
UNSTRESSED VOWELS IN GERMAN LEARNER ENGLISH: AN INSTRUMENTAL STUDY

LUKAS SÖNNING

University of Bamberg

lukas.soenning@uni-bamberg.de

Abstract

This study investigates the production of vowels in unstressed syllables by advanced German learners of English in comparison with native speakers of Standard Southern British English. Two acoustic properties were measured: duration and formant structure. The results indicate that duration of unstressed vowels is similar in the two groups, though there is some variation depending on the phonetic context. In terms of formant structure, learners produce slightly higher F_1 and considerably lower F_2 , the difference in F_2 being statistically significant for each learner. Formant values varied as a function of context and orthographic representation of the vowel.

Keywords: vowel reduction, acoustic analysis, SLA, L1 German, L2 English

1. Introduction

In English, vowels occurring in unstressed syllables are reduced – they are articulated with a more central position of the tongue, a narrower jaw-opening and a loss of lip rounding (Delattre 1982). Acoustically, this is reflected in their duration and formant structure. One characteristic feature of German Learner English (GLE) is a lack of vowel reduction in unstressed syllables (e.g. Parkes 2001). This paper gives an exploratory account of the production of vowels in unstressed syllables by advanced German Learners of English in comparison with native speakers of Standard Southern British English. Two acoustic properties are investigated: duration and formant structure.

The paper will start out with an overview of instrumental approaches to speech rhythm, a prosodic construct that is closely related to the properties of unstressed vowels and has received considerable attention by researchers in the past two decades. Though located at the same end of the rhythmic continuum, English and German differ with respect to the properties of unstressed vowels. A brief account of relevant contrasts will be given. With transfer from L1 being an important factor in L2 phonological acquisition (e.g. Major 2008), these contrasts are assumed to play a role in the interlanguage of German learners. While research on unstressed vowels in GLE remains scarce, studies including other L1-L2 combinations have shed light on possible universal tendencies in learner speech. Based on the contrastive account as well as the findings of previous studies, two hypotheses are formulated and tested in this study.

2. Background

2.1 Speech rhythm

The occurrence of unstressed vowels in connected speech is closely related to the rhythmical properties of a language or accent. Speech rhythm is generally defined as “a perceived regularity of prominent units in speech” (Crystal 2008: 417). Acoustic correlates of prominence are duration, vowel quality, pitch, and intensity. A long-standing claim is that languages can be grouped into two rhythmic classes: syllable-timed and stress-timed (Pike 1945, Abercrombie 1967). This rhythm-class hypothesis still awaits empirical verification (e.g. Roach 1982), and the binary opposition has given way to the view that languages vary along a continuum. The notion of isochrony has been questioned by Dauer (1983), who claims that rhythmic differences between languages reflect a number of phonological properties, such as syllable structure, length as a distinctive feature in vowels, and the (non-)existence of vowel reduction.

Recent attempts to quantify the rhythmic properties of languages have relied on durational measurements to capture the rhythmic qualities of speech. These rhythm metrics can be grouped into global and local measurements. Global measurements include the calculation of (i) the proportion of vocalic intervals (%V) in an utterance and (ii) the standard deviation of vocalic and consonantal interval durations, either as a raw (ΔV and ΔC , Ramus et al. 1999) or rate-normalized measure (VarcoV and VarcoC, Dellwo and Wagner 2003). Low %V values indicate a high degree of vowel reduction and/or high complexity of syllable structure (i.e. stress-timing properties). Syllable structure complexity and vowel reduction are positively correlated with (Varco) ΔC and (Varco) ΔV respectively.

While global measurements pay no attention to the linear arrangement of events, local measures rely on the ratio of successive interval durations (i.e. successive syllables or vocalic/consonantal intervals). Metrics that have been proposed differ in terms of formulae and, crucially, in the choice of intervals for comparison. Low and Grabe’s (1995) Pairwise Variability Index (PVI) and Gibbon and Gut’s (2001) Rhythm Ratio include all successive units, while Gut’s (2003) Syllable Ratio only covers stressed/unstressed syllable pairs. Generally, local rhythm metrics calculate a single value per speaker or speaker group by averaging durational ratios of successive intervals. A higher degree of temporal vowel reduction is reflected in a higher ratio between successive syllables or vowels.

2.2 Unstressed vowels: English vs. German

German and English are both considered stress-timed languages (Kohler 1995, Giegerich 1992); they share a number of typical properties. Apart from having a complex syllable structure, both languages (i) distinguish stressed and unstressed syllables in terms of quality and quantity; (ii) have the short central vowel [ə] and (iii) show schwa deletion and syllabic consonants as extreme forms of reduction. These features allow the com-

pression of syllable nuclei to (theoretically) achieve isochrony between stressed syllables.

However, the distribution of schwa vowels in German is more restricted (Kaltenbacher 1998). In simple lexemes, they are only found in stem-final syllables (*Hase*, [ˈha:zə]) or inflectional affixes (*ge-dacht*, [gəˈdaxt]; *denk-e*, [ˈdɛŋkə]). In English, there are no morphological restrictions. In contrast to English, German has a second, more open schwa vowel [ɐ], which occurs in contrast with [ə] (*bitte* [ˈbitə], *bitter* [ˈbitɐ]). The distribution of schwa also differs in complex lexemes. In both languages, morphophonological processes apply to derived words such as *photography* and *Fotografie*, which differ from their base in terms of primary and/or secondary stress placement. In German, vowel reduction in such cases can be observed as a shortening of long vowels (*Foto* [ˈfo:to] – *Fotograf* [fotoˈgra:f] – *Fotografie* [fotoɡraˈfi:]). A change of vowel quality from tense to lax is less frequent, and vowels are never reduced to schwa. Through productive morphophonological processes in English, on the other hand, unstressed vowels are shortened and centralized to(wards) schwa (*photo* [ˈfəʊtəʊ] – *photograph* [ˈfəʊtəɡrɑ:f] – *photography* [fəˈtɒɡrəfi]).

In general, the quality of unstressed vowels in polysyllabic words shows a higher degree of reduction in English. In an acoustic analysis of derivational word pairs in four languages, Delattre (1981) compared the first two formants of stressed vowels (base form) with the respective unstressed ones (derivative). In terms of distance in the $F_1 \times F_2$ vowel space, the degree of vowel reduction in German was much smaller than in English.

In connected speech, closed-class function words (e.g. determiners, pronouns, conjunctions, auxiliaries) can undergo reduction in both languages (*und* [ʊnt]→[ənt]→[ən]→[n]; *and* [ænd]→[ənd]→[ən]→[n]). In German, however, these reduction processes are stylistically marked; they only occur in informal speaking styles (Kohler 1995, Wesener 1999). In clear speech, syllable nuclei in monosyllabic function words are not reduced to [ə]. In English, the weak form of function words (which involves [ə] in many cases) is the unmarked variant, even in formal speech. Thus, while both languages show reduction in function words, a centralization of vowel quality is much more common in English, which is primarily due to stylistic differences.

2.3 Unstressed vowels in L2 acquisition

Past research on the presence/absence of reduction phenomena in learner speech has shown this to be an area of difficulty in L2 acquisition. Studies differ in scope and methodology. In an instrumental study of the pronunciation of four derivational word pairs, Flege and Bohn (1989) found that Spanish learners of English lack vowel reduction in unstressed syllables in terms of quality and quantity. Lee et al. (2006) investigated acoustic properties of unstressed vowels in late Japanese and late Korean bilinguals. Compared with native speakers, less vowel reduction and an influence of orthography was found in the productions of the bilinguals, whose unstressed vowels were scattered wider in the $F_1 \times F_2$ space. Gut (2006) analyzed the acoustic properties of several affixes in reading passages and retellings produced by English, Chinese and

Italian learners of German. In post-stress syllables of the type C+<en>, F₁ and F₂ values differed significantly from the native speakers (English learners: F₁ only).

A number of studies have looked at the pronunciation of function words in connected speech. Ghazali and Bouchhioua (2003) auditorily analyzed 15 sentences read by Tunisian learners of English. The strong form of function words (e.g. *for*, *to*, *that*) was produced in 92.5% of the cases. Comparing Japanese learners with native speakers, Aoyama and Guion (2007) analyzed the duration of 3 function words embedded in sentences; these were significantly longer in the speech of learners. Similar findings are reported by Porzuczek (2010) for Polish learners of English. Compared to native speakers, they produced longer vowels in the function word *to* in a reading passage.

Acoustic studies on unstressed vowels in GLE have applied rhythm metrics, thus producing a more general assessment of its rhythmical properties. Comparing German learners with native speakers of British English, Gut (2009) analyzed durational differences of successive stressed/unstressed syllable pairs in a reading passage. The native speakers' Syllable Ratio (2.50:1) was larger than that of the learners (2.23:1), indicating that, while German learners transfer quantity reduction, they do not reach the level of native speakers. Ordin et al. (2011) investigated the timing patterns of GLE at various proficiency levels. The VarcoV and vocalic PVI measurements showed that the variation of vowel durations increases with language proficiency. This might indicate that more advanced learners show a higher durational reduction of unstressed vowels. Due to the lack of a native speaker control group, however, their results cannot be compared with SSBE speech.

Transfer from L1 plays an important role in L2 phonology (e.g. Major 2008). Rhythmic interference in unstressed syllables was observed by Gut (2003), who identified transfer of L1 properties in the speech of Polish, Chinese and Italian learners of German. In a review of past research on L2 stress patterns, Broselow and Kang (2013) conclude that prosodic similarity between L1 and L2 facilitates acquisition; errors tend to reflect L1 influence. A general feature of non-native speech seems to be a tendency to overarticulate compared to native speakers (Barry 2007).

3. Aims and method

3.1 Aims of the study

Research on unstressed vowels in German Learner English has mostly relied on auditory descriptions (e.g. Pascoe 1996, Parkes 2001, Dretzke 2006). This study aims to identify the acoustic properties of unstressed vowels in advanced German Learner English. Based on the prosodic similarities and differences of unstressed vowels described above, German learners are expected to show negative transfer in the reduction of vowel quality. Positive transfer of the reduction of vowel duration is expected. Based on the findings reported by Gut (2009), however, durational reduction in GLE is expected to be smaller than in SSBE. Thus, this study sets out to test two hypotheses: Compared to native speakers, learners produce unstressed vowels with (i) a longer duration and (ii) different F₁ and F₂ values.

3.2 Method and data

Recordings of 4 advanced learners (female, aged 19-29, university students) and 3 native speakers of Standard Southern British English (female, aged 19-30) were analyzed. There were two reading tasks. Task 1 consisted of 66 short phrases, which served to elicit canonical vowels. Structured in a similar way (*I said ..., not ...*), they contained two monosyllabic words. These either formed a minimal pair differing in the vowel or they rhymed, the second slot eliciting (nonsense) words with the syllable frames [hVd] and [hVt] (Peterson and Barney 1952). Task 1 produced 7-12 tokens of each monophthong. These measurements served as reference values for the reduction of vowel quality. Task 2 consisted of 28 sentences, which were designed to elicit (a) stressed vowels (5-7 tokens of each monophthong per speaker) and (b) unstressed vowels in three contexts: in function words (e.g. *to, from, some, have*; 23 tokens per speaker); in polysyllabic lexemes in pre-stress (e.g. *consider, again, information*; 20 tokens per speaker) and post-stress position (e.g. *number, probably, attention*; 17 tokens per speaker).

To arrive at comparable speech rates in task 2, native speakers were asked to imagine they were reading to a non-native speaker; learners were asked to read in a way they felt comfortable. The speech rate (syllables/second) was calculated for each speaker using the 5% Winsorized mean to avoid the influence of outliers (Wilcox 2012). Pre-pausal syllables were excluded from analysis. The groups differed in speech rate: learners read at a slower rate ($M = 5.15$, $SD = 0.26$) than native speakers ($M = 5.53$, $SD = 0.64$).

The acoustic analysis was carried out in Praat. Vowel duration was determined using the onset and offset of F_2 . Deleted vowels were included in the analysis (duration = 0). Due to the differences in speech rate, the vowel durations were standardized to z-scores for each speaker. This standardization was based on the stressed and unstressed vowels measured in task 2. Z-scores express the duration of vocalic segments in terms of standard deviations away from the mean. To avoid the influence of outliers, the standardization was based on the 5% Winsorized mean and the 5% Winsorized standard deviation. The vowel target was determined visually at the point of maximal displacement (Di Paolo et al. 2011). F_1 and F_2 were transformed to Bark (Traunmüller 1997) using the package `vowels` in R (Kendall and Thomas 2013).

4. Results

This section starts with a descriptive summary of the duration and formant measurements, followed by the results of the hypothesis tests. The boxplots in Figure 1 compare the distribution of unstressed vowel durations in the two groups across the different contexts. Looking at all contexts, there appears to be no difference between the groups. The variation between the three contexts is similar in both groups: vowels in post-stress position were on average longer, those in pre-stress position shorter. Differences between the groups emerge in the three contexts. While the patterns in the weak form words are very similar, there is a clear difference in pre-stress position, with learners producing shorter vowels. In post-stress position, native speakers produce shorter segments.

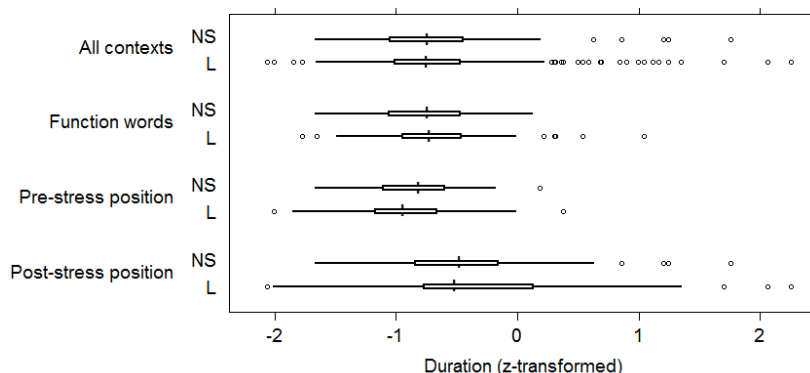


Figure 1: Duration of unstressed vowels by group and context
(NS = native speakers; L = learners)

Figures 2a-c show the Bark-transformed formant measurements. The mean values of 8 canonical monophthongs elicited in task 1 are represented as IPA symbols. They serve as reference points of vowel quality. The boxplots at the margins give more detailed information about the distribution of the measurements in the two groups.

Figure 2a plots F_1 and F_2 of all unstressed vowels by group. The location of the centroids reveals that, on average, the learners produced higher F_1 and lower F_2 values. The difference in F_2 is more pronounced, which is clearly illustrated by the boxplots. The ellipses around the centroids, which were drawn with the function `dataEllipse` in the package `car` in R (Fox 2013), contain 50% of the tokens. The area of the two ellipses shows that the variation in the learner group was higher. A comparison of the 3 different contexts reveals that this tendency is less pronounced in post-stress syllables, where the groups produce more similar formant values. The F_1 and F_2 values measured in the weak form words are shown in Figure 2b. The distribution is similar to Figure 2a, but the differences are slightly larger. Higher F_1 and lower F_2 values seem to reflect the influence of the strong form of the function words (e.g. *has, have, than, some; was, of, from, for*). The shape of the data ellipses and the F_2 boxplots show that the variation in F_2 was larger in the learner group. To detect a possible influence of orthography (cf. Lee et al. 2006) the set of vowels in pre-stress position was split by their orthographic representation – <o> vs. <a>. While pre-stress <a>-vowels (9 tokens per speaker) differ only slightly, the difference in pre-stress <o>-vowels (8 tokens per speaker) is remarkable. Pre-stress <o> vowels are shown in Figure 2c. The ellipses do not overlap. This is due to the striking difference in F_2 .

In the inferential analysis, Wilcoxon tests were used to determine whether the individual learners differed significantly from the group of native speakers regarding duration, F_1 and F_2 of all unstressed vowels. Table 1 lists the results. There was no significant difference in vowel duration. Only Learner 3 had a significantly higher F_1 . The differences in F_2 , which were described above, were statistically significant for all learners.

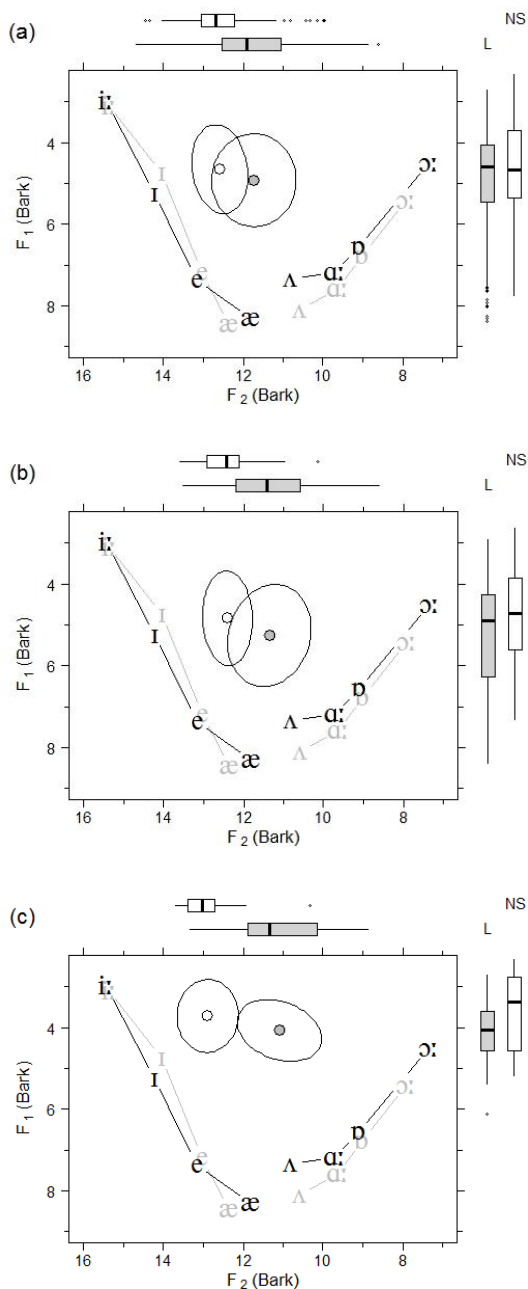


Figure 2: F₁-by-F₂ plot of unstressed vowels (a) in all contexts; (b) in function words; (c) in pre-stress position with orthography $\langle o \rangle$. 50% data ellipses are drawn around the centroids (white: NS, grey: L); boxplots show the distribution of F₁ and F₂ by group; IPA symbols represent the centroids of canonical vowels elicited in task 1 (grey: L; black: NS)

	Duration		F ₁ (Bark)		F ₂ (Bark)	
	<i>Mdn</i>	<i>p</i>	<i>Mdn</i>	<i>p</i>	<i>Mdn</i>	<i>p</i>
L1	-.77	n.s.	4.2	n.s.	12.1	*
L2	-.69	n.s.	4.6	n.s.	11.8	***
L3	-.73	n.s.	5.8	***	11.9	***
L4	-.77	n.s.	4.3	n.s.	11.9	***
NS	-.75		4.7		12.7	

Table 1: Results of Wilcoxon tests: comparison of each learner with the native speaker group (significance levels: n.s. $p > .05$; * $p < .05$; ** $p < .01$; *** $p < .001$)

5. Discussion

This study investigated two acoustic properties of unstressed vowels in advanced GLE. No significant differences between learners and native speakers were found in the duration of unstressed vowels. Contrary to hypothesis (i), the descriptive analysis showed that the durational measurements were very similar. As predicted by hypothesis (ii), differences were found in vowel quality: learners produced slightly higher F₁ and considerably lower F₂ values, the difference in F₂ being statistically significant for each learner. An exploration of the formant values in the different contexts showed that in post-stress position differences between the groups were less pronounced, while in pre-stress position there was a clear influence of orthography. These findings indicate that in polysyllabic lexemes, the lack of deprecipitating of unstressed vowels is higher in pre-stress position than in post-stress position. Differences in vowel quality were also found in weak forms words.

These findings could be explained in terms of transfer from the native language. L1 reduction processes are transferred to L2, leading to positive transfer of vowel duration and negative transfer of vowel quality in unstressed syllables. However, this interpretation overlooks two findings of previous research, which were stated above (cf. section 2.3.): (1) there appears to be a universal tendency of learners to overarticulate; (2) a lack of durational reduction in GLE was reported by Gut (2009). A possible explanation is that the advanced learners in this study have developed beyond the stage where the duration of unstressed vowels is influenced by universal factors. This would be in line with Major's (2001) Ontogeny and Phylogeny Model of second language acquisition. Major claims that the influence of universals first increases and then decreases in the chronological course of second language acquisition. It is possible that this universal feature surfaced at an earlier stage in the interlanguage of the learners investigated in this study. Possibly, further development resulted in a substitution of these universal structures with target language structures. It has to be kept in mind, however, that the findings of this study are not directly comparable to Gut (2009), since vowel reduction was operationalized in different ways.

The use of a reading task for data elicitation is a methodological weakness of this study. First, it is questionable whether this speaking style is representative of learner speech – especially for an investigation of reduction phenomena. Second, the clear

influence of orthography revealed in the descriptive analysis might be a characteristic feature of reading style, but not necessarily learner speech.

Duration and vowel quality are acoustic correlates of prominence, and thus contribute to the rhythmic properties of speech. The findings of this study seem to suggest that the differences in speech rhythm between advanced German learners of English and native speakers of British English might concern vowel quality rather than vowel duration. It must be kept in mind, however, that the measurement of selected vocalic intervals in utterances does not allow generalizations about the rhythmic properties of speech. Future research will attempt to bridge the gap between this “local” approach, i.e. the comparison of selected intervals across speakers and groups, and a more “global” approach, i.e. an assessment of rhythmic properties that includes all vocalic intervals in the speech signal.

References

- Abercrombie, D. 1967. *Elements of general phonetics*. Edinburgh: Edinburgh University Press.
- Aoyama, K. and S. Guion. 2007. Prosody in second language acquisition: Acoustic analyses of duration and F0 range. In M. Munro and O. Bohn (eds) *Language experience in second language speech learning*. Amsterdam: Benjamins. 281-297.
- Barry, W. 2007. Rhythm as an L2 problem: How prosodic is it? In J. Trouvain and U. Gut (eds) *Non-native prosody: phonetic description and teaching practice*. Berlin: Mouton de Gruyter. 97- 120.
- Broselow, E. and Y. Kang. 2013. Phonology and speech. In J. Herschensohn and M. Young-Scholten (eds) *The Cambridge handbook of SLA*. Cambridge: CUP. 529-553.
- Crystal, D. 2008. *A dictionary of linguistics and phonetics*. Malden: Blackwell.
- Dauer, R. 1983. Stress-timing and syllable-timing reanalysed. *Journal of Phonetics* 11: 51-62.
- Delattre, P. 1981. An acoustic and articulatory study of vowel reduction in four languages. In P. Delattre (ed) *Studies in comparative phonetics*. Heidelberg: Julius Groos. 63-93.
- Dellwo, V. and P. Wagner. 2003. Relations between language rhythm and speech rate. *Proceedings of the 15th ICPHS*, Barcelona, Spain. 471-474.
- Di Paolo, M., Yaeger-Dror, M. and A. Beckford Wassink. 2011. Analyzing vowels. In M. Di Paolo and M. Yaeger-Dror (eds) *Sociophonetics. A student's guide*. New York: Routledge. 87-106.
- Dretzke, B. 2006. Ausspracheschulung im Fremd-sprachenunterricht. In U. Jung (ed) *Praktische Handreichung für Fremdsprachenlehrer*. Frankfurt: Peter Lang. 132-140.
- Flège, J. and O. Bohn. 1989. An instrumental study of vowel reduction and stress placement in Spanish-accented English. *Studies in Second Language Acquisition* 11: 35-62.
- Fox, J. 2013. *Package 'car'*. R-package version 2.0-17.
<http://cran.r-project.org/web/packages/car/car.pdf>.

- Ghazali, S. and N. Bouchhioua. 2003. The learning of English prosodic structures by speakers of Tunisian Arabic: word stress and weak forms. *Proceedings of the 15th ICPHS*, Barcelona, Spain. 961-964.
- Gibbon, D. and U. Gut. 2001. Measuring speech rhythm. *Proceedings of Eurospeech*, Aalborg, Denmark. 91-94.
- Giegerich, H.J. 1992. *English Phonology*. Cambridge: Cambridge University Press.
- Gut, U. 2003. Non-native speech rhythm in German. *Proceedings of the 15th ICPHS*, Barcelona, Spain. 2437-2440.
- Gut, U. 2006. Unstressed vowels in non-native German, *Speech Prosody 2006. Proceedings of the 3rd International Conference on Speech Prosody*, eds. R. Hoffmann and H. Mixdorff. Dresden, Germany.
- Gut, U. 2009. *Non-native speech. A corpus-based analysis of phonological and phonetic properties of L2 English and German*. Frankfurt: Peter Lang.
- Kaltenbacher, E. 1998. Zum Sprachrhythmus des Deutschen und seinem Erwerb. In H. Wegener (ed) *Eine zweite Sprache lernen*. Tübingen: Narr. 21-38.
- Kendall, T. and E.R. Thomas. 2013. *Package 'vowels'*. R-package version 1.2. <http://ncslaap.lib.ncsu.edu/tools/norm/>.
- Kohler, J. 1995. *Einführung in die Phonetik des Deutschen*. Berlin: Erich Schmidt Verlag.
- Lee, B., Guion, S.B. and T. Harada. 2006. Acoustic analysis of the production of unstressed English vowels by early and late Korean and Japanese bilinguals. *Studies in Second Language Acquisition* 28(3): 487-513.
- Low, E.-L. and E. Grabe. 1995. Prosodic patterns in Singapore English. *Proceedings of the 13th ICPHS*. 636-639.
- Major, R. 2001. *Foreign accent: The ontogeny and phylogeny of second language phonology*. Mahwah: Erlbaum.
- Major, R. 2008. Transfer in second language phonology: A review. In J. Hansen Edwards and M. Zampini (eds) *Phonology and second language acquisition*. Amsterdam: Benjamins. 63-94.
- Ordin, M., Polyanskaya, L. and C. Ulbrich. 2011. Acquisition of timing patterns in second language. *Proceedings of INTERSPEECH 2011*, 1129-1132.
- Parkes, G. 2001. *The Mistakes Clinic for German-speaking Learners of English*. Southampton: Englang.
- Pascoe, G. 1996. Pronunciation Analysis and Teaching. In W. Barry and A. Addison (eds) *Phonus 2*. Saarbrücken: Institut für Phonetik. 109-122.
- Peterson, G.E. and H.E. Barney. 1952. Control methods used in a study of vowels. *Journal of the Acoustical Society of America* 24: 175-184.
- Pike, K.L. 1945. *The intonation of American English*. Ann Arbor: Michigan University Press.
- Porzuczek, A. 2010. The weak forms of TO in the pronunciation of Polish learners of English. In E. Waniek-Klimczak (ed) *Issues in Accents of English 2: Variability and norm*, ed.. Newcastle upon Tyne: Cambridge Scholars Publishing. 309-322.
- Ramus, F., Nespors, M. and J. Mehler. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73: 265-292.

-
- Roach, P. 1982. On the distinction between 'stress-timed' and 'syllable-timed' languages. In D. Crystal (ed) *Linguistic controversies. Essays in linguistic theory and practice*. London: Arnold. 73-79.
- Traunmüller, H. 1997. *Auditory scales of frequency representation*.
<http://www.ling.su.se/staff/hartmut/bark>.
- Wesener, T. 1999. The phonetics of function words in German spontaneous speech. In K. Kohler (ed) *Phrase-level phonetics and phonology*. Kiel: Universität Kiel. 327-377.
- Wilcox, R. 2012. *Modern statistics for the social and behavioral sciences*. Boca Raton: CRC Press.