in the context of speaker identification, verification or elimination. Of course, this will have to be taken with a grain of salt as the International Association for Forensic Phonetics and Acoustics (IAFPA) states in its *Code of Practice*, rightly so, that '[m]embers should exercise particular caution with cross-language comparisons' when carrying out forensic speaker identification/elimination work. Nevertheless, the results might help forensic caseworkers by adding an additional piece of evidence to a given case, especially if dialectologically rare features are involved.

1.5. Research question and scope

The following question guided the research: Keeping in mind a speaker's ALM dialect background, how does interspeaker variability regarding SSG consonant quantity, vowel quality, and temporal features manifest itself?

It is important to understand that the variables used to explore speaker individuality were chosen because of statistically significant average interdialectal differences in Zihlmann's previous two studies (2020a, 2021) and, in the case of temporal variables, because previous research (Wiget et al., 2010; Leemann et al., 2014b; Dellwo et al., 2015; Asadi et al., 2018) has reported that *AR* and *%V* are valuable variable to explore interspeaker differences. No claims are made that other variables not scrutinised in the current study are not useable to address the research question or that the variables scrutinised in this study are the best ones to explore speaker individuality.

2. DATA AND METHODS

2.1. Speakers

32 speakers from four regions were recorded (8 each from BE, GR, VS, and ZH; 50% female; age range: 17–32 years; mean = 22.5; standard deviation (SD) = 3.42). All subjects held a Matura degree (higher secondary education) except one, and only three did not hold a university degree. One (or both) of each speakers' parents had to have grown up with the same dialect, and if only one parent spoke the same dialect, the other parent could not have grown up in a country where another German variety is spoken. To reduce the amount of possible dialect contact, the participants either had to have still been residing in the city in which they had grown up or they could not have lived elsewhere for more than three years. 2 GR and 7 VS speakers used to study in Bern or Zurich but they reported that their primary social group consisted of people speaking the same dialect. Generally, it was made sure that the speakers in the four groups were as homogeneous as possible to reduce potential effects of social factors as, e.g., education level and age.

2.2. Wordlists and recording procedure

Two wordlists were used, (1) for the assessment of consonant (and vowel) quantity as well as temporal variables, and (2) for the assessment of vowel quality.

Wordlist 1 consisted of disyllabic words with one of the vowels /i a u/ as the nucleus of the first syllable plus the consonants /p b̥ t d̥ k g̊ l n s z̥/ as the onset of the second syllable in the four phonotactically permissible vowel-consonant sequences VC, VC:, V:C, and V:C:. Due to variety-specific phonotactic constraints (see Zihlmann, 2020a, pp. 13–14), 61 words were used for BE, 65 for GR, 59 for VS, 64 for ZH, and 62 for SSG.

Wordlist 2 consisted of monosyllabic words representing the variety-specific long-vowel inventories. Due to each variety's historical development (see Zihlmann,

2021, pp. 3–4), different numbers of vowels exist (specified in Table 1).

The words from both lists were put into variety-specific generic carrier phrases whose sound immediately before the target word was a vowel (see Table 2). All ALM words were written in Dieth's (1938) spelling system, and all SSG words were written in SSG orthography.

Subsequently, for each variety blocks were created containing all variety-specific words embedded in the carrier phrases in randomised order. These blocks were arranged as ALM-SSG-ALM-SSG-ALM-SSG, resulting in three repetitions of each word. To familiarise the participants with the recording situation and elicit more naturalistic speech, prior to the recording session an interview was conducted in which metadata on the participants was collected. While most of the insights of the interviews were irrelevant to the purpose of the study, it provided an overview of frequency of SSG use and self-evaluated SSG proficiency. However, the analysis of frequency of use and the self-evaluation yielded no statistically significant correlations, which is why they are not mentioned in the result section.

For the recording of the stimuli, *SpeechRecorder* version 3.28.0 (Draxler & Jänsch, 2004) was used, so the participants could read the words from a screen sentence by sentence. They were instructed to speak as naturally as possible without artificial hyperarticulation. If possible, the recordings took place 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 a participant could not come to Zurich (which was the case for 19 subjects), the recording was conducted in a quiet furnished room either at the University of Bern or at their homes with portable recording equipment consisting of an identical interface model and the microphone *Opus 54.16/3* by *BeyerDynamic* (at 16-bit/44.1 kHz in mono, stored as .WAV). The collected corpora contained about 12,000 tokens for wordlist 1, and about 2,000 tokens for wordlist 2.

Table 1: Long vowel inventories of the four ALM dialects and SSG.

Variety	Long vowels	Number of vowels
BE	/iː yː eː øː ɛː œː æː ɑː ɔː oː uː/	11
GR	/iː yː eː øː ɛː aː ɔː oː uː/	9
VS	/iː yː eː æː ɒː oː/	6
ZH	/iː yː eː øː ɛː œː æː ɒː oː uː/	10
SSG	/iː yː eː øː ɛː aː oː uː/	8

Table 2: Variety-specific carrier phrases of the four ALM dialects as well as SSG. Each ALM carrier phrase can be translated into English as 'I said ___ too.', and the SSG one means 'I said ___.'

Variety	Carrier phrase
BE	<Ig ha o ___ gsèit.>
GR	<I han au ___ gsait.>
VS	<Ich ha öi ___ gsèit.>
ZH	<Ich han äu ___ gsäit.>
SSG	<Ich habe ___ gesagt.>

2.3. Data preparations and selection

The recordings were automatically segmented via an *R* script (2019; courtesy of Markus Jochim, Ludwig Maximilian University of Munich) using the *Munich AUtomatic Segmentation (MAUS) System* (Schiel, 1999; Kisler et al., 2017) with the language setting *General Swiss German* for ALM and *standard German* for SSG. Subsequently, the sentences were uploaded to the *EMU Speech Database Management System* (Winkelmann et al., 2017), where they were manually corrected by Zihlmann (78%) and, due to reasons of time, two additional researchers (22%).

The analysis of phonological quantity included PVD, amongst other variables not discussed here (see Zihlmann, 2020a, pp. 15–16). PVD has the advantage of being a value that disposes of articulation-rate differences, which makes varieties comparable. As mentioned in section 1.2.1, given that the average SSG PVD values of *Bullen* 'cops' were only statistically significantly different for BE and GR speakers (Zihlmann, 2020a, p. 30), VS and ZH SSG speakers are excluded from the analysis, and only the subset of the 16 BE and GR SSG speakers is analysed.

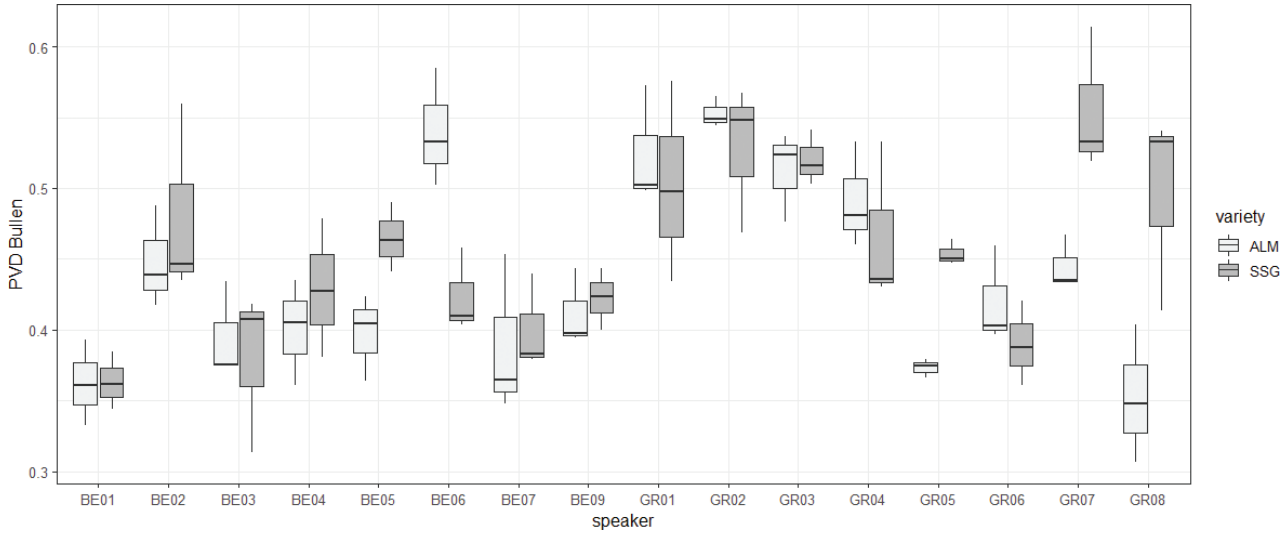
As there was a great amount of intradialectal interspeaker variation observed with regard to vowel quality for the BE SSG realisations of <ä>, the vowel-quality analysis will focus on the 8 BE SSG speakers. The assessment of their realisations of SSG <ä> was done by Zihlmann auditorily (grouped categorically by phoneme) and visually (by means of vowel plots) due to the relatively small amount of data per speaker. The vowel plots were made in *R* (2019) with *phonR* (McCloy, 2016) using frequencies normalised with Lobanov's (1971) procedure with *NORM* (Thomas & Kendall, 2007).

Regarding temporal variables, the current study will analyse the data of all 32 speakers given that the factors *AR* and *%V* were reported to be good indicators for interspeaker variation (see, e.g., Leemann, 2017; Asadi et al., 2018).

2.4. Statistical analyses

All statistical analyses were conducted in *R* (2019) and involved linear mixed-effects models (LMM) with *lme4* (Bates et al., 2015). No obvious deviations from homoscedasticity or normality were observed for the residual plots. For post-hoc pairwise comparisons (Tukey method), *lsmeans* (Lenth, 2016) was used. Regarding the LMMs for temporal variables, I opted for a model with a double interaction rather than a triple interaction

Figure 2: PVD of the word *Bulle*/*Bullen* 'cops' in ALM/SSG by speaker.



as *dialect* turned out not to be a statistically significant fixed factor for *AR* and *SR* (Zihlmann, 2020b, p. 622). Concerning *%V*, the fixed factor *dialect* was only statistically significant for ALM but not for SSG in Zihlmann's (2020b) study. However, due to convergence issues, a double interaction was opted for. For correlation analyses, *ggpubr* (Kassambara, 2019) was employed using the Pearson method.

3. RESULTS

3.1. Analysis of consonant quantity

Figure 2 depicts each BE and GR speaker's PVD of the word *Bulle/Bullen* 'cops' in ALM/SSG.

For the consonant-quantity results, the speakers' behaviour will be categorised in three patterns. Pattern 1 refers to SSG speakers behaving how it would be expected in typical SSG. Pattern 2 refers to SSG speakers behaving how it would be expected in their ALM dialects, i.e., showing ALM interferences. Pattern 3 is used for anything that is not covered by patterns 1 and 2.

In the context of the word *Bulle/Bullen* 'cops', pattern 1 corresponds to SSG speakers using long realisations of /l/, resulting in smaller PVD values. This is, incidentally, also pattern 2 for BE SSG speakers, given that BE ALM produces the word *Bulle* 'cops' with a geminated /l/. Pattern 2 for GR speakers corresponds to shorter SSG /l/, i.e., higher PVD values, due to their ALM dialect equivalent having a singleton /l/.

Zihlmann (2020a, p. 29) reports that on average, speakers behaved according to pattern 2. This, however, does not completely hold true on the speaker level. Most BE SSG speakers, i.e., BE01, BE03, BE06, BE07, and BE09, behave as expected (though in this case, patterns 1 and 2 are congruent). However, BE02, BE05, and to some degree also BE04 show a tendency to produce shorter /l/, in a way that resembles, e.g., GR04 or GR05. Thus, it can be argued that these five speakers behave according to pattern 3. As for the remaining GR speakers, most of them show pattern 2, i.e., GR01, GR02, GR03, GR07, and GR08. However, GR06 has a PVD value that is commensurate with typical SSG, i.e., comparable to the BE speakers, thus showing pattern 1.

The LLM with *PVD* as the dependent variable, and *speaker* and *variety* as independent variables (with interaction term) shows that the interaction ($F(1,15)=3.45$, $p<.001$) is statistically significant. The pairwise comparison reveals that there are intra- and interspeaker as well as intra- and interdialectal differences and similarities. Regarding intraspeaker differences, only one speaker showed statistically significantly different ALM and SSG PVDs, namely GR08 ($p=.017$). BE06, although in the opposite direction, came very close as well and narrowly missed the 5% threshold ($p=.068$). Moreover, GR07 ($p=.151$) shows a weak tendency to produce a shorter, i.e., nongeminated, consonant in SSG.

The statistically significant interspeaker differences in SSG PVDs are summarised in Table 3. The values are to be interpreted in a way that, e.g., GR02's PVD was statistically significantly different from BE01's, BE03's, and GR06's PVDs.

Seven SSG speakers, i.e., BE02, BE04, BE05, BE06, GR04, GR05, and GR08, did not show any statistically significant interspeaker PVD differences.

3.2. Analysis of vowel quality

The speaker-specific vowel plots are visualised in Figure 3.

When auditorily (by grouping the realisations categorically) and visually (based on Figure 3) inspecting the production of the BE speakers' SSG /e:/ (*See* 'lake') and /ɛ:/ (*Bär* 'bear'), several phenomena can be identified. Firstly, some SSG vowels are pronounced according to the typical SSG language convention. Secondly,

Table 3: Interspeaker differences regarding PVDs for the SSG word *Bullen* 'cops'.

Speaker	differs statistically significantly from:
BE01	GR01: p=.029; GR02: p=.002; GR03: p=.005; GR07: p<.001
BE02	-
BE03	GR02: p=.012; GR03: p=.025; GR07: p<.001
BE04	-
BE05	-
BE06	-
BE07	GR07: p=.006
BE09	GR07: p=.048
GR01	BE01: p=.029
GR02	BE01: p=.002; BE03: p=.012; GR06: p=.023
GR03	BE01: p=.005; BE03: p=.025
GR04	-
GR05	-
GR06	GR02: p=.023; GR07: p=.003
GR07	BE01: p<.001; BE03: p<.001; BE07: p=.006; BE09: p=.048; GR06: p=.003
GR08	-

some SSG vowels are produced with ALM interference. Thirdly, some speakers pronounced their SSG vowels inconsistently, resulting in intraspeaker variation, which is evident when the SSG mean values are between two ALM mean values. Lastly, although it was only observed once, hypercorrection occurred, i.e., BE09 producing SSG /ɛː/ as [eː], which, however, is also accepted in German Standard German (König, 1989, pp. 44–46).

To bring some order in these results, the speakers were grouped by shown SSG behaviour. Behaviour 1, i.e., typical SSG pronunciation, was only shown by BE02. Behaviour 2, i.e., showing ALM interferences in both SSG vowels, could be observed for BE03 and BE05 as they produced either vowel in a statistically significantly lower manner. Producing one SSG vowel consistently according to the SSG language convention and the other one consistently with ALM interference was regarded as behaviour 3. Specifically, this was observed for BE06 and BE07. However, while BE06 showed ALM interferences for SSG /eː/, BE07 produced SSG /ɛː/ in a lowered fashion. Therefore, two manifestations exist for this behaviour, i.e., producing SSG /eː/ as [ɛː] and producing SSG /ɛː/ as [æː], while pronouncing the other vowel compliant with typical SSG. Similarly, behaviour 4 describes speakers who show ALM interferences in a vowel, yet they do it inconsistently, i.e., they show intraspeaker variation. This was the case for BE01 and BE04, who sometimes produced SSG /ɛː/ as [ɛː] and sometimes as [æː], thus showing a varying pronunciation of the word *Bär* 'bear'. However, the two speakers differed in their production of SSG /eː/. While BE01 realised it as [eː], which is typical with regard to Hove's (2002) language convention, BE04 showed ALM interferences and produced it lower as [ɛː]. Therefore, regarding the intraspeaker-variation behaviour, speakers may either pronounce the other vowel according to the SSG language convention, or with ALM interference. Lastly, behaviour 5 was observed for speaker BE09, who hypercorrected SSG /ɛː/ to [eː], which resulted in an overlap with the SSG phoneme /eː/, which the speaker produced typically.

The LMM regarding vowel height with the first formant *(F1)* as dependent variable, *variety, speaker* and *vowel quality* as fixed factors (with interaction term), and random intercepts for *word* shows that the triple interaction *variety * speaker * vowel quality* is statistically significant (F(7,119)=14.71, p<.001). With regard to vowel backness, the LMM with the second formant *(F2)* as dependent variable, *variety, speaker* and *vowel quality* as fixed factors (with interaction term), and random intercepts for *word* shows the triple interaction *variety * speaker * vowel quality* to be statistically significant (F(7,119)=15.32, p<.001).

3.3. Temporal variability

Figure 4 shows each speaker's *AR* in ALM and SSG in syl/s. The LMM with *AR* as dependent variable, *variety* and *speaker* as fixed factors (with interaction term), and random intercepts for *word* shows the interaction (F(31,8830.70)=77.10, p<.001) to be statistically significant.

Due to the large amount of data, pairwise comparisons were only conducted for each speaker's ALM and SSG differences. Only two speakers showed no statistically significant differences in *AR* between their ALM and SSG varieties, namely BE02 and GR03. All other speakers had a statistically significantly higher *AR* in SSG.

As for *%V*, Figure 5 depicts the ALM and SSG results by speaker. The LMM with *%V* as dependent variable, *variety* and *speaker* as fixed factors (with interaction term), and random intercepts for *word* shows the interaction (F(31,1146)=7.99, p<.001) to be statistically significant.

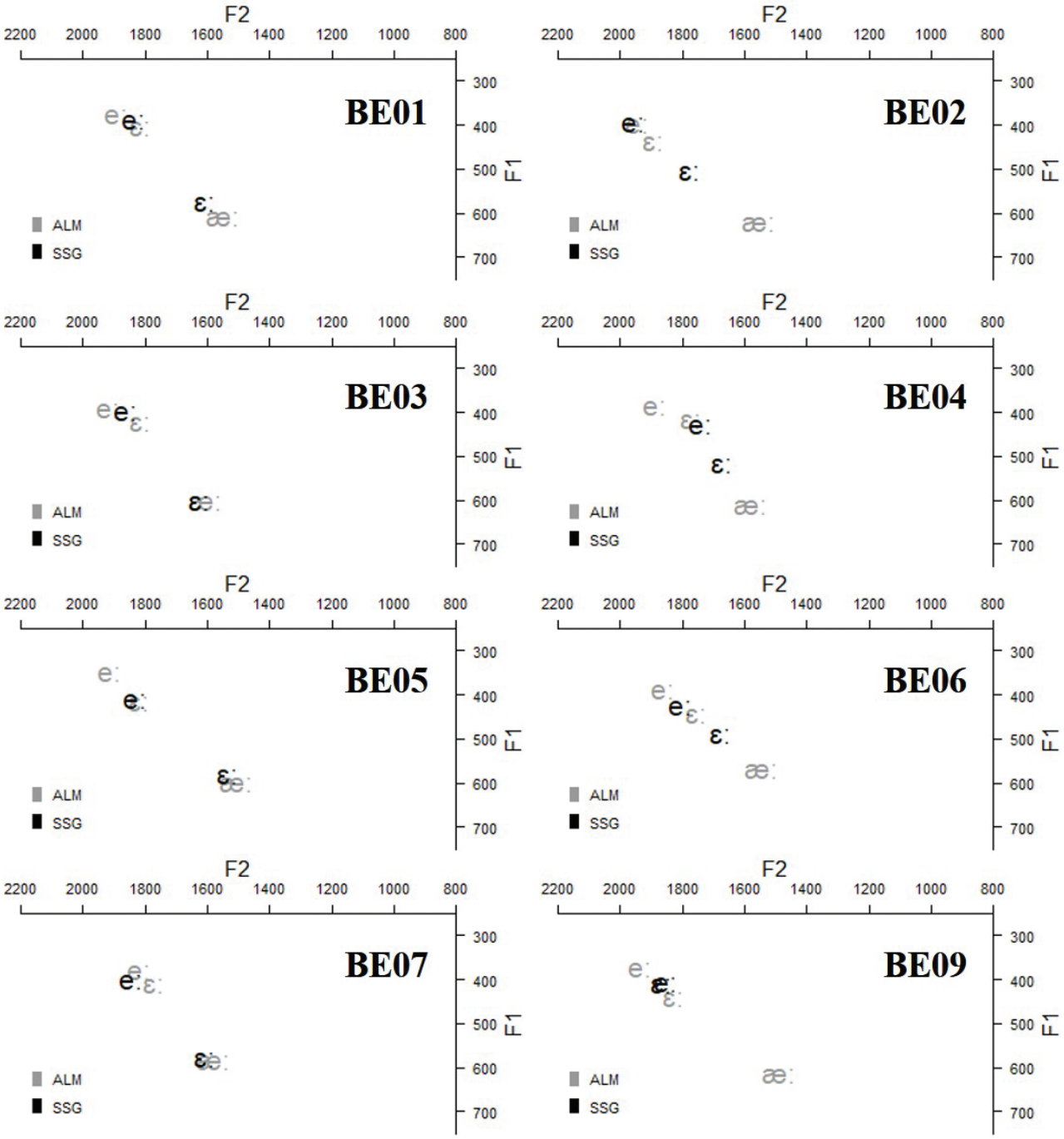
Here as well, pairwise comparisons were limited to intraspeaker differences between the two varieties. The results suggest that five speakers show no statistically significant differences between their ALM and SSG *%V* values, i.e., VS01, VS02, ZH02, ZH05, and ZH08. All other speakers had a statistically significantly lower *%V* value in SSG.

The correlation analysis between *AR* and *%V* shows a statistically significantly negative correlation between the two variables (R=-.37; p=.038). In other words, the quicker the speech, the more vowel reduction occurs.

4. DISCUSSION

Let us start with consonant quantity. Zihlmann (2020a, pp. 26–28) reports that overall, BE and GR SSG show statistically significantly different mean PVD values for

Figure 3: Mean normalised ALM and SSG F1/F2 (in Hz) of unrounded front mid vowels in BE by speaker.



some words with a quantity mismatch between ALM and SSG. It would thus be surprising not to find more GR SSG speakers who pronounced the word *Bullen* 'cops' with a comparatively shorter /l/ compared to BE SSG speakers. However, this is not to say that SSG speakers with the same ALM dialect background necessarily speak SSG more similarly than SSG speakers with a different ALM dialect background. As a matter of fact, within the group of GR SSG speakers, GR06's PVD shows more similarities to the BE SSG mean PVD, and even differs statistically significantly from speakers with the same dialectal background, namely GR02's and GR07's ones in that GR06 geminated /l/ more than the other GR SSG speakers tested. But in the BE group as well, some individual speakers stand out. Specifically, BE02 and BE05 show more similar PVD values to GR04 or GR05 than to, e.g., BE01. This shows that the mean is indeed just an abstraction of how all speakers behave, yet some speakers

Figure 4: *AR* in ALM and SSG by speaker (in syl/s).

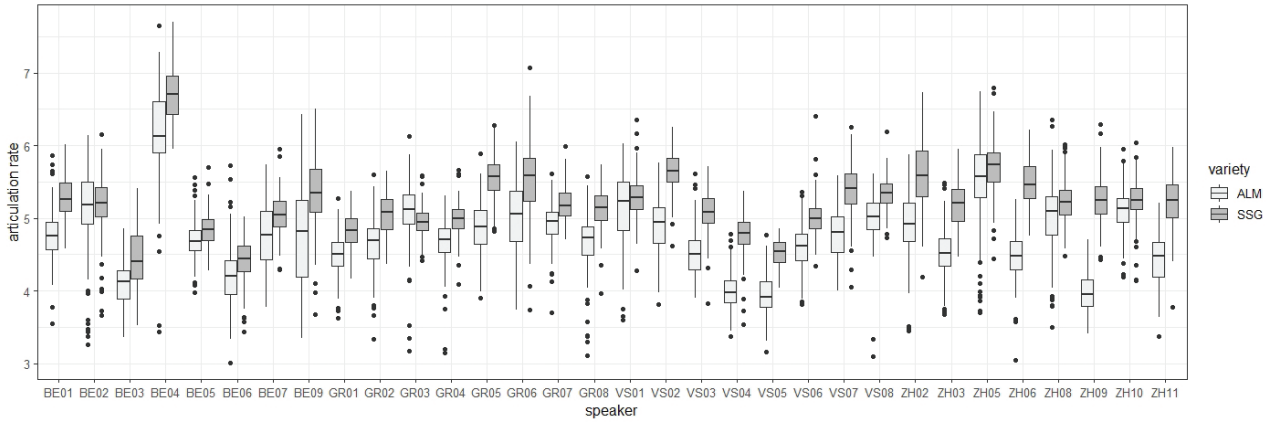
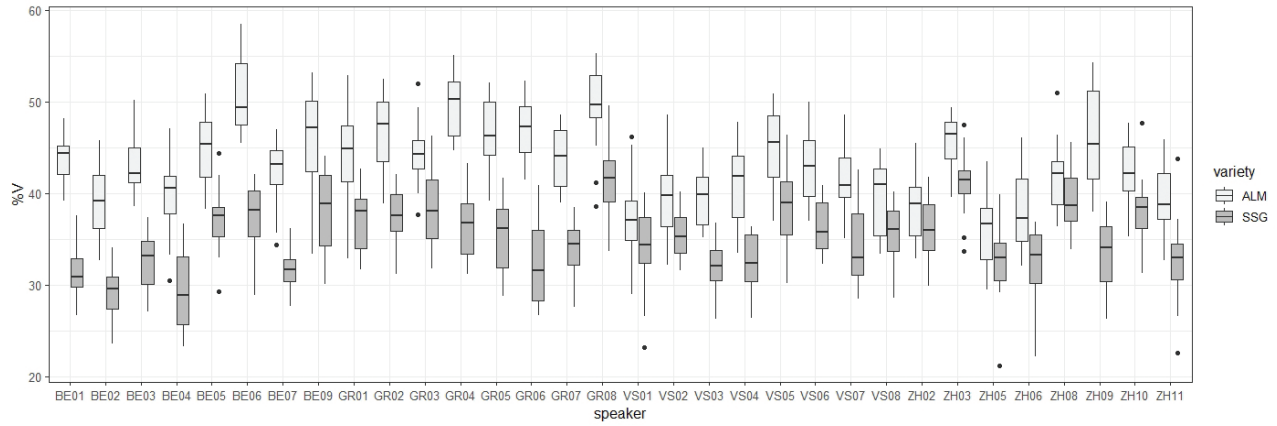


Figure 5: *%V* in ALM and SSG by speaker.



might still pronounce words in a way that the average would not predict. Furthermore, the fact that seven speakers did not differ statistically from any of the other speakers indicates that there is indeed a somewhat neutral zone that does not allow for any conclusions regarding a speaker's ALM dialect background.

Even in ALM intradialectal differences were observed, both for BE and for GR speakers. Among the BE speakers, especially BE06 behaves differently as almost all BE speakers articulated the BE ALM word *Bulle* 'cops' with a relatively short vowel and a relatively long consonant. The same is true for GR ALM speakers, where two broad types can be identified: those who geminated /l/ and those who did not. Now, of course, this could be due to the fact that GR ALM shows some less rigid rules regarding vowel and consonant length, meaning that the quantity opposition may be in the process of ceasing to exist as claimed by Eckhardt (1991, pp. 36–38); this could explain why differences in SSG consonant duration were observed for GR speakers but not for BE speakers. However, if this were true, it would also have to be the case that the same SSG speaker performs similarly in ALM, which is not always the case, even though both BE and GR used their variety-specific ALM equivalent for the word *Bullen* 'cops'.

In fact, when we compare the same speaker's ALM and SSG, it becomes apparent that their ALM performance cannot always predict their SSG performance. While about 11 out of the 16 speakers scrutinised show comparable values in ALM and SSG, 5 show rather different values, i.e., BE05, BE06, GR05, GR07, and GR08. These GR speakers as well as BE05 produced the /l/ in ALM longer than in SSG, which manifest itself with a higher PVD value in SSG compared to ALM. In contrast, BE06 showed a shorter /l/ in ALM compared to SSG, as evident by the lower SSG PVD value. This suggests (1) that, for this case, a speaker's SSG performance is not necessarily influenced by their ALM dialect, and (2) that even within a dialect, speakers can perform diametrically opposed to one another, as it is the case for BE05 and BE06.

With regard to vowel quality, as only BE SSG speakers were scrutinised, no between-dialect comparison can be made. The eight speakers showed four broad behaviours,

yet which one they showed differed so greatly that five behavioural patterns could be identified, implying that there is indeed much intradialectal interspeaker variability. Yet, it was not only the speakers that differed in comparison to one another. Intraspeaker variation was also found in that two speakers (BE01 and BE04) did not read the word *Bär* 'bear' with the same vowel quality throughout the approximately 70 minutes of testing. On top of that, these two speakers showed intraspeaker variation to varying degrees; BE01 tended to pronounce SSG with more ALM interference than BE04. All other speakers stayed consistent, even though the auditory assessment pointed towards some variation in the degree to which /eː/ or /ɛː/ are lowered. This could, of course, be due to the relatively low amount of data per vowel. In fact, both SSG /eː/ and SSG /ɛː/ were only realised three times per speaker. Therefore, to verify the generalisability of this claim, more recordings will have to be analysed. Nevertheless, the fact that a relatively frequent word shows variation to such a degree demonstrates (1) that the 'language convention' (Hove, 2002, p. 6) is very much active, and (2) that some people seem to pronounce the words in a more monitored fashion than others, like, e.g., BE09 who even overcorrected, or BE02, who was the only one who produced the words in a consistently typical manner.

When we look at temporal variables, things become a bit complicated. While previous research has found that ALM dialects differ statistically significantly regarding rhythmic measurements (Leemann et al., 2012; 2014a; 2014b; 2018), this was not confirmed for SSG (Zihlmann, 2020b). What is more, on the speaker level, already Dellwo et al.'s (2015) study showed that interspeaker variation amongst ALM speakers is rather great, which has also been found in this study for both ALM and SSG regardless of their dialect background. On top of this, regarding the evidence that some temporal measurements stay stable when speaking a foreign language (de Jong, 2018; de Jong & Mora, 2019), this could not be confirmed for ALM and SSG articulation rates (*AR*) and the share of vocalic speech in percent (*%V*) as only 2 and 4 speakers, respectively, showed similar values in the 2 varieties. On average, neither *AR* nor *%V* showed statistically significant region-specific values in SSG (Zihlmann, 2020b). Furthermore, the speaker-specific evaluation reveals that some speakers even show more similarities with speakers of other dialects than with speakers of their own one, as it is the case for, e.g., the *AR* of BE01 and BE02 that are much more similar with the ones of VS01, GR09 or ZH08 than with the ones of BE03 or BE06. When looking at previous studies who included more than just four dialects (e.g., Leemann, 2017), we can see that even though there are *AR* differences amongst dialects, there are also distinct dialects that show very comparable values. Moreover, when looking at Figures 4 and 5, we can see that not all speakers behave the same within their own dialect, as, e.g., BE06, whose ALM *%V* value is more comparable to the ones of GR08 or ZH09 than the one of BE04. Therefore, speakers with different dialect backgrounds do not necessarily show differences in temporal variables.

These insights all point towards the conclusion that in general, speakers with the same dialect background do not necessarily pronounce SSG comparably. This is not the case because general trends for SSG speakers with the same ALM dialect background cannot be identified; rather, some individuals within a group of SSG speakers with the same dialect background speak in a way that is more associated with SSG speakers of another dialect background. This suggests that other factors must also influence the way in which SSG is spoken, and that in particular the above-mentioned SSG language convention seems to play a major role. It is, however, unclear to what extent other factors correlate with the results. In the following, some of these factors are introduced and discussed.

Possible explanations for the variation observed can be found in the sociolinguistic domain. For instance, the lowering of BE SSG mid vowels is very salient to both speakers from BE and speakers of other dialects. It is stigmatised to some degree being considered a clear ALM interference, which is associated with a low education level (Hove, 2002, p. 20). Anecdotal evidence from BE speakers suggests that teachers told their students not to lower their mid vowels in SSG as it is considered too dialectal and thus wrong. Consequently, BE speakers are very aware of their SSG pronunciation and try not to lower said vowels to avoid being associated with a stigmatised pronunciation. This even goes so far that hypercorrection occurs, as it was the case with BE09, who raised /ɛː/ and pronounced it as [eː], which suggests that BE speakers try to avoid sounding as if they were poorly educated.

However, constant monitoring of one's speech is rather energy-demanding and so occasionally, speakers fall into old habits as evident by intraspeaker variation. Guntern (2012) portrays this situation by introducing a continuum of variants among dialect speakers can choose when speaking SSG. This continuum ranges from variants that clearly have their origin in ALM to variants that clearly stem from German Standard German; in between are 'ambiguous' or 'neutral' variants. She claims that the more variants that clearly have their origin in ALM are employed, the more the SSG variety of a speaker will be perceived as being Swiss, or rather, showing ALM interferences (p. 106). However, the exact manifestation of said variation appears to be unpredictable, even though it is very likely that variation occurs in general.

Variation can also be caused by many other factors, as, e.g., differences in how often someone speaks SSG. Although in the context of this study, the inspection of the questionnaires resulted in no evidence for a correlation between frequency of SSG use and the participants' self-evaluated SSG proficiency, it could undoubtedly influence the results. BE speakers reported to speak SSG the least amongst the four regions tested, and some BE speakers tended to show more significant non-standard (i.e., atypical) results than speakers of other dialect backgrounds. This especially applies to vowel quality where

BE shows the most salient deviations from the SSG norms. However, there were some speakers (as, e.g., BE02), who use SSG rarely, rated their SSG as only 'sufficient', and still produced all vowels abiding by the SSG language convention. Simultaneously, the only BE speaker to use SSG multiple times daily (i.e., BE09) engaged in hypercorrection. It is hence uncertain to what degree the factor *frequency of SSG use* influenced the variation observed.

Another factor that could influence SSG performance is the speakers' attitude towards German Standard German: Are people actively trying to sound more German, are they opposed to such behaviour, or do they have a neutral opinion? Guntern (2012, pp. 107–109) states that there exist different ideals as to how SSG should be spoken, ranging from 'as the Germans do', which is the case for the 'urban in-crowd' who find it embarrassing that SSG is linguistically different from German Standard German, to 'consciously stressing ALM interferences', which is the case for, e.g., nationalist politicians, who want to highlight the differentness of Switzerland and Germany. Obviously, this has implications for many aspects, as, e.g., for consonant quantity; while gemination is generally expected in typical SSG, this is not the case in German Standard German, where no gemination takes place.

The context in which SSG is spoken also influences pronunciation. Specifically, this can include the interlocutor (e.g., one's boss vs. one's friend), the setting (e.g., while being interviewed on television vs. talking to a tourist on the street), and the set, i.e., the personal state of mind (well rested, focused, or at ease vs. fatigued, unconcentrated, or hungry). Lastly, there always exist the possibility that a speaker simply makes a pronunciation mistake. All factors influence how well speakers can monitor their SSG performance, which affects pronunciation.

In sum, the research question, i.e., how interspeaker variability regarding SSG consonant quantity, vowel quality, and temporal features manifests itself while keeping in mind a speaker's ALM dialect background, can be answered as follows. While overall trends for the four ALM dialect regions could be identified for consonant quantity if there is a consonant quantity mismatch between a given ALM word and its direct translation in SSG, this does not apply to the speaker level. Indeed, some speakers performed very differently in ALM and SSG, to the extent that they showed more similarities with the SSG mean of a different ALM dialect. This indicates that the ALM dialect origin does not necessarily influence a speaker's SSG performance. Vowel quality and temporal variables showed a lot of interspeaker variation too. As for vowel quality, some speakers pronounced the SSG vowels scrutinised in a typical fashion, some showed clear ALM interferences, and some even showed intraspeaker variation. Thus, many behaviours were observed to the extent that no clear generalisations can be drawn for how BE SSG mid vowels are pronounced. Similarly, temporal features show a great amount of intra- and interdialectal interspeaker variation as well.

Of course, these conclusions are drawn solely by analysing three features while, as previously mentioned, no claims are made that they are the best indicators to explore speaker individuality in SSG. While, due to the relatively small number of speakers, the results have to be taken with a grain of salt, they nevertheless do allow a glimpse at how speakers of with the same dialect background speak ALM as well as SSG.

5. CONCLUSION AND SUGGESTIONS FOR FUTURE RESEARCH

This study has explored previously reported interdialectal differences in SSG consonant quantity and vowel quality, as well as two temporal variables that turned out to show much interspeaker variation in other languages on the speaker level. The analyses revealed that average tendencies found for dialects might not be true for individual speakers. Specifically, it was found that while some speakers behave very much like their dialectal mean, others performed rather dialectally atypically, occasionally even resembling other dialects' means better. This was especially the case for consonant quantity and temporal variables in SSG. In other words, there is much intra- and interdialectal interspeaker variation observable for all three variables scrutinised. The reasons for this are not straightforward. Indeed, many factors may play a role, including stigmatisation of ALM dialect interferences in SSG, frequency of SSG use, attitude toward German Standard German, or fatigue at the point of speaking.

To understand the links between ALM and SSG on various levels better, future research should include more speakers and/or speakers of other dialects and level of self-evaluated SSG proficiency. One could also analyse other parameters of vowel and consonant quantity (e.g., more words whose vowel and consonant quantity differs in ALM and SSG), vowel quality (e.g., the quality of short vowels, which can vary in tenseness and were found to be helpful for identifying the ALM dialect origin of SSG speakers in Guntern's (2011) perception experiment), or prosodic variables (see, e.g., Pellegrino et al., 2019) to shed more light on the interrelations of a single speaker's ALM and SSG varieties. Lastly, one could also investigate to what degree the factor *lexical item* plays a role with regard to variability.

6. ACKNOWLEDGEMENTS

This study was funded by the Swiss National Science Foundation (SNSF), grant Nr. 164377. First and foremost, I would like to thank Stephan Schmid and Volker Dellwo for their feedback. Furthermore, I thank Sandra Schwab, who was of great help during the statistical analysis. Very much appreciated is also the help of Marie-Anne Morand and Seraina Nadig, who assisted me during the arduous work of manually correcting the automatically segmented data. Moreover, without the valued help of Markus Jochim (Ludwig Maximilian University of Munich) and his expertise in R, the data analyses would not have been as smooth. I would also like to thank the two anonymous reviewers for their suggestions and advice. Lastly, I am

also very grateful for the help of the following three people, who each helped me find appropriate words for their native dialect: Christa Schneider (University of Bern) for BE, Oscar Eckhardt (Grisons Institute for Cultural Studies) for GR, and Sandro Bachmann for VS

REFERENCES

Ammon, U., Bickel, H., Ebner, J., Esterhammer, R., Gasser, M., Hofer, L., Kellermeier-Rehbein, B., Löffler, H., Mangott, D., Moser, H., Schläpfer, R., Schlossmacher, M., Schmidlin, R. & Vallaster, G. (2004). Variantenwörterbuch des Deutschen. Die Standardsprache in Österreich, der Schweiz und Deutschland sowie in Liechtenstein, Luxemburg, Ostbelgien und Südtirol. Unter Mitarbeit von Kyvelos, R., Nyffenegger, R. & Oehler, T. Berlin & New York: Mouton de Gruyter.

Asadi, H., Nourbakhsh, M., He, L., Pellegrino, E. & Dellwo, V. (2018). Between-speaker rhythmic variability is not dependent on language rhythm, as evidence from Persian reveals. *International Journal of Speech*, *Language and the Law*, *25*(2), 151–174.

Bates, D. M., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. http://dx.doi.org/10.18637/jss.v067.i01

Christen, H., Glaser, E. & Friedli, M. (2013). *Kleiner Sprachatlas der deutschen Schweiz.* Frauenfeld: Huber.

Christen, H., Guntern, M., Hove, I. & Petkova, M. (2010). *Hochdeutsch in aller Munde. Eine empirische Untersuchung zur gesprochenen Standardsprache in der Deutschschweiz.* Stuttgart: Steiner.

de Jong, N. H. (2018). Fluency in second language testing: Insights from different disciplines. *Language Assessment Quarterly*, *15*(3), 237–254. http://dx.doi.org/10.1080/15434303.2018.1477780

de Jong, N. H. & Mora, J. C. (2019). Does having good articulatory skills lead to more fluent speech in first and second languages? *Studies in Second Language Acquisition*, *41*(1), 227–239. http://dx.doi.org/10.1017/S0272263117000389

Dellwo, V., Leemann, A. & Kolly, M.-J. (2015). Rhythmic variability between speakers: Articulatory, prosodic, & linguistic factors. *Journal of the Acoustical Society of America*, *137*(3), 1513–1528.

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

Draxler, C. & Jänsch, K. (2004). SpeechRecorder – a universal platform independent multi-channel audio recording software. *Proceedings of the 4th International Conference on Language Resources and Evaluation*, Portugal, 559–562.

Duden (2015). *Aussprachewörterbuch. Wörterbuch der deutschen Standardaussprache*. S. Kleiner, R. Knöbl & M. Mangold (Eds). 7th ed. Berlin: Dudenverlag.

Eckhardt, O. (1991). *Die Mundart der Stadt Chur*. Zurich: Phonogram Archives of the University of Zurich.

Ferguson, C. (1959). Diglossia. *Word*, *15*, 325–340.

Grabe, E. & Low, E. L. (2002). Durational variability in speech and the Rhythm Class Hypothesis. In V. Gussenhoven & N. Warner (Eds), *Laboratory Phonology*, *7*, 515–545. Berlin & New York: Mouton de Gruyter.

Guntern, M. (2011). Erkennen von Dialekten anhand vom gesprochenen Schweizerhochdeutsch. *Zeitschrift für Dialektologie und Linguistik*, *78*(2), 155–187.

Guntern, M. (2012). Dialekt und gesprochene Standardsprache: Wie Laien gesprochenes Schweizerhochdeutsch beurteilen. *Sociolinguistica, 26*, 102–119.

Haas, W. (2000). *Die deutschsprachige Schweiz*. In H. Bickel & R. Schläpfer (Eds), *Die viersprachige Schweiz*, 57–138. 2nd ed. Aarau: Sauerländer.

Ham, W. (2001). *Phonetic and Phonological Aspects of Geminate Timing.* London: Routledge.

Hove, I. (2002). *Die Aussprache der Standardsprache in der deutschen Schweiz*. Berlin & Boston: Max Niemeyer.

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

International Association for Forensic Phonetics and Acoustics. (n.d.). Code of Practice. Retrieved on 7 November 2020 from https://www.iafpa.net/code.htm

Kassambara, A. (2019). ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.2.4. https://CRAN.R-project.org/package=ggpubr

Kisler, T., Reichel, U. D. & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, *45,* 326–347.

Kleber, F. (2017). Complementary length in vowel-consonant sequences: acoustic and perceptual evidence for a sound change in progress in Bavarian German. *Journal of the International Phonetics Association*, *50*(1), 1–22. http://dx.doi.org/10.1017/S0025100317000238

Kohler, K. J. (1979). Dimensions in the perception of fortis and lenis plosives. *Phonetica*, *36*, 332–343.

König, W. (1989). *Atlas zur Aussprache des Schriftdeutschen in der Bundesrepublik Deutschland* (Band 1: Text). Ismaning: Hueber.

Leemann, A. (2017). Analyzing dialectal variation in articulation rate using crowdsourced speech data. *Journal of Linguistic Geography*, *4*(2), 76–96.

Leemann, A., Dellwo, V., Kolly, M.-J. & Schmid, S. (2012). Rhythmic variability in Swiss German dialects. *Proceedings of the 6th Speech Prosody*, China, 607–610.

Leemann, A., Dellwo, V., Kolly, M.-J. & Schmid, S. (2014a). Disentangling sources of rhythmic variability between dialects. *Proceedings of the 7th Speech Prosody*, *Ireland*, 693–697.

Leemann, A., Kolly, M.-J. & Dellwo, V. (2014b). Speaker-individuality in suprasegmental temporal features: Implications for forensic voice comparison. *Forensic Science International, 238*, 59–67.

Leemann, A., Kolly, M.-J., Nolan, F. & Li, Y. (2018). The role of segments and prosody in the identification of a speaker's dialect. *Journal of Phonetics*, *68*, 69–84.

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, *69*(1), 1–33. http://dx.doi.org/10.18637/jss.v069.i01

Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *Journal of the Acoustical Society of America*, *29*, 88–104.

McCloy, D. R. (2016). phonR: tools for phoneticians and phonologists. R package version 1.0-7. https://CRAN.R-project.org/package=phonR

Pellegrino, E., Kathiresan, T., Roswandovitz, C., Frühholz, S. & Dellwo, V. (2019). Can prosody be the key to spot fake voices? Acoustic and automatic speaker verification analyses on digital and natural voices [Conference Poster]. *28th International Conference of the International Association for Forensic Phonetics and Acoustics*, *Turkey*.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing (version 3.6.0). Retrieved on 30 April 2019 from https://www.R-project.org/

Ramus, F., Nespor, M. & Mehler, J. (1999). Correlates of linguistic rhythm in speech signal. *Cognition*, *73*, 268–292.

Rose, P. (2002). *Forensic Speaker Identification.* London: Taylor & Francis.

Schiel, F. (1999). Automatic Phonetic Transcription of Non-Prompted Speech. *Proceedings of the 14th International Congress of Phonetic Science*, USA, 607–610.

SDS (1962). *Sprachatlas der deutschen Schweiz Volume I*. R. Hotzenköcherle & R. Trüb (eds). Bern & Basel: Francke.

Siebs, T. (1969). *Deutsche Aussprache. Reine und gemässigte Hochlautung mit Aussprachewörterbuch*. H. de Boor, H. Moser & C. Winkler (Eds). 19th ed. Berlin: Walter de Gruyter.

Thomas, E. R. & Kendall, T. (2007). NORM: The vowel normalization and plotting suite. Retrieved on 23 July 2019 from http://ncslaap.lib.ncsu.edu/tools/norm/.

Wiget, L., White, L., Schuppler, B., Grenon, I., Rauch, O. & Mattys, S. L. (2010). How stable are acoustic metrics of contrastive speech rhythm? *Journal of the Acoustical Society of America*, *127*(3), 1559–1569. http://dx.doi.org/10.1121/1.3293004

Winkelmann, R., Harrington, J. & Jänsch, K. (2017). EMU-SDMS: Advanced speech database management and analysis in R. *Computer Speech & Language*, *45*, 392–410.

Zihlmann, U. (2020a). 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, U. (2020b). Temporal variability in four Alemannic dialects and its influence on the respective varieties of Swiss Standard German. *Speech Prosody 2020*, 620–624. http://dx.doi.org/10.21437/SpeechProsody.2020-127

Zihlmann, U. (2021). Vowel quality in four Alemannic dialects and its influence on the respective varieties of Swiss Standard German. *Journal of the International Phonetic Association,* 1–28. http://dx.doi.org/10.1017/S0025100320000377