

PRE-FORTIS SHORTENING IN CZECH ENGLISH: A PRODUCTION AND REACTION-TIME STUDY*

RADEK SKARNITZL

Univerzita Karlova v Praze
radek.skarnitzl@ff.cuni.cz

PAVEL ŠTURM

Univerzita Karlova v Praze
pavel.sturm@ff.cuni.cz

Abstract

This study focuses on the production and perception of English words with a fortis vs. lenis obstruent in the syllable coda. The contrast is mostly cued by the duration of the preceding vowel, which is shorter before fortis than before lenis sounds in native speech. In the first experiment we analyzed the production of 10 Czech speakers of English and compared them to two native controls. The results showed that the Czech speakers did not sufficiently exploit duration to cue the identity of the word-final obstruent. In the second experiment we manipulated C and V durations in target words to transplant the native ratios onto the Czech-accented speech, enhancing the fortis–lenis contrast, and vice versa. 108 listeners took part in a word-monitoring task in which reaction times were measured. The hypothesized advantage to items in which the target word (with a fortis or lenis obstruent) was semantically congruent with the following context was not confirmed, and subsequent analyses showed that the words' frequency of use and the collocations they enter into strongly affect speech processing and correlate to a large degree with the reaction times.

Keywords: pre-fortis shortening, vowel duration, voicing, Czech English, reaction times

1. Introduction

Voicing is a phenomenon which has been investigated to a great extent in English as a native, as well as foreign language, especially when it comes to its implementation in word-initial plosives: languages differ substantially in the use of voice onset time (VOT) to cue the identity of initial plosives, and these differences have often been shown to lead to foreign-accentedness in non-native speakers (see Zampini 2008: 221–226 for a review of this line of research). This study focuses on the implementation of the voicing contrast in the word-final

* This study was supported by the Czech Science Foundation, GA14-08084S, and by the Programme of Scientific Areas Development at Charles University in Prague (PRVOUK), subsection 10 – Linguistics: Social Group Variation.

position, a context which has received comparably less attention, in the speech of Czech learners of English.

Czech belongs to the many languages which apply word-final devoicing (see, e.g., Blevins 2006), so that the contrast between voicing cognates is (near-)neutralized word-finally: in isolation, the word pairs *plot* (“fence”) and *plod* (“fruit”), or *les* (“forest”) and *lez* (“climb” [IMP]) are pronounced identically, as [plot] and [lɛs], respectively. On the other hand, English belongs to those languages in which underlying voicing distinctions are preserved word-finally: word pairs like *face* and *phase*, or *sack* and *sag* are therefore not homophones. However, since keeping full phonetic voicing at the end of a word would be very challenging, native speakers of English typically pronounce word-final lenis obstruents with partial or even complete devoicing: [sæǰ], [ferǰ]. In order to maintain the contrast between words like *sack* and *sag*, English exploits the duration of the preceding vowel. In what has come to be described as *pre-fortis shortening* or *clipping*, vowels before tautosyllabic fortis (voiceless) obstruents are shortened in English, well beyond the universal tendency for vowels to be shorter before fortis obstruents (Chen 1970). With this universal tendency greatly enhanced, vowel duration before fortis and lenis consonants has become phonologized in English (Kohler 1984).

In other words, the fortis–lenis contrast is preserved in the word-final position in English, unlike in Czech, but it is phonetically implemented not as a contrast in (phonetic) voicing but predominantly as a contrast in the duration of the preceding vowel. We may regard pre-fortis shortening in English as a form of cue trading (Repp 1982), which is also illustrated in the allophonic transcription of the above-mentioned word pairs: *sack* [sæ̃k] vs. *sag* [sæǰ], *face* [fɛ̃s] vs. *phase* [ferǰ]. In addition, the duration of the word-final sound may also contribute to the perception of voicing.

Since this way of implementing the voicing contrast in the word-final position is highly unusual in the world’s languages, it is to be expected that foreign learners of English will produce the syllable rhymes of English words differently from native speakers of English. Indeed, Smith, Hayes-Harb, Bruss and Harker (2009) showed that German speakers of English neutralized voicing word-finally in L2 English, although not to such an extent as in their native German. Interestingly, these relatively advanced speakers were able to exploit temporal cues to some extent. Preliminary studies suggest that Czech learners of English with a relatively strong foreign accent do not sufficiently exploit duration to cue the identity of the English word-final obstruents (Fejlová 2013). As for the degree of voicing, they tend to produce both the fortis and lenis obstruents without any phonetic voicing (with the exception of assimilatory contexts; see Skarnitzl and Šturm 2014).

In the present research, the English fortis–lenis distinction will be examined from the perspective of both production by Czech learners of English and perception. The production study (section 2) will analyze the temporal structure of target word pairs differing in the voicing status of the coda: Czech speakers of

English will be compared with native speakers of British English. The two groups are expected to differ considerably in the way vowel duration is used to cue the difference between word-final fortis and lenis sounds. The perception study (section 4) will examine the changes in reaction times due to manipulations performed on the duration of pre-fortis and pre-lenis vowels in recordings of native English and Czech English speakers.

2. Experiment 1 – Production

2.1. Method

For the purposes of this analysis, we recorded 10 Czech speakers of English (5 females, 5 males, aged between 20 and 32 years). They were all identified as having a relatively strong Czech accent in English (*cf.* Skarnitzl, Volín and Drenková 2005), but they were able to read a text fluently. In order to obtain reference values, we also recorded 2 native speakers of Standard British English (1 female, 1 male).

All the speakers were asked to read a set of 32 sentences. These were 16 pairs which contained target words like *rich* and *ridge*, *dock* and *dog*. All the words were monosyllabic and each pair differed only in the voicing status of the coda consonant. The sentences were designed so that the target word would not be predictable from the preceding text; some of the same sentences were used in the perception experiment (see section 4 for more details), where this was a necessary condition for the reaction time measurements to be reliable. The recordings were obtained in the sound-treated studio of the Institute of Phonetics at the Faculty of Arts in Prague with a 32-kHz sampling rate, using a high-quality condenser microphone, AKG C4500 B-BC.

The boundaries of the speech sounds in target words were segmented manually, following the guidelines by Machač and Skarnitzl (2009). In total, the analyses are based on 384 items (320 of which were produced by the Czech students of English and 64 by the native speakers of English). We measured the raw duration of the sounds in milliseconds, and also computed their relative duration within the word with the aim of neutralizing differences in speech rate. In the following section, we will only report on these latter, normalized duration values. We will consider the normalized duration of the vowel and of the coda consonant; since it is well known that voiced obstruents are shorter than voiceless ones (see, e.g., Jessen 1998), the latter value may also indirectly reveal the different treatment of the word-final obstruents by the two speaker groups in terms of phonetic voicing.

Statistical analyses were performed in the Statistica software and results displayed using the R package *ggplot2* (Wickham 2009).

2.2. Results and Discussion

As shown in Figure 1, the Czech and English speakers differed in the implementation of the phonological voicing contrast; an analysis of variance indicates that the difference was statistically significant: $F(1, 380) = 7.51, p < 0.01$. The native English speakers use vowel duration to cue the fortis–lenis distinction: as predicted, the vowel’s relative duration was significantly higher before lenis consonants than before fortis consonants (Tukey HSD test: $p < 0.001$). On the other hand, the Czech learners of English failed to produce the distinction adequately, beyond the universal tendency mentioned above ($p > 0.05$).

It is interesting to compare the durational implementation of the fortis–lenis contrast in individual speakers. It is clear from Figure 2 that the Czech learners of English do not form a homogenous group, especially with speaker F5 enhancing the contrast more than the others. It is clear that attainment, as far as foreign language pronunciation is concerned, may be comparable globally – i.e., with all the speakers manifesting a relatively strong accent in their L2 English – but differ when it comes to the realization of individual phonetic features. This is in accordance with the Dynamic Systems Theory (Larsen-Freeman 2011): second language acquisition cannot be treated as a uniform process, with learners placed on different places along one single trajectory.

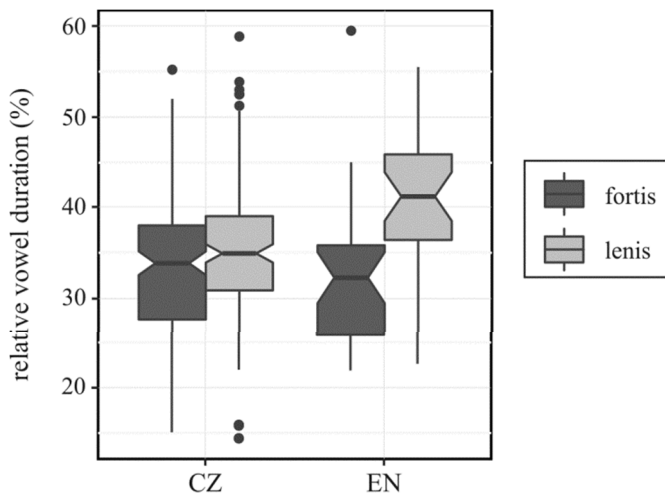


Figure 1. Relative vowel duration before fortis and lenis obstruents in Czech learners of English (CZ) and in native speakers of English (EN). Boxes indicate quartile ranges, whiskers denote ± 1.5 IQR from the quartiles, and outliers are plotted as individual points.

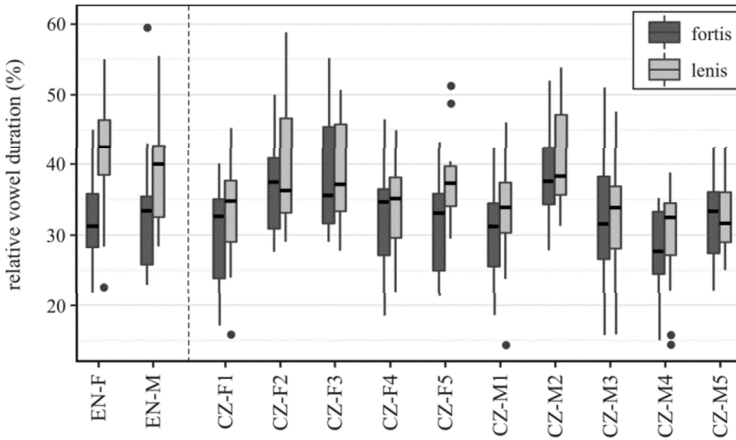


Figure 2. Relative vowel duration before fortis and lenis obstruents in the individual Czech (CZ) and native (EN) speakers. Boxes indicate quartile ranges, whiskers denote ± 1.5 IQR from the quartiles, and outliers are plotted as individual points.

The lack of differentiation of word-final fortis and lenis obstruents in Czech English is obvious not only with respect to vowel duration, but also in the duration of the final consonant itself (Figure 3). While the native speakers of English clearly maintained the contrast, with lenis (phonologically voiced) obstruents being relatively shorter than fortis (phonologically voiceless) obstruents, the Czech speakers of English produced these two classes of sounds in virtually an identical way. Therefore, it appears that the word-final fortis–lenis distinction is neutralized in their L2 English, as it is in L1 Czech: word pairs like *dock* and *dog* are then realized simply as [dok].

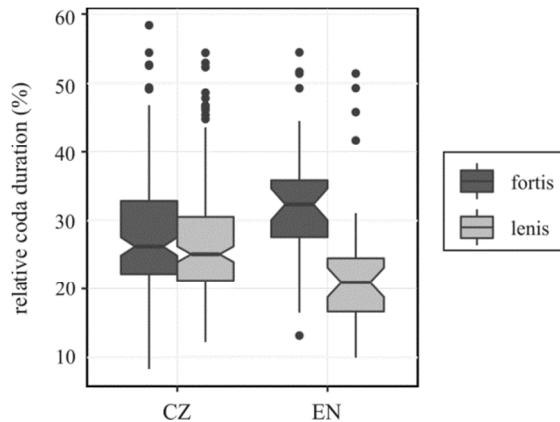


Figure 3. Relative duration of the coda obstruent in Czech learners of English (CZ) and in native speakers of English (EN). Boxes indicate quartile ranges, whiskers denote ± 1.5 IQR from the quartiles, and outliers are plotted as individual points.

Our production data thus confirmed the hypothesis regarding the implementation of the word-final fortis–lenis contrast in Czech learners of English: in general, they do not employ pre-fortis shortening to cue the difference, and the almost identical coda duration values suggest that the phonologically voiced obstruents were substituted with neutralized, fortis obstruents. The phonetic voicing of the final obstruent was not analyzed in this study, but it might provide additional support to the results presented here. On the other hand, the native English speakers maintained the contrast in the duration of the vowel, as well as of the coda consonant.

3. Reaction time measurements and comprehensibility

Before describing the perception study, it is necessary to introduce reaction time measurements and what they have been related to in the past. Foreign-accented speech has been related to various concepts. Many early studies focused on *accentedness* and asked listeners to assess how strong the speakers' accent is. However, the research of Derwing and Munro showed that foreign-accented speech needs to be addressed from the viewpoint of its intelligibility and comprehensibility (Munro and Derwing 1995; Derwing and Munro 2011), because even strongly accented speech can be quite intelligible and easy to listen to. *Intelligibility* refers to an objective indicator of how much listeners have understood of the speaker's message, while *comprehensibility* concerns the amount of effort the listener has to exert to understand the speaker. Comprehensibility is typically measured using self-reported assessments. In the experiment reported in section 4, we wanted to try out another way of examining the cognitive load associated with the processing of foreign-accented speech (in other words, *comprehensibility*) – one which would be more direct than subjective statements, namely reaction times of listeners to acoustic stimuli.

Measurements of reaction times (RTs) have been successfully applied in various psycholinguistic experiments, typically related to the recognition of given speech units – words, syllables, phonemes or phoneme sequences – in a stream of speech (see Kilborn and Moss 1996 for more information about word monitoring). Reaction time measurements are based on the straightforward assumption that a longer reaction to a given stimulus corresponds to a more demanding cerebral processing. In such experiments, the RTs and error rates are the measured dependent variables, and some manipulated or controlled factors serve as independent variables. These involve the nature of stimuli to which respondents are reacting, e.g. syllable complexity of the target words, lexical status of the target words, natural speech *vs.* speech with distorted rhythmic patterns or frequency of occurrence of the target words. Grosjean and Frauenfelder (1996) gathered 18 experimental paradigms which may be used to investigate the recognition of words from the stream of speech and showed that 14 of them rely on reaction times.

In a typical experimental study, Sendlmeier (1995) asked his subjects to monitor continuous speech and press a button as quickly as possible when they heard a target phone or syllable. The results of his study suggest that RTs to syllables were significantly shorter than to phones. Similarly, the subjects in the study conducted by Šturm and Volín (2012) were asked to react to a word printed on the screen; their reactions were shorter when the speech was segmented into syllables than when it was segmented into phones. An earlier study (Mehler et al. 1981) also showed a syllable effect for sequence monitoring in French, as fragments of speech were processed faster when they matched syllable boundaries (e.g. /pa/ in /pa.'lɛ/ and /pal/ in /pal.'mje/) and slower when the syllabification was incongruent with the target (e.g., /pa/ in /pal.'mje/). These and other studies point to the syllable rather than the segment as a crucial unit in speech perception.

The objective of the following experiment is to examine reaction times to native English and Czech English stimuli in which cues to the distinction between fortis and lenis obstruents in the word-final position have been manipulated.

4. Experiment 2 – Reaction times

4.1. Material and Stimuli

We used recordings of two native speakers of standard British English (the same ones as in the production study) and two Czech speakers of English with a strong Czech accent (which were also part of the Czech learner group in the production study), one female and one male in each group. The speakers were asked to read 10 sentence pairs which contained target words with a word-final fortis and lenis obstruent (see Appendix for the list of all sentences). The requirement was that the target word should not be predictable from the preceding context (i.e., from the words which occur before the target word itself), as illustrated by this example sentence (with the target words capitalized): *You really know to CEASE/SEIZE an opportunity when you see one.* To prepare stimuli for the subsequent perception experiment, durations of individual speech sounds of the target words were manipulated in Praat (Boersma and Weenink 2015) using PSOLA (Moulines and Charpentier 1990).

As mentioned in section 2, the Czech speakers of English (CzE) did not produce any significant differences in the vowel duration before fortis and lenis obstruents. Our goal was to enhance the duration contrast in the words pronounced by the Czech speakers of English. So as not to disturb the rhythmical structure of the sentences, the duration of all speech sounds (i.e., not only the vowels) was manipulated. The total duration of each target word was thus preserved, only its internal temporal structure was modified. The temporal

structure of the CzE words was copied from the corresponding words pronounced by the British speaker of the same gender.

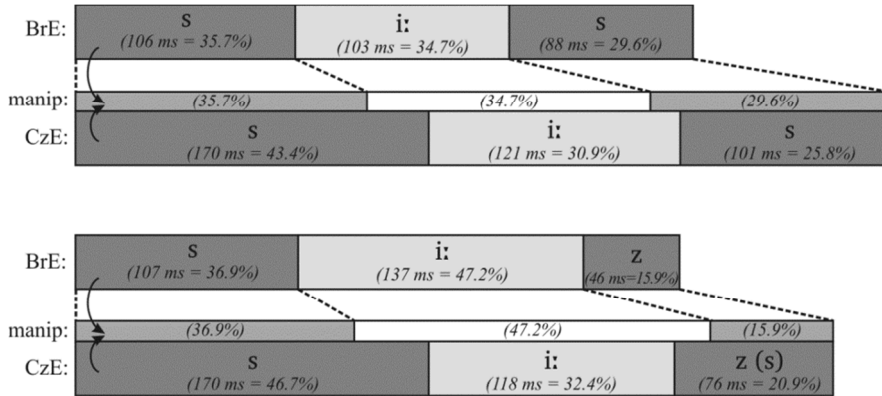


Figure 4. Illustration of manipulations of speech sound durations: the temporal structure of BrE words transplanted onto the corresponding CzE words.

Figure 4 illustrates these manipulations on the words *cease* /si:s/ and *seize* /si:z/, as pronounced by the female speakers. We can see that the /i:/ in both CzE words is very similar in the absolute and relative duration (121 ms in *cease*, corresponding to 30.9% of the word's duration, and 118 ms in *seize*, corresponding to 32.4% of the word's duration). In the British English (BrE) female speaker's word *cease*, the vowel represents 34.7% of its duration, while in *seize* it corresponds to 47.2% of its duration. As the figure shows, the BrE temporal structure was transplanted onto the CzE word: in the manipulated *cease* item, the vowel was slightly lengthened so that it equalled 34.7% of the word's duration; in the manipulated *seize* item, it was lengthened considerably more. The consonant durations were changed in the same way. The originally non-existent duration contrast between *cease* and *seize* was thus enhanced. We hypothesized that listeners' reactions to the target would be facilitated by these modifications and that reaction times to the modified CzE words would thus be shorter.

Conversely, the temporal structure of the native English words was modified according to the structure of the corresponding Czech words, so that the duration contrast was diminished in the British speakers' productions. This was expected to make reaction times to the modified stimuli longer.

As the example sentence mentioned above indicates, one of the sentence pairs was always semantically "correct" (*You really know to SEIZE an opportunity when you see one*). Although the target word was not supposed to be predictable from the preceding context, we hypothesized that semantic congruence in the later part of the sentence (*SEIZE an opportunity; SURGE of*

food prices) may yield shorter reaction times than the semantically incongruous version (*CEASE an opportunity; SEARCH of food prices*).

4.2. Test procedure

The perception test included 20 target items focusing on the fortis–lenis distinction (see Appendix for the target words and carrier sentences). The target items were interspersed with fillers so as to conceal the research objective from the subjects. In total, the test contained 96 items targeting different phonetic phenomena. The test was divided into four blocks, each taking approximately 5 minutes, with a short break in between during which the respondents chatted with the experimenter. In demanding perception tasks like this, it is crucial that the test is not too long, that subjects do not waver in their concentration. A training session preceded the test, during which subjects familiarized themselves with the procedure and speech material. The test items were presented to the listeners in four different, pseudo-random orders.

The perception test was administered to 108 Czech college-level students of English via the DMDX software (Forster and Forster 2003) run from a high-quality laptop. To guarantee as precise RT measurements as possible, a special device called the BlackBox ToolKit was used; this device has, unlike regular computer or laptop keyboards, minimal hardware latency. The subjects were asked to press a button on the device as soon as they heard the target word (e.g., “DOG”) printed on the screen; their response also launched a new item: a new target word appeared on the screen after the current item’s completion. When they could not identify the target word, they pressed another button to simply launch a new item. Each respondent was tested individually in the sound-treated recording studio of the Institute of Phonetics to ensure there were no disturbances. Respondents received partial credit for their participation.

4.3. Results

In total, 2160 responses were obtained (108 listeners \times 20 target items). Of these, 514 items (24%) had to be discarded because the RT values were out of range, i.e. longer than 1500 ms (which means that the reaction is not spontaneous) or shorter than 150 ms (which means that the listener reacted too quickly for it to really be a reaction to the target item). Researchers differ somewhat in the specific boundaries for identifying RT outliers, which is natural also due to varying task complexities. In the end, 1646 items therefore remained for analysis.

Figure 5 shows the overall results. The reactions to native English speakers, with diminished fortis–lenis contrasts (left panel), yielded no RT differences when the lenis target was semantically congruent with the sentence, in accordance with the hypothesis, but there was an advantage for the fortis targets when the fortis obstruent was expected. In contrast, the reactions to Czech

English speakers (right panel of Fig. 5) were expected to show an RT advantage for the congruent version of the target sentence because the fortis–lenis differences were enhanced by the manipulation. However, when the lenis obstruent was semantically appropriate, lenis items were associated with slower reaction times (713 ms) than the fortis items (676 ms), countering the expectation. When the fortis obstruent was congruent with the sentence, there was no difference between the fortis and lenis targets (643 ms × 646 ms, respectively). These results therefore suggest that other factors seem to interact with the semantic congruence, overriding the relatively small temporal differences between the fortis vs. lenis tokens.

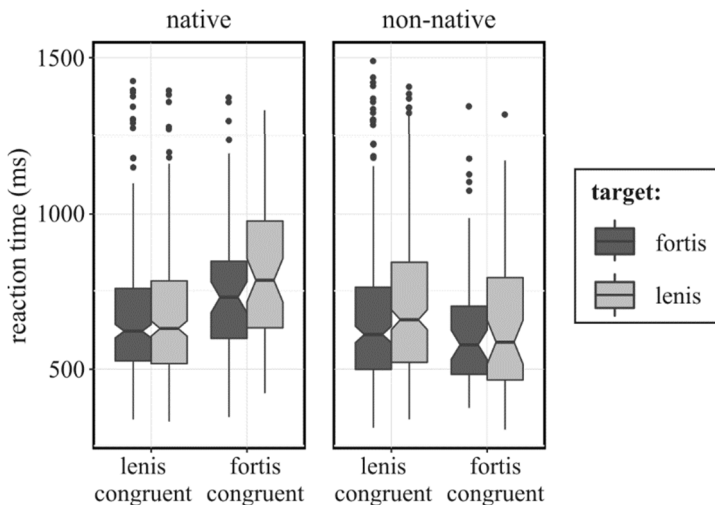


Figure 5. Reaction times (ms) to items spoken by native speakers of English (left panel) and Czech learners of English (right panel) when the word-final obstruent was fortis vs. lenis (boxes) and depending on the semantic congruence of the target word with the sentence (x-axis). Boxes indicate quartile ranges, whiskers denote ± 1.5 IQR from the quartiles, and outliers are plotted as individual points.

Two factors lend themselves as potential sources of facilitating influence in word monitoring: *frequency of use* of the target item and *collocation strength*. Words that are used frequently have been shown to be easier to process, and words that often appear together (i.e., words that are strongly collocated) are similarly associated with shorter RTs (Marslen-Wilson 1990; Bybee 2001). Therefore, we extracted absolute token frequency of the targets and logDice collocation measures between the targets and the previous words from the extensive *Araneum Anglicum Maius corpus* (ca. 1,200,000,000 tokens; Benko 2014).

Table 1 shows the relevant results for six selected items (three sentences, two target words) from each speaker group. In the native English group, the fortis vs. lenis items were expected to yield similar reactions because of the diminished

temporal differences. However, it is clear that in all three carrier sentences, the observed RT advantage was associated with items that manifest higher frequency of occurrence and (simultaneously) a higher measure of collocation strength. This happened irrespective of the target's semantic congruence with the end of the sentence. The lower part of the table shows the data for Czech speakers, whose speech *did* manifest differences between the target words with fortis and lenis obstruents, so semantic congruence should play a significant role. In all three cases both frequency of occurrence and collocation strength favoured the incongruent version; however, unlike in the native group, the reaction times were not faster to these favoured items (the last carrier sentence is an exception that *does* have an RT advantage, in the direction of the incongruent version slightly favoured by the two parameters). Nevertheless, the remaining four items (not shown in the table) cannot be so easily interpreted with respect to these variables, which is to be expected since the responses are likely to be influenced by other, unknown factors as well.

Table 1. Reaction times (RT), frequency of occurrence and collocation strength (expressed as a sum of logD measures) for semantically congruent and incongruent fortis/lenis target words in selected native British English and Czech non-native stimuli.

Group	Target word	RT (ms)	Frequency	Collocation (logD sum)
native	CODE (congruent)	592	134,000	7.69
	COAT	661	21,000	1.43
native	STACK (congruent)	760	20,000	2.72
	STAG	819	762	< 0
native	SURGE (congruent)	709	10,000	< 0
	SEARCH	675	195,000	1.62
non-native	GRAZE (congruent)	647	3,000	0.97
	GRACE	644	41,000	1.67
non-native	DOCK (congruent)	644	11,000	< 0
	DOG	646	104,000	1.14
non-native	SEIZE (congruent)	709	17,000	< 0
	CEASE	648	21,000	1.33

Finally, the influence of frequency of use and collocation strength is evident from their correlations with RTs. In most of the items, when the RT difference between the fortis–lenis targets was large, the difference between them in terms of the two measures was also large. Regarding the items in Table 1, RTs were negatively correlated with both target frequency ($r = -0.40$) and collocation strength ($r = -0.46$). However, when all items were analyzed, the correlations were weaker and less transparent ($r = -0.27$ for frequency and $r = 0.26$ for collocation strength).

5. Discussion

This study focused on production and perception aspects of the fortis–lenis contrast in Czech English. The production experiment reported in section 2 confirmed our hypothesis: Czech learners of English fail to exploit vowel duration to cue the contrast. In other words, pre-fortis shortening does not extend beyond the universal tendency for pre-fortis vowels to be slightly shorter than pre-lenis vowels. In addition, the relative duration of the coda was essentially the same for underlying fortis and lenis consonants, suggesting the transfer of coda voicing neutralization from L1 Czech. At the same time, it should be pointed out that the implementation of the word-final voicing contrast in English is very unusual in the world’s languages, and it is thus not surprising that learners with a relatively strong foreign accent have not acquired it.

The perceptual experiment reported in section 4 did not confirm the hypothesized advantage to items in which the target word was semantically congruent with the following context, and subsequent analyses showed that the words’ frequency of use and the collocations they enter into strongly affect speech processing. Although the objective was for the target words not to be predictable from the preceding context, the differing collocability scores of the target words with the preceding words indicate that much greater care must be taken of this aspect when designing reaction time experiments.

Moreover, the absence of significant results in the RT experiment, which was based on reactions to one single point of manipulation, may be overridden by the listeners’ processing of the whole sentence; in other words, the global comprehensibility of a speaker may have overridden the reduction of vowel duration contrast in native speakers and its enhancement in non-native speakers. It might well be, therefore, that in combination with higher and lower global comprehensibility of the native and non-native speakers, respectively, the net effect of the temporal manipulations is negligible.

It may also be the case that the manipulation of segmental contrasts – or, specifically, the duration differences which cue the fortis–lenis distinction in the final position – does not, on average, lead to faster or slower RTs, simply because segmental or sub-segmental characteristics are less conspicuous, especially when there is little danger of creating a confusion in the meaning of the whole sentence. It is also conceivable that native English listeners would be more sensitive to the relatively fine manipulations of the internal word structure. The suggestion that the manipulations of segmental contrasts do not yield significantly different reaction times seems to be supported by comparable data from Spanish (Černíková and Čermák, in print). This is in accordance with current models of phonology (Bybee 2001; Wedel 2011) and speech processing (Grossberg 2003; Goldinger and Azuma 2003), which no longer see the role of segments as central in word recognition and mental representations of speech. From another perspective, research findings from the past 15 years (e.g., Derwing and Rossiter 2003; Hahn 2004) provide evidence that it is the prosodic

features of speech which impact aspects like intelligibility and comprehensibility much more than segmental contrasts. That does not mean, however, that segmentals should not be targeted in pronunciation instruction, especially not in the case of contrasts which manifest a high functional load (Derwing and Munro 2015).

References

- Benko, V. 2014. Aranea: Yet Another Family of (Comparable) Web Corpora. In P. Sojka, A. Horák, I. Kopeček and K. Pala (eds.), *Text, Speech and Dialogue. 17th International Conference*, TSD 2014, LNCS 8655, 257–264.
- Blevins, J. 2006. A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics* 32, 117–166.
- Boersma, P. and D. Weenink. 2015. Praat: doing phonetics by computer. Version 5.4.20. [Computer programme]. Available from: <http://www.praat.org>
- Bybee, J. 2001. *Phonology and Language Use*. Cambridge: CUP.
- Chen, M. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22, 129–159.
- Černíková, Š. and P. Čermák. in print. Rysy cizineckého přízvuku ve španělštině českých mluvčích. *Časopis pro moderní filologii*.
- Derwing, T. M. and M. J. Munro. 2011. Intelligibility, comprehensibility and accent: Their relevance to pronunciation teaching. *Speak Out! IATEFL Pronunciation Special Interest Group Newsletter X*: 1–5.
- Derwing, T. M. and M. J. Rossiter. 2003. The effects of pronunciation instruction on the accuracy, fluency, and complexity of L2 accented speech. *Applied Language Learning* 13, 1–17.
- Fejlová, D. 2013. Pre-fortis shortening in fluent read speech: A comparison of Czech and native speakers of English. *AUC Philologica 1/2014, Phonetica Pragensia XIII*, 101–112.
- Forster, K. I. and J. C. Forster. 2003. DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, and Computers* 35/1, 116–124.
- Goldinger, S. and T. Azuma. 2003. Puzzle-solving science: The quixotic quest for units in speech perception. *Journal of Phonetics* 31, 305–320.
- Grosjean, F. and U. H. Frauenfelder. 1996. A Guide to Spoken Word Recognition Paradigms: Introduction. *Language and Cognitive Processes* 11/6, 553–558.
- Grossberg, S. 2003. Resonant neural dynamics of speech perception. *Journal of Phonetics* 31, 423–445.
- Hahn, L. D. 2004. Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly* 38, 201–223.
- Jessen, M. 1998. *Phonetics and phonology of tense and lax obstruents in German*. Amsterdam: John Benjamins.
- Kilborn, K. and H. Moss. 1996. Word Monitoring. *Language and Cognitive Processes* 11/6, 689–694.
- Kohler, K. J. 1984. Phonetic explanation in phonology: the feature fortis/lenis. *Phonetica* 41, 150–174.
- Larsen-Freeman, D. 2011. A complexity theory approach to second language development/acquisition. In D. Atkinson (ed.), *Alternative Approaches to Second Language Acquisition*, 48–72. London: Routledge.
- Machač, P. and R. Skarnitzl. 2009. *Principles of Phonetic Segmentation*. Praha: Epoque.

- Marslen-Wilson, W. D. 1990. Activation, competition, and frequency in lexical access. In G. T. M. Altmann (ed.), *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*, 148–172. Cambridge, MA: MIT Press.
- Mehler, J., Dommergues, J., Frauenfelder, U. H. and J. Segui. 1981. The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior* 20, 298–305.
- Moulines, R. and F. Charpentier. 1990. Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication* 9, 453–467.
- Munro, M. J. and T. M. Derwing. 1995. Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning* 45, 73–97.
- Repp, B. H. 1982. Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin* 92, 81–110.
- Sendlmeier, W. F. 1995. Feature, phoneme, syllable or word: How is speech mentally represented? *Phonetica* 52, 131–143.
- Skarnitzl, R. and P. Šturm. 2014. Assimilation of voicing in Czech speakers of English: The effect of the degree of accentedness. *Research in Language* 12(2), 199–208.
- Skarnitzl, R., Volín, J. and L. Drenková. 2005. Tangibility of Foreign Accents in Speech: the Case of Czech English. In: A. Grmelová, L. Dušková and M. Farrell (eds.), *2nd Prague Conference on Linguistics and Literary Studies Proceedings*, 11–20. Praha: PedF UK.
- Smith, B. L., Hayes-Harb, R., Bruss, M. and A. Harker. 2009. Production and perception of voicing and devoicing in similar German and English word pairs by native speakers of German. *Journal of Phonetics* 37, 257–275.
- Šturm, P. and J. Volin. 2012. Syllables versus phones in a word-monitoring experiment. *Studies in Applied Linguistics* 3(1-2), 25–34.
- Wedel, A. 2011. Self-organization in phonology. In M. van Oostendorp, C. Ewan, E. Hume and K. Rice (eds.), *The Blackwell Companion to Phonology*, 130–147. Oxford: Blackwell.
- Wickham, H. 2009. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.
- Zampini, M. L. 2008. L2 speech production research: Findings, issues, and advances. In J. G. H. Edwards and M. L. Zampini (eds.), *Phonology and second language acquisition*, 219–249. Philadelphia: John Benjamins Publishing.

Appendix

Carrier sentences used in the perception test (see section 4), with the target words printed in capitals.

I agreed with the SEARCH/SURGE of food prices.

The nasal RICH/RIDGE extends from the nasal root to its very tip.

For such situations they had to devise some COAT/CODE of conduct.

You're preventing yourself from staying on CLOUT/CLOUD nine.

There were some tattered books and a STACK/STAG of old newspapers on the shelf.

The sculpture of Diana riding a STACK/STAG accompanied by Cupid was priceless.

The descending DOCK/DOG was already a foot under water.

He felt the sudden pressure to like a DOCK/DOG on a tether.

These creatures tended to GRACE/GRAZE the grasslands near the forest.

You really know to CEASE/SEIZE an opportunity when you see one.