2014 ACTA UNIVERSITATIS CAROLINAE PAG. 91–100
PHILOLOGICA 1 / PHONETICA PRAGENSIA XIII

PRE-FORTIS SHORTENING IN FLUENT READ SPEECH: A COMPARISON OF CZECH AND NATIVE SPEAKERS OF ENGLISH

DITA FEJLOVÁ

ABSTRACT

This paper inspects the details of a phenomenon called *pre-fortis shortening*, the existence of which is widely acknowledged by phoneticians. It occurs in VC sequences where the final consonant is voiceless (fortis). For English, the difference in the duration before fortis and lenis consonants is recognized as a cue of the consonant's voicing, since the actual voicing tends to be missing. The study compares the extent to which *pre-fortis shortening* is employed by native speakers of English and by Czech students of English with different degrees of foreign accent. The results suggest that the difference in the duration of pre-fortis and pre-lenis vowels is considerably lower in connected speech than in previously reported results, even in native speakers, with the difference more pronounced in long (tense) vowels.

Key words: vowel duration, fortis and lenis consonants, pre-fortis shortening, Czech English

1. Introduction

The duration of individual segments in speech has been researched from several perspectives. Leaving aside higher-level cognitive decisions determined by the communicative intent, it is possible to identify several factors which affect the ultimate duration of vowels and consonants in connected speech. The following paragraphs will focus on vowels only, not only because it is their duration that will be investigated in this paper, but also because studies on vowels are more numerous. Van Santen (1992) lists seven factors which have quantitative effects on vowel durations, some of which will be discussed here in some detail.

The effect of the position of the word in a sentence concerns mostly those words in the vicinity of phrase boundaries. It has been demonstrated that a word-final syllable (or its nucleus, the vowel) is longer when followed by a syntactic pause than when it is located within a phrase, a phenomenon called *prepausal lengthening* or, more appropriately, *final deceleration* (Volín and Skarnitzl, 2007). The increase in duration found by Crystal and

House (1988b) was 40%, and the speakers in Byrd's (2000) study manifested more than a 50% increase in the duration of vowels before major prosodic boundaries.

The second aspect playing a role in vowel duration relates to the notion of word stress. Crystal and House (1988b) analyzed vowels in stressed and unstressed syllables, and they showed that the difference between the duration of an English vowel in a stressed and unstressed syllable is 70 ms on average, which is quite substantial. Vowel duration is thus one of the key correlates of stress in English (Cruttenden, 2008: 237; Roach, 2009: 74).

On top of the position of the word within a sentence and of the syllable within a word, each vowel can be said to have its inherent duration. Klatt (1976: 1213) pointed out that the vowels /ɪ ɛ ʌ ʊ/ are generally shorter than other English vowels. Crystal and House (1988a) specify these other vowels as /i e æ ɑ o ɔ u/ and call them long (tense), as opposed to the short ones (lax). It is worth noting that the vowel system presented in their study relates to American English; their /e/ and /o/ correspond to the diphthongs /eɪ/ and /əʊ/, respectively in British English. The term inherent duration, however, is typically associated with another observation, namely that vocalic duration tends to vary inversely with vowel height (see, e.g., Lehiste, 1970: 18 or Crystal and House, 1988a and studies quoted therein). This finding has a natural explanation: low vowels are intrinsically longer because it takes more time for the articulators (tongue, jaw) to move downwards into the target position for /æ ɑ/ and back up again. Crystal and House put this hypothesis to test on long vowels and found out that this finding holds only for the high-nonhigh distinction: vowels in the mid section were not shorter than open vowels.

Consonants have their inherent durations as well, mainly with respect to their voicing status. In what appears to be a linguistic universal, voiceless obstruents have been shown to be longer than voiced obstruents (Klatt, 1976; Jessen, 1998; a summary of studies dealing with consonant duration in Czech can be found in Skarnitzl, 2011: 104ff.).

Having mentioned findings about inherent segmental durations, let us turn to the next factor, namely the effect of combining speechsounds in a sequence. Vowels were discovered to behave in a certain way before certain consonants and that gave rise to the notion of *compensatory shortening*. The original underlying assumption stated that the duration per syllabic unit is relatively constant (Chen, 1970: 146); in other words, the inherent segment durations would be compromised so that the even flow of syllables is not disrupted (also called the *isochrony hypothesis*). Thus, obstruents and vowels are known to mutually affect each other in duration, as summarized by Machač and Skarnitzl (2007).

On the one hand, obstruents appear to be affected by the (inherent) duration of the tautosyllabic vowel, being longer in the neighbourhood of close vowels like [i u], which are shorter, and shorter in the context of (inherently) longer open vowels like [a]. On the other hand, the duration of vowels has been demonstrated to vary depending on the voicing status of the neighbouring tautosyllabic obstruent: vowels are shorter when flanked by fortis consonants (which are longer) than when flanked by lenis consonants (which are shorter). It is this second type of temporal compensation which will be the subject of our study. In Czech, it manifests itself more in CV sequences, as these are tautosyllabic, unlike Czech VC sequences (see Machač and Skarnitzl, 2007). In English, however, where the intervocalic consonant in words like *many*, *discover*, or *latent* is syllabified with the preceding (stressed) syllable (see Krakow, 1999 or Wells, 2008: XXVII), the effect is substantially greater in VC sequences. That is why this compensatory shortening has been

called specifically *pre-fortis* shortening: other things being equal, vowels will be shorter in English words like *litre* [liˑtə] and *latter* [lǽtə] than in *leader* [liːdə] and *ladder* [lædə], respectively. It is the voicing feature of the consonant following the vowel in question which is the key factor used for interpreting the changes in vowel duration. Van Santen (1992) showed that the difference in the duration of pre-fortis and pre-lenis vowels can be up to 120 ms.

The last century has seen several explanations for this phenomenon, which appears to constitute a linguistic universal. A very influential theory was voiced by Simon Belasco (1953; in Chen, 1970: 140): Given that a fortis (voiceless) obstruent takes more force to produce (as the term "fortis" implies), the anticipation of spending more energy on it shortens the preceding vowel. Similarly to the notion of *compensatory shortening,* it is implied here that each syllabic unit takes a constant amount of energy to produce, so that the energy expended to produce the vowel varies inversely with the energy consumed by the following consonant.

Chen (1970) proposes his own theory, according to which the difference in the duration of vowels preceding consonants depends on the speed of the transition from the vowel to the consonant closure (152). This transition is faster when the speaker anticipates greater effort in holding the articulators together (e.g. lips with /p/). Such greater effort is necessary in the case of fortis (voiceless) plosives, because pronunciation with an open glottis leads to greater building up of intraoral pressure behind the articulators. Lenis (voiced) plosives do not require such effort, therefore the transition can be slower. Compare the duration of the transition phases in Figure 1 below.

Chen investigated various languages and, as Figure 2 shows, different languages exploit *pre-fortis shortening* to different extents. While Korean, Russian and French manifest ratios of mean vowel durations before voiceless/voiced consonants between 0.78 and 0.89, the same ratio is substantially lower for English (0.61). This is because in English the distinction between vowel duration has become a cue of the final consonant's voicing. English does not neutralize voicing in the word-final position (unlike Czech, for example), but voicing is difficult to control word-finally and is often lost. In order for the

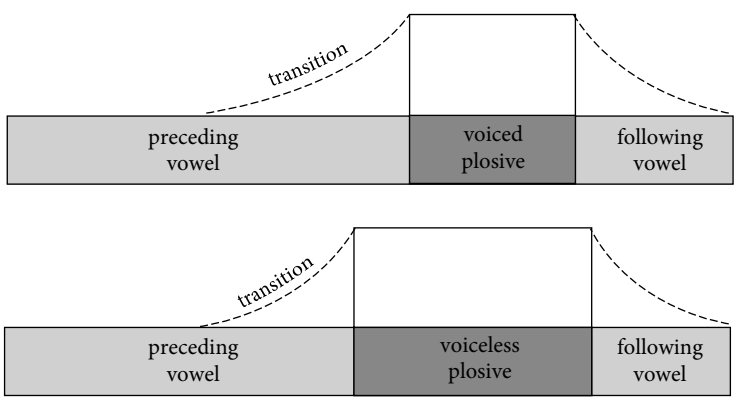

Figure 1. Two VCV schemes showing the durations of transition phases from a vowel to a voiced and voiceless plosive. From Skarnitzl, 2011: 109; adapted from Chen, 1970: 153.

contrast to be preserved, the difference in the duration of the preceding vowel is greater. Although there are no data for Czech (also because, as we have mentioned, a consonant in a VCV sequence typically belongs to the following syllable), it is assumed that Czech vowels behave similarly to the other languages depicted in Figure 2, with the ratio lying in the 0.8–0.9 region.

The objective of the present paper, which is partially based on a preliminary analysis by Fejlová (2012), is, therefore, to compare native speakers of English with Czech students of English, and specifically, the degree to which they employ *pre-fortis shortening* in English VC sequences. Our hypothesis is that this degree will correlate inversely with the strength of foreign accent in Czech English: the distinction between pre-fortis and pre-lenis vowels will be less apparent in speakers with a strong Czech accent, while native-like and native speakers of English will demonstrate a clear-cut difference between vowel duration preceding a voiced and a voiceless obstruent.

2. Method

For the purpose of this paper, we examined recordings of BBC World Service news bulletins read by 13 female speakers, Czech students of English and American studies at the Faculty of Arts in Prague, as well as original recordings by four female newsreaders from the BBC. The recordings of the Czech students were obtained in a soundproof booth with a studio electret microphone IMG ECM 2000 and digitized at the sampling rate of 22,050 Hz. These speakers had been previously placed in three separate categories, A (native-like or near native-like pronunciation), B (discernible but not strong Czech accent) and C (a strong Czech accent); see Skarnitzl et al. (2005) for more detail.

The nature of the speech material in our study differs largely from studies which focused on pre-fortis shortening: earlier studies made use of the target words – cognate

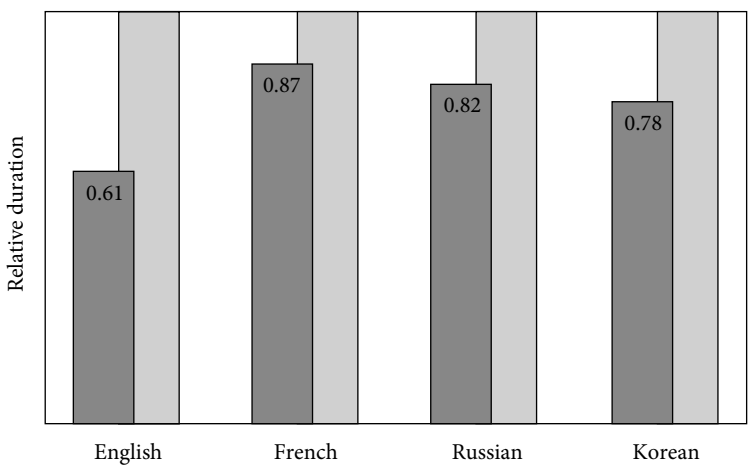

Figure 2. The results of Chen's comparison of vowel duration in minimal word pairs in English, French, Russian and Korean (see text). Adapted from Chen, 1970: 138.

pairs of words like /biˑt/ and /biːd/ – embedded in carrier phrases, or even pronounced in isolation (e.g., Chen, 1970; Wardrip-Fruin, 1984). Our words were extracted from communicatively meaningful texts, and the vowels thus appeared in numerous consonantal contexts and prosodic positions. Naturally, however, this may turn out to be a crucial factor which may render the results more fuzzy than predicted in our hypothesis.

The recordings were automatically segmented using the Penn Phonetics Lab Forced Aligner (P2FA, Yuan and Liberman, 2008), and the boundaries of the target speech-sounds were then manually adjusted following the principles listed in Machač and Skarnitzl (2009).

In order to reduce the effect of speech rate on vowel duration, the raw vowel durations were normalized against local speech rate. For each target word, one stress group on each side was marked and the duration of the target vowel normalized against the number of syllables contained in the three stress groups.

All the necessary data was then extracted from the material with the help of a script in Praat (Boersma and Weenink, 2012). We analysed a total of 637 vowel tokens from the 13 Czech speakers of English and 206 tokens from the newsreaders. These were only vowels in stressed syllables and breath-group non-final positions (in order to eliminate the effect of final deceleration). The number of tokens per vowel quality was not equalized since the material had not been designed with this phonetic study in mind. For the same reason, the number of tokens in pre-fortis contexts slightly prevailed over the number of tokens in pre-lenis contexts.

Statistical analyses with the normalized vowel durations as the dependent variable were carried out in Statistica, version 7; the phonological voicing status of the following obstruent (fortis or lenis) always served as one of the independent variables. Only the underlying fortis/lenis identity was decisive, disregarding the sometimes imperfect physical realisation of voicing, especially in speakers with a Czech accent (groups B and C).

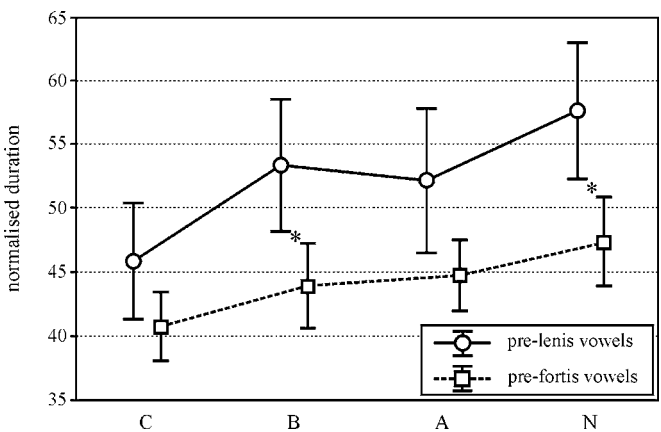

Figure 3. Mean normalized durations of vowels preceding lenis and fortis obstruents for the individual categories of speakers: C, B, A, and N (referred to as speaker "quality"). Vertical lines indicate 95% confidence intervals of vowel durations before lenis and fortis obstruents. The asterisk indicates a significant result at the $p < 0.05$ level.

3. Results and discussion

The first point of interest in our analysis was whether the (normalized) vowel duration data truly are distinct for pre-fortis and pre-lenis contexts. The results presented in Figure 3 show that they are, with the difference reaching statistical significance ($p < 0.05$) for native speakers (group N) and for the Czech group with a recognizable but not strong Czech accent (group B).

Concerning the way durational difference is pronounced depending on the degree of foreign accent, we hypothesized that the "gap" between pre-lenis and pre-fortis data would gradually widen proceeding from category C (strongest Czech accent) to N (native speakers). Mean vowel durations were indeed most similar in category C (with a ratio of 0.85, and therefore close to Chen's data on French and Russian, which we mentioned in the Introduction). However, category B speakers yielded a lower duration ratio (0.81) than speakers in the native-like category A (0.83). A follow-up inspection revealed that this was caused by a single speaker in category B who achieved remarkably significant results, with the duration ratio of pre-fortis/pre-lenis vowels being as low as 0.67. Had it not been for this speaker, the ratio for category B would be 0.88. The given values support the claim that Czech vowel duration ratios are similar to French (see Figure 2). Category N performed best, although the ratio obtained (0.73) was only halfway between the values of Czech speakers and the value given by Chen for English (0.61).

The following step was to examine in more detail the categories with the best pronunciation, A and N, to find whether there were any strong influencing factors as in category B. It is plainly visible from Figure 4 that speakers BCA and HNA in category A essentially do not distinguish the duration of pre-fortis and pre-lenis vowels, while the data of speakers JLA and KCA are comparable to those of the native speakers. However, in none of the speakers can we find a clear-cut, statistically significant distinction between pre-fortis and pre-lenis vowel durations. The lowest ratios of means were achieved by speakers JIA (0.76) from category A and EMA (0.75) from category N.

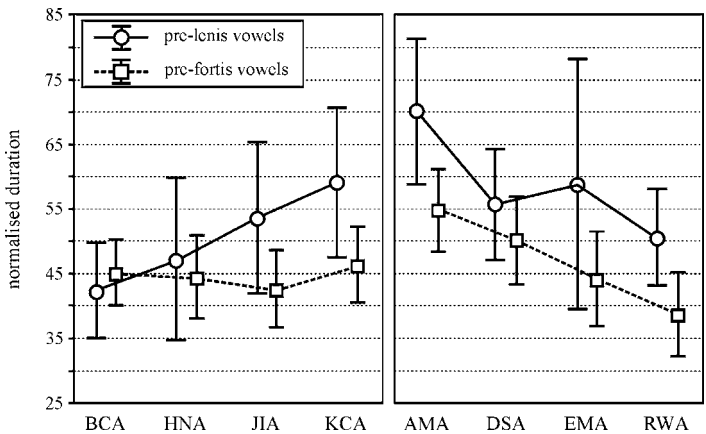

Figure 4. Mean vowel durations (fortis/lenis contexts) for speakers within category A (left side) and N (right side).

These results seem to suggest that *pre-fortis shortening* is hard to detect in fluent speech. In other words, the differences in vowel durations before fortis and lenis consonants reported from carefully controlled speech are, naturally, somewhat blurred in communicative contexts, even for native speakers. Speaker MILA, who had caused the significance of results for category B, was special in that she spoke very slowly and carefully.

In the next step, we wanted to examine the possibility that the duration difference between pre-fortis and pre-lenis contexts may be more manifested in some vowels than in others. Specifically, we considered it as conceivable that the difference might become more pronounced in phonologically long (i.e., tense) vowels than in phonologically short (i.e., lax) vowels. The results in Figure 5 show one noteworthy result, namely the tendency for the duration of short vowels to increase across categories (from C to N). Within one category (except cat. A), it seems that short vowel durations are greater in pre-fortis than pre-lenis contexts (which runs against the rule of *pre-fortis shortening*), but the difference is statistically insignificant. More importantly, however, the difference in vowel duration before fortis and lenis obstruents is greater in the long vowels in all speaker categories, though reaching statistical significance only in categories B and N. It therefore appears easier for the phenomenon of *pre-fortis shortening* to manifest itself on tense vowels with a longer temporal span; it is possible that the lax character of short vowels does not allow for fine temporal manipulations, especially when faster speech rates are involved.

4. General discussion and conclusions

As was argued in the preceding section, the fact that some of the results did not confirm our original hypothesis appears to owe much to the nature of fluent speech. Category

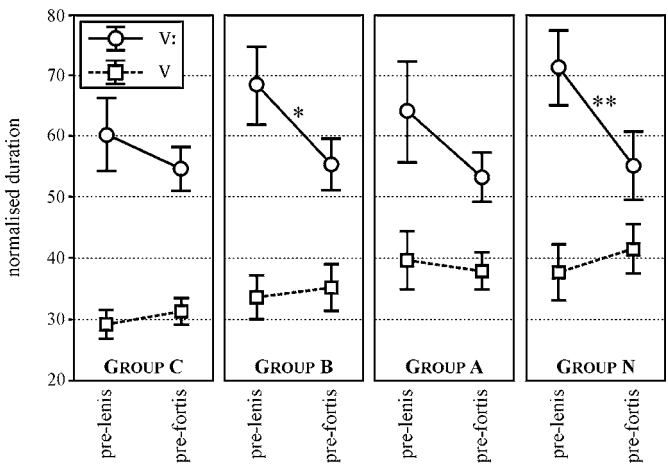

Figure 5. Mean vowel durations for categories C, B, A, and N, depending on the "phonological" length of the vowel. The circles represent tense, long vowels (/ɑ: i: u:/) including the diphthongs (/aɪ aʊ eɪ ɔɪ əʊ/), and the squares correspond to lax vowels (/ʌ æ e ɪ ɒ ʊ/). * = a significant result at the $p < 0.05$ level; ** = a highly significant result at the $p < 0.001$ level.

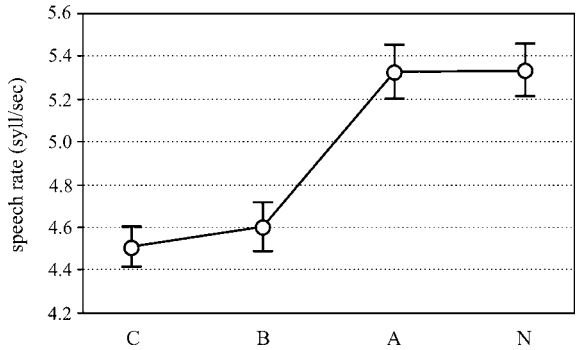

Figure 6. Speech rate in syllables per second given for categories C, B, A, and N, organised in order of proficiency.

A speakers with native-like English pronunciation did not yield the lowest pre-fortis/pre-lenis vowel duration ratios from the Czech speakers, presumably because the differences between vowel durations in the fortis and lenis contexts were obscured due to faster speech in proficient speakers; as can be seen in Figure 6, speakers in groups A and N did speak considerably faster.

Category N proved to display the greatest effect of *pre-fortis shortening*, which was best visible on its long vowels (Figure 5). However, this category also contained pre-fortis tokens of the longest duration (Figure 3). It may be that strong emphasis put on key words by the newscasters, for the sake of pointing out important information, affected the overall target vowel durations.

It is clear that in natural, fluent speech, there are many factors which affect the duration of vowels; some of these were mentioned in the Introduction. Moreover, van Santen (1992: 533) asserts in his study that when examining vowel duration, two syllable-position factors have to be taken into account: the *stress-interval based* factor and the *word-based* factor. Integrating the *stress-interval based* factor agrees with our method of normalising vowel duration against local speech rate. However, to include the *word-based* factor would mean to make separate classes of vowels in word-initial, -medial, and -final syllables because, according to van Santen, vowel duration depends on the number of syllables from the word's end (534).

The same author argues elsewhere in the text that it is the manner of articulation of the following consonant that matters more than its voicing distinction (*ibid*: 527). He shows that the two factors interact, and arranges consonant classes by how the vowel before them gradually increases in duration: *Voiceless stops → Voiceless affricates → (Liquids) → Voiceless fricatives → (Nasals) → Voiced stops → Voiced affricates → Voiced fricatives* (*ibid*: 529). However, to include such factors in our study would require a considerably larger speech material.

In conclusion, *pre-fortis shortening* is a phenomenon easily found when we compare cognate pairs pronounced in isolation; however, in connected speech and in vowels extracted from various prosodic contexts, the difference in the duration of pre-fortis and pre-lenis vowels is much lower. The speech rate of more experienced speakers blurred

the pre-fortis/pre-lenis difference, so that the data of nearly all speakers maintained some degree of overlap. The pre-fortis/pre-lenis distinction was manifested in a more stable manner in tense (long) vowels; the question remains whether it is their tense character which makes it possible for the distinction to appear, or whether it is the physical shortness of lax vowels which does not provide for further shortening.

If we were to suggest directions for future work, it would be interesting to include van Santen's criteria when selecting target words for analysis and to come another step towards disentangling the effect of various factors which define *pre-fortis shortening* in commonly spoken English. This could also be done on data from Czech speakers of English extracted from fluent speech, with respect to the question of how the presence of *pre-fortis shortening* better correlates with the speakers' proficiency in English.

ACKNOWLEDGEMENTS

I would like to give my thanks to Radek Skarnitzl, who provided me with many useful comments on an earlier version of this paper. This research was supported by the Programme of Scientific Areas Development at Charles University in Prague (PRVOUK), subsection 10 – Linguistics: Social Group Variation.

REFERENCES

Byrd, D. (2000). Articulatory vowel lengthening and coordination at phrasal junctures. Phonetica, 57, pp. 3–16.

Boersma, P. & Weenink, D. (2012). Praat: doing phonetics by computer [Computer program]. Version 5.3.14, retrieved on April 28, 2012 from <http://www.praat.org>.

Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. Phonetica, 22, pp. 129–159.

Cruttenden, A. (2008). Gimson's Pronunciation of English. London: Hodder Education.

Crystal, T. H. & House, A. S. (1988a). Segmental durations in connected-speech signals: Current results. Journal of the Acoustical Society of America, 83, pp. 1553–1573.

Crystal, T. H. & House, A. S. (1988b). Segmental durations in connected-speech signals: Syllabic stress. Journal of the Acoustical Society of America, 83, pp. 1574–1585.

Jessen, M. (1998). Phonetics and phonology of tense and lax obstruents in German. Amsterdam: John Benjamins.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. Journal of the Acoustical Society of America, 59, pp. 1208–1221.

Krakow, R. A. (1999). Physiological organization of syllables: A review. Journal of Phonetics, 27, pp. 23–54.

Lehiste, I. (1970). Suprasegmentals. Cambridge, Mass.: MIT Press.

Machač, P. & Skarnitzl, R. (2007). Temporal compensation in Czech? In: Proceedings of the 16th ICPhS. Saarbrücken: ISPhS, pp. 537–540.

Machač, P. & Skarnitzl, R. (2009). Fonetická segmentace hlásek. Praha: Epocha.

Roach, P. (2009). English Phonetics and Phonology. Cambridge: Cambridge University Press.

Skarnitzl, R. (2011). Znělostní kontrast nejen v češtině. Praha: Epocha.

Skarnitzl, R., Volín, J. & Drenková, L. (2005). Tangibility of Foreign Accents in Speech: The Case of Czech English. In: A. Grmelová, L. Dušková & M. Farrell (Eds.), 2nd Prague Conference on Linguistics and Literary Studies Proceedings. Praha: UK PedF, pp. 11–20.

Van Santen, J. P. H. (1992). Contextual effects on vowel duration. Speech Communication, 11, pp. 513–546.

Volín, J. & Skarnitzl, R. (2007). Temporal downtrends in Czech read speech. In: Proceedings of Interspeech 2007. Antwerpen: ISCA, pp. 442–445.
Wardrip-Fruin, C. (1982). On the status of temporal cues to phonetic categories: Preceding vowel duration as a cue to voicing in final stop consonants. Journal of the Acoustical Society of America, 71, pp. 187–195.
Wells, J. C. (2008). Longman Pronunciation Dictionary. Harlow: Pearson Longman.

KRÁCENÍ VOKÁLU PŘED NEZNĚLÝM KONSONANTEM VE ČTENÉ ŘEČI: SROVNÁNÍ ČESKÝCH A RODILÝCH MLUVČÍCH ANGLIČTINY

Resumé

Dloužení vokálu před znělým konsonantem je fonetický jev, který byl pro angličtinu popsán již v 50. letech minulého století: rané studie ukázaly, že hraje roli při rozpoznání znělostní povahy daného konsonantu, jelikož znělost jako taková často není fyzicky zachována. Čím delší je trvání vokálu, tím více je následující konsonant vnímán jako znělý. Pokud přistoupíme na hypotézu, že v řeči existuje tendence k vyrovnávání trvání v rámci slabiky, vyplývá z toho i doplňkový jev krácení vokálu před neznělým konsonantem, neboť se tak kompenzuje delší trvání neznělého konsonantu. Tento druhý jev byl tradičně testován a potvrzen na sadách odděleně vyslovovaných slov, která tvořila minimální páry. Zdá se však, že není zdaleka tak stabilní v plynulé řeči. Při zkoumání čtených vět od českých a rodilých mluvčích angličtiny se ukázalo, že rozdíl mezi trváním vokálu v neznělém (fortisovém) a znělém (lenisovém) kontextu není příliš markantní. Nejlépe šel pozorovat v případě, kdy se zvlášť zkoumaly fonologicky dlouhé (napjaté) vokály. Je tedy důvod se domnívat, že by pro lepší postižení krácení vokálu před konsonantem bylo užitečné vytvořit složitější vnitřní klasifikaci daných hlásek v sekvencích, v nichž ke krácení dochází. Taková klasifikace by pak mohla být použita při zkoumání plynule mluvené angličtiny českých mluvčích s cílem zjistit, zda se projeví jednoznačnější přímá úměrnost mezi zdatností v angličtině a mírou výskytu krácení před neznělým konsonantem.