INTERSPEECH 2023
20-24 August 2023, Dublin, Ireland

The emergence of obstruent-intrinsic f0 and VOT as cues to the fortis/lenis contrast in West Central Bavarian

Jasmin Pöhnlein[1], *Felicitas Kleber*[1]

[1]Institute for Phonetics and Speech Processing (IPS), LMU Munich, Germany
{jasmin.poehnlein|kleber}@phonetik.uni-muenchen.com

Abstract

This study examines the effects of underlying voicing in word initial and medial stops on f0 at the onset of the following stressed and unstressed vowel (CF0), respectively, in Standard German (SG) and West Central Bavarian (WB). As opposed to SG, WB hitherto did not contrast fortis and lenis stops by VOT, but the importance of this cue increases in younger WB speakers. The replacement of VOT by f0 as acoustic cue in connection with voicing mergers and tonogenesis is well-studied but not the emergence of CF0 effects together with an evolving VOT contrast. An acoustic analysis of twenty SG speakers as well as ten older and ten younger WB speakers showed higher f0 after fortis compared to lenis stops in SG but only in initial position where VOT was much longer. Younger but not older WB speakers showed signs of developing CF0 effects in initial stops as found in SG which may forecast VOT differences in this position at the population level possibly due to speaker-specific cue enhancement.

Index Terms: CF0 and VOT, sound change, German varieties, initial vs. prosodic strengthening, group vs. individual level

1. Introduction

Contextual or obstruent-intrinsic f0 (hereafter CF0; [1]) describes the raising of the fundamental frequency (f0) at the vocalic onset following upon a voiceless (fortis) stop in comparison to voiced (lenis) stops (or nasals for that matter). CF0 effects emerge in both true voicing [2] and aspirating languages [3], i.e., regardless the use of VOT as cue to stop voicing. Although these CF0 effects are microprosodic in nature, i.e., irrelevant for the macroprosodic f0 structure (e.g., intonation), they have been shown to be used as perceptual cues to the fortis/lenis contrast in various languages (e.g., [4]).

Whether CF0 effects are best explained by (i) an automatic biomechanical consequence of a laryngeal voicing inhibition gesture or (ii) an enhancement of the voicing contrast intended by the speaker is still a matter of debate. At least to some extent, however, CF0 raising after voiceless stops occurs independently of VOT as results in a purely closure duration based fortis/lenis contrast in Swiss German suggest, lending some support for a biomechanical cause (cf. [5]). On the other hand, similar values along the VOT continuum [6] between languages do not automatically imply similar CF0 effects (cf. [7]).

The existence of CF0 effects has also been shown for Standard German (SG; [8, 9]), an aspirating language with a fortis/lenis contrast cued primarily by long vs. short lag VOT in both initial (prevocalic) and medial (postvocalic) position. Phonetic voicing plays no role. [10] confirmed generally higher f0 and longer VOT for SG fortis compared to lenis stops. In line with enhanced CF0 effects found in high pitch environments for mostly non-tonal languages (e.g., [7]), [10] additionally showed that most speakers produced greater VOT and CF0 differences as a function of underlying voicing in strong prosodic contexts (both lexically stressed and phrasally accented) than in prosodically weak contexts (i.e., in unstressed and unaccented syllables). But since [10] used stress and accent as combined cues to prosodic prominence and the position of the analyzed stops within the target words was not fully controlled, it remains an open question whether these differences in SG are (i) caused by stress or by accent, and (ii) influenced by an unbalanced distribution of stop position across prosodic contexts.

In West Central Bavarian (WB), a dialect spoken in the southeast of Germany, fortis stops are neutralized towards the lenis variant in initial position [11] and signaled by closure duration instead of VOT in medial position [12, 13]. However, recent apparent-time studies on WB suggest that under the influence of SG (cf. [14]) the voicing merger in initial position might currently be reversed and the acoustic cues to the contrast in medial position are being reweighted such that VOT is increasingly used by younger WB speakers (in medial stops particularly after short vowels; [13, 15, 16]).

While the decrease of VOT contrasts in connection with increasing f0 differences have been discussed extensively in the light of tonogenesis by which a voicing contrast is eventually giving way to a tonal contrast (e.g., [17, 18]), the study of emerging CF0 effects in the context of an evolving VOT contrast remains thusfar a research desideratum.

The main aims of this study are thus to shed light on the interplay of CF0 and VOT in an (emerging) fortis/lenis contrast at two word positions: initial before stressed vowels and medial before unstressed vowels. This allows us to systematically investigate the influence of lexical stress in phrasally accented German words (though, possibly reinforced by differences in domain-initial strengthening; [19]). Moreover, we directly compare SG and two generations of WB speakers, i.e., three speaker groups known to differ with respect to their primary stop voicing cue, by obtaining acoustic measures for both CF0 and VOT. Our CF0 hypotheses are:

- (H1) In initial position, we expect a raised f0 after fortis stops but neither after lenis stops nor nasals (which are fully voiced and thus suitable controls for the underlying, unperturbed f0 contour; cf. [3]) in SG. Due to presumed increasing VOT differences, attenuated CF0 effects occur in younger but not in older WB speakers.
- (H2) CF0 effects are greater in initial than in medial position, as a function of domain-initial strengthening and/or lexical stress.
- (H3) In case of emerging CF0 effects in medial position, these are greater in younger vs. older WB speakers and greatest in SG speakers due to the emerging VOT contrast in younger but not older WB speakers and an existing contrast in SG speakers.

 10.21437/Interspeech.2023-1879

2. Methods

2.1. Participants

The analysis was based on acoustic recordings obtained from 20 SG (aged 21–82, mean 41.0 years, 10 female) as well as 10 younger (aged 18–30, mean 24.6 years, 5 female) and 10 older (aged 50–65, mean 56.4 years, 5 female) WB speakers. Since SG speakers functioned as control group, they were equally balanced with respect to age as WB speakers, yet not assigned to two different age groups. While SG speakers were largely from the urban areas of Munich and self-identified as non-dialectal users of the local Southern SG variety, WB speakers stemmed from the rural surroundings of Munich (mostly Upper Bavaria) and acquired WB as first variety of German. Their dialect competence was assessed by the experimenters, who were native WB speakers.

2.2. Speech materials

We selected 69 lexical items (5 monosyllables, 64 trochaic disyllables) from a larger corpus comprising a total of 115 words balanced, among others, for lexical frequency. The data set contained slightly more lexically low than high frequent items (39 of 69) and more phonemically long than short vowels preceding medial stops (20 of 33). However, these predictors were not in focus of our analyses. The selected items contained either nasal (phonetically voiced) or phonologically voiced (lenis, short voiceless) and voiceless (fortis, long voiceless) oral stops in word-initial or -medial position. The stops were either alveolar or bilabial, e.g., *Nager* /ˈnaː.ɡɐ/ ‘rodent’, *Mars* /ˈmaʁs/ ‘mars’, *Dame* /ˈdaː.mə/ ‘lady’, *Park* /ˈpʰaʁkʰ/ ‘park’, *Kette* /ˈkʰɛt̩ʰə/ ‘necklace’, … (target segments underlined). Six of these words had target segments in both position (e.g., *Dame*). Only words with mid or low vowels were considered for this analysis to avoid effects of both vowel-intrinsic f0 and transconsonantal coarticulation (cf. [5]). Note that while vowels following word-medial stops were always unstressed, only stops after long stressed vowels (e.g., *Kater* ‘tomcat’) are part of the unstressed syllable. Stops following short stressed vowels (e.g, *Modder* ‘mud’) are considered ambisyllabic in German.

All target words are part of both the SG and the WB lexicon and were embedded in word-specific, yet syntactically similar carrier phrases with a preferably equal number of syllables. WB sentences (e.g., *Ea mog de* ***Dame*** *treffa.*) were non-standardized translations from SG orthography (e.g., *Er will die* ***Dame*** *treffen.*). Sentence constructions triggered nuclear pitch accents to occur on the target word (cf. [20, 21]). Thus, target segments differed in lexical stress (due to the respective position within the target word) but not in phrasal accent.

2.3. Recordings and data pre-processing

Recordings were made at a minimum sampling rate of 44.1 kHz using SpeechRecorder [22], either in a sound-attenuated booth with a condenser microphone or remotely with headsets via Wikispeech [23]. We carefully monitored sufficient recording quality, especially in the latter cases. Carrier phrases were presented in randomized order on a screen and were read by participants in silence. After the respective prompt had disappeared, speakers reproduced it from memory. Participants were asked to repeat carrier phrases whenever the experimenter noted obvious mispronunciations. We included only one – if possible the second – out of three repetitions per sentence into the present analysis, for a total of 2,760 utterances (40 speakers × 69 target words).

The recorded utterances were automatically segmented into words and phones using WebMAUS [24] and stored as two interval levels of an EMU database [25, 26]. We manually corrected all relevant segment boundaries, namely the start and end of (i) the utterance, (ii) the target word, and (iii) the target word’s phonemes. Whenever possible, we added a further segment boundary right before the target stop’s burst to mark the end of its closure and the simultaneous beginning of its aspiration phase (always positive and equivalent to VOT).

2.4. Acoustic analysis

All acoustic measures were retrieved with emuR [27], using R [28] (v. 4.2.0) in RStudio [29] (v. 2.3.492); all subsequent statistical analyses were also conducted in R. Prior to analysis, we replaced the second with the third repetition for a total of 14 items due to bad recording quality. Diverging dialectal realisations by WB speakers (e.g., *Kettn* instead of *Kette* with a nasal and not an oral release) were discarded completely (7 %, 197 of 2,760 utterances). Stops without a segmented aspiration phase were included in the analyses; their VOT was set to 0 ms. VOT duration was normalized to the duration of the entire target word minus the target stop’s duration.

We calculated f0 trajectories separately for males and females using the *ksvF0()* pitch estimator of *wrassp* [30] and extracted the corresponding Hertz (Hz) values over the vowel following the respective target stop, i.e., the stressed vowel after initial stops (hereafter V_1) and the unstressed vowel after medial stops (hereafter V_2). After the removal of missing Hz values (i.e., equal to zero), we length-normalized each target vowel to 21 equidistant time points. For the remaining data set, speaker-scaled f0 values (*z*-scores) were calculated across all voiced frames to control for individual pitch level (cf. [10]).

To relate the following results more easily to the various statistics applied to the acoustic measures, we will detail each statistical model alongside with its corresponding results in Section 3.

3. Results

3.1. Effects on obstruent-intrinsic F0

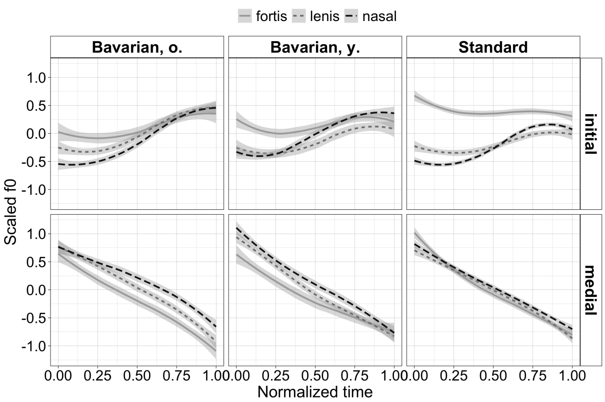

Figure 1: *Loess-smoothed f0 contours by speaker group and consonant type, averaged across speakers and items. Top panels: initial position, bottom panels: medial position.*

Fig. 1 shows the speaker-scaled, time-normalized f0 contours over the respective target vowel. For statistical modeling, we used scaled f0 values averaged over the first 25 % of the vowel as dependent variable. Due to considerable macroprosodic differences between V_1 (mostly rising f0) and V_2 (falling f0), we fitted one linear mixed-effects model per consonant position in the target word using *lmer()* of *lme4* [31]. In each model, we included *consonant type* and *speaker group* as fixed effects (main effects and interaction) and added random intercepts for *speaker* and *item* as well as random slopes for both *consonant type* (within-*speaker*) and *speaker group* (within-*item*).

3.1.1. *Initial position*

A total of 1,301 mean scaled f0 values were analyzed (WB, older: 311; WB, younger: 298; SG: 692). To improve model convergence, random slopes for *speaker group* were removed from the full model using the default settings of *step()* (*lmerTest*; [32]). The general model outcomes were robust against this modification. We further excluded 21 outliers based on model residuals. The results of the most parsimonious model showed a significant main effect of *consonant type* (F(2, 37) = 5.3, p < .01), an approaching significant effect of *speaker group* (F(2, 38) = 3.1, p = .05) as well as a significant interaction effect between the two (F(4, 38) = 5.9, p < .001). Pairwise comparisons of the estimated marginal means (*emmeans*; [33]) including p-value adjustments (Tukey method) returned significant contrasts between fortis stops (F) and both lenis stops (L) and nasals (N) for SG speakers as well as between fortis stops and nasals for younger WB speakers (cf. Tab. 1).

Table 1: *Pairwise comparisons of estimated marginal means for CF0, separately for speaker group*

	Contrast	Est.	SE	df	t-ratio	p-value
WB, o.	F – L	0.180	.261	58.9	0.691	.770
	F – N	0.424	.243	50.4	1.745	.199
	L – N	0.244	.240	49.0	1.013	.572
WB, y.	F – L	0.459	.262	59.5	1.755	.194
	F – N	0.605	.243	50.7	2.487	.042
	L – N	0.146	.241	49.6	0.605	.818
SG	F – L	0.907	.236	46.2	3.840	.001
	F – N	1.063	.226	40.4	4.701	<.001
	L – N	0.157	.225	39.8	0.697	.767

This finding suggests the presence of CF0 effects in SG speakers and the emergence of such effects in younger WB speakers at the population level: f0 was higher in the beginning of V_1 following fortis stops than following nasals, while lenis stops seem to pattern with the latter. Note that only the significant CF0 differences between SG speakers' fortis and lenis stops indicate a systematic use of this cue in the signaling of this phonemic contrast in this group. Younger WB speakers' CF0 raising after fortis stops in comparison to nasals only, on the other hand, suggests the emergence of CF0 that is not yet systematically used as a cue. For older WB speakers, none of the contrasts between consonant types was significant, suggesting no systematic CF0 effects after initial stops in this speaker group.

3.1.2. *Medial position*

Following consonants in medial position, a total of 1,167 mean scaled f0 values for V_2 were available for analysis (WB, older: 280; WB, younger: 269; SG: 618). Again, we applied *step()* to address model convergence issues. However, none of the entered effects and interactions reached significance at the 5 % level, suggesting no systematic CF0 effects after medial stops in any of the speaker groups.

3.2. Effects on Voice Onset Time

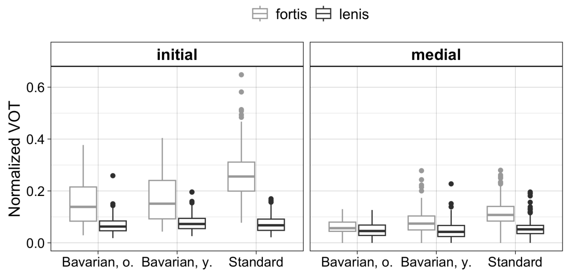

Figure 2: *VOT normalized to word duration by speaker group, consonant type (grey: fortis, black: lenis) and word position (left: initial, right: medial).*

Fig. 2 illustrates the effects of an oral stop's position in the target word on its normalized VOT. A linear mixed-effects model with normalized VOT as dependent variable and *consonant type*, *speaker group* and *position* as fixed effects (main effects and interactions) was fitted to a total of 886 fortis and lenis stops in initial position (WB, older: 206; WB, younger: 200; SG: 480) as well as 837 fortis and lenis stops in medial position (WB, older: 191; WB, younger: 186; SG: 460). We also included random intercepts for *speaker* and *item* as well as random slopes for *consonant type*, *position* (both within-*speaker*) and *speaker group* (within-*item*). The application of *step()* returned the full model as being the most parsimonious one. The exclusion of a total of 52 outliers based on model residuals again did not affect the general outcomes of the model, which revealed highly significant main effects for *consonant type* (F(1, 64) = 104.5, p < .001), *speaker group* (F(2, 53) = 11.4, p < .001) and *position* (F(1, 85) = 71.3, p < .001) as well as significant interaction effects between *consonant type* and both *position* (F(1, 61) = 40.9, p < .001) and *speaker group* (F(2, 59) = 14.5, p < .001). The three-way interaction between *consonant type*, *speaker group* and *position* was also significant (F(2, 99) = 4.9, p < .01). Commensurate with Fig. 2, VOT was significantly longer in initial compared to medial stops – regardless of *speaker group* and *consonant type.* Pairwise *emmeans*-comparisons between *speaker group* and *consonant type*, separately for *position* (cf. Tab. 2), showed that all speaker groups distinguish fortis (F) and lenis (L) stops in initial position by means of VOT, but only younger WB and SG speakers did so in medial position. Further pairwise comparisons between speaker groups showed comparable normalized VOT values in initial fortis stops of both WB age groups, which were significantly lower than in SG.

Table 2: *Pairwise comparisons of estimated marginal means for VOT per speaker group and consonant type*

	Contrast	Est.	SE	df	t-ratio	p-value
initial	F WB, o. – WB, y.	-.020	.017	39.6	-1.135	.8636
	F WB, o. – SG	-.108	.017	62.9	-6.282	<.001
	F WB, y. – SG	-.089	.018	66.1	-5.056	<.001
	WB, o. F – L	.072	.013	75.7	5.375	<.001
	WB, y. F – L	.085	.014	75.2	6.156	<.001
	SG F – L	.177	.012	64.4	14.468	<.001
medial	F WB, o. – WB, y.	-.014	.012	45.6	-1.181	.844
	F WB, o. – SG	-.052	.014	82.0	-3.828	.003
	F WB, y. – SG	-.038	.014	87.5	-2.687	.088
	WB, o. F – L	.020	.013	76.6	1.503	.664
	WB, y. F – L	.038	.014	78.1	2.713	.084
	SG F – L	.068	.012	65.2	5.536	<.001

All lenis stops showed comparable short lag VOT values across *speaker group* and *position* with none of the pairwise comparisons reaching significance. Taken together, these findings are indicative of (i) a generally more pronounced fortis/lenis distinction based on VOT in SG compared to WB as well as (ii) in initial compared to medial position, and (iii) a population level sound change in progress affecting medial stops in WB such that younger but not older WB speakers have started to use VOT for signaling underlying voicing differences.

3.3. Correlation of CF0 and VOT in fortis stops

In a last step, we directly looked at the relation between CF0 and VOT in fortis stops to investigate whether (i) the age-induced WB differences in terms of CF0 in initial position are determined by individual cue enhancement via VOT in younger WB speakers, and whether (ii) emerging VOT differences between WB age groups in medial position lead to speaker-specific CF0 contrasts, potentially covered at the group level by generally lower VOT values in this position.

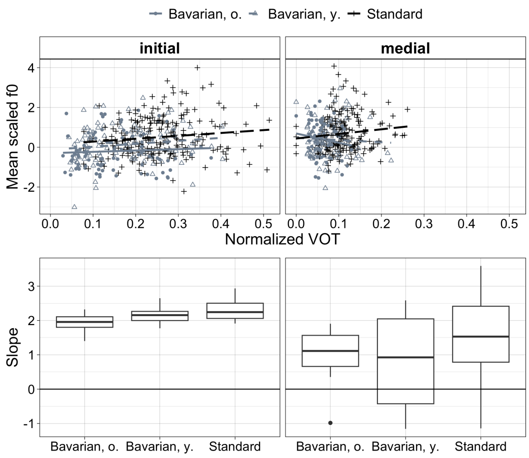

Figure 3: *Upper panels: VOT-CF0 correlation in fortis stops by position and speaker group with linear trend lines at group level. Lower panels: distribution of corresponding individual slopes.*

In the upper panels of Fig. 3, normalized VOT values are plotted against the corresponding speaker-scaled CF0 means over the first 25 % of the vowel. Each data point represents one item per speaker. Only items with values available for both variables were included, resulting in 424 (WB, older: 101; WB, younger: 97; SG: 226) fortis stops in initial and 342 (WB, older: 79; WB, younger: 69; SG: 194) fortis stops in medial position. We excluded 9 outliers based on visual inspection prior to analysis and added linear trend lines at group level to facilitate the interpretation of individual slopes. In initial position, these lines indicate no meaningful correlation between VOT and CF0 for older WB speakers (Pearson's r = .06) but weak positive correlations for both younger WB (r = .15) and SG speakers (r = .13). In medial position, the group level regression lines show a trend from a weak negative correlation in older WB (r = -.15) over no meaningful correlation in younger WB (r = -.05) to a weak positive correlation in SG (r = .12).

To assess potential individual differences in the relationship of VOT to CF0, we fitted one *lmer()* per stop position. In each model, we entered *scaled CF0 mean* as dependent variable, *normalized VOT* as fixed factor and within-*speaker* and within-*item* random slopes for *normalized VOT*. The distribution of only the speaker-specific slopes in the lower panel of Fig. 3 (which are of greatest interest here) suggests generally higher CF0 effects with increasing VOT, as they were mostly positive. Separate pairwise repeated-measures ANOVAs [34] for each stop position and speaker group pairing and with *speaker-specific slope* as dependent variable, *speaker group* as independent variable and *speaker* as random factor revealed for the initial position a significant difference between older WB and SG speakers ($F(1, 28) = 11.4$, $p = .002$). Neither the difference between the two WB age groups nor that between younger WB and SG speakers reached significance, suggesting intermediate values for younger WB speakers. In medial position, we removed further seven outliers based on model residuals. Here, the speaker groups showed no systematic slope differences accompanied by a larger range at individual level.

4. Discussion and conclusion

This study yields four main findings. First, and consistent with H1, clear and emerging CF0 effects were found for SG and younger WB speakers, respectively, in initial position. No such effects emerged for older WB speakers or in medial position. The non-existent CF0 effects in medial position constitute an extreme form of the predictions made in H2 and render the potential group differences alluded to in H3 irrelevant. Second, in initial position, all speaker groups used VOT differences to cue the fortis/lenis contrast, but SG more than WB speakers, who showed no age groups differences at the group level in this context. Third, in medial position before an unstressed vowel the VOT contrast was in general diminished (possibly such that CF0 effects were prevented altogether), but here younger as opposed to older WB speakers approached VOT values typical of SG in fortis stops. Forth, CF0 and VOT were positively correlated at the individual level in all but six speakers such that fortis-typical CF0 effects increased with longer VOT values.

In particular, the group level correlations for younger WB and SG speakers' initial fortis stops suggest that the magnitude of the effect rises in tandem with VOT, which contradicts the assumption of a purely biomechanical cause [3] and lends support for the idea of fortis stop enhancement [35], though presumably not deliberately. The idea of a causal relationship between VOT and CF0 was also put forward in [10], but the diminished CF0 effects described therein for unstressed syllables of unaccented words differ from the totally absent CF0 effects at the onset of unstressed vowels of accented words in the present study. This difference may support the idea that CF0 effects are driven by domain-initial strengthening [19] rather than prosodic strength, given that in [10] four out of twelve stops in prosodically weak contexts occurred in initial position.

Intermediate correlation results at the individual level further suggest an increased importance of CF0 effects together with an increase of VOT in younger WB speakers' initial fortis stops (concurrent with [13]), despite the (unexpected) lack of significant age-induced VOT differences in this position at the group level. CF0 effects seem to play no role in the emergence of the medial VOT contrast which is otherwise, and contrary to the assumption in [10] that prosodically weak contexts provide the condition for sound changes to arise, more advanced in medial than in initial position as the significant age differences at the group level suggest.

5. Acknowledgements

R scripts and data frames are available at [36]. The study was funded by DFG grant KL 2697/1-2 awarded to F. Kleber.

6. References

[1] J. Kingston, "Segmental influences on F0: Automatic or controlled?," in *Tones and tunes*, vol. 2, C. Gussenhoven and T. Riad Eds., (Experimental studies in word and sentence prosody. Berlin: Mouton de Gruyter, 2007, pp. 171-201.

[2] J. P. Kirby and D. R. Ladd, "Effects of obstruent voicing on vowel F0: Evidence from “true voicing” languages," *The Journal of the Acoustical Society of America,* vol. 140, no. 4, pp. 2400-2411, 2016.

[3] H. M. Hanson, "Effects of obstruent consonants on fundamental frequency at vowel onset in English," *The Journal of the Acoustical Society of America,* vol. 125, no. 1, pp. 425-441, 2009.

[4] D. H. Whalen, A. S. Abramson, L. Lisker, and M. Mody, "F0 gives voicing information even with unambiguous voice onset times," *The Journal of the Acoustical Society of America,* vol. 93, no. 4, pp. 2152-2159, 1993.

[5] D. R. Ladd and S. Schmid, "Obstruent voicing effects on F0, but without voicing: Phonetic correlates of Swiss German lenis, fortis, and aspirated stops," *Journal of Phonetics,* vol. 71, pp. 229-248, 2018.

[6] T. Cho and P. Ladefoged, "Variation and universals in VOT: evidence from 18 languages," *Journal of Phonetics,* vol. 27, no. 2, pp. 207-229, 1999.

[7] J. P. Kirby, "Onset pitch perturbations and the cross-linguistic implementation of voicing: Evidence from tonal and non-tonal languages," *Journal of Phonetics,* vol. 71, pp. 326-354, 2018.

[8] K. J. Kohler, "F0 in the production of fortis and lenis plosives," *Phonetica,* vol. 39, no. 4-5, pp. 199-218, 1982.

[9] M. Jessen, *Phonetics and phonology of tense and lax obstruents in German*. Amsterdam: John Benjamins Publishing Company, 1999.

[10] J. Kirby, F. Kleber, J. Siddins, and J. Harrington, "Effects of prosodic prominence on obstruent-intrinsic F0 and VOT in German," presented at the 10th International Conference on Speech Prosody (SP), Tokyo, Japan (virtual), 2020.

[11] P. Wiesinger, "The Central and Southern Bavarian Dialects in Bavaria and Austria," in *The Dialects of Modern Germany. A Linguistic survey*, C. V. J. Russ Ed. Abingdon: Routledge, 1990, pp. 438-519.

[12] G. Seiler, "On the development of the Bavarian quantity system," *Interdisciplinary Journal for Germanic Linguistics and Semiotic Analysis,* vol. 10, no. 1, pp. 103-129, 2005.

[13] F. Kleber, "VOT or quantity: what matters more for the voicing contrast in German regional varieties? Results from apparent-time analyses," *Journal of Phonetics,* vol. 71, pp. 468-486, 2018.

[14] P. Trudgill, *Dialects in Contact*. Oxford: Blackwell, 1986.

[15] F. Kleber, "Complementary length in vowel–consonant sequences: Acoustic and perceptual evidence for a sound change in progress in Bavarian German," *Journal of the International Phonetic Association,* vol. 50, no. 1, pp. 1-22, 2020.

[16] M. Jochim and F. Kleber, "Fast-Speech-Induced Hypoarticulation Does Not Considerably Affect the Diachronic Reversal of Complementary Length in Central Bavarian," *Language and Speech,* vol. 0, no. 0, 2022.

[17] H.-Y. Bang, M. Sonderegger, Y. Kang, M. Clayards, and T.-J. Yoon, "The emergence, progress, and impact of sound change in progress in Seoul Korean: Implications for mechanisms of tonogenesis," *Journal of Phonetics,* vol. 66, pp. 120-144, 2018.

[18] A. W. Coetzee, P. S. Beddor, K. Shedden, W. Styler, and D. Wissing, "Plosive voicing in Afrikaans: Differential cue weighting and tonogenesis," *Journal of Phonetics,* vol. 66, pp. 185-216, 2018.

[19] T. Cho, D. J. Kim, and S. Kim, "Prosodic strengthening in reference to the lexical pitch accent system in South Kyungsang Korean," *The Linguistic Review,* vol. 36, no. 1, pp. 85-115, 2019, doi: doi:10.1515/tlr-2018-2008.

[20] D. R. Ladd, I. Mennen, and A. Schepman, "Phonological conditioning of peak alignment in rising pitch accents in Dutch," *The Journal of the Acoustical Society of America,* vol. 107, no. 5, pp. 2685-2696, 2000.

[21] D. R. Ladd, A. Schepman, L. White, L. M. Quarmby, and R. Stackhouse, "Structural and dialectal effects on pitch peak alignment in two varieties of British English," *Journal of Phonetics,* vol. 37, no. 2, pp. 145-161, 2009.

[22] C. Draxler and K. Jänsch, "SpeechRecorder – a Universal Platform Independent Multi-Channel Audio Recording Software," presented at the IVth International Conference on Language Resources and Evaluation (LREC), Lisbon, Portugal, 2004.

[23] C. Draxler and K. Jänsch, "Wikispeech – a content management system for speech databases," presented at the 9th Annual Conference of the International Speech Communication Association (INTERSPEECH), Brisbane, Australia, 2008.

[24] T. Kisler, U. D. Reichel, F. Schiel, C. Draxler, B. Jackl, and N. Pörner, "BAS Speech Science Web Services – an Update on Current Developments," presented at the Xth International Conference on Language Resources and Evaluation (LREC), Portorož, Slovenia, 2016.

[25] R. Winkelmann, J. Harrington, and K. Jänsch, "EMU-SDMS: Advanced speech database management and analysis in R," *Computer Speech & Language,* vol. 45, pp. 392-410, 2017.

[26] M. Jochim, "Extending the EMU Speech Database Management System: Cloud Hosting, Team Collaboration, Automatic Revision Control," presented at the 18th Annual Conference of the International Speech Communication Association (Interspeech), Stockholm, Sweden, 2017.

[27] *emuR: Main Package of the EMU Speech Database Management System. R package version 2.3.0*. (2021).

[28] *R: A language and environment for statistical computing*. (2022). R Foundation for Statistical Computing, Vienna, Austria. [Online]. Available: https://www.R-project.org/

[29] *RStudio: Integrated Development Environment for R*. (2022). RStudio, PBC, Boston, MA. [Online]. Available: http://www.rstudio.com/

[30] *wrassp: an R wrapper to the ASSP Library. R package version 1.0.1*. (2021).

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

[32] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest Package: Tests in Linear Mixed Effects Models," *Journal of Statistical Software,* vol. 82, no. 13, pp. 1-26, 2017.

[33] *emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.7.4-1*. (2022). [Online]. Available: https://CRAN.R-project.org/package=emmeans

[34] *ez: Easy Analysis and Visualization of Factorial Experiments. R package version 4.4-0*. (2016). [Online]. Available: https://CRAN.R-project.org/package=ez

[35] J. Kingston and R. L. Diehl, "Phonetic knowledge," *Language,* vol. 70, no. 3, pp. 419-454, 1994.

[36] J. Pöhnlein and F. Kleber. *The emergence of obstruent-intrinsic f0 and VOT as cues to the fortis/lenis contrast in West Central Bavarian.* (2023). Open Data LMU. [Online]. Available: https://doi.org/10.5282/ubm/data.384