INCREASING ASPIRATION OF SWISS GERMAN PLOSIVES: A SOUND CHANGE IN PROGRESS?

Franka Zebe, Camille Watter, and Stephan Schmid

University of Zurich
franka.zebe@uzh.ch, camille.watter@uzh.ch, stephan.schmid@uzh.ch

ABSTRACT

In Alemannic dialects of German-speaking Switzerland, the primary cue between lenis and fortis plosives is closure duration, with lenis plosives having a shorter closure than fortis plosives, while both are phonetically voiceless. Recently, it has been called into question whether there is an increasing tendency for speakers to additionally produce aspirated fortis plosives, possibly due to the contact to German Standard German. To investigate this, we recorded word-initial and word-medial fortis plosives produced by 24 older and 24 younger speakers of Zurich German and analysed their normalized VOT values.

Results show that, although the word-medial plosives can probably all be considered unaspirated, younger speakers overall produced significantly longer VOT values compared to older speakers. Word-initial plosives differed much more between age groups and also resulted in considerable variability for certain words. These new results hint at a possible sound change in progress.

Keywords: VOT, aspiration, sound change, Swiss German, Zurich German

1. INTRODUCTION

It is general consensus that the main acoustic cue to the phonemic contrast between homorganic lenis and fortis plosives in Alemannic dialects of German-speaking Switzerland is closure duration: lenis plosives display shorter closures compared to fortis plosives, while both are phonetically voiceless [1, 2, 3, 4]. Recently, it has been observed that another category is evolving, namely aspirated plosives [5, 6]. Hence, we can assume three phonetic categories of Swiss German plosives: lenis, unaspirated fortis, and aspirated fortis, where the phonemic status of the latter category is unclear [6]. Using the apparent-time construct [7, 8], the aim of this study is to investigate how older and younger speakers of Zurich German differ in their VOT productions in word-medial and word-initial fortis plosives.

1.1. Swiss German plosives

Recent studies have revealed an increasing amount of word-initial aspirated fortis plosives in Alemannic dialects of German-speaking Switzerland [5, 6]. [6] investigated a set of young speakers of Zurich German. Results show that the difference between lenis and unaspirated fortis plosives is stable in that its acoustic correlate primarily lies in closure duration, not in VOT [6]. There is, however, also evidence for a third category of aspirated fortis plosives in word-initial position. First and foremost, it is clear from their study that aspiration only occurs in a certain set of words, in particular loanwords from German Standard German (GSG) or English [6]. Furthermore, low frequency words seem to be aspirated more often than high frequency words [6]. Secondly, findings indicate a high degree of variability in the words that are aspirated [6].

The goal of the current study is to add onto the evidence from [6] by including a group of older speakers and investigating not only word-initial but also word-medial plosives.

1.2. Sound change

When dealing with a sound change in progress, the question usually arises whether the change is internally or externally driven. Traditionally, it has been argued that a regular sound change takes place in a phonetically gradual manner and is lexically independent, while an irregular sound change occurs in a phonetically abrupt manner due to the process of dialect levelling driven by external factors, e.g. when dialect speakers produce features similar to that of the standard variety [9]. These theoretical implications raise an issue for the current situation in German-speaking Switzerland. Since it has been shown that aspirated plosives occur in loanwords from GSG or even English, it would be logical to assume that the sound change is driven by external factors, since in those two languages aspiration is used as the main cue to distinguish between lenis and fortis plosives [10, 11, 12]. In addition, recent studies have shown that in other southern German varieties with high contact to GSG VOT has become increasingly important for distinguishing between

lenis and fortis plosives, namely in Standard Austrian German [13] and in Western Central Bavarian [14].

On the other hand, there is evidence for both lexical diffusion (in the sense of [15]) and a certain amount of interspeaker variability in Zurich German [6]. These two aspects would be an indication for an internally driven sound change in progress. In recent years, however, the evidence of continuous sound changes due to dialect levelling has been increasing [16, 17, 18].

Therefore, the results of the current study will have to be discussed also in relation to the general nature of this (possible) sound change in progress.

2. METHODS

2.1. Participants and stimuli

A corpus of read speech was collected from 24 older speakers (12 female) aged between 53 and 81 years (mean=66.54, SD=6.34) and 24 younger speakers (12 female) aged between 20 and 34 years (mean=27.96, SD=3.52) at the time of the recording.

Stimuli consisted of 24 words, 13 with a word-initial fortis plosive and 11 with a word-medial fortis plosive as target. Target words were either disyllabic trochees or monosyllables; thus, they did not differ substantially in phonological word length. This set of stimuli is part of a larger study with the same participants. For the current production study, however, the stimuli were limited to 24 words containing fortis plosives. All words were embedded in a carrier sentence. All in all, we recorded and analyzed a corpus of 1152 items (48 speakers x 24 stimuli/plosives).

2.2. Recording procedure

Participants were recorded in a soundproof booth at the Phonetics Laboratory of the University of Zurich, using a personal computer with the interface USBPre® 2 (Sound Devices) and the microphone NT2-A (RØDE) with the recording software SpeechRecorder [19]. The speakers were instructed to read each sentence once. All recordings had a sample rate of 16-bit/44.1 kHz and were saved as .wav-files. Speakers were reimbursed for their participation.

2.3. Data preparation and measurements

Each recorded sentence was automatically segmented using WebMAUS [20], selecting *German Dieth (CH)* in the language annotation settings. The recordings were then manually adjusted using the EMU-webApp [21] according to the the following procedure: The beginning and end of each sentence was adjusted as well as the boundaries of each segment of the target word. In addition, we annotated two phases within the plosives, i.e. the closure phase and the VOT, which was positive in all cases. The VOT values and target word durations were obtained in milliseconds using the emuR package [22] in R [23]. The VOT values were then normalized by dividing their absolute duration by the duration of the entire word to account for differences in articulation rate. The same normalization procedure has been adopted by [18]. In addition, the word frequency in terms of logSubtlex values (log lexical frequency per million) for each word were obtained from the Subtlex corpus [24].

2.4. Statistical analysis

For the statistical analysis, we fitted two linear mixed-effects models using the R packages lme4 [25] and lmerTest [26]. The first model was fitted for the word-medial plosives with the normalized VOT values *VOTnorm* as dependent variable. Fixed effects consisted of *word* and *age* (old vs. young). Lastly, random intercepts for *speaker* were added. The second model for the word-initial plosives had the same structure.

To obtain the Chi-square and p values reported in the results section, a Type II ANOVA for each model was calculated using the R package car [27]. In case of a significant interaction, pairwise comparisons using Tukey's test were calculated using the R package emmeans [28].

3. RESULTS

The word frequencies in terms of logSubtlex values are shown in Table 1a and 1b.

Word	logSubtlex
buute ('to boot')	0
Sitte ('manner')	1.292
räppe ('to rap')	1.415
Sippe ('clan')	1.531
Wappe ('coat of arms')	1.663
Matte ('mat')	1.851
Huupe ('car horn')	1.954
Latte ('lath')	1.964
Kööter ('dog')	2.121
Kaater ('tomcat')	2.297
Schatte ('shadow')	2.873

Table 1a: logSubtlex values for target words with plosives in word-medial position.

As can be seen, the values for the word-medial target plosives are slightly lower compared to those with word-initial target plosives. Still, we assume that the differences between word-medial and word-initial plosives are not due to the slightly different word frequencies.

Word	logSubtlex
Torf ('peat')	0.477
paare ('to mate')	1.748
Part ('part')	2.025
Paan (Greek god 'Pan')	2.041
Pèèrli ('couple', dim.)	2.041
Paar ('couple')	2.072
Poole ('Poland')	2.176
Tipp ('tip')	2.553
Tuur ('tour')	2.651
Päch ('bad luck')	2.786
Taate ('acts')	3.032
Tee ('tea')	3.176
Pass ('passport')	3.629

Table 1b: logSubtlex values for target words with plosives in word-initial position.

3.1. Word-medial plosives

	Chisq	Df	Pr(>Chisq)
word	177.1698	10	< 2e-16 ***
age	4.6053	1	.03187 *
word:age	15.5738	10	.11250

Table 2: Statistical output of the linear mixed-effects model for the word-medial plosives.

The linear mixed-effects model for the word-medial plosives revealed significant main effects for both word and age, while the interaction between the two did not turn out to be significant, which can also be seen in Table 2. Younger speakers produced longer VOT values compared to older speakers in all words, as is illustrated in Figure 1.

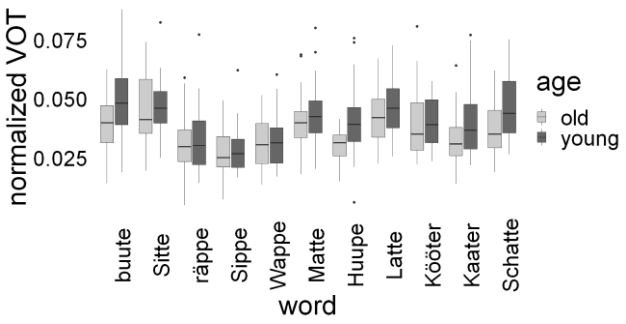

Figure 1: Normalized VOT values (y axis) for the words containing word-medial target plosives (x axis), ordered by word frequency; older speakers in light grey, younger speakers in dark grey.

Even though younger speakers produced significantly longer VOT values, it must be emphasized that these values are overall very short and can probably all be classified as unaspirated, especially when compared to the initial plosives.

3.2. Word-initial plosives

	Chisq	Df	Pr(>Chisq)
word	1061.1147	12	< 2.2e-16 ***
age	8.5101	1	.003532 **
word:age	26.8246	12	.008189 **

Table 3: Statistical output of the linear mixed-effects model for the word-initial plosives.

The model for the word-initial plosives revealed significant main effects for *word* and *age*. The interaction between the two also turned out to be significant, as is shown in Table 3. Again, younger speakers produced overall longer VOT values, although there were some exceptions.

Since the interaction turned out to be significant, pairwise comparisons between age groups for each word were calculated. The following words were significantly different between age groups: *Torf* ('peat') (z=-3.187, p=.001), *Part* ('part') (z=-3.308, p<.001), *Paan* (Greek god 'Pan') (z=-2.149, p=.032), *Poole* ('Poland') (z=-2.773, p=.006), and *Tee* ('tea') (z=-3.225, p=.001). In all cases, younger speakers produced longer VOT values compared to older speakers. This can also be seen in Figure 2.

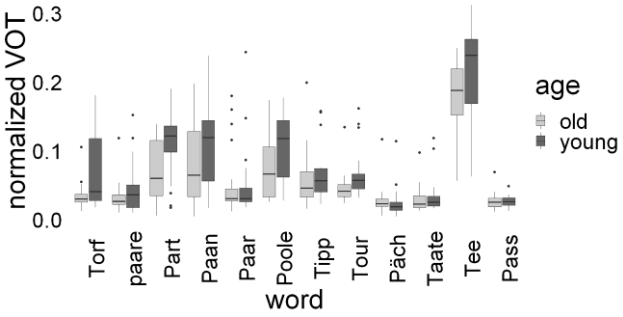

Figure 2: Normalized VOT values (y axis) for the words containing word-initial target plosives (x axis), ordered by word frequency; older speakers in light grey, younger speakers in dark grey.

It is also important to note that for the words that significantly differ between age groups (which are arguably undergoing a sound change), there is a much higher amount of variability compared to the other words, whose plosives can be considered unaspirated. Interestingly, this is not only the case for the younger but also the older speakers, indicating that the older generation is also part of the sound change in progress. Furthermore, it is a clear

indication that the change is gradual rather than abrupt.

4. DISCUSSION

4.1. Discussion of the results

First of all, our results indicate that there is a sound change in progress for word-initial fortis plosives, some of which are produced with aspiration, primarily by younger speakers. As for the word-medial plosives, although younger speakers produce significantly longer VOT values compared to older speakers, ultimately one cannot speak of a sound change in progress, as the values are overall very short. Still, word-medial plosives should also not be disregarded in future investigations.

As already pointed out by [6], the current results provide evidence for the fact that only certain words are affected by the sound change in progress. We can also confirm that, in general, lower frequency words are more aspirated than higher frequency words. Even so, it is important to note that the word frequencies of the items with initial target plosives are slightly higher with respect to those with medial target plosives. Still, within the category of word-initial plosives, words with lower logSubtlex values seem to be affected more than those with higher values. Furthermore, there is a considerable amount of variability not only for younger but also for older speakers in the words with significantly different VOT values between age groups.

This brings us back to the above mentioned question of whether this sound change in progress is internally or externally driven. Even though the change is clearly not of an abrupt manner, which would be typical for an externally driven sound change, we cannot conclude that the change is internal. The evidence that we deal with lexical diffusion seems to be the stronger argument here, leading us to the conclusion that this sound change in progress is, at least in part, due to the contact with GSG.

The issue then relates to the process of lexical diffusion itself. On the one hand, it must be emphasized that one cannot rule out completely that a sound change is both internally and externally driven. On the other hand, there is an increasing amount of evidence that the specific progress of dialect levelling is not necessarily abrupt but can also be of a gradual manner [16, 17, 18]. Based on the results of the current study, we conclude, at least for now, that the latter is also the case for the increasing aspiration in Zurich German, which allows for the possibility of a sound change being both externally driven and continuous.

4.2. Outlook

The results of the current study provide evidence on the specific sound change in progress taking place in Alemannic dialects of German-speaking Switzerland on the one hand and on the general phenomenon of sound change, specifically of the lexical diffusion type, on the other hand.

Still, there remain some important issues that need to be addressed in future research. Firstly, perceptual data is needed in order to better understand the sound change in progress in Swiss German dialects. Secondly, not only VOT but also its relationship to closure duration should be investigated. Lastly, a larger amount of words containing fortis plosives is needed both to confirm that mostly loanwords are produced with aspiration and to gather more information on how word frequency plays a role in this sound change in progress.

5. ACKNOWLEDGEMENTS

This study was funded by the Swiss National Science Foundation (SNSF), grant Nr. 190005.

6. REFERENCES

[1] Kraehenmann, A. 2001. Swiss German Stops: Geminates All over the Word. *Phonology* 18(1), 109–145.

[2] Kraehenmann, A., Lahiri, A. 2008. Duration Differences in the Articulation and Acoustics of Swiss German Word-Initial Geminate and Singleton Stops. *J. Acoust. Soc. Am.* 123(6), 4446–4455.

[3] Fleischer, J., Schmid, S. 2006. Zurich German. Journal of the International Phonetic Association 36(2), 243–253.

[4] Willi, U. 1996. *Die segmentale Dauer als phonetischer Parameter von 'fortis' und 'lenis' bei Plosiven im Zürichdeutschen: Eine akustische und perzeptorische Untersuchung*. Franz Steiner.

[5] Schifferle, H. 2010. Zunehmende Behauchung. Aspirierte Plosive im modernen Schweizerdeutsch. In: H. Christen et al. (eds.) *Alemannische Dialektologie: Wege in die Zukunft*. Steiner, 43–55.

[6] Ladd, R., Schmid, S. 2018. Obstruent Voicing Effects on F0, but without Voicing: Phonetic Correlates of Swiss German Lenis, Fortis, and Aspirated Stops. *Journal of Phonetics* 71, 229–248.

[7] Labov, W. 1994. *Principles of language change: Internal factors*. Blackwell.

[8] Bailey, G., Wikle, T., Tillery, J., Sand, L. 1991. The apparent time construct. *Language Variation and Change* 3, 241–264.

[9] Beckman, M. E., Li, F., Kong, E. J., Edwards, J. 2014. Aligning the timelines of phonological acquisition and change. *Laboratory Phonology* 5(1), 151–194.

[10] Lisker, L., Abramson, A. S. 1964. A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. *WORD* 20(3), 384-422.
[11] Jessen, M. 1998. *Phonetics and phonology of tense and lax obstruents in German*. John Benjamins Publishing.
[12] Beckman, J., Jessen, M., Ringen, C. 2013. Empirical evidence for laryngeal features: Aspirating vs. true voice languages. *Journal of linguistics* 49(2), 259-284.
[13] Luef, E. M. 2020. Development of voice onset time in an ongoing phonetic differentiation in Austrian German plosives: Reversing a near-merger. *Zeitschrift für Sprachwissenschaft* 39(1), 79–101.
[14] Kleber, F. 2018. VOT or quantity: What matters more for the voicing contrast in German regional varieties? Results from apparent-time analyses. *Journal of Phonetics* 71, 468–486.
[15] Labov, W. 2007. Transmission and diffusion. *Language* 83(2), 344–387.
[16] Harrington, J., Kleber, F., Reubold, U. 2012. The production and perception of coarticulation in two types of sound changes in progress. In: S. Fuchs, M. Weirich, D. Pape, & P. Perrier (eds.) *Speech planning and Dynamics*. Peter Lang, 39–62.
[17] Bukmaier, V., Harrington, J., Kleber, F. (2014). An analysis of post-vocalic /s-ʃ/ neutralization in Augsburg German: Evidence for a gradient sound change. *Frontiers in Psychology* 5. https://doi.org/10.3389/fpsyg.2014.00828
[18] Jochim, M., Kleber, F. 2022. Fast-Speech-Induced Hypoarticulation Does Not Considerably Affect the Diachronic Reversal of Complementary Length in Central Bavarian. *Language and Speech*. https://doi.org/10.1177/00238309221127641
[19] Draxler, C., Jänsch, K. 2004. SpeechRecorder – a Universal Platform Independent Multi-Channel Audio Recording Software. *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)*, Lisbon, Portugal. European Language Resources Association (ELRA). http://www.lrecconf.org/proceedings/lrec2004/pdf/242.pdf
[20] Schiel, F. 1999. Automatic Phonetic Transcription of Non-Prompted Speech. *Proc. of the 14th ICPhS*. San Francisco, 607–610.
[21] Winkelmann, R., Raess, G. 2014. *Introducing a Web Application for Labeling, Visualizing Speech and Correcting Derived Speech Signals*.
[22] Winkelmann, R., Jaensch, K., Cassidy, S., Harrington, J. 2021. *emuR: Main Package of the EMU Speech Database Management System*.
[23] RStudio Team. 2020. *RStudio: Integrated Development for R. RStudio*. PBC, Boston.
[24] Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., Böhl, A. 2011. The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology* 58(5), 412–424.
[25] Bates, D., Mächler, M., Bolker, B., Walker, S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1), 1–48.
[26] Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B. 2017. lmerTest Package: Tests in Linear Mixed Effects Models, *Journal of Statistical Software* 82(13), 1–26.
[27] Fox, J., Weisberg, S. 2018. *An R companion to applied regression*. Thousand Oaks, CA: Sage publications.
[28] Lenth, R. V. 2021. *emmeans: Estimated Marginal Means, aka Least-Squares Means*.