FINAL LENGTHENING ACROSS VARIOUS SOUND CLASSES IN A LANGUAGE DOCUMENTATION CORPUS OF LOWER SORBIAN

Ludger Paschen

Leibniz-Zentrum Allgemeine Sprachwissenschaft
paschen@leibniz-zas.de

ABSTRACT

This paper investigates segmental lengthening before speech pauses in Lower Sorbian, an endangered Slavic language. The paper compares duration of vowels, sonorants, fricatives and other obstruents at various distances to speech pauses, based on a sample of 15,000 word tokens from a corpus of spontaneous speech. Results indicate lengthening of segments directly adjacent to boundaries for vowels, sonorants, and fricatives, but not for stops and affricates. The effect of lengthening was stronger for more sonorous sounds (vowels, sonorants) than for less sonorous sounds (fricatives). No evidence for lengthening of stressed vowels in (ante-)penultimate syllables was found, suggesting a narrow domain of lengthening that does not extend beyond the final syllable.

Keywords: Final lengthening, Lower Sorbian, corpus phonetics, elasticity.

1. INTRODUCTION

Lower Sorbian is an endangered West Slavic language spoken in the state of Brandenburg in Germany. While its sound inventory shares many common features with other Slavic languages, its phonetics and phonology are severely underdescribed, also compared to the closely related Upper Sorbian language [1, 2, 3, 4, 5]. The goal of this paper is to describe the durational properties of vowels and consonants before pauses in LS to test the scope and degree of *final lengthening* (FL), a common cross-linguistic process which lengthens segments before prosodic boundaries [6, 7, 8, 9]. More specifically, the paper tries to answer the following questions: Does FL in LS proceed progressively, i.e. are segments closer to a boundary more strongly affected than segments with greater distance to a boundary, as described in [10, 9]? Are all types of sounds (vowels, sonorants, fricatives, other obstruents) equally affected by FL or are there differences, as suggested by [11, 9]? Is the effect of FL stronger for stressed syllables compared to unstressed syllables [6, 12]?

2. DATA AND METHODS

2.1. The Lower Sorbian DoReCo dataset

DoReCo is a corpus containing time-aligned speech data from 51 endangered or minority languages from around the world [13]. Audio and annotations are freely available under CC-BY or similar licenses from the website: https://doreco.huma-num.fr/. The Lower Sorbian DoReCo dataset used in this study [14] derives from language documentation data within the DoBeS project archived at TLA [15]. The corpus includes more than 100 hours of transcribed interviews conducted between 2010 and 2015. The Lower Sorbian DoReCo dataset is a subset of the one archived at TLA consisting of 15 recordings of narrative texts from four speakers. It contains a total of 95 minutes of time-aligned transcriptions and 14,597 word tokens, or 41,923 individual phones.

2.2. Annotation and labeling

All recordings from the DoReCo core set feature time alignment on the level of phones, words, and inter-pausal units (IPUs). Time alignment was created using the MAUS forced aligner [16], with manual corrections of word and pause boundaries. DoReCo also includes manually annotated labels for special speech events such as hesitations, false starts, filled pauses, but also code-switching, singing, or unidentifiable stretches of speech (e.g. due to background noise). For the purpose of this study, all labeled tokens were excluded from the analysis.

2.3. Data preparation

A CSV file with phone start and end times from the Lower Sorbian DoReCo dataset was read in RStudio [17]. First, subsets were created for four major sound classes: vowels, sonorants, fricatives, and non-continuant obstruents, i.e. stops and affricates. Then, a number of segments were excluded:

- Labeled segments (see Sec. 2.2)
- Segments with a duration $\leq$ 30 ms
- Segments with a duration $>$ 3sd above the mean of the sound group
- Word-initial segments

The decision to set a global lower duration threshold of $>$ 30 ms was motivated by the fact that 30 ms was set as the minimum segment duration of MAUS, and segments of 30 ms are sometimes the result of last resort, and hence less accurate, alignments. Excluding segments with a duration $>$ 3sd above the mean in the respective sound group was done to remove outliers, again mostly corresponding to alignment errors. Word-initial phones were filtered in order to avoid effects of word-initial strengthening and/or weakening [18, 19]. This exclusion step also accounts for the fact that the onset of stop consonants following silent intervals cannot be determined with sufficient accuracy from the acoustic signal. After filtering, the remaining dataframe contained a total of 23,861 phones, of which there were 10,847 vowels, 6259 sonorants, 2666 fricatives, and 4089 plosives and affricates.

2.4. Statistical modeling

To test the hypothesis that segments are lengthened in the vicinity of speech pauses in Lower Sorbian, linear mixed models were run in R for each sound class (1). Models were run repeatedly with different values of the `position` variable to capture all possible combinations of positional effects, and the obtained significance levels were Bonferroni corrected (α*3).

(1)
```
log(duration) ~ position
+ (1|speaker) + (1|phone)
+ (1|left_context)
+ (1|speech_rate)
+ (1|pause_duration)
```

To account for speaker-specific properties, intrinsic segment duration, and coarticulatory effects, random intercepts for *speaker*, *phone*, and *left context* were included. Local *speech rate* was calculated as phones per second for each IPU. The factor *pause duration* was added to the model because pause duration may reflect the strength of prosodic boundaries, with shorter pauses correlating with weaker boundaries or greater textual cohesion [20].

To test the effect of lexical stress, another model was run on the vowel subset. This model had the same basic structure as in (1), but had the interaction between `position*stress` as the predictor variable, with `stress` being a binary variable distinguishing stressed and unstressed vowels.

3. RESULTS

3.1. An example of FL in Lower Sorbian

This subsection presents an example of context-dependent lengthening of the same word in two excerpts taken from the recording doreco_lowe1385_MEW-113-20121214, which is part of the Lower Sorbian DoReCo dataset [14]. The word *kutšu* 'carriage:ACC.F' appears twice in the examples, once in phrase-medial position (2) and once before a pause (3).

(2) doreco_lowe1385_MEW-113-20121214, 00:32–00:40
ten kněcht to (0.3) jo tu ***kutšu*** *pśigótował a farař se sednuł […] a pón su (.) z kónjom jěli (0.3) tam a cas (0.5)*
"The servant would prepare the carriage and the pastor would sit on it and then they went horseback riding from time to time."

(3) doreco_lowe1385_MEW-113-20121214, 01:48–02:04
ten jo (.) šeł flink domoj (0.5) jo holował (.) jogo (.) ***kutšu*** *(0.6)*
"He quickly went home, fetched his carriage […]."

Figure 1 illustrates the temporal properties of *kutšu* in different prosodic contexts. Image a. on the left corresponds to the token in (2), with a total duration of 520 ms. The duration of the word-final /u/ is 130 ms. In contrast, image b. on the right shows the *kutšu* token in (3), which displays significant lengthening of the final /u/ to 250 ms, nearly twice as much as the phrase-medial vowel in (2). The affricate /tʂ/ and the /u/ in the preceding syllable exhibit only minor traces of lengthening. Note that the onset of the plosive /k/ cannot be reliably determined in b. because it is preceded by a short silent pause.

We thus observe prosodically-driven differences in the acoustic realizations of the same word, with lengthening occurring mostly in the final vowel, and to a lesser extent also in the preceding two segments. The following section will provide a quantitative analysis of lengthening in the Lower Sorbian DoReCo dataset.

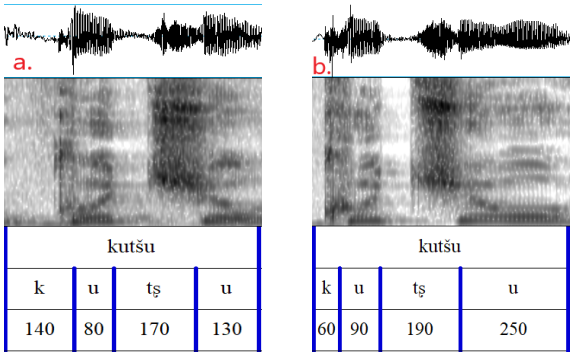

Figure 1: Praat editor windows showing time alignments for *kutšu* 'carriage' a. in phrase-medial position (2) and b. in pre-pausal position (3).

3.2. Quantitative results

Figure 2 shows the duration of vowels in the Lower Sorbian dataset, where "-n" refers to the n-th position before an IPU boundary and "Non-final" refers to all positions outside the prepausal 3-segment window. The boxplot reveals a gradual increase in segmental duration towards the prosodic boundary, and a notably longer duration of prepausal vowels (-1) compared to all other positions (-2, -3, Non-final).

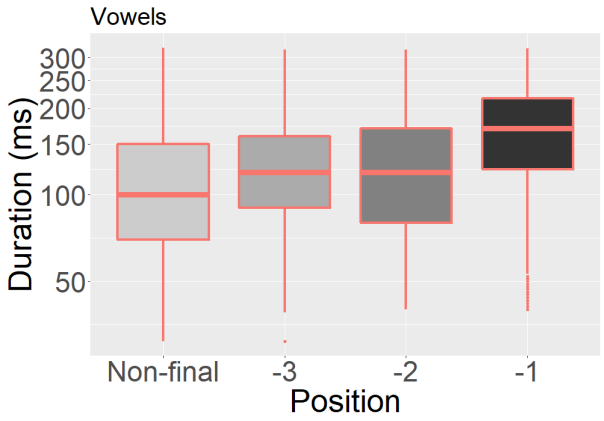

Figure 2: Vowel duration in various positions before IPU boundaries.

Differences between -1 and all other positions pause were highly significant ($p < 0.001$***). No significance was reached for differences between Non-final and -2 ($p = 0.180$); however, differences between Non-final and -3 were significant ($p < 0.01$**).

Sonorants show fairly similar durational patterns to vowels (Fig. 3). Again, a gradual increase in duration towards an IPU boundary can be observed, and segments in close proximity to a boundary (-1) were significantly longer than in any of the other tested positions ($p < 0.001$***). The interactions between other neighboring positions were also significant, albeit not as strongly (-2 vs. -3: $p < 0.05$*, Non-final vs. -3: $p < 0.05$*).

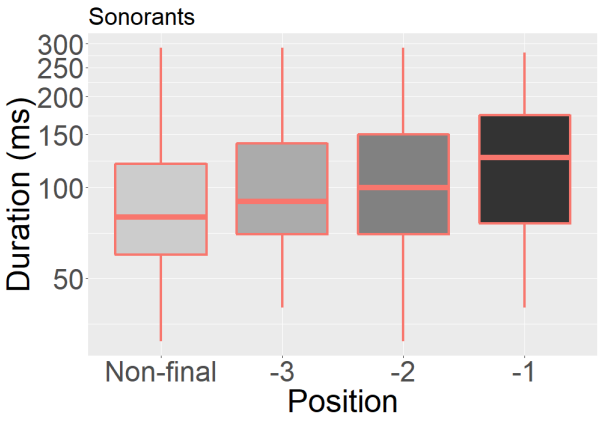

Figure 3: Sonorant duration in various positions before IPU boundaries.

Fricatives behave in a similar way to sonorants. Mean differences between positions were lower, however, suggesting a weaker effect of final lengthening (Fig. 4). Differences between positions reached significance in all but two cases: -2 vs. Non-final ($p = 0.160$) and -2 vs. -3 ($p = 0.864$).

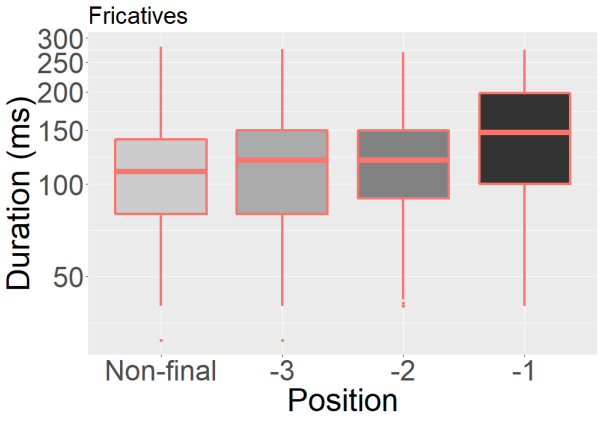

Figure 4: Fricative duration in various positions before IPU boundaries.

A divergent picture is presented by non-continuant obstruents (Fig. 5). No clear trend for durational differences is visible and durational differences did not reach significance, with one exception (-2 vs. -3, $p < 0.05$*). It is not clear why the -3 position shows a longer duration than the -2 position. One would expect that if any effect of FL exists at all for stops or affricates, it should be between Non-final and -1 rather than between -2 and -3. A potential explanation could be the phonotactics of Lower Sorbian, which heavily restricts the types of consonants that can appear word-finally. In the sample analyzed in this study, only three stops /p, t, k/ and two affricates /ts, tɕ/ occur in word-final position. The absence of voiced and palatalized sounds, which are generally associated with more

complex laryngeal and articulatory movements and thus greater intrinsic duration, may explain the lack of an effect for the -1 position in this group of sounds.

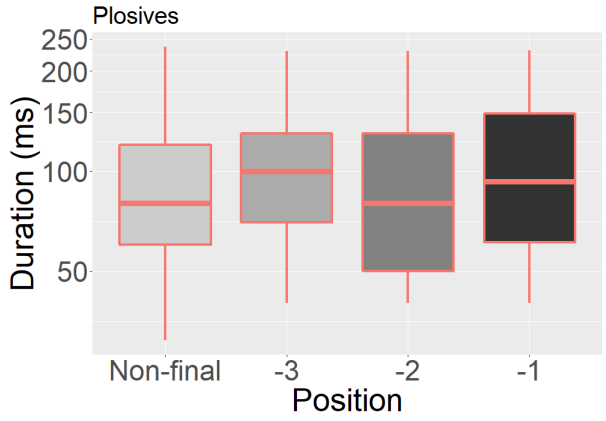

Figure 5: Stop and affricate duration in various positions before IPU boundaries.

Fig. 6 shows the duration of vowels in various syllable positions, split into unstressed and stressed syllables. Unsurprisingly, vowels in boundary-adjacent syllables (σ-1) were on average longer than in other syllables.Linear mixed models revealed several significant interactions both within and across position and stress groups. Depending on the base level, though, not all interactions involving stress reached significance, especially for the σ-2 and σ-3 positions. Estimates for σ-2 and σ-3 were shorter than for σ-1 and Non-final, suggesting that final lengthening in Lower Sorbian does not extend to the penultimate and antepenultimate syllables even when a stressed syllable occurs as a potential attractor in those positions.

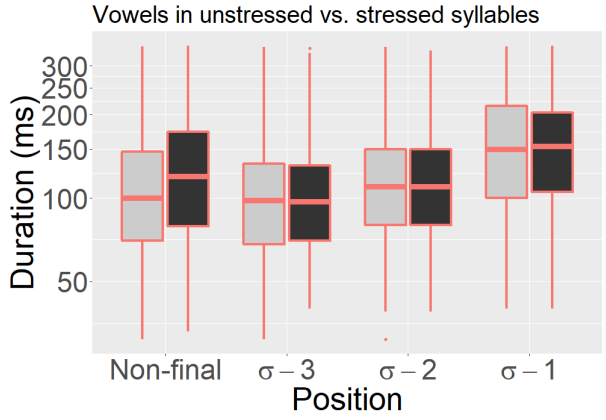

Figure 6: Vowel duration in unstressed (grey boxes) and stressed (black boxes) syllables in various positions before an IPU boundary.

4. DISCUSSION

Final lengthening in Lower Sorbian can be compared to other Slavic languages and to German, its major contact language. [21] reports longer syllable duration before prosodic boundaries in Polish. [22] describe lengthening of vowels and onset consonants in the two pre-boundary syllables in Czech, and an intricate interaction between final lengthening and phonological length. [23, 24] observe vocalic lengthening before prosodic boundaries and an increased effect before boundaries accompanied by pauses in Russian. They also report differences between stressed and unstressed vowels. [12] note systematic lengthening of IP-final vowels in Serbian.

For German, [9] present articulatory and acoustic evidence for lengthening of vowels and onset consonants, whereby lengthening is stronger in the ultimate syllable and more sonorous sounds are lengthened to a greater extent than less sonorous segments. This is well in line with the results obtained for Lower Sorbian, and confirms previous findings on general segmental elasticity [25, 26] as well as the relevance of those features for FL [11]. It should be noted, though, that the above mentioned studies on Slavic and German are experimental studies based on stimulus-based reading tasks, while the Lower Sorbian dataset consists of interviews and narrative texts. Ideally, as more experimental and corpus data become available in the future, these two types of data can complement each other by highlighting different aspects of the temporal organization of speech. Another promising point of departure for future research could be a closer look at various types of pauses such as silent, filled, and breath pauses in order to shed more light on the relation between speech planning, discourse structure, and the durational properties of segments.

5. CONCLUSION

This paper has offered an account of final lengthening in Lower Sorbian. Final Lengthening affects vowels, sonorants, and fricatives, and its effect is contingent on both the sonority of a segment and on the distance to an utterance boundary. The Lower Sorbian data show similarities with patterns reported for other Slavic languages and German. The study exemplifies the reusability of speech corpora originally created with the aim of language documentation.

6. ACKNOWLEDGEMENTS

This research was supported by a grant from the German Research Foundation (DFG-PA2368/1-1).

7. REFERENCES

[1] P. J. Howson, "An acoustic examination of the three-way sibilant contrast in Lower Sorbian," in *Proc. Interspeech 2015*, 2015, pp. 2670–2674.

[2] ——, "Upper Sorbian," *Journal of the International Phonetic Association*, vol. 47, no. 3, pp. 359—367, 2017.

[3] ——, "Rhotics and palatalization: an acoustic examination of Upper and Lower Sorbian," *Phonetica*, vol. 75, pp. 132—150, 2018.

[4] P. J. Howson and A. Kochetov, "Lowered F2 observed in uvular rhotics involves a tongue root gesture: Evidence from Upper Sorbian," *The Journal of the Acoustical Society of America*, vol. 147, p. 2845, 2020.

[5] I. Kraljevski, M. P. Bissiri, F. Duckhorn, C. Tschoepe, and M. Wolff, "Glottal stops in Upper Sorbian: a data-driven approach," in *Proc. Interspeech 2021*, 2021, pp. 1001–1005.

[6] A. E. Turk and S. Shattuck-Hufnagel, "Multiple targets of phrase-final lengthening in American English words," *Journal of Phonetics*, vol. 35, no. 4, pp. 445–472, 2007.

[7] J. Fletcher, "The Prosody of Speech: Timing and Rhythm," in *The Handbook of Phonetic Sciences, Second Edition*, W. J. Hardcastle, J. Laver, and F. E. Gibbon, Eds. Chichester: Blackwell, 2010, pp. 521–602.

[8] S. Nakai, A. E. Turk, K. Suomi, S. Granlund, R. Ylitalo, and S. Kunnari, "Quantity constraints on the temporal implementation of phrasal prosody in Northern Finnish," *Journal of Phonetics*, vol. 40, no. 6, pp. 796–807, 2012.

[9] M. Belz, O. Rasskazova, J. Krivokapić, and C. Mooshammer, "Interaction between phrasal structure and vowel tenseness in German: An acoustic and articulatory study," *Language and Speech*, pp. 1–32, 2022.

[10] D. Byrd, J. Krivokapić, and S. Lee, "How far, how long: On the temporal scope of prosodic boundary effects," *The Journal of the Acoustical Society of America*, vol. 120, no. 3, pp. 1589–1599, 2006.

[11] J. P. van Santen, "Assignment of segmental duration in text-to-speech synthesis," *Computer Speech & Language*, vol. 8, no. 2, pp. 95–128, 1994.

[12] B. Jakovljević and M. Marković, "Properties of I-boundary lengthening of vowels in English and Serbian," *Annual Review of the Faculty of Philosophy, Novi Sad*, vol. 45, no. 5, 2020.

[13] F. Seifart, L. Paschen, and M. Stave, Eds., *Language Documentation Reference Corpus (DoReCo) 1.1*. Berlin & Lyon: Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), 2022.

[14] H. Bartels and M. Szczepański, "Lower sorbian doreco dataset," in *Language Documentation Reference Corpus (DoReCo) 1.1*, F. Seifart, L. Paschen, and M. Stave, Eds. Berlin & Lyon: Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2), 2022. [Online]. Available: https://doreco.huma-num.fr/languages/lowe1385

[15] J. Měškank, M. Elikowska-Winklerowa, G. Wieczorek, K. Thorquindt-Stumpf, S. Krawcojc, L. Jocz, S. Elikowski-Winkler, J. Kaulfürstowa, M. Kóńcaŕ, F. Kaulfürst, and J.-F. Käthlitz, *Collection "Lower Sorbian"*. The Language Archive, 2011–2015. [Online]. Available: https://hdl.handle.net/1839/b7cb7b41-8362-4488-b8cb-86fbfea6c7c7

[16] F. Schiel, C. Draxler, and J. Harrington, "Phonemic segmentation and labelling using the maus technique," in *Workshop New Tools and Methods for Very-Large-Scale Phonetics Research*, 2011.

[17] RStudio Team, *RStudio: Integrated Development Environment for R*, RStudio, PBC, Boston, MA, 2020. [Online]. Available: http://www.rstudio.com/

[18] L. White, S. Benavides-Varela, and K. Mády, "Are initial-consonant lengthening and final-vowel lengthening both universal word segmentation cues?" *Journal of Phonetics*, vol. 81, no. 100982, 2020.

[19] C. DiCanio and J. Sharp, "Initial weakening in Mixtecan languages," *The Journal of the Acoustical Society of America*, vol. 148, p. 2273, 2020.

[20] B. Andreeva, B. Möbius, and J. Whang, "Effects of surprisal and boundary strength on phrase-final lengthening," in *Proc. of Speech Prosody 10*, Tokyo, Japan, 2020, pp. 146–150.

[21] A. Wagner, "Description of polish speech rhythm using rhythm metrics and the time-delay approach: a comparative study," in *Proceedings of the 7th International Conference on Speech Prosody*, Dublin, Ireland, 2014, pp. 366–370.

[22] N. Spina and F. Schubö, "Vowel length affects pre-boundary lengthening in Czech," in *Proc. 1st International Conference on Tone and Intonation (TAI), Sonderborg, Denmark*, 2021.

[23] T. Kachkovskaia, N. B. Volskaya, and P. A. Skrelin, "Final lengthening in Russian: a corpus-based study," in *INTERSPEECH 2013, 14th Annual Conference of the International Speech Communication Association*. Lyon, France: ISCA, 2013, pp. 1438–1442.

[24] T. Kachkovskaia, "Phrase-final lengthening in Russian: Pre-boundary or pre-pausal?" in *Speech and Computer*, A. Ronzhin, R. Potapova, and V. Delic, Eds., pp. 353–359.

[25] D. H. Klatt, "Synthesis by rule of segmental durations in English sentences," *Frontiers of Speech Communication*, pp. 290–297, 1979.

[26] N. Campbell, "Segmental elasticity and timing in Japanese speech," in *Speech perception, production, and linguistic structure*, Y. Tohkura, E. Vatikiotis-Bateson, and Y. Sagisaka, Eds. Tokyo: Ohmsha, 1992, pp. 403–418.