ACOUSTIC CORRELATES OF WORD STRESS AS A CUE TO ACCENT STRENGTH

JAN VOLÍN
Metropolitan University Prague, Czech Republic
jan.volin@ff.cuni.cz

LENKA WEINGARTOVÁ
Institute of Phonetics, Prague, Czech Republic
lenka.weingartova@ff.cuni.cz

Abstract
Due to the clear interference of their mother tongue prosody, many Czech learners produce their English with a conspicuous foreign accent. The goal of the present study is to investigate the acoustic cues that differentiate stressed and unstressed syllabic nuclei and identify individual details concerning their contribution to the specific sound of Czech English. Speech production of sixteen female non-professional Czech and British speakers was analysed with the sounds segmented on a word and phone level and with both canonical and actual stress positions manually marked. Prior to analyses the strength of the foreign accent was assessed in a perception test. Subsequently, stressed and unstressed vowels were measured with respect to their duration, amplitude, fundamental frequency and spectral slope. Our results show that, in general, Czech speakers use much less acoustic marking of stress than the British subjects. The difference is most prominent in the domains of fundamental frequency and amplitude. The Czech speakers also deviate from the canonical placement of stress, shifting it frequently to the first syllable. On the other hand, they seem to approximate the needed durational difference quite successfully. These outcomes support the concept of language interference since they correspond with the existing linguistic knowledge about Czech and English word stress. The study adds specific details concerning the extent of this interference in four acoustic dimensions.

Keywords: Czech English, word stress, duration, F0, SPL, spectral slope

1. Introduction

A foreign accent is a multidimensional phenomenon: its manifestations can be explored with respect to vowels or consonants, intonation or rhythm, or, from the acoustic point of view, in the domains of frequency, intensity, timing, or spectral properties of speech units. Interestingly, the term *accent* itself refers not only to specific pronunciation patterns, but in many languages also to specific prominence lent to a syllable or a word (e.g., Polish *akcent*, English *accent*, French *accent*, German *Akzent*, or Czech *přízvuk*, which also bears both meanings). Although traditional dialectology focused mainly on segmental aspects of regional accents in its descriptions, the term itself is motivated

prosodically as if to suggest that one of the most conspicuous features of 'pronunciation other than the reference standard' is prominence distribution and prominence manifestation. This aspect of foreign accentedness is also the focus of our present study.

Despite the generally accepted awareness of the differences between native and non-native treatment of stress patterning, their detailed descriptions, let alone applicable models are still largely missing for most languages and accents. On the other hand, there is a growing body of research reflecting many aspects of the problem. One of the strands of the research shows that the consequences of deviations from expected prominence patterns can be quite important. Crystal (1996: 9) reports cases in which unusual prominence distribution caused communicational problems. In less extreme cases it could be predicted that the foreign speaker is understood, but the extra processing demands may perhaps cause irritation on the part of the listener. It follows that, among many other things, regular stress patterns in speech may be linked to a positive acceptance of the speaker.

Several reliable acoustic correlates of word stress are recognized for English and other languages: F0, duration and intensity were experimented with already in the 1950s (e.g., Fry, 1955; Fry 1958), but also in the following decades (e.g., Klatt, 1976 or Beckmann, 1986). The parameters of spectral slope were acknowledged as important and added later (e.g., Sluijter and van Heuven, 1996). However, stressed and unstressed syllables in Czech do not differ systematically in any of these characteristics (Janota and Palková, 1974; Palková and Volín, 2003; Volín, 2008); although the results concerning spectral slope are still preliminary.

Discussing prominence naturally invites some attention to the non-prominence which provides the indispensable background to perceptually salient elements. Vowel reduction processes are partially responsible for the very specific sound of native English. Standard Czech, on the other hand, requires all vowels to materialize in their full, unreduced forms regardless of their position in the word or phrase. Czech English as an interlanguage is expected to produce a mixture of these two tendencies. Vowel strengthening and weakening in polysyllabic chains creates contrasts that our current study intends to map.

2. Data and method

Recordings of 16 female non-professional speakers aged 20–25 years were used, eight native Czech speakers and eight native English speakers of Southern British Standard. None of them reported any hearing disorder or speech impediment. They were asked to read out an English BBC news bulletin text of 4–4.5 minutes in duration. The speakers were instructed to familiarize themselves with the text beforehand and to read it as naturally as possible. The Czech subjects were recorded in the sound-treated studio of the Institute of Phonetics in Prague with an electret microphone IMG ECM 2000, soundcard SB Audigy 2 ZS, 32-kHz sampling frequency and 16-bit resolution. The British subjects were recorded with a portable professional device Edirol HR-09, with the sampling frequency of 48 kHz and 16-bit resolution. These recordings were afterwards downsampled to 32 kHz to match the sampling frequency of the Czech items.

All the vowels in the recorded utterances were manually labelled in Praat (Boersma and Weenink, 2013) by experienced phoneticians. Apart from the identity and boundaries of the vowels, we carefully annotated the positions of both canonical (according to Wells, 2008) and actually observed word stresses. Not all vowels were selected for analysis: diphthongs and vowels confined to foreign proper names (such as "Arafat") or in words pronounced dysfluently were discarded. Eventually, a total number of 10044 vowels were analyzed, 32 % of them stressed and 68 % unstressed. The identities of the vowels were clustered into six types: /i/, /e/, /a/, /o/, /u/, /ə/ in order to allow for direct comparison of Czech and British vowels. The factual status of the vowel with regard to word stress will be referred to with capital S for stressed and capital U for unstressed vowels.

After the labelling process, the following acoustic measurements were extracted with the help of Praat scripts for each of the vowels:
- duration (in ms)
- fundamental frequency (F0, in semitones relative to 100 Hz)
- sound pressure level (SPL, in dB)
- spectral slope (measure α, in dB)

F0, SPL and spectral slope were measured in the middle third of each vowel to reduce transitional effects from neighbouring consonants or possible annotation inaccuracies. The spectral slope was expressed with the measure α, i.e., the difference between the spectral energy in the frequency bands 0–1000 Hz, and 1000–16000 Hz.

Statistical significance of the results was assessed by two-way analyses of variance (ANOVAs) for independent measures with the two-level factors of LANGUAGE (BrE, CzE) and STRESS (S, U), this being the actual realization of the stress by the given speaker. In order not to inflate the significance of the results, the items were grouped according to speaker, language, vowel type and real stress status. In this way, 180 degrees of freedom instead of 10038 for the ungrouped set were obtained.

The tangibility of the Czech speakers' foreign accent was verified by a perception test. 20 native Czech listeners with working knowledge of English heard two utterances from each of our sixteen speakers in a random order and were asked to judge their native-likeness on a five-point Likert scale, where at one edge number 1 signified "certainly a native English speaker" and at the other edge number 5 meant "certainly a Czech speaker of English". The utterances were selected to satisfy the following conditions: no dysfluencies, no less common proper names, and comparable length of around 5-6 seconds.

3. Results

Figure 1 shows the listeners' answers in the perception test. It is obvious that the group of Czech speakers (left) is clearly separated from the British speakers (right). Also, the fact that the mean scores of the Czech speakers are in all cases between 4 and 5 indicates that their accent is audible and distinguishable even for lay listeners. Even though speaker Cz4 exhibits the most native-like accent, her speech is not at all confusable with native production.
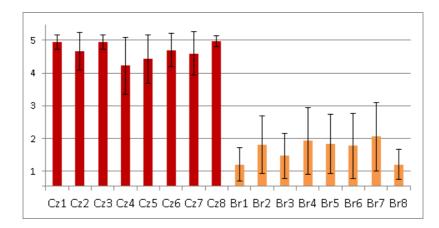
Figure 1: Results of the perception test showing means and standard deviations of the listeners'
answers: 1 = "certainly a native speaker", 5 = "certainly a Czech speaker"

Since the information about canonical word stress was available, the alterations in stress
placement in the production of both groups of speakers could be assessed (see Figure 2).
As expected, the Czech speakers displayed far more alterations than the British speakers.
From a total of over 3000 words, 331 words uttered by the Czech speakers and 92 words
in the case of British speakers exhibited some form of alteration. Most of them can be
attributed to shift of stress to the first syllable which corresponds to the stress pattern
which is obligatory in Czech. Interestingly, British speakers also exhibited several shifts
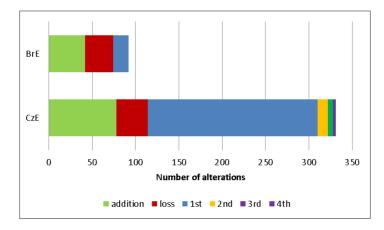to the first syllable (18 words).

Figure 2: Number of alterations in the material. Addition: a canonically unstressed word received
stress; loss: a canonically stressed word does not receive stress; 1st, 2nd, 3rd or 4th: stress shift to the
respective syllable.

The results concerning durations of stressed vs. unstressed syllables are displayed in Figure 3. Both Czech and British speakers modify their vowel duration according to the stress status of the vowel; the stressed ones are significantly longer than the unstressed. The two-way ANOVA returns a highly significant effect of STRESS: $F(1, 180) = 80$, $p < 0.001$. However, the difference between Czech and British speakers is not significant.
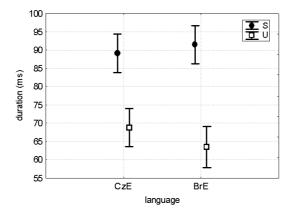

Figure 3: Means of stressed (S) and unstressed (U) vowel durations for Czech (CzE) and British (BrE) speakers. Whiskers indicate 95% confidence interval.

Turning to F0, the results look quite different (see Figure 4). The interaction LANGUAGE*STRESS is significant: $F(1, 180) = 4.42$, $p = 0.037$, which indicates that the Czech speakers treat F0 in stressed and unstressed vowels differently from the British speakers. The trend in Czech speakers appears reversed – unstressed display higher F0 than stressed – but this particular difference is not statistically significant.
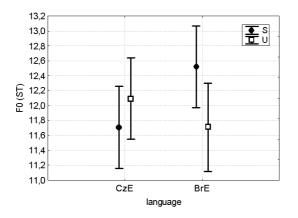

Figure 4: F0 means stressed (S) and unstressed (U) vowels for Czech (CzE) and British (BrE) speakers. Whiskers indicate 95% confidence interval.

Figure 5 shows the differences in spectral slope. It should be noted that caution is needed in this case since the α measure used here is sensitive to speaker and vowel identity (see Weingartová and Volín, 2014) and the vowels of all speakers are pooled together. Even so, both effects were significant, STRESS: $F(1, 180) = 4.01$, $p = 0.047$ and LANGUAGE: $F(1, 180) = 9.68$, $p = 0.002$, while the interaction was not. The significance of the STRESS difference is contributed to mainly by the unstressed British vowels which show the steepest spectral slope. SPL (shown in Figure 6) grants the most clear-cut results from all acoustic correlates of word stress and shows the largest difference between both speaker groups. The effects are significant, STRESS: $F(1, 180) = 62.76$, $p < 0.001$ and LANGUAGE: $F(1, 180) = 9.89$, $p = 0.002$, as well as the interaction: STRESS* LANGUAGE: $F(1, 180) = 16.79$, $p < 0.001$.
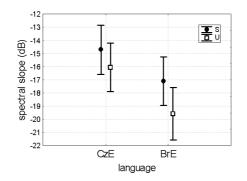

Figure 5: Spectral slope means stressed (S) and unstressed (U) vowels for Czech (CzE) and British (BrE) speakers. Whiskers indicate 95% confidence interval.

While the Czech speakers produce their stressed vowels with only slightly higher SPL (approximately 0.8 dB on average), the British speakers' difference is evidently much more conspicuous (around 2.5 dB on average). It is noteworthy that the unstressed vowels' level is more or less identical for both speaker groups.
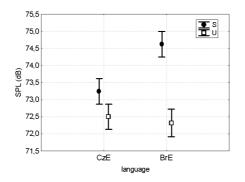

Figure 6: Sound pressure level (SPL) means stressed (S) and unstressed (U) vowels for Czech (CzE) and British (BrE) speakers. Whiskers indicate 95% conf. interval.

4. Discussion and conclusions

We have seen that Czech speakers realize English word stress differently from British speakers in several ways. First, they deviate more substantially from the canonical placement of stress, frequently shifting it to the first syllable of a word. This could be attributed to interference from their native language where the first syllable is stressed obligatorily. Interestingly enough, the British speakers also display some alterations in stress placement. Stress addition or loss on some words may be due to individual rhythmic patterning. When a stress shift on a word in our sample did occur, it was always in favour of the first syllable. As Cutler (2005) already noted, this could be caused by distributional asymmetry of the English word stress patterns.

Concerning the individual acoustic correlates of stress it can be said that Czech speakers in general use much less acoustic marking of stressed vowels than the British. The difference is most prominently seen in F0 (Figure 4) – where the Czech trend is in fact reversed (that is, lower F0 in stressed vowels). As Volín has demonstrated, the predominant F0 pattern in Czech stress-groups is a post-stress rise or L*+H (Volín, 2008). In news reading this is true not only for non-final, but also for final stress-groups, i.e., the post-stress rise occurs very often even in nuclear positions.

Similarly conspicuous difference between Czech and British speakers of English was found in their treatment of SPL (Figure 6), with Czech speakers showing little difference between stressed and unstressed syllables. In this case it could be argued that the perceptual importance of the English free stress is greater than that of the fixed Czech stress. Hence, such a physiologically expensive feature as SPL is habitually avoided by the Czech speakers.

On the other hand, the Czech speakers seem to achieve the required durational difference quite successfully (Figure 3). It can be inferred from the statistical analysis that they are almost native-like in their treatment of the temporal difference between stressed and unstressed vowels even though duration is not typically used as marking of Czech stressed vowels. However, the Czech language does exploit vowel duration to achieve the phonological quantity difference between long and short vocalic phonemes. It is perhaps possible to speculate that the speakers are somehow more sensitive to the temporal differences due to their phonological importance and, therefore, are able to transfer them to a different function in a foreign language. This speculation would, however, require corroboration from other languages.

When interpreting spectral slope differences, caution is needed, since all metrics (including α used here) are sensitive to vowel and speaker identity (Weingartová and Volín, 2014). If we compare the results with our earlier study on *schwa* (Volín *et al.*, 2013), the current results regarding spectral slope are in agreement but much clearer, due to the fact that the identity of the vowel was controlled for. Nevertheless, our findings show a significant effect of stress caused primarily by the unstressed vowels of British speakers, which exhibit the steepest spectral slope. This can be attributed to a lower vocal effort when pronouncing these vowels. Czech speakers display, on the one hand, a smaller difference between stressed and unstressed vowels and, on the other hand, a flatter spectral slope in general. We hypothesize that the experimental task of reading a relatively difficult text in a foreign language could elicit a tenser voice quality which is indeed associated with a less steep spectral tilt (e.g., Hammarberg *et al.*, 1980).

The differences between Czech and British realizations of word stress in English can be ordered according to their significance as follows: The most prominent marker of Czech-accented speech is SPL (in relation to perceived loudness). F0 (in relation to perceived pitch) and spectral slope (in relation to vocal effort) are also treated differently by the native and non-native speakers, but not to the same extent as SPL. Durational differences, on the other hand, show little divergence between both groups and seem to be easy to acquire for the Czech group of speakers: they behave native-like in this respect.

In conclusion, we can summarise the outcome of this study as follows: while the native speakers of British English behave in accordance with earlier literature on acoustic correlates of word stress (e.g. Fry, 1955, Beckman, 1986 or Sluijter and van Heuven, 1996), the Czech group produces significantly different stress manifestations. Since the detail of Czech word stress is not satisfactorily described as yet, our findings might help to illuminate the Czech prominence patterns that cause interference with the native ones. From a didactic point of view, this may be eventually useful for foreign language teaching and learning purposes.

References

Beckman, M. E. 1986. *Stress and Non-Stress Accent*. Dordrecht: Foris.

Boersma, P. and Weenink, D. 2013. *Praat: doing phonetics by computer* (version 5.3.41). Retrieved from http://www.praat.org/.

Crystal, D. 1996. The past, present and future of English rhythm. In M. Vaughan-Rees (ed) *Changes in Pronunciation*. Whitstable: IATEFL Pronsig.

Cutler, A. 2005. Lexical stress. In D. B. Pisoni and R. E. Remez (eds) *The handbook of speech perception*. Oxford: Blackwell: 264-289.

Fry, D. B. 1955. Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustic Society of America* 27: 765-768.

Fry, D. B. 1958. Experiments in the perception of stress. *Language and Speech 1*: 126-152.

Hammarberg, B., Fritzell, B., Gauffin, J., Sundberg, J. and Wedin, L. 1980. Perceptual and acoustic correlates of abnormal voice qualities. *Acta Otolaryngologica* 90: 441-451.

Janota, P. and Palková, Z. 1974. Auditory evaluation of stress under the influence of context. *AUC Philologica 2/1974, Phonetica Pragensia,* 4: 29-59.

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

Palková, Z. and Volín, J. 2003. The role of F0 contours in determining foot boundaries in Czech. *Proceedings of the 15th ICPhS* 2: 1783-1786. Barcelona: UAB.

Sluijter, A. and van Heuven, V. 1996. Acoustic correlates of linguistic stress and accent in Dutch and American English. *Proceedings of the 4th International Conference on Spoken Language Processing*, 630-633. Philadelphia.

Volín, J. 2008. Z intonace čtených zpravodajství: výška první slabiky v taktu [Intonation in newsreading: pitch of the first syllable in a stress group]. *Čeština doma a ve světě* 3-4: 89-96.

Volín, J., Weingartová, L. and Skarnitzl, R. 2013. Spectral Characteristics of Schwa in Czech Accented English. *Research in Language* 11(1): 31-39. DOI: 10.2478/v10015-012-0008-6

Weingartová, L. and Volín, J. 2014. Short-term spectral slope measures and their sensitivity to speaker, vowel identity and prominence. *Akustické listy* 20(1): 5-12.

Wells, J. C. 2008. *Longman Pronunciation Dictionary (3rd ed.)*. Harlow: Pearson Education Limited.