Revisiting Pfalz's law for two Viennese varieties: on speaker group differences in the implementation of vowel+stop sequences

Felicitas Kleber – Nicola Klingler – Markus Jochim – Michael Pucher – Stephan Schmid – Urban Zihlmann

Abstract. In this paper we report on first results from a combined speech production and speech perception experiment conducted within one of Sylvia Moosmüller's latest research projects – a joint DFG/FWF/SNF-funded so-called D-A-CH project on the synchronic implementation of phonemic vowel and post-vocalic consonant quantity in southern German varieties and potential diachronic changes within these quantity contrasts. While the project investigates a total of six varieties, two each from Austria, Germany, and Switzerland, the focus of the present paper is on the two Viennese varieties – the Viennese standard variety and the East Central Bavarian Viennese dialect – which are compared to two varieties from Germany. The project builds among others upon a previous study by Moosmüller and Brandstätter (2014) that suggested the presence of long vowels before (long) fortis stops in the phonological systems of the two varieties (although differently implemented) despite the prevailing assumption that this combination is illegal in Central Bavarian varieties and merged either with long vowel + lenis stop or with short vowel + fortis stop sequences. The paper specifically extends this previous study by investigating data from two age groups within each variety. Acoustic measurements of vowels and post-vocalic stops in the speech materials of the newly collected speech data indeed suggest the emergence of a third category of long vowel + fortis stop sequences. These sequences are acoustically clearly separated from long vowel + lenis stop sequences, on the one hand, and from short vowel + fortis stop combinations, on the other, in all four Viennese groups who implement the contrast differently from German speakers of the German standard variety and a West Central Bavarian dialect. While there was no evidence of a more standard like performance in younger speakers, the separation was less pronounced in Viennese dialect than in Viennese standard speakers, suggesting the diachronically stable existence of dialectal traces of the Bavarian quantity system particularly in the Viennese dialect and only to a lesser extent in the Viennese standard speakers. All four Austrian speaker groups, however, did not differ from the German standard group in the perceptual categorization of an acoustic continuum from /hˈaːgən/ to /hˈakən/ encompassing /hˈaːkən/.

PREAMBLE

Our present understanding of the phonetics and phonology of Austrian German varieties has been substantially informed by the many contributions by Sylvia Moosmüller (Moosmüller, 1984, 1987, 1996, 2015, 2016;

Moosmüller & Ringen, 2004; Moosmüller & Scheutz, 2013; Moosmüller et al., 2015), from her 1984 dissertation on Viennese German to the 2015 illustration of Standard Austrian German in the 45th volume of the *Journal of the International Phonetic Association*. They provided the empirical basis for sociophonetic analyses, the re-evaluation of phonological phenomena in Bavarian and the phonetic classification of Standard Austrian German as a major language variety. One of her latest research projects was concerned with a large-scale cross-linguistic investigation of the development of segmental quantity in Bavarian varieties and southern German regional standard varieties, addressed in a so-called DFG/FWF/SNF-funded D-A-CH project, a joint collaboration between the Viennese Acoustics Research Institute at the Austrian Academy of Sciences, the Institute of Phonetics and Speech Processing at the LMU Munich, and the Phonetics Laboratory at the University of Zurich. The project draws, among others, on production results presented in Moosmüller and Brandstätter (2014), suggesting a prosodic change in Viennese varieties regarding segmental timing (cf. below). This article presents both production and perception results arising from the follow-up studies conducted within the D-A-CH project which Sylvia Moosmüller so sadly did not live to see. We dedicate this article to her.

INTRODUCTION

In 1913 Anton Pfalz first described for the East Central Bavarian (hereafter ECB) dialect spoken in the Marchfeld area east of Vienna the interdependency between the length of a syllable-final consonant and the length of the preceding stressed vowel: long vowels can only precede lenis obstruents (which are phonetically shorter) and short vowels only occur before (phonetically) longer fortis stops (Pfalz, 1913). This phenomenon of complementary length, often referred to as Pfalz's law, has been described as a more general characteristic of the larger Bavarian dialect area, including West Central Bavarian (hereafter WCB, e.g. Bannert, 1976; Hinderling, 1980), but also Southern Bavarian (see e. g. Wiesinger, 1990) and North Bavarian (Rowley, 1990) varieties. According to this phenomenon , speakers of Bavarian varieties realize words that contain long vowels before fortis stops (e.g. *beten*, /bˈeːtən/, 'to pray') and short vowels before lenis stops (e.g. in words borrowed from standard German such as *Widder* /vˈɪdɐ/), respectively, in free variation either as e.g. [bˈetən], [bˈeːdən] and [vˈɪtɐ], or [vˈɪːdɐ]. The restriction to long+short and short+long combinations, respectively, was for a long time conside-

red uncontroversial; the focus of the discussion was instead on the phonemic status of vowel length and consonant strength (see Seiler, 2005 for an overview). While most researchers argued in favour of a phonemic fortis/lenis contrast and allophonic vowel length (e.g. Hinderling, 1980; Wiesinger, 1990) that is predictable by the underlying consonant strength category of the following consonant, Bannert (1976) argued that prosodic quantity determines whether the vowel or the consonant is long and the adjacent segment then predictably short. This issue shall not be discussed here further; instead, the focus will once again be on the nature of segmental quantity and the validity of Pfalz's law in two varieties spoken in Vienna, including the ECB Viennese dialect, and a WCB variety spoken in Germany.

The motivation for the present study comes from a previous, partly apparent-time[1] study by Moosmüller and Brandstätter (2014) where they showed on the basis of acoustic measurements that long vowel+long fortis stop combinations appear possible. More precisely, they measured the duration of the stressed vowel and medial stop in words like *Siebe* ('sifters', /sˈiːb̥ə/, where underlyingly V_1 is a long vowel and C_2 a short lenis stop, henceforth V:C type sequence), *Siepe* (a proper name, /sˈiːpə/, where underlyingly V_1 is again a long vowel and C_2 a long unaspirated fortis stop, henceforth V:C:), and *Sippe* ('clan', /sˈipə/, where underlyingly V_1 is a short vowel[2] and C_2 again a long unaspirated fortis stop, henceforth VC:) in nine older speakers of the ECB Viennese dialect and 13 older and 13 younger Viennese speakers of the phonologically ECB-based Austrian standard variety (cf. Moosmüller, 1996) and found that sequences of the type /V:C:/, considered illegal in Bavarian, were not merged – in terms of identical duration patterns – with one of the legal sequences but realized with a duration pattern that lay in between that of /V:C/ and /VC:/ type sequences (cf. Fig 6, p. 87). More specifically, the proportion of the vowel duration in a vowel+stop sequence (henceforth V/(V+C) ratio; where the mostly unaspirated Viennese stop corresponds approximately to the stop's

[1] The apparent-time approach builds upon the assumption that the speech of a comparatively homogenous group of speakers (homogenous particularly in age and variety) reflects the stage of a variety at which they have acquired it prior to the age of dialect acquisition. Any deviation in the speech of a later generation of speakers of the same speech community is consequently indicative of a diachronic linguistic change (cf. Labov, 1994: 45–54).

[2] In the Viennese varieties of German, front, high vowels differ largely in vowel quantity and much less in quality (Schmid & Moosmüller, 2017) than in standard German as spoken particularly in the north of Germany.

closure phase) was longest for /V:C/ and shortest for /VC:/ type sequences reflecting the complementary vowel length and stop length pattern described above in words with these combinations. The intermediate proportional vowel duration for /V:C:/ type sequences mirrors, in turn, the combination of two equally long segments, in this case two phonemically long segments. The acoustic divide into three characteristic V/(V+C) ratio patterns was, however, less pronounced in older speakers of the Viennese dialect than in Viennese standard speakers. Moreover, older more than younger speakers of the Viennese standard variety varied absolute durations as a function of adjacent segments. Taken together, the data then suggest that the restrictions of Pfalz's law on the temporal organization of vowel+stop sequences must have become loosened over time – first in the standard variety, where this prosodic change of segmental quantity appears to be still in progress, and subsequently in the dialect. A more direct investigation of (1) age group differences within the Viennese dialect group and (2) potential differences in the phonetic implementation of vowel and stop length contrasts between speakers of the Austrian standard variety, on the one hand, and speakers of the German standard variety, on the other, remains to be done. The first aim was therefore to replicate the study described in Moosmüller and Brandstätter (2014) with different materials and speakers, by extending the apparent-time approach to dialect speakers and by a comparison of the temporal patterns found in ECB with those in WCB and standard German.

Further evidence for an eventual break-up of Pfalz's law in Central Bavarian (presumably due to dialect levelling) comes from an apparent-time study on WCB (Kleber, 2017) which showed that when speaking standard German with a Bavarian accent older more than younger speakers of WCB realize standard German words with V:C: sequences (i.e. those considered illegal in Bavarian) in greater compliance with Pfalz's law in that they additionally vary stop duration. More specifically, older WCB speakers tended to realize words like /bˈiːtən/, 'to offer' either as [bˈiːdən] with a long vowel preceding a short lenis stop or as [bˈɪtən] with a short vowel preceding a longer fortis stop. Younger WCB speakers, on the other hand, realized these words more standard-like as /bˈɪtən/, which stands in clear contrast to trochaic words that contain /V:C/ and /VC:/ type sequences. Such a three-way contrast between /V:C/, V:C:/, and /VC:/ type sequences also emerged in the analysis of spontaneous speech data obtained from a different group of younger speakers from the WCB area (Kisler & Kleber, 2019). Interestingly, however, younger

ECB[3] speakers from Austria who were recorded for the same corpus (the *Deutsch Heute*-Corpus, cf. Brinckmann et al., 2008) did not show this three-way contrast but a tendency to merge /V:C:/ type sequences with /VC:/ type sequences (cf. Fig. 4 in Kisler & Kleber, 2019: 291). These findings suggest that Pfalz's law appears to be giving way to the three-way contrast found in standard German only in WCB but possibly not in ECB.[4] One fundamental difference between the studies by Moosmüller and Brandstätter (2014) and Kisler and Kleber (2019) that is most relevant in the present context is that the former study focused on Viennese varieties while the latter included primarily ECB speakers from other Austrian regions. The second aim of the project was therefore to investigate in more detail whether or not and if so to what extent younger ECB speakers from Vienna differ in their temporal organization of vowel+stop sequences from younger WCB speakers from Upper Bavaria when they are speaking the dialect (and not just regional varieties as in Kleber, 2017, and Kisler & Kleber, 2019).

The change in the production of the WCB regional variety of standard German is also manifested in WCB listeners' perception of standard German vowel+stop sequences. A follow-up perception experiment in Kleber (2017) showed that the same WCB speakers turned listeners perceived instances of standard German *Haken* (/hˈaːkn̩/, 'hook') more often as *hacken* (/hˈakn̩/, 'to chop') when they were part of the older group and when they had been asked to judge these instances in a two-alternative forced choice (AFC) task between *hacken* and *Hagen* (/hˈaːgn̩/, a proper name).[5] Younger WCB listeners, on the other hand, judged the *Haken*-like

[3] While the study only included speakers from areas where Central Bavarian dialects are spoken, the study erroneously allocated some speakers from Western Austria to the ECB group who, from a dialectal point of view, should have been part of the WCB group. On the assumption that WCB is less conservative than ECB this would imply that the difference between ECB and WCB should have been even more pronounced if these speakers had been allocated correctly. Alternatively, the group differences may have been enlarged, namely if the Austrian WCB speakers were more conservative than the German WCB and the Austrian ECB speakers.

[4] Such linguistic meaningful interpretations of the analysis in Kisler and Kleber (2019) need to be treated with caution given that the speech materials were not designed for a duration analysis of vowel + stop sequences in trochaic words and thus were not well balanced. The focus of the paper was on automatic segmentation of large databases and the duration analysis of vowel + stop sequences in polysyllabic (but not necessarily trochaic) words was a means to evaluate the reliability of automatically segmented data.

[5] No such group difference emerged in a control 3AFC task where *Haken* was among the response options.

stimuli equally often as *hacken* or *Hagen* in the two-AFC task (henceforth 2AFC), i.e. as opposed to the older group they showed no such bias towards one of the two Bavarian legal patterns. This study was replicated for the three regional standard varieties spoken in Vienna/Austria, Munich/Southern Germany, and Zurich/Switzerland by Klingler et al. (2019) using the same minimal set and the same procedure of only manipulating the proportional vowel duration but with more stimuli. They found that speakers of the standard varieties of Vienna and Munich clustered together in rating *Haken*-like stimuli as *hacken* or *Hagen* at chance level while speakers from Zurich almost always judged the same stimuli as *hacken.* Our third aim in the present study was to test the same continuum used in Klingler et al. (2019, see below for details regarding stimulus creation) with dialect and standard speakers from Vienna, this time taking age into account.

SPECIFIC AIMS AND HYPOTHESES

A specific aim of the present paper is to present first of all an overview of the Viennese production data obtained within the project. It extends this previous study by including (1) production data also from younger speakers of the Viennese dialect and (2) perception data to investigate the perceptual status of these combinations which were considered illegal in East Bavarian for a long time. The present study is preliminary in that most of the analyses are based on a subset of data for which the automatically set segment boundaries have already been checked. The specific hypotheses for speech production were that all groups produce all three combinations by means of acoustic duration differences but that dialect speakers do so less than standard speakers. If this was a sound change in progress then younger speakers should be – commensurate with the apparent time construct (cf. footnote 1) – in between older speakers, on the one hand, and speakers of the more standard-like variety, on the other.

If Viennese speakers of the dialect differed from Viennese speakers of the Austrian standard variety and if older and younger speakers differed then we would predict that any such differences also materialized in perceptual judgments to variants of the various vowel+stop sequences in the form of a bias towards the more dialectal variant. In light of the sound change hypothesis we also need to consider the possibility of a mismatch between the perception and production results given some evidence in favour of a misalignment of the two modalities during a sound change in progress (see Kleber et al., 2012, for discussion). Whether the more

conservative variant dominates in perception or in production appears to depend on he progression of a change which might take several decades (Pinget, 2015).

METHODS

Participants

A total of 45 speakers for whom data from a speech production task and a follow-up perception experiment were available were selected for the present analyses. 25 of these participants were from Vienna and 20 from Upper Bavaria in Germany. The Viennese group comprised again twelve speakers of the ECB Viennese Dialect (henceforth VD) and thirteen speakers of the Viennese standard variety (henceforth VS). Six VD speakers were allocated to a younger VD group (mean age = 24.7, sd = 3.6, three female) and seven VD speakers to the older VD group (mean age = 55.7, sd = 4.0; five female). Among the VS speakers were six older (mean age = 59.2, sd = 5.5; four female) and seven younger (mean age = 24.9, sd = 3.8; three female) speakers, who were then likewise grouped into an older and a younger VS group. The German group comprised seven younger speakers of the standard variety spoken in Munich (henceforth MS, mean age = 23.9, sd = 2.5, four female) and twelve younger speakers of WCB (mean age = 25, sd = 2.7, six female). Since almost no dialect is spoken in the city of Munich, WCB speakers were recruited from the rural surrounds of Munich.

Production

To analyse the phonetic implementation of various vowel+stop sequences in production we created a new corpus with target words containing among others the three vowel+stop type sequences V:C, V:C:, and VC:. All target words were then embedded in the narrow focus position of appropriate carrier phrases. These carrier phrases varied depending on the speaker group: speakers of the standard variety were presented with sentences in standard orthography; dialect speakers were presented with sentences written in an orthographic proxy of the dialect (since no standardized orthography is available here). To minimize influences of orthography, speakers did not read out loud the respective sentences presented to them one after another and in random order on a screen directly, but were asked to read each sentence silently and to repeat it as soon as

the sentence disappeared from the screen and within a pre-set speaker-individual time frame.

The corpus contains a total of 21 such real words which have been produced five times each at two different speech rates (a normal and a speaker-specific faster rate). The recordings from the faster rate condition were, however, lumped together with those of the normal rate condition in the present study, since here speech rate was excluded as a predictor variable. All recordings have been automatically segmented among others into words and phonemes using WebMAUS (Kisler et al., 2017) and stored as an EmuDB (Winkelmann et al., 2017). These automatic segment boundaries were (and currently still are) checked and whenever necessary hand corrected, complying with the following criteria: the boundaries that mark the on- and offset of a stressed vowel were placed at the first and last, respectively, clearly visible glottal pulse. A vowel's offset concomitantly marks the onset of the postvocalic stop. The stop's offset was again set to the first clearly visible glottal pulse of the following voiced segment (usually a vowel). All stops were further segmented into a closure and aspiration phase, the latter beginning with the burst. Regardless of how short the aspiration phase was (e.g. in tokens with lenis stops) it was marked unless it was not detectable (usually cases with no or a very weak burst). For a set of 3 sequence types × 2 words the segment boundary check has been completed; this set is referred to as corrected data set. This set contained the words *Hagen*, *Haken*, *hacken* (see above for transcriptions and English glosses), *wieder* (/vˈiːdɐ/, 'again'), *Bieter* (/bˈiːtɐ/, 'bidder'), and *bitter* (/bˈɪtɐ/, 'bitter'). Tab. 1 summarizes the data points available for analysis. The item *Hagen* is missing most tokens because in this context /g/ was often, particularly by dialect speakers, assimilated to the following /n/.

Table 1: Distribution of analysed tokens per speaker group.

Group	corrected data		uncorrected data		
	alveolar	velar	labial	alveolar	velar
younger MS speakers	239	234	*not analysed*		
older VS speakers	150	150	300	656	286
younger VS speakers	209	198	349	765	323
older VD speakers	210	167	239	547	233
younger VD speakers	159	130	250	546	212
younger WCB speakers	359	313	*not analysed*		

We then retrieved acoustic vowel and closure duration from the EmuDB in R (version 3.6.0 R Core Team, 2019) using RStudio environment (version 1.2.1335, RStudio Team, 2018) and the emuR package (Winkelmann et al., 2017) and calculated the V/(V+C) ratio where V corresponds to the vowel duration and C to the closure phase duration. The dependent variable V/(V+C) (or a derived measure) was subjected to the statistical analyses described below in the RESULTS section.

PERCEPTION

The same speakers from whom production data were obtained also participated in a speech perception experiment. The perception experiments always took place after the production experiment – sometimes right after the recording session, sometimes with a delay of several days. In the present experiment we will only present the results from all four Viennese groups and compare them to the results of the MS group given that they were listening to the exact same resynthesized versions of a natural *Hagen* production of a middle-aged male model talker of standard German who is Austrian by nationality but acquired the German standard variety in the north of Germany.

More precisely, following the method in Kleber (2017),[6] the model talker produced first ten repetitions of the target words *Hagen* and *hacken* embedded in the carrier sentence *Maria hat* target word *gesagt.* (lit. 'Maria has target word said'). We then measured vowel and stop duration in the two target words and calculated the V/(V+C) ratio. To create the *Hagen–hacken* continuum that naturally encompasses *Haken*, we selected from the ten *Hagen* repetitions the token with the V/(V+C) value closest to the mean V/(V+C) ratio across all ten repetitions. This token was labelled stimulus 1 and corresponds to one endpoint of the continuum. We then shortened the vowel duration successively 16 times while simultaneously lengthening the closure phase of the velar stop using Praat's Manipulate and Get resynthesis (overlap and add) functions (Boersma & Weenik, 2017; cf. Tab. 2 for subsequently measured vowel and closure duration in the resulting stimuli). Thus, the duration of the vowel+stop sequence was identical across all stimuli of the continuum; only the V/(V+C) ratio varied systematically from stimulus to stimulus. The ratio of the final stimulus 17 corresponded to the mean V/(V+C) ratio across

[6] The model talker was the same as in Kleber (2017) but the measurements and stimuli were based on newly obtained recordings.

all ten *hacken*-repetitions available from the model talker. Prior to the experiment the stimuli were auditorily cross-checked by the authors to validate a change in perception along the continuum from an unambiguous *Hagen*-percept to a *Haken*-percept somewhere in the middle of the continuum to an unambiguous *hacken*-percept.

Table 2: Vowel (V) and stop closure (C) duration [ms] for the stimuli of the Hagen –hacken continuum.

	Stimulus																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
V	160	154	148	142	136	130	124	118	112	106	100	94	88	82	76	70	64
C	55	61	67	73	79	85	91	97	103	109	115	121	118	124	130	136	142

Listeners rated the 17 stimuli in a 2AFC and subsequently in a control 3AFC task both conducted in Praat. They listened once to a stimulus presented over headphones which was in the 2AFC task immediately to be judged as sounding either more like *hacken* or more like *Hagen* by clicking on the respective word which was presented on a computer screen. In the control task *Haken* was added as a third response option. Each stimulus was presented ten times and in randomized order. The focus will be on the results from the 2AFC task. The responses from this task were the dependent variable in a generalized mixed effects model with stimulus and group as fixed factors.

RESULTS

Production

Measurements of the V/(V + C) ratio in the subset of the data where segment boundaries have already been checked and corrected wherever necessary indicate that all speakers differentiate between three combinations by means of characteristic V/(V + C) ratios (cf. Fig. 1). These observations are supported by mixed effects models with ratio as the dependent variable and word and speaker as random factors showing significant effects for both predictor variables, i.e. sequence type ($F[2,35] = 925.5$, $p < 0.001$) and group ($F[5,36] = 4.3$, $p < 0.01$) as well as a significant interaction between them ($F[10,35] = 4.1$, $p < 0.001$). The most striking group differences were firstly the difference within the V:C: type (Fig. 1 middle panel) between younger WCB and younger MS

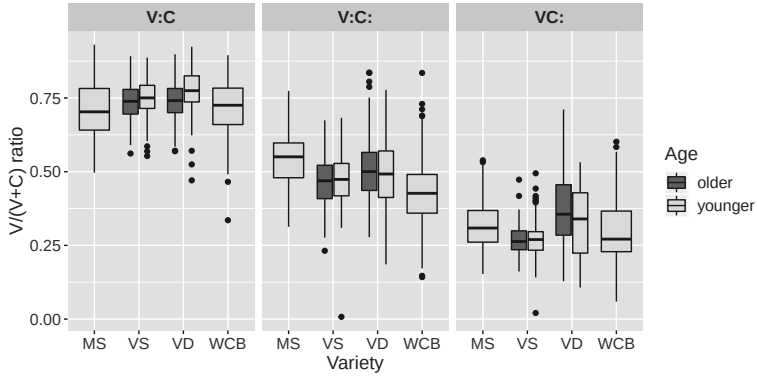

Figure 1: V/(V+C) ratio as a function of vowel+stop sequence (left panel: V:C, middle panel: V:C:, right panel: VC:) separately for older (dark grey) and younger (light grey) speakers of the Munich (MS) and Viennese (VS) standard varieties, respectively, and speakers of the WCB and Viennese dialect (VD).

speakers[7] with all four Viennese groups taking up an intermediate position and secondly, the difference between the two Viennese varieties regarding the VC: type sequence. Two separate models fitted (1) to the Viennese V:C: category and (2) to the Viennese VC: category revealed – commensurate with Fig. 1 – no significant group differences (i.e. neither age nor variety) for the V:C: category and no significant age but a significant effect of variety in the VC: category ($F[1,22] = 14.5$, $p < 0.001$). Irrespective of a speaker's age, Viennese dialect speakers realized words like *hacken* with a greater vowel proportion, indicating some form of lenition and as a result a less sharp contrast between V:C: and VC: type sequences.

Fig. 1 in fact suggests for all groups except the MS group a greater tendency for V:C: type sequences to cluster more with VC: sequences than with V:C sequences, which is, however, more pronounced in dialect speakers. This then suggests a general bias towards adjusting vowel duration and not that of the stop, at least for the subset of words selected for the present analysis.

This pattern also emerges quite clearly in Fig. 2, which shows the distribution of V/(V+C) ratios as a function of sequence type in the automatically segmented and not yet segment boundary corrected data that

[7] A post-hoc test with Bonferroni correction revealed that only these two groups differed significantly ($p < 0.001$) in their V/(V+C) ratios of the V:C: type sequence.

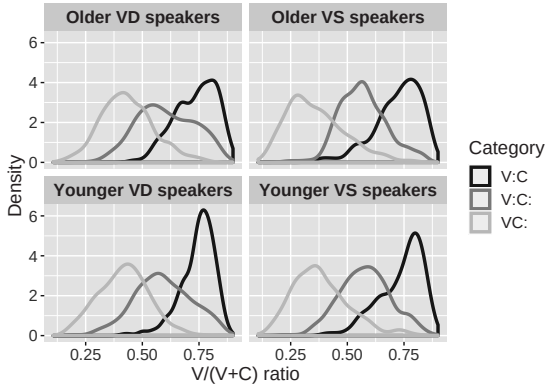

Figure 2: Density distribution of V/(V+C) ratios in all words of the uncorrected set as a function of sequence type (black = V:C, dark grey = V:C:, and light grey = V:C) separately for older (top panels) and younger (bottom panels) VD (left panels) and VS (right panels) speakers.

included multiple target words (cf. above Tab. 1, i.e. the data most likely include errors in measurement but see Kisler & Kleber, 2019, for the reliability of such data). In this MAUS-based analysis C corresponds to the entire stop including the aspiration phase. The amount of overlap between V:C: type sequences with either one of the legal sequences is obviously greater for Viennese dialect compared to Viennese standard speakers, and in particular with the VC: type sequence. Most interestingly, older (but not younger) VD speakers' distribution of V/(V+C) ratios for V:C: type sequences show a right-skewed distribution, indicating a greater tendency for V:C: type sequences to overlap with VC: type sequences and not with V:C type sequences.

The purpose of the last analysis of the corrected production data was to further quantify the degree of acoustic difference between V:C: type and VC: type sequences. To this end, we excluded all V:C type sequences from the analysis and calculated for each speaker and each token the difference between the V/(V+C) ratio of each repetition of a word (e.g. *Haken*) and the mean V/(V+C) ratio of this minimal pair's counterpart word (i.e. in this case *hacken*) aggregated across all repetitions.[8] This left us with forty values per speaker. We will refer to these values as ratio difference. Visual inspection and statistical analyses revealed no signifi-

[8] Two VD speakers had to be excluded from this analysis because no repetition was available for one word (these had been excluded because the stop was missing, cf. methods above).

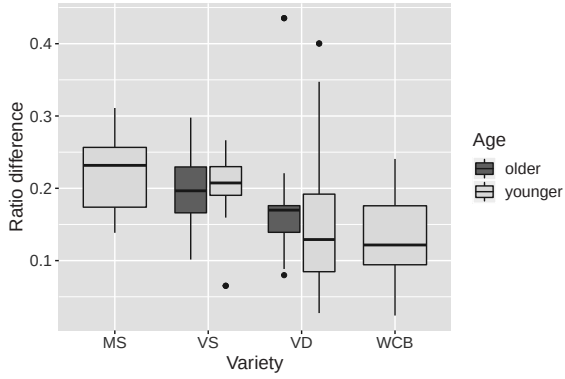

Figure 3: Aggregated ratio differences between V:C: type and VC: type sequences for the corrected data set separately for the Munich (MS) and Viennese standard (VS) varieties, the Viennese (VD) and the WCB dialect as well as older (dark grey) and younger (light grey) speakers.

cant influence of word or sequence type (i.e. whether V:C: tokens were subtracted from the mean of V:C tokens and vice versa) on the ratio difference. We therefore aggregated the ratio difference values across word and sequence type for the final analysis. Fig. 3 shows a clear separation between standard (MS and VS) and dialect (VD and WCB) speakers. The latter groups' smaller difference values indeed suggest a greater amount of overlap between V:C: type and VC: type sequences than in the two standard groups. These observations were again supported by statistical analysis: a linear mixed effect model with the ratio difference as the dependent variable, order of subtraction and group as fixed factors and word and speaker as random factors showed a significant effect for group ($F[1,37] = 5.6$, $p < 0.05$) but none for order of subtraction. Subsequent models with the same variables above but fitted to VD and the VS group, respectively, again showed no effects for age.

The most important findings from the analyses of the production data are that (1) Viennese speakers of the dialect do not differ from Viennese speakers of the standard in terms of a clearly stable three-way contrast (in line with Moosmüller & Brandstätter, 2014), but (2) in the implementation of fortis stops after short vowels which in turn leads to a greater tendency for V:C: type and VC: type sequences to overlap. The non-existing age differences (3) do not support the idea of a sound change in progress in Viennese. The dialect-standard divide regarding the Bavarian illegal pattern was (4) most pronounced for the two German groups with younger WCB speakers showing a V/(V+C) ratio well below those of

same-aged younger standard speakers, suggesting that WCB speakers still shorten underlying long vowels before fortis stops to a greater extent (i.e. in accordance with the dialectal background). The Viennese dialect-standard divide became apparent only when taking into account the difference between the V:C: type and VC: type sequences.

Perception

The results for the control 3AFC task showed that – commensurate with production – listeners from all speaker groups perceived the stimuli from the middle of the continuum as *Haken.*

The aggregated response curves (aggregated across the speakers within a group) to stimuli judged in the 2AFC task in Fig. 4 likewise do not differ greatly between speaker groups and show a similar progression. The five listener groups thus did not differ substantially in their judgment of the stimuli from the *Hagen–hacken* continuum (encompassing *Haken*) in the 2AFC task. A generalized mixed effects model with response as dependent variable, stimulus and group as contrast coded and centred on zero fixed factors (i.e. older VD listeners = −2, younger VD listeners = −1, older VS listeners = 0, younger VS listeners = 1, younger MS listeners = 2, stimulus 1 = −8, stimulus 2 = −7, … stimulus 9 = 0, stimulus 10 = 1, … stimulus 17 = 8 so that the fixed effects can be interpreted as in an ANOVA, cf. Eger & Reinisch, 2019), and listener as random factor corroborates this visually apparent result as it revealed only a significant effect for stimulus ($z = 15.3$, $p < 0.001$) but none for group.

DISCUSSION AND CONCLUSION

The aims of the paper were to investigate cross-linguistic and apparent time differences in the implementation of quantity contrasts in three types of vowel+stop sequences between different-aged speakers of different varieties spoken in Vienna and to compare their data to that of German younger speakers of the Munich standard variety and WCB, respectively. The following findings arose from the analyses above. First, irrespective of age or variety all 24 speakers of Viennese varieties realized words containing the three vowel+stop sequence types by means of three characteristic and unambiguous V/(V+C) ratios – similarly to speakers of standard German as spoken in the north of Germany (Kohler, 1979). This finding indicates that the V:C: type sequence considered illegal in previous descriptions of Central Bavarian varieties is legal both in the system

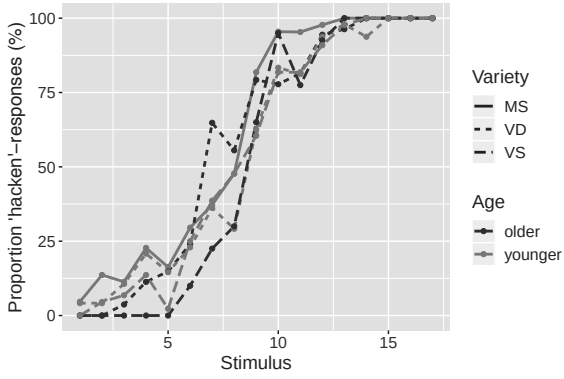

Figure 4: Response curves to the Hagen – hacken continuum aggregated across listeners' responses within one group separately for older (black) and younger (light grey) Viennese dialect (VD) and standard (VS) speakers as well as younger speakers of the Munich standard (MS) variety.

of the Viennese standard variety and that of the Viennese ECB variety. It is therefore in line with the findings described in Moosmüller and Brandstätter (2014). The result further suggests that the observation of a merger between V:C: type and VC: type sequences in Kisler and Kleber (2019) must have been due to the vast majority of ECB speakers from other Austrian regions than Vienna.

Second, both Viennese varieties showed, on the one hand, similar ratio values in V:C: type sequences that were, on the other hand, in between those of German standard speakers (showing the highest values) and those of younger WCB speakers (showing the lowest values). This again suggests (a) a greater standard-dialect divide for the two German varieties than for the two Viennese varieties and (b) that dialectal traces of Pfalz's law are still visible in the temporal organization of V:C: type sequences in all groups of Viennese speakers (and WCB speakers for that matter). Only the two German varieties differed significantly in the V/(V+C) ratio of V:C: type sequences, with MS speakers showing values similar to those of northern German speakers (cf. Kohler, 1979) and WCB speakers showing the greatest tendency towards vowel shortening before fortis stops (i.e., the /aː/ in *Haken*, for example, became more [a]-like though not as short as in *hacken*). Although vowel shortening before fortis stops is not per se predicted for Central Bavarian varieties (given the free variation), the lower V/(V+C) ratios are in line with one of two possibilities of the Central Bavarian phonology, namely that of vowel shortening before fortis stops (instead of stop lenition

after long vowels as in e.g. /fˈɒd̥ɐ/, ‘father’, the other possibility).[9] This pattern – which is in accordance with the observations in Kisler and Kleber (2019) – indicates a bias towards adjusting vowel duration and not that of the stop. Taken together this can then be taken as tentative support in favour of phonemic stop length (cf. e.g. Wiesinger, 1990; Seiler, 2005). In any case, the intermediate Viennese values can be interpreted as representing remnants of this more dialectal ECB pattern as described on the basis of auditory judgements (e.g. Pfalz, 1913; Wiesinger, 1990) and acoustic analyses (Kisler & Kleber, 2019).

Third, although the two Viennese varieties did not differ in the temporal implementation of V:C: type sequences, they did so in the degree of temporal contrast between V:C: type sequences, on the one hand, and VC: type sequences, on the other. Speakers of the Viennese dialect realized VC: type sequences with higher V/(V+C) ratios than speakers of the Viennese standard variety, thereby minimizing the difference between ratio values of V:C: type and VC: type sequences in the former compared to the latter group. That is, dialectal differences between Viennese speakers of the standard variety and the dialect, respectively emerged in the analyses, but not in terms of the implementation of V:C: sequences but in that of VC: sequences and – as a result of this – in the degree to which the contrast is pronounced. Our findings for Viennese variety differences regarding the V/(V+C) ratios in particular in VC: type sequences differ from those described in Moosmüller and Brandstätter (2014), where group difference were more apparent in V:C: type than in VC: type sequences. Under the assumption of the possibility of lexically dependent vowel shortening and stop lenition, this between-study difference may then be related to the different words investigated in the present and the 2014 study, respectively. Further, although Moosmüller and Brandstätter did not directly investigate the degree to which the contrast between V:C: type and VC: type sequences was implemented by means of the V/(V+C) ratio, a visual inspection of Fig. 6 in their article suggests a less pronounced contrast in the VD speakers compared to the Viennese standard speakers (Moosmüller & Brandstätter, 2014: 91).

Our findings from this cross-linguistic study furthermore showed that, as opposed to speakers of the Viennese standard variety, speakers of the ECB Viennese dialect did not differ from WCB speakers in the degree of temporal contrast between the two sequence types with fortis stops. This then fourthly supports retrospectively the more general validity of Pfalz’s

[9] Note, however, the possibility of vowel shortening being lexically dependent.

law in various Central Bavarian varieties that is still present in the form of traces of Pfalz's law.

The perception test showed that all groups divided the *Hagen-hacken* continuum into two perceptually clearly distinct categories, but apart from that revealed no group differences either between varieties or between generations. This suggests that Viennese listeners' perceptual category boundaries along such a V/(V+C) ratio continuum simply fall in the same range of V/(V+C) ratio values as that of MS listeners. These results extend the findings for the Viennese standard variety reported in Klingler et al. (2019) in that they show that Viennese dialect speakers do not differ from Viennese standard speakers (and Munich standard speakers for that matter) in their perceptual categorization of the three-way contrast between V:C type, V:C: type and VC: type sequences, although they differ in production. For the following reasons we will not discuss this mismatch between group differences in production, on the one hand, and no such differences in perception, on the other, in terms of implications for a sound change in progress (as alluded to in the SPECIFIC AIMS AND HYPOTHESES section above). First, the production results did not support the idea of a sound change in progress in the form of age differences within groups (discussed in more detail below). Second, although previous findings for a very similar continuum and an identical task showed significant regional and age group differences in the 2AFC responses to a *Hagen-hacken* continuum (encompassing *Haken*) based on the same model talker's production (Kleber, 2017), the paradigm of applying a 2AFC task to a three-way contrast with the prediction of a greater bias towards a regional variant (e.g. a *hacken-p*ercept of a *Haken*-like stimulus) may have disguised potential perceptual group differences in the present study because the regional differences in production only emerged in the form of a diminished ratio contrast between V:C: type and VC: type sequences but not in the ratio values for V:C: type sequences. In the light of the present findings for speech production future studies should directly test the contrast between V:C: type and VC: type sequences given that differences between Viennese groups were greatest for this contrast.

Except for the trend towards a greater right-skewed distribution of V:C: type sequences in the uncorrected data of older VD speakers in Fig. 2, none of the analyses of both the production and the perception data supported the original idea that – based on the results in Moosmüller and Brandstätter (2014) and on the assumption of a general trend towards dialect levelling in German varieties (Wagener, 2002; Lameli, 2004; Bukmaier et al., 2014; Harrington et al., 2012) – younger Viennese dialect

speakers behaved more standard like than older Viennese dialect speakers. Thus, the results, lastly, do not indicate a sound change in progress regarding the temporal organization of vowel+stop sequences either in the Viennese dialect or the Viennese standard variety. A first interpretation of this finding would be that we are observing mere variety differences, with each variety representing a stable system of its own. Our finding does not contradict in principle the result in Moosmüller and Brandstätter (2014) as we only analysed the V/(V+C) ratio (for reasons to do with providing a combined measure of the quantity relations in vowel+stop sequences) and not the absolute durations, for which the previous study found age group differences. But the present study does not arrive at the same general conclusion that younger Viennese standard speakers behaved more like German speakers of the German standard variety as put forward in the previous analysis (see also Schmid & Moosmüller, 2017).

Another potential explanation for the non-emerging age effects in the Viennese varieties might be that the change from a two-way contrast to a three-way contrast, in which the strict pattern of complementary length in Pfalz's law has been suspended, is already complete. Such a view is based on the very general assumption that Pfalz's law must have applied to the varieties spoken in the city of Vienna at some point (as stated e.g. in Koekkoek, 1953). Belated acoustic support for this idea comes from the present study in the form of the intermediate V/(V+C) ratio values found in all Viennese speakers' realization of V:C: type sequences and the diminished V:C: –VC: contrast in the Viennese dialect. The latter observation may then be taken to suggest that the change may have started in the Viennese standard speakers, where it was still observable as apparent-time differences in absolute durations in Moosmüller and Brandstätter (2014). It remains to be seen whether Viennese dialect speakers will eventually realize the vowel length contrast before fortis stops like (Viennese) standard speakers.

Yet another explanation – that draws less on the observed temporal patterns – is that the general trend for dialect levelling currently observable in the studies mentioned above may not be so pronounced in Austria as it appears to be the case for Germany. With the exception of the studies by Moosmüller and Brandstätter (2014) and Schmid and Moosmüller (2017), where the observations for age group differences were indeed not very pronounced, all of the studies above are on regional varieties spoken in Germany. These in turn may not allow for generalizations to German varieties spoken outside Germany, although the prerequisite for dialect levelling, increasing contact with another variety (often the standard),

can be assumed for Austrian varieties as well, given increasing speaker mobility and the influence of the media, etc. Of course, it cannot be ruled out that Austrian varieties such as the Viennese dialect level out at other varieties than the German standard variety. But this is yet another question to be addressed in future studies, preferably in tandem with – as stated above – refined perception experiments that would then test more directly the relation between production and perception during dialect-levelling induced sound changes (Kleber et al., 2012; Pinget, 2015; Coetzee et al., 2018).

One of the main conclusions from this study therefore is that speakers of the ECB Viennese dialect do not show the same levelling tendencies as observed for speakers of WCB varieties (Kleber, 2017; Kisler & Kleber, 2019). This may be, of course, a mere regional difference: while Pfalz's law of complementary length in Bavarian vowel+stop sequences is still more pronounced in WCB speakers (despite the levelling tendencies mentioned above), it has already given way to a three-way system in Viennese varieties. Another likely and perhaps additional (i.e. not necessarily contradictory) explanatory factor for the observed greater standard-dialect divide in German compared to Austrian varieties lies in the very fundamental difference of all Austrian dialect speakers in the study stemming from the city of Vienna and all German dialect speakers stemming from the rural surrounds of Munich. That is, the greater standard-dialect divide may also be a consequence of an urban-rural divide present in the German but not the Austrian data. The study nevertheless very clearly shows the existence of long vowel + fortis stop sequences in Viennese varieties just as first described in the study by Moosmüller and Brandstätter (2014).

REFERENCES

Bannert, Robert. 1976. *Mittelbairische Phonologie auf akustischer und perzeptorischer Grundlage*. Ph.D. dissertation, Lund University.

Boersma, Paul and Weenik, David. 2017. Praat: Doing phonetics by computer [Computer program]. Version 6.0.24. Available online: http://www.fon.hum.uva.nl/praat/, last accessed 08.11.2019.

Brinckmann, Caren, Kleiner, Stefan, Knöbl, Ralf, and Berend, Nina. 2008. German today: An areally extensive corpus of spoken standard German. *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC), Marakech*, 3185–3191.

Bukmaier, Véronique, Harrington, Jonathan, and Kleber, Felicitas. 2014. An analysis of post-vocalic /s-ʃ/ neutralization in Augsburg German: Evidence for a gradient sound change. *Frontiers in Psychology* 5, Article 828.

Coetzee, Andries W., Beddor, Patrice Speeter, Shedden, Kerby, Styler, Will, and Wissing,

Daan. 2018. Plosive voicing in Afrikaans: Differential cue weighting and tonogenesis. *Journal of Phonetics* 66, 185–216.

Eger, Nikola A. and Reinisch, Eva. 2019. The impact of one's own voice and production skills on word recognition in a second language. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 45, 552–571.

Harrington, Jonathan, Kleber, Felicitas, and Reubold, Ulrich. 2013. The effect of prosodic weakening on the production and perception of trans-consonantal vowel coarticulation in German. *The Journal of the Acoustical Society of America* 134, 551–561.

Hinderling, Robert. 1980. Lenis und Fortis im Bairischen. Versuch einer morphophonemischen Interpretation. *Zeitschrift für Dialektologie und Linguistik* 47, 25–51.

Kisler, Thomas and Kleber, Felicitas. 2019. Zur Validität automatisch segmentierter Daten: Eine akustische Analyse der mittelbairischen Lenisierung im *Deutsch Heute*-Korpus. Sebastian Kürschner, Mechthild Habermann, Peter O. Müller (ed.), *Methodik moderner Dialektforschung: Erhebung, Aufbereitung und Auswertung von Daten am Beispiel des Oberdeutschen.* Hildesheim: Olms. (Germanistische Linguistik ; 241–243)

Kisler, Thomas, Reichel, Uwe D., and Schiel, Florian. 2017. Multilingual processing of speech via web services. *Computer Speech & Language*, 45, 326–347.

Kleber, Felicitas. 2017. Complementary length in vowel-consonant sequences: Acoustic and perceptual evidence for a sound change in progress in Bavarian German. *Journal of the International Phonetics Association*, 1–22. Available online: DOI: https://doi.org/10.1017/S0025100317000238, last accessed: 08.11.2019.

Kleber, Felicitas, Harrington, Jonathan, and Reubold, Ulrich. 2012. The relationship between the perception and production of coarticulation during a sound change in progress. *Language and Speech* 55, 383–405.

Klingler, Nicola, Kleber, Felicitas, Jochim, Markus, Pucher, Michael, Schmid, Stephan, and Zihlmann, Urban. 2019. Temporal organization of vowel plus stop sequences in production and perception: Evidence from the three major varieties of German. *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS), Melbourne*, 825– 829.

Koekkoek, Byron J. 1953. *Phonetik und Phonologie des Wiener Dialektes*. PhD thesis, University of Vienna.

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

Labov, Wiliam. 1994. *Principles of language change: Internal factors.* Oxford: Blackwell.

Lameli, Alfred. 2004. Hierarchies of dialectal features in a diachronic view: Implicational scaling of real time data. In: Britt-Louise Gunnarsson (ed.), *Language variation in Europe: Papers from the 2nd International Conference on Language Variation in Europe* (ICLaVE 2), 253–266.

Moosmüller, Sylvia. 1984. *Soziale und psychosoziale Sprachvariation: eine quantitative und qualitative Untersuchung zum gegenwärtigen Wiener Deutsch.* PhD Thesis, University of Vienna.

Moosmüller, Sylvia. 1987. Soziophonologische Variation bei österreichischen Politikern. *Zeitschrift für Germanistik* 4, 429–439.

Moosmüller, Sylvia. 1996. Die österreichische Variante der Standardaussprache. In: Eva-Maria Krech and Ulrich Püschel (eds.), *Beiträge zur deutschen Standardaussprache*.

Hanau/Halle: Dausien, 204–214.
Moosmüller, Sylvia. 2015. The interaction of temporal organization and phonotactics: Resyllabification in three varieties of German. *Italian Journal of Linguistics* 27, 111–132.
Moosmüller, Sylvia. 2016. The strength of stereotypes in the production and perception of the Viennese dark lateral. In: Jennifer Cramer and Chris Montgomery (eds.), *Cityscapes and Perceptual Dialectology: Global perspectives on non-linguists' knowledge of the dialect landscape.* Boston/Berlin: de Gruyter Mouton, 121–138.
Moosmüller, Sylvia and Brandstätter, Julia. 2014. Phonotactic information in the temporal organization of Standard Austrian German and the Viennese dialect. *Language Sciences* 46, 84–95.
Moosmüller, Sylvia and Ringen, Catherine. 2004. Voice and aspiration in Austrian German plosives. *Folia Linguistica* XXXVIII/1–2, 43-62.
Moosmüller, Sylvia and Scheutz, Hannes. 2013. Chain shifts revisited: The case of Monophthongisation and E-confusion in the city dialects of Salzburg and Vienna. In: Peter Auer, Caro Reina, Javier, and Kaufmann, Göz (eds.), *Language variation - European Perspectives IV.* Amsterdam: Benjamins, 173–186.
Moosmüller, Sylvia, Schmid, Carolin, Brandstätter, Julia. 2015. Standard Austrian German. *Journal of the International Phonetic Association* 45, 339–348.
Pfalz, Anton. 1913. *Die Mundart des Marchfeldes*. Wien: Hölder. (Sitzungsberichte der kaiserlichen Akademie der Wissenschaften in Wien. Philosophisch-historische Klasse, vol. 170, 6th Abhandlung)
Pinget, Anne-France. 2015. *The actuation of sound change*. Utrecht: LOT.
R Core Team. 2019. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/, last accessed 27.11.2019.
RStudio Team. 2018. *RStudio: Integrated Development for R.* RStudio, Inc., Boston, MA URL http://www.rstudio.com/, last accessed 27.11.2019.
Rowley, Anthony. 1990. North Bavarian. In: Charles V. J. Russ (ed.), *The dialects of Modern German. A linguistic survey.* London: Routledge, 417–437.
Schmid, Carolin and Moosmüller, Sylvia. 2017. An acoustic comparison between stressed and unstressed vowels in Standard Austrian German and Standard German German. In: Sylvia Moosmüller, Carolin Schmid, Manfred Sellner (eds.), *Phonetik in und über Österreich.* Wien: Verlag der Österreichischen Akademie der Wissenschaften, 45–59.
Seiler, Guido. 2005. On the development of the Bavarian quantity system. *Interdisciplinary Journal of Germanic Linguistic and Semiotic Analysis* 10, 103–129.
Wagener, Peter. 2002. German dialects in real time change. *Journal of Germanic Linguistics* 14, 271–285.
Wiesinger, Peter. 1990. The Central and Southern Bavarian Dialects. In: Charles V. J. Russ (ed.), *The dialects of Modern German. A linguistic survey.* London: Routledge, 438–519.
Winkelmann, Raphael, Harrington, Jonathan, and Jänsch, Klaus. 2017. EMU-SDMS: Advanced speech database management and analysis in R. *Computer Speech & Language* 45, 392–410.