Investigating the relationship between accentuation, vowel tensity and compensatory shortening

Jessica Siddins, Jonathan Harrington, Ulrich Reubold, Felicitas Kleber

Institute of Phonetics and Speech Processing, Ludwig-Maximilians-Universität Munich, Germany

`jessica | jmh | reubold | kleber@phonetik.uni-muenchen.de`

Abstract

The aim of this study was to investigate the relationship between compensatory shortening and coarticulation in German tense and lax vowels and to determine whether this relationship was influenced by prosodic accentuation. While previous studies focussed on temporal vowel reduction due to compensatory shortening, and often found conflicting results, our study extends previous results by including a formant analysis of spatial reduction in two types of compensatory shortening. Specifically, we tested for polysyllabic shortening (monosyllabic vs. disyllabic words) and incremental coda shortening (words with final singletons vs. final clusters). Speakers produced minimal pairs differing in vowel tensity in accented and deaccented contexts for both shortening conditions. Vowel duration was influenced primarily by vowel tensity as well as by accentual lengthening for tense but not lax vowels. While vowel duration was not affected by compensatory shortening, formant analyses revealed an effect of coda cluster for tense vowels as well as clear effects of accentuation and vowel tensity. There was no effect of polysyllabic shortening on formants.

Further to previous studies on compensatory shortening, these results reveal that compensatory shortening is not limited to temporal reduction, but can have an impact on vowel quality as well.

Index Terms: coarticulation, speech timing, compensatory shortening, accentual lengthening, target undershoot, sound change

1. Introduction

This study investigated the effects of two types of compensatory shortening on the acoustic duration and formant values of vowels in rhythmically strong syllables. Our general goal within a larger series of experiments is to better understand the relationship between prosodic weakening, coarticulation and sound change.

The first type of compensatory shortening we investigated, polysyllabic shortening, refers to the compression of a vowel spoken in a polysyllabic compared with a monosyllabic word. Thus, the vowel in English *sleep* is longer than the same vowel in English *sleepy* [1]. Polysyllabic shortening has been well-documented in Germanic languages such as English [1, 2, 3, 4, 5, 6], Dutch [7] and Swedish [8, 9, 10, 11, 12], although [13] found no evidence of polysyllabic shortening in a study of read speech in English.

The second type of compensatory shortening we investigated, incremental coda shortening, refers to the compression of a vowel spoken before a consonant cluster compared with a consonant singleton. [14] found acoustic shortening of vowels before consonant clusters in their study of three speakers of

English, and [15] replicated these effects for obstruent clusters (but not for sonorant clusters in syllable-final position). A recent study on German by [16] found no results of incremental coda shortening in production, but they did find that listeners expected shorter vowel durations before complex clusters in a perception experiment.

Prior research has largely failed to establish the effect of accentuation on compensatory shortening (see [17]), in that studies have mostly examined stressed syllables or accented words only. In German, as in other Germanic languages, a word is accented when a pitch accent is associated with the rhythmically strongest syllable [18]. In addition to the f0 changes caused by a pitch accent and any adjacent boundary tones, pitch accented syllables in many Germanic languages are also hyperarticulated and produced with greater duration [19].

Thus, this study aimed to investigate vowel reduction and the extent to which compensatory shortening interacts with prosodic accentuation and vowel tensity. Based on previous studies, we had three main hypotheses for the durational analysis as well as analogous hypotheses for the formant analysis. Firstly, we expected acoustic vowel shortening and thus also vowel undershoot before coda clusters compared with coda singletons [14, 15] as well as in disyllabic compared with monosyllabic words [11, 20]. Secondly, we expected compensatory shortening and undershoot induced by this shortening to be more marked in accented than in deaccented contexts [4, 17, 21, 22]. Thirdly, we expected more shortening and hence more shortening-induced undershoot of tense than of lax vowels [4, 23, 24].

2. Method

2.1. Participants

Twenty-nine L1 speakers of Standard German (11 male, 18 female) were recorded at a sampling rate of 44 100 Hz in a sound-attenuated booth. Speakers had no known speech or language impairments and were paid for their participation.

2.2. Stimuli

We chose a tense-lax target vowel pair believed to vary mostly in quantity and only minimally in quality in order to study the effects of compensatory shortening. In German, tense vowels are phonologically long, while lax vowels are phonologically short, for example /ˈbiːtn̩/ (to offer) vs. /ˈbɪtn̩/ (to request). The German vowel that differs least in quality between its tense and lax versions is the open central /a/, which varies mainly in vowel height (F1), for example a lower F1 for lax *Kamm* /kam/ (comb) and a higher F1 for tense *kam* /kaːm/ (came) [23, 24]. [23, p341] claim that even the durational contrast is minimised in prosodically weak contexts. Our stimuli were real German words with

lax /a/ and tense /aː/. We restricted the target vowels to one consonantal context in order to avoid varying effects of CVC coarticulation. The target words were embedded in phrase-medial position in the carrier sentence *Anna hatte [target word] verstanden* (Anna understood [target word]) in order to avoid utterance or phrase-final lengthening.

	Cluster Shortening		
Lax Vowels	/zak/	/zakt/	/zaktə/
Tense Vowels	/zaːk/	/zaːkt/	/zaːktə/
		Polysyllabic Shortening	

Table 1: The 6 target words, spoken in both accented (A) and deaccented (U) contexts (= 12 target words).

A minimal pair paradigm was created using one consonantal context adjusted for Syllabicity (monosyllabic vs. disyllabic), Coda (final singleton vs. final cluster), Accentuation (accented vs. deaccented) and Vowel Tensity (tense vs. lax). The monosyllabic level of the Syllabicity condition overlapped with the cluster level of the Coda condition (see Table 1), leading to a total of 12 target items. Each speaker produced 10 repetitions of each item, leading to a total of 120 utterance tokens for each speaker (29 subjects x 120 utterances = 3 480 utterances). Filler words and sentences were also included to disguise the object of the experiment.

2.3. Procedure

The stimuli were presented and recorded in randomised order using SpeechRecorder software [25]. Participants were first presented with a question designed to elicit a narrow focus on the target word for the accented context and a broad focus for the deaccented context: either *WAS hatte Anna verstanden?* (WHAT did Anna understand?) or *WER hatte [target word] verstanden?* (WHO understood [target word]?). The target sentence was then presented with the accented word in capital letters. The experimenter asked subjects to repeat the sentence if they made a mistake (either segmentally or suprasegmentally).

Three subjects were eliminated from further analysis as they were unable to elicit the correct accentuation patterns of the stimuli. The remaining 26 speakers (9 males, 17 females; mean age 24 years) were included in the analysis.

The entire corpus was automatically segmented and labelled using the Munich Automatic Segmentation System [26] and corrected manually. During this procedure several rare cases of incorrect accentuation were discovered and removed from the database.

2.4. Analysis

For the durational analysis, the dependent variable was normalised vowel duration. We normalised the vowel durations by dividing the duration of each target vowel by the duration of each /a/ in the utterance-final word *verstanden* in order to control for changes in speech tempo throughout the experiment. *Verstanden* was chosen as it was considered least likely to be affected by the varying accentuation pattern, which was confirmed by a visual analysis of the data. The normalised vowel duration was then averaged per speaker and per condition.

For the formant analysis, the dependent variable was maximum F1 measured in Hz from the central third of the vowel. We calculated formants with a frame shift of 5 ms and a window length of 12.5 ms for female speakers and 20 ms for male

speakers using Emu [27]. Formant errors were hand-corrected in Emu when necessary.

Both dependent variables were tested individually alongside within-subjects factors Vowel Tensity, Accentuation and Syllabicity (for the polysyllabic shortening analyses) or Coda (for the cluster shortening analyses) and random factor Speaker in a repeated measures ANOVA design using the `ez` package in R [28, 29].

3. Results

3.1. Cluster Shortening

For the cluster shortening condition, we compared the monosyllables with and without final clusters (the stimuli coloured light grey in Table 1). Disyllabic stimuli were excluded from analysis.
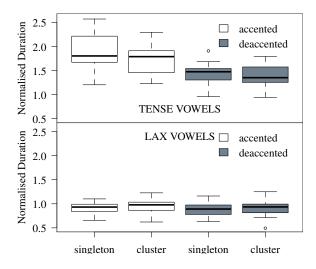

Figure 1: Effects of incremental coda shortening, accentuation and vowel tensity on normalised vowel duration

The durational analysis revealed main effects of both Vowel Tensity ($F[1, 25] = 278.7; p < .001$) (top vs. bottom of Figure 1) and Accentuation ($F[1, 25] = 54.6; p < .001$) (white vs. grey in Figure 1), but there was no effect of Coda on the normalised vowel duration. We found significant interactions between Vowel Tensity and Coda ($F[1, 25] = 6.3; p < .05$) and Vowel Tensity and Accentuation ($F[1, 25] = 82.2; p < .001$). Post-hoc Bonferroni-corrected t-tests revealed accentual lengthening of tense ($p < .001$) but not lax vowels. The post-hoc tests did not reveal the reason for the interaction between Vowel Tensity and Coda, but Figure 1 shows a slight tendency toward compensatory shortening of tense vowels which cannot be seen for lax vowels.

The formant analysis revealed main effects of Vowel Tensity ($F[1, 25] = 105.7; p < .001$) (black vs. grey in Figure 2), Accentuation ($F[1, 25] = 102.8; p < .001$) (solid vs. dashed in Figure 2) and Coda ($F[1, 25] = 5.8; p < .05$) on the maximum F1, with a significant interaction between Vowel Tensity and Coda ($F[1, 25] = 8.5; p < .01$). Pairwise post-hoc comparisons for the Vowel Tensity vs. Coda interaction indicated that the effect of Coda on F1 is restricted to tense vowels only, although the effect is weak (see Figure 3).
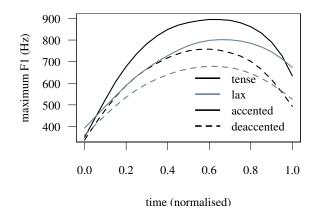
Figure 2: F1 trajectories (linear time-normalised) for the cluster shortening experiment as a function of vowel tensity and accentuation
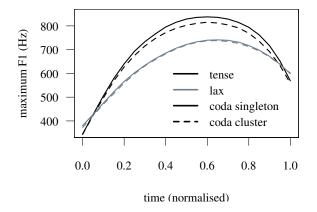

Figure 3: F1 trajectories (linear time-normalised) for the cluster shortening experiment as a function of vowel tensity and coda condition

Figure 4: Effects of polysyllabic shortening, accentuation and vowel tensity on normalised vowel duration

Figure 5: F1 trajectories (linear time-normalised) for the polysyllabic shortening experiment as a function of vowel tensity and accentuation

3.2. Polysyllabic Shortening

For the polysyllabic shortening condition, we excluded all monosyllabic words with final singleton from the analysis and compared the monosyllabic and disyllabic words with clusters at the end of the first syllable (the stimuli coloured dark grey in Table 1).

While we found main effects of both Vowel Tensity ($F[1, 25] = 33.2; p < .001$) (top vs. bottom of Figure 4) and Accentuation ($F[1, 25] = 401.9; p < .001$) (white vs. grey in Figure 4), there was no effect of Syllabicity on the normalised vowel duration. We also found a significant interaction between Vowel Tensity and Accentuation ($F[1, 25] = 31.6; p < .001$). Post-hoc Bonferroni-corrected t-tests revealed a significant effect of accentual lengthening on tense vowels ($p < .001$), but not on lax vowels.

The F1 analysis found main effects of Vowel Tensity ($F[1, 25] = 100; p < .001$) (black vs. grey in Figure 5) and Accentuation ($F[1, 25] = 103.7; p < .001$) (solid vs. dashed in Figure 5), but not of Syllabicity. There was a significant interaction between Vowel Tensity and Accentuation ($F[1, 25] = 9.3; p < .01$). In view of Figure 5, this interaction is likely due to a slightly larger difference between tense and lax vowels in accented contexts than in deaccented contexts (see Figure 5).
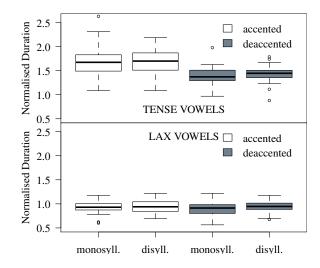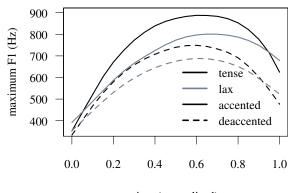
4. Discussion

We expected compensatory shortening and target undershoot of vowels before coda clusters (incremental coda shortening) and in disyllabic words (polysyllabic shortening), and for these effects to be more marked in accented than in deaccented words. In addition, we expected tense vowels to undergo more compensatory shortening and shortening-induced vowel reduction than lax vowels.

4.1. Vowel Tensity

We confirmed that the tense-lax distinction (at least of low vowels) is defined by both vowel quantity and quality in German.

The quantity distinction was maintained even in deaccented contexts, and there was accentual lengthening of tense but not lax vowels. In terms of quality, we found lower first formants in lax than in tense vowels and in deaccented tense vowels than in accented tense vowels (black vs. grey in Figures 2, 3 & 5), contrary to [24, p14]'s conclusion that "tense and lax low vowels [i.e. /a, aː/] in [Standard] German differ consistently only in duration, but not in formant structure". A clear distinction in F1 between tense and lax low vowels may be necessary because the

durational distinction between tense and lax vowels lessens (but is not neutralised) in deaccented contexts (see Figures 1 and 4; see also [23, 24]).

4.2. Accentuation

We found clear effects of accentual lengthening of tense but not lax vowels. In line with [30], who found larger jaw movements in stressed than in unstressed syllables, there was a significant effect of accentuation on the F1 of both tense and lax vowels, and this effect was slightly stronger for tense vowels (see also [23]). As a result, there is less difference in duration between tense and lax /aː/ and /a/ in deaccented contexts. In Figures 2 and 5, a large amount of overlap between tense and lax vowels is visible: /a/ has a higher first formant (indicating greater jaw opening and greater tensity) in the accented condition than /aː/ in the deaccented condition.

4.3. Cluster Shortening

Contrary to our hypothesis and some previous studies [14, 15], we found no effect of complex coda on the duration of the target vowel. Our findings thus match those of [16] and are compatible with the c-centre hypothesis [31], which predicts incremental onset shortening but not incremental coda shortening.

The conflicting results of incremental coda shortening emphasise the need to investigate whether shortening in fact induces other types of reduction. Indeed, we did find significant spatial reduction of F1 in tense vowels before complex clusters. That is, incremental coda shortening (at least in German) appears to induce F1 undershoot of primary-stressed tense vowels, even if there is no durational shortening. This provides support for our hypothesis that there is greater undershoot before clusters than before singletons, thus leading to a smaller difference between tense and lax vowels before coda clusters (see Figure 3).

There is no evidence for our hypothesis that undershoot is greater in accented than in deaccented contexts: that is, the degree of separation in F1 between tense and lax vowels before singletons or clusters is largely unaffected by accentuation. Thus, while accentuation does influence vowel quality and quantity, it does not interact with incremental coda shortening.

4.4. Polysyllabic Shortening

In this study, we found no effect of polysyllabic shortening on vowel duration. This result is contrary to the main body of research on polysyllabic shortening, but in line with [13], who found no polysyllabic shortening in connected speech. However, their study differed from others in that it was based on syllable shortening in stress groups rather than in polysyllabic words. In addition, they counted syllables with secondary stress as stressed syllables.

According to [4]'s incompressibility theory, "a vowel that is shortened by one rule becomes less compressible to additional shortening influences" [4, p1103]. As a result, one might argue that there was no polysyllabic shortening of the target vowel because both the monosyllabic and the disyllabic condition contained a final coda cluster, which can induce incremental coda shortening. However, as there was no effect of incremental coda shortening on vowel duration, it is unlikely that the final coda cluster prevented shortening of the target vowel in the disyllable compared with the monosyllable.

In addition, there was no effect of polysyllabic shortening on the first formant, which is known to be correlated with jaw opening. There is no evidence from our results that polysyllabic shortening induces greater durational compression in accented contexts than in deaccented contexts and in tense vowels than in lax vowels. As a result, there is no evidence from this study that polysyllabic shortening induces any type of vowel reduction, neither temporal nor spatial.

4.5. Implications for Sound Change

The results of this study show slightly less durational contrast between tense and lax vowels in deaccented contexts (white vs. grey in Figures 1 and 4) and a slight decrease in the quality contrast before coda clusters (see Figure 3). In general, there is a great amount of overlap in the first formant of accented lax vowels and deaccented tense vowels (see Figures 2 and 5).

If tense and lax vowels are more difficult to distinguish before clusters based on their vowel quality, this might be the reason the durational contrast between tense and lax vowels was maintained in our analysis of acoustic shortening before coda clusters. Alternatively, the tense-lax contrast may be diminishing in German [32], and this might be a context in which listeners could misperceive the vowel, leading to a mini sound change [33].

4.6. Outlook

In view of the above, there may be a greater risk of confusing /aː/ and /a/ before clusters, not because of acoustic shortening, but because of F1 undershoot. In order to determine how listeners parse this undershoot, we are now running perception experiments in which a synthetic vowel continuum is created and spliced into two contexts: /zaːk/-/zak/ and /zaːkt/-/zakt/. For the target vowel, we chose an ambiguous vowel duration (the mean of tense and lax tokens of a speaker) and varied only the height of the first formant. If listeners attribute vowel undershoot before coda clusters to compensatory shortening, we would expect a higher F1 at the category boundary for the continuum ending in coda cluster than the one ending in coda singleton. Alternatively, if listeners incorrectly attribute vowel undershoot before coda clusters to (lax) vowel quality, i.e. they are not aware of the effects of compensatory shortening on vowel quality, we would expect the continuum before coda singleton and the continuum before coda clusters to have the same category boundary.

5. Conclusions

This study examined the relationship between vowel tensity, accentuation and compensatory shortening and in particular how these factors affect vowel quantity and quality. At least for the vowel pair we tested, vowel tensity in German is clearly marked by both duration and quality. In addition, accentuation affects both duration and quality, but not equally for tense and lax vowels. This asymmetry leads to considerable overlap between tense and lax vowels. We found no effects of compensatory shortening on duration, but we did find F1 undershoot before coda clusters. Thus, the quality contrast is diminished before clusters. A perception experiment is being carried out to determine whether or not listeners correctly attribute cluster-induced vowel undershoot to its source. If not, this context could be a source of sound change.

6. Acknowledgements

7. References

[1] I. Lehiste, "The timing of utterances and linguistic boundaries," *JASA*, vol. 51, pp. 2018–2024, 1972.

[2] D. Jones, "Chronemes and tonemes," *Acta Linguistica*, vol. 4, no. 1, pp. 11–10, 1944.

[3] T. P. Barnwell, "An algorithm for segment durations in a reading machine context," Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts, Tech. Rep. 479, 1971.

[4] D. H. Klatt, "Interaction between two factors that influence vowel duration," *The Journal of the Acoustical Society of America*, vol. 54, no. 4, pp. 1102–1104, 1973.

[5] ——, "Linguistic uses of segmental duration in English: Acoustic and perceptual evidence," *The Journal of the Acoustical Society of America*, vol. 59, no. 5, pp. 1208–1221, 1976.

[6] R. F. Port, "Linguistic timing factors in combination," *The Journal of the Acoustical Society of America*, vol. 69, no. 1, pp. 262–274, 1981.

[7] S. G. Nooteboom, "Production and perception of vowel duration: a study of durational properties of vowels in Dutch." Ph.D. dissertation, University of Utrecht, 1972.

[8] S. Öhman, "Syllabic function of vowel length," *STL-QPSR*, vol. 1, pp. 7–9, 1961.

[9] C. C. Elert, *Phonologic Studies of Quantity in Swedish*. Stockholm: Almqvist & Wiksell, 1964.

[10] B. Lindblom, "Temporal organization of syllable production," *Speech Transmission Lab. Quarterly Progress Status Report*, vol. 2, no. 3, pp. 1–5, 1968.

[11] B. Lindblom and K. Rapp, "Reexamination of the compensatory adjustment of vowel duration in Swedish words," *Occasional Papers, University of Essex*, vol. 13, pp. 204–224, 1972.

[12] ——, "Some temporal regularities of spoken Swedish," *Papers in Linguistics from the University of Stockholm*, vol. 21, pp. 1–59, 1973.

[13] T. H. Crystal and A. S. House, "Articulation rate and the duration of syllables and stress groups in connected speech," *Journal of the Acoustical Society of America*, vol. 88, pp. 101–112, 1990.

[14] K. Munhall, C. H. Fowler, S. Hawkins, and E. Saltzman, ""Compensatory shortening" in monosyllables of spoken English." *Journal of Phonetics*, vol. 20, no. 2, pp. 225–239, 1992.

[15] S. Marin and M. Pouplier, "Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model," *Motor Control*, vol. 14, no. 3, pp. 380–407, 2010.

[16] S. Peters and F. Kleber, "Compensatory vowel shortening before complex coda clusters in the production and perception of german monosyllables," *The Journal of the Acoustical Society of America*, vol. 134, no. 5, pp. 4202–4202, 2013.

[17] L. White, "English speech timing: a domain and locus approach," Ph.D. dissertation, University of Edinburgh, 2002.

[18] J. Pierrehumbert and M. Beckman, *Japanese Tone Structure*. Cambridge: MIT Press, 1988, vol. Linguistic Inquiry Monograph 15.

[19] K. de Jong, "The supraglottal articulation of prominence in english: Linguistic stress as localized hyperarticulation," *Journal of the Acoustical Society of America*, vol. 97, no. 1, pp. 491–504, 1995.

[20] I. Lehiste, *Suprasegmentals*. Cambridge, Massachusetts: The MIT Press, 1970.

[21] A. E. Turk and S. Shattuck-Hufnagel, "Word-boundary-related duration patterns in English," *Journal of Phonetics*, vol. 28, no. 4, pp. 397 – 440, 2000.

[22] L. White and A. E. Turk, "English words on the Procrustean bed: Polysyllabic shortening reconsidered," *Journal of Phonetics*, vol. 38, pp. 459–471, 2010.

[23] C. Mooshammer and S. Fuchs, "Stress distinction in German: simulating kinematic parameters of tongue-tip gestures," *Journal of Phonetics*, vol. 30, no. 3, pp. 337 – 355, 2002.

[24] M. Jessen, "Stress conditions on vowel quality and quantity in German," *Working Papers of the Cornell Phonetics Laboratory*, vol. 8, pp. 1–27, 1993.

[25] C. Draxler and K. Jänsch, "SpeechRecorder - a Universal Platform Independent Multi-Channel Audio Recording Software," in *Proc. of the IV. International Conference on Language Resources and Evaluation*, Lisbon, Portugal, 2004, pp. 559–562.

[26] F. Schiel, C. Draxler, and J. Harrington, "Phonemic segmentation and labelling using the MAUS technique," Philadelphia, PA, USA, 2011. [Online]. Available: http://epub.ub.uni-muenchen.de/13684/

[27] J. Harrington, *Phonetic Analysis of Speech Corpora*. Wiley Publishing, 2010.

[28] M. A. Lawrence, "Package 'ez'," Sep. 2013. [Online]. Available: http://cran.r-project.org/web/packages/ez/ez.pdf

[29] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2011, ISBN 3-900051-07-0. [Online]. Available: http://www.R-project.org/

[30] R. Kent and R. Netsell, "Effects of stress contrasts on certain articulatory parameters," *Phonetica*, vol. 24, pp. 23–44, 1971.

[31] C. P. Browman and L. Goldstein, "Competing constraints on intergestural coordination and self-organization of phonological structures," *Bulletin de la Communication Parlee*, vol. 5, pp. 25–34, 2000.

[32] G. Seiler, "How contrastive vowel quantity can become non-contrastive," in *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, vol. 40, no. 1. Chicago Linguistic Society, 2004, pp. 349–363.

[33] J. J. Ohala, "The listener as a source of sound change," in *Papers from the Parasession on Language and Behavior*, C. S. Masek, R. A. Hendrick, and M. F. Miller, Eds. Chicago: Chicago Linguistic Society, 1981, pp. 178–203.