A perceptual investigation into the laryngeal contrast for word-initial labial stops in the non-standard German variety, Upper Saxon

Franziska Kruger

Indiana University
fkruger@indiana.edu

ABSTRACT

Standard German (SGer) maintains a contrast between lenis (voiceless unaspirated) stops and fortis (voiceless aspirated) stops. For the non-standard German variety Upper Saxon (USax), traditional accounts report that only lenis stops are retained. Thus, <Bass> (*bass*) and <Pass> (*passport*) are a minimal pair in SGer but homophones in USax. However, recent phonetic investigations suggest that USax speakers produce fortis stops with aspiration.

This study examines word-initial labials in USax speakers' utterances. It also uses a discrimination task to explore if SGer and USax listeners can distinguish USax fortis from lenis stops.

The analysis revealed that few USax speakers produced a clear fortis/lenis distinction. On the perceptual task, the SGer listeners perceived aspiration on fortis stops more reliably than the USax listeners. These results support traditional descriptions that claim aspiration is not phonemic in USax. They also indicate that USax speakers might be less accustomed to using this cue in perception.

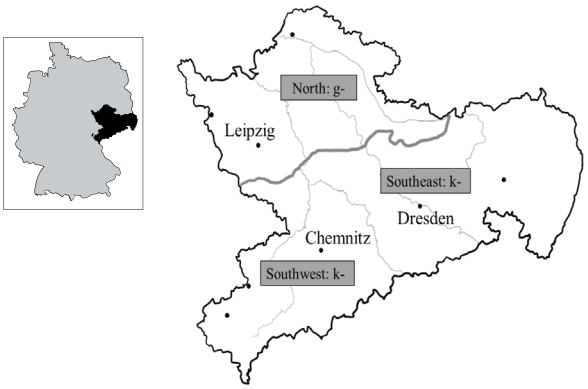
Keywords: Upper Saxon, laryngeal contrast, VOT, aspiration, speech perception.

1. INTRODUCTION

1.1. Background

Upper Saxon is spoken in western and central parts of modern-day Saxony, Germany. Traditional accounts report an USax stop system where fortis /p, t, k/—voiceless and aspirated in SGer—are lenited and produced as lenis (voiceless unaspirated) /b, d, g/. Labial and alveolar stops are voiceless and unaspirated in most areas, while lenition of word-initial velar stops is restricted to the northern regions [1]. Varieties in the southern areas retain the velar stop as voiceless, aspirated /k/ [10][14]. Thus, /k/ creates an isogloss between the northern and southern regions of USax.

Figure 1: Saxony in Germany (small map, black) and USax regions (large map) with /k-/ isogloss (adapted from [3]).



This paper focuses on the lenition of word-inital labial stops in speakers from the southern regions. Unlike alveolar stops, they have not yet been phonetically investigated [13][21], but see [15].

In 1961, [9] observed that stop lenition was still an obligatory feature in the regiolect of many speakers. Recent investigations contradict his observation, although the results are inconsistent. In [15], speakers from Dresden produced word-initial fortis stops with aspiration. [21] also reports speakers from Dresden producing 60% of their /t/-initial tokens as aspirated [i]. This is surprising, given the assertions that labial and alveolar fortis and lenis stops have merged in this region [14]. [13] investigated word-medial /t/ and /d/ in speakers from Dresden and Chemnitz. Most of her participants did not create a fortis/lenis distinction. Several speakers, however, produced lenis stops with voicing during closure, differentiating them from fortis stops, which remained voiceless. Other speakers used stop closure durations to distinguish the fortis and lenis category. [13] excluded aspirated fortis tokens as hyperarticulated from her analysis. However, as USax speech incorporates phonological features from the standard variety [1][9][10] or might be undergoing a sound change that is undoing the lenition process [15][17], these instances of aspiration may have been speaker-specific strategies to create the fortis/lenis contrast.

Since the word onset carries the heaviest burden of lexical access and retrieval, neutralization in this position is typologically rare [4]. It is thus likely that speakers would strive to re-introduce the contrast in word-initial position. Speakers of other German varieties that reportedly lenite their fortis stops have been found to implement a word-initial laryngeal contrast [2][6][20]. In [23], some speakers of Tyrolean German—a variety where word-initial fortis and lenis stops are still merged—produced lenis stops with voicing lead. Their fortis stops were neither voiced nor aspirated. These studies indicate that different varieties use different acoustic cues to re-introduce a phonemic contrast. However, [23] acknowledge that it remains to be seen if their findings indicate a change in progress and if the observed distinction is perceptually salient. While the studies of [13], [15] and [21] suggest that USax speakers might now be producing the laryngeal contrast by means of aspiration, the phonemic function of aspiration in USax fortis stops is still an empirical question.

1.2. Research Questions

This study was conducted in two stages. First, we recorded speakers from the southern region and investigated if they produce a laryngeal contrast. Specifically, we investigated voice onset time differences (VOT, [6]) in word-initial labial stops. We then chose a subset of speakers to be included in stage II. Stage II explored if USax and SGer listeners could identify phonemically relevant information in USax speech that would allow them to discriminate USax lenis and fortis stops.

2. METHOD

2.1. Participants

2.1.1. Participants (speakers, stage I)

All speakers (n = 6) had grown up and lived in or around the larger area of Chemnitz until after high school graduation. Three speakers had never left (IK, KN, CX), while one had moved away for 15 years and later returned (MI). Two had permanently left Saxony to live in Brandenburg (ES, GS). Speaker GS was male, all others were female. None had grown up bilingually.

2.1.2. Participants (listeners, stage II)

For stage II, we recruited 38 participants in Berlin/Brandenburg and Saxony. Eighteen listeners were speakers of USax, and 20 listeners spoke a standard variety of German and had had no systematic exposure to any USax dialects. We included both USax and SGer listeners to explore dialect familiarity effects. We hypothesized that if subtler acoustic cues are used to indicate a contrast, they might be perceivable only to speakers who had grown up speaking the dialect [8]. No participant had grown up bilingually, and none indicated having hearing, visual, speech, or cognitive impairments. All passed a hearing test.

2.2. Stimulus materials

2.2.1. Production stimuli (stage I)

In a word list, 20 /b/- and /p/-initial target word pairs were randomly interspersed between distractor items (N_{total} = 120, 1:3 ratio). All target word pairs were matched in syllable number, stress pattern and the following vowel (see Table 1 below). Speakers read the entire word list twice. Recordings were made with a portable ZOOM H2n recorder at 44.1kHz and saved as WAV files.

2.2.2. Perception stimuli (Stage II)

Stimuli for the discrimination task were word-initial syllables from a subset of stage I utterances. The task included two target contrasts (/bal-pal/, /bum-pum/) and a control contrast (/ties-nies/). Recordings from KN, IK, and ES where chosen to be included in this task. Table 1 lists the target contrasts and the words from which they were cut.

Table 1: Word pair examples (from stage I) and discrimination task target stimuli (stage II).

Syllable	Original item *(gloss)*	
bal-pal	Ballast *(ballast)*	Palast *(palace)*
bum-pum	Bummel *(stroll)*	Pumpe *(pump)*

Syllables were arranged as triplets (*same*: BBB, PPP, *different*: BPP, PBP, PPB, PBB, BPB, BBP). In *same*-triplets, all speakers said the same syllable. In *different*-triplets one speaker said a different syllable than the other speakers, e.g., /bal-pal-pal/. During a training phase with 5 trials, participants responded to the contrast /lø:v/-/mø:v/. The use of three different voices, feedback during training, and training repeat if participants did not reach the accuracy threshold of 90% ensured that they focus on the phonologically relevant information and discard phonetically irrelevant differences (e.g., quality of recording, pitch, stimulus length, etc.) [22].

2.3. Procedure

2.3.1. Procedure (production, stage I)

Speakers recorded themselves at home in a quiet room without the experimenter being present. The materials were mailed to them, and instructions for setting up the microphone and for the procedure were given in written form. Participants were instructed to read the word list twice in "Sächsisch" (Saxon) and to repeat a word for a third time if there was an unexpected background noise.

2.3.2. Procedure (perception, stage II)

The contrasts were embedded in an oddity discrimination task [22] (implemented in a web browser with jsPsych [7]). Participants wore headphones while working on an HP laptop. They saw 4 buttons on the screen (3 robots, 1 large X). In each trial, participants heard a triplet where each robot said one syllable (e.g., bal_{KN}-pal_{IK}-pal_{ES}). They clicked on the robot that had said something different or indicated that all robots had said the same thing by clicking on the X. Trials (N_{total} = 72; 12 triplets x 3 contrasts x 2 iterations [ii]) were separated by 1000 ms. The inter-stimulus-interval was set at 600 ms and response timeout at 6000 ms. Participants were instructed to respond as quickly as possible. The task took 10-12 minutes to complete.

3. ANALYSIS AND RESULTS

3.1. Stage I: USax labial stop production

The analysis focused on word-initial VOT duration for /b/ and /p/ as indicator for aspiration. Recordings where the burst or voice onset could not be unequivocally determined were excluded. A total of 392 tokens with word-initial labials were analyzed in Praat [5]. Table 2 shows the number of analyzed lenis and fortis tokens and the mean (*M*) VOT and standard deviation (SD) for each speaker.

Table 2: Number of word-initial lenis & fortis labial stops and mean VOT (in ms) per speaker.

Speaker	# Token (lenis/fortis)	VOT_{lenis} *M* (SD)	VOT_{fortis} *M* (SD)
GS	18/19*	22.9 (5.5)	27.2 (15.2)
ES	36/40	12.3 (3.1)	25.3 (19.4)
MI	35/36	20.0 (5.0)	48.5 (14.3)
IK	34/35	13.9 (5.1)	38.4 (13.0)
CX	40/39	12.0 (3.6)	19.7 (17.1)
KN	33/27	22.9 (6.5)	23.7 (10.2)

*Note: * GS read the list only once.*

A Mann-Whitney-U test showed that mean VOTs for lenis stops (N = 196, M = 17.4ms, SD = 5.2) were significantly shorter than for fortis stops (N = 196, M = 30.5ms, SD = 11, U = 15971.5, p < .001, Figure 2 top). No speaker produced pre-voiced stops.

Figure 2: Boxplots with VOT durations (in ms) for word-initial /p/ and /b/: all stops averaged (top) and arranged by speaker (bottom).

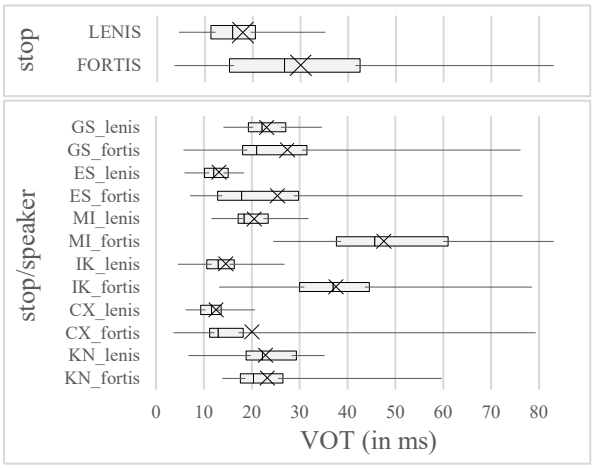


Figure 2 (bottom) shows that GS and KN did not categorically distinguish between lenis and fortis stops by means of aspiration, while the distinction was significant for the other speakers (ES: U = 361, p < .001; MI: U = 7, p < .001; IK: $t(67)$ = -10.10[iii], p < .001; CX: U = 525, p < .05).

It has been suggested that listeners rely on a variety of cues to distinguish fortis from lenis stops [12][16]. For SGer, aspiration is the most relevant phonetic indicator to discriminate the word-initial stop contrast [12][20], and VOTs for fortis labial stops reportedly range from 48-54 ms [18]. Only MI's VOT durations for USax fortis /p/ reached that value. For other USax speakers who produced the distinction, fortis mean VOT ranged from 20-38 ms.

3.2. Stage II: USax labial stop perception

The minimum degree of aspiration necessary to distinguish lenis and fortis stops in German varieties is an empirical question. Furthermore, it has yet to be established whether some varieties maintain the contrast through other means (e.g., voicing-lead [23]). The discrimination task in stage II explored if USax word-initial fortis /p/ and lenis /b/ could be perceived as fortis and lenis, respectively, based on VOT duration or other acoustic cues. The speakers that were selected for this task produced their fortis stops with different degrees of aspiration. VOT durations for each target syllable are given in Table 3.

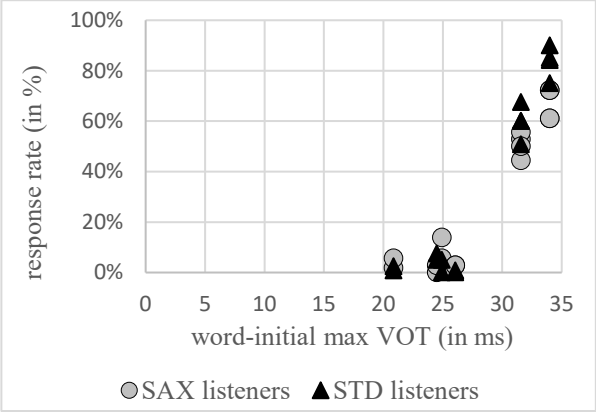
Table 3: VOT (ms) per speaker for lenis and fortis word-initial stops from the discrimination task.

Item	Speaker	Lenis /b/	Fortis /p/
/C-al/	KN	20.9	24.5
	IK	16.4	34.0
	ES	9.7	15.0
/C-um/	KN	26.0	24.9
	IK	13.1	31.6
	ES	15.6	15.9

We calculated *d'* [19] to establish how likely participants where to detect a fortis stop among lenis stops and vice versa. While discrimination of the control contrast was high (USax: d' = 3.71, SD = 0.53, SGer: d' = 4.00, SD = 0.25), d' for the target contrasts was low for both listener groups (USax: d' = -0.16, SD = 0.16; SGer: d' = -0.02, SD = 0.12). The groups' performance indicates that there were no cues in the utterances that allowed listeners to reliably identify the fortis or lenis stops as intended by the speaker. A triplet like /pal$_{\text{-KN}}$-bal$_{\text{-IK}}$-bal$_{\text{-ES}}$/ was frequently heard as /bal-bal-bal/.

Nevertheless, a detailed analysis of the response patterns revealed that both groups were sensitive to VOT duration. Figure 3 illustrates the response rates where the groups identified the token with the longest VOT duration in a triplet as different.

Figure 3: Scatterplot with response rate (in %) for tokens with max VOT in a triplet identified as different, USax (circle), SGer (triangle).



Essentially, listeners approached this task as a *fortis*-detection task. That is, they interpreted longer VOTs (>30 ms) as indicator for 'different.' Both groups rarely chose KN or ES stops as 'different' (6%). Even in the triplet /pal$_{\text{-KN}}$-pal$_{\text{-IK}}$-bal$_{\text{-ES}}$/, they marked IK's token as different, despite ES's token being acoustically the most distant. Listeners in both groups heard IK's fortis token in the /-al/ context as different (i.e., fortis) significantly more often, than not (SGer: $M_{\text{NO\_DIFF}}$ = 0.12, $SD_{\text{NO\_DIFF}}$ = 0.3, $M_{\text{IK\_DIFF}}$ = 0.84, $SD_{\text{IK\_DIFF}}$ = 0.3, $t(158)$ = 1.98, p < .001, USax: $M_{\text{NO\_DIFF}}$ = 0.3, $SD_{\text{NO\_DIFF}}$ = 0.4, $M_{\text{IK\_DIFF}}$ = 0.6, $SD_{\text{IK\_DIFF}}$ = 0.4, $t(142)$ = 1.98, p < .001). For the SGer listeners this was also true for the /-um/ context (SGer: $M_{\text{NO\_DIFF}}$ = 0.35, $SD_{\text{NO\_DIFF}}$ = 0.4, $M_{\text{IK\_DIFF}}$ = 0.6, $SD_{\text{IK\_DIFF}}$ = 0.5, t(*158*) = 1.98, p < .001). Furthermore, SGer listeners perceived IK's /C-al/ fortis token significantly more often as different than her fortis stop in /C-um/ ($M_{\text{-al/}}$ = 0.44, $SD_{\text{-al/}}$ = 0.45, $M_{\text{-um/}}$ = 0.31, $SD_{\text{-um/}}$ = 0.43, $t(318)$ = 1.97, p < .01). USax listeners' detection patterns did not significantly change based on the context.

4. DISCUSSION

The production data (stage I) revealed that VOTs in word-initial fortis /p/ and lenis /b/ largely overlapped for the USax speakers. Even where speakers produced fortis stops with aspiration, VOT did not commonly reach the durations previously described for SGer [11] or even other USax speakers [15]. It has been suggested that the contrast might be more pronounced in younger speakers [15]. The data from this study does not support this hypothesis. Speaker ages varied (KN, CX = 29; MI = 57; IK = 66; GS, ES = 75) and did not correlate with their fortis stop VOT duration. Instead, the data align with traditional descriptions of USax, where fortis stops in word-initial position are unaspirated.

The results from the discrimination task (stage II) showed that listeners tried to rely on aspiration duration to discriminate fortis and lenis stops. Fortis stops with a VOT below 30 ms were not perceived as different from USax lenis stops by either group. Furthermore, no other acoustic cues (e.g., burst intensity, formant transitions [11][16]) seem to have emerged in USax to establish a fortis/lenis contrast.

Even though listeners from both groups perceived stops with VOT above 30 ms as different, the USax group was less likely than the SGer group to label them as 'different'. With increasing VOT, the SGer group's identification of the 'different' token—but not the USax group's —became more and more robust. This suggests that aspiration—a well-established, distinctive cue in the phonological grammar of SGer speakers—was salient to all listeners, but its distinctive function was less familiar to the USax listeners. Contrary to previous phonetic investigations on USax stops [15][21], these data seem to confirm traditional descriptions where aspiration is not a contrastive marker in USax. Further research is needed to determine the status of the laryngeal contrast for alveolars and velars in different positions and to establish a complete picture of the modern USax stop system.

5. ACKNOWLEDGMENTS

I am grateful for comments and feedback from Isabelle Darcy and three anonymous reviewers, as well as for statistical consulting from my colleagues Ryan Lidster and Danielle Daidone.

6. REFERENCES

[1] Albrecht, K. 1965. *Die Leipziger Mundart: Grammatik und Wörterbuch der Leipziger Volkssprache. Zugleich ein Beitrag zur Schilderung der Volkssprache im Allgemeinen.* Leipzig: Zentral-Antiquariat der Deutschen Demokratischen Republik.

[2] Barry, W. J., Pützer, M. 1997. Zur phonetischen Basis der Fortis-Lenis-Opposition bei Plosiven in moselfränkischen und rheinfränkischen Dialekten sowie in Übergangsgebieten im germanophonen Lothringen (Frankreich). *Zeitschrift für Dialektologie und Linguistik* 64, 155-178.

[3] Becker, H., Bergmann, G. 1969. *Sächsische Mundartenkunde: Entstehung, Geschichte und Lautstand der Mundarten des obersächsischen Gebietes* (2. ed.). Halle/Saale: Niemeyer.

[4] Beckman, J. 1998. *Positional faithfulness*. PhD dissertation, University of Massachusetts, Amherst. Published 1999, New York: Garland.

[5] Boersma, P., Weenink, D. 2017. Praat: doing phonetics by computer [Computer program]. Version 6.0.25 from http://www.praat.org/.

[6] Braun, A. (1988). *Zum Merkmal "Fortis/Lenis" - Phonologische Betrachtungen und instrumentalphonetische Untersuchung an einem mittelhessischen Dialekt*. Wiesbaden: Steiner. (Zeitschrift für Dialektologie und Linguistik; Beiheft 55).

[7] De Leeuw, J. R. 2015. jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior research methods* 47, 1-12.

[8] Dufour, S., Nguyen, N., Frauenfelder, U. H. 2007. The perception of phonemic contrasts in a non-native dialect. *J. Acoust. Soc. Am*. 121, EL131-EL136.

[9] Fleischer, W. 1961. *Namen und Mundart im Raum von Dresden*. Berlin: Akademie Verlag.

[10] Grosse, R. 1955. *Die meissnische Sprachlandschaft: dialektgeographische Untersuchungen zur obersächsischen Sprach- und Siedlungsgeschichte,* 15. Halle (Saale): Niemeyer.

[11] Jessen, M. (1998). *Phonetics and phonology of tense and lax obstruents in German* 44. Amsterdam: John Benjamins Publishing.

[12] Jessen, M., Ringen, C. 2002. Laryngeal features in German. *Phonology* 19, 189-218.

[13] John, T. 2004. *Eine akustische Analyse der Lenis/Fortis—Opposition in Varietäten des Sächsischen.* University of Kiel, (Unpublished MA thesis).

[14] Khan, S. u. D., Weise, C. 2013. Upper Saxon (Chemnitz dialect). *Journal of the International Phonetic Association* 43, 231-241. doi:10.1017/S0025100313000145.

[15] 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.

[16] Kohler, K. J. 1979. Parameters in the production and the perception of plosives in German and French. *Arbeitsberichte Kiel* 12, 261-291.

[17] Lameli, A. 2004. Hierarchies of dialect features in a diachronic view: implicational scaling of real time data. *Papers from the Second International Conference on Language Variation in Europe* (ICLaVE 2).

[18] Lein, T., Kupisch, T., van de Weijer, J. 2016. Voice onset time and global foreign accent in German–French simultaneous bilinguals during adulthood. *International Journal of Bilingualism,* 20, 732-749.

[19] Macmillan, N. A., Creelman, C. D. 2004. *Detection theory: A user's guide.* Oxford: Routledge.

[20] Moosmüller, S., Ringen, C. 2004. Voice and aspiration in Austrian German plosives. *Folia Linguistica* 38, 43-62.

[21] Rocholl, M. J. 2015. *Ostmitteldeutsch—eine moderne Regionalsprache?: eine Untersuchung zu Konstanz und Wandel im thüringisch-obersächsischen Sprachraum*. Hildesheim: Georg Olms Verlag.

[22] Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Sung, H., Flege, J. E. 2005. A Developmental Study of English Vowel Production and Perception by Native Korean Adults and Children. *Journal of Phonetics* 33, 263-290.

[23] Vietti, A., Alber, B., Vogt, B. 2018. Initial laryngeal neutralisation in Tyrolean. *Phonology* 35, 79-114.

[i] Her speakers did not produce enough instances of word-initial /p/ to report them in the analysis.

[ii] For each contrast, listeners heard 6 *different*-triplets and 6 *same*-triplets (3 x BBB, 3 x PPP). The task further included another distractor contrast, as well as two target contrasts for the velar stop. Altogether, 144 trials were presented to the participants.

[iii] We ran a *t*-test for IK, as her data were normally distributed.