
ASSIMILATION OF VOICING IN CZECH SPEAKERS OF ENGLISH: THE EFFECT OF THE DEGREE OF ACCENTEDNESS

RADEK SKARNITZL

Metropolitan University Prague
skarnitzl@mup.cz

PAVEL ŠTURM

Institute of Phonetics, Faculty of Arts, Charles University in Prague
Pavel.Sturm@ff.cuni.cz

Abstract

Czech and English are languages which differ with respect to the implementation of voicing. Unlike in English, there is a considerable agreement between phonological (systemic) and phonetic (actual) voicing in Czech, and, more importantly, the two languages have different strategies for the assimilation of voicing across the word boundary. The present study investigates the voicing in word-final obstruents in Czech speakers of English with the specific aim of ascertaining whether the degree of the speakers' foreign accent correlates with the way they treat English obstruents in assimilatory contexts. L2 speakers, divided into three groups of varying accentedness, were examined employing *categorization* and a *voicing profile* method for establishing the presence/absence of voicing. The results suggest that speakers with a different degree of Czech accent do differ in their realization of voicing in the way predicted by a negative transfer of assimilatory habits from Czech.

Keywords: foreign accent, voicing, assimilation of voicing, Czech, English

1. Introduction

Voicing is a phenomenon which distinguishes members of speech sound pairs like /t/-/d/ or /s/-/z/. A cursory look at the IPA chart may suggest that voicing is a trivial binary contrast, accomplished by a simple “switch”, the vocal folds. A number of studies showed that this is not the case: the laryngeal muscles are activated and suppressed in a complex manner (see, e.g., Hirose & Gay, 1972; Löfqvist & Yoshioka, 1980; Ridouane, Fuchs & Hoole, 2006).

On the linguistic level, it is necessary to distinguish, on the one hand, voicing from the perspective of the phonological system of the given language, the so-called **phonological voicing**, and, on the other hand, voicing defined by the actual presence of vocal fold vibration, or what we call **phonetic voicing**. We may then talk about the phonetic implementation of phonological voicing, and this is a characteristic in which

individual languages may differ. Thus, while no one would dispute the existence of a two-way voicing contrast – in other words the existence of phonological pairs like /p/-/b/ or /s/-/z/ – in both Czech and English, the voicing contrast in these two languages is implemented differently from the phonetic point of view.

In Czech, we observe a considerable agreement between phonological and phonetic voicing (Skarnitzl, 2011). That means that phonologically voiceless obstruents like /p/ or /s/ are really pronounced as voiceless, without vocal fold vibration, while phonologically voiced /b/ or /z/ are typically produced with vibration of the vocal folds. Although devoicing of voiced obstruents does occur in Czech (Skarnitzl, 2011: Chapter 9), it may be regarded as online, momentary reduction of laryngeal gestures, rather than a systematic phenomenon. In addition, Czech neutralizes voicing word finally, so that words like *les* (*forest*, NOM, SG) and *lez* (*crawl*, IMPERATIVE, SG-2) will have identical surface phonetic representation, /les/.

In English, phonologically voiceless obstruents are pronounced as phonetically voiceless (and plosives are accompanied by aspiration in many positions), but phonologically voiced obstruents become partially or completely devoiced in many contexts (see, e.g., Ogden, 2009, p. 99ff.; Roach, 2009, p. 26ff.). For example, in the words *dog* or *buzz*, none of the consonants will be pronounced with full voicing but only with partial voicing; they will be transcribed [d̥ɒɡ̊] and [b̥ʌz̥]. As a descriptive category, voicing is therefore not very useful in some languages, and **tenseness** has come to be exploited to account for differences between phonologically voiced and voiceless obstruents (Ladefoged & Maddieson, 1996, p. 95; Butcher, 2004; Kohler, 1984). The former are called *lenis* or *lax*, suggesting a weaker pronunciation, while the latter are called *fortis* or *tense*, suggesting stronger articulation.

In the context of second language acquisition, the differences in the implementation of phonological voicing in Czech and English are likely to result in negative transfer from the learner's native language into the second language. We have to realize that the laryngeal settings associated with the distinction between fully and partially voiced obstruents are extremely fine. We believe that, due to this subtle difference, a fully voiced obstruent like [z] and a partially voiced or devoiced obstruent like [z̥] will be categorized as similar sounds in Flege's Speech Learning Model (Flege, 1987; 1995); the same would probably apply for a devoiced [z̥] and voiceless [s]. It is well known that our speech perception system is tuned into the phonological system of our native language (L1) and that we have learned to ignore the "ballast" of acoustic differences which are irrelevant for the distinguishing of L1 phonemic categories (see, e.g., Jusczyk, 1993; Kuhl & Iverson, 1995). It is precisely the relatively small acoustic differences between similar L1 and L2 sounds which are, according to Flege, rather difficult to acquire for foreign learners. Furthermore, pronunciation instruction in English as a foreign language only seldom targets voicing (Skarnitzl, 2002), and at least Czech students of English are rarely aware that voicing functions differently in the two languages.

On the one hand, therefore, we are talking about subtle distinctions on the laryngeal level and we might want to dismiss voicing as a marginal, irrelevant phenomenon in the acquisition of English as a second language. It is true that, to our knowledge, voicing has not been the target of much research from the perspective of second language

acquisition. However, we believe that when these L1–L2 differences in the implementation of voicing are combined with divergent strategies for the **assimilation of voicing** across the word boundary, the transfer effect may become much more perceptible for native listeners.

In Czech, (mostly) regressive assimilation of voicing is very pervasive: *pes[s] plaval* (the dog was swimming) – *pes[z] běžel* (the dog was running). In contrast, voicing assimilation is very rare in English: according to Cruttenden, only phonologically voiced fricatives may be realized as voiceless by some speakers “if the two words form part of a close-knit group” (Cruttenden, 2008, p. 299f.). He mentions *with[θ] thanks*, *off[f] course* or *was[s] sent* as examples. As the examples indicate, voicing assimilation concerns only grammatical words and, most importantly, voicing may only be lost, never gained. The pronunciation of the phrase *Have a nice day* with [nɑɪz deɪ] rather than [nɑɪs deɪ] leads to an impression of a foreign accent (cf. Roach, 2009, p. 212).

In a preliminary study, Skarnitzl & Poesová (2008) investigated the pronunciation of word-final obstruents by five Czech speakers of English who had been evaluated as having a strong Czech accent (see Skarnitzl, Volín & Drenková, 2005 for more detail). The results showed that, for instance, word-final neutralization of voicing was nearly ubiquitous in voiced–voiceless contexts across the word boundary (e.g., *called Peter*), or that in voiceless–voiced contexts (e.g., *back door*) the voiceless obstruent assimilated to the following voiced obstruent in approximately one half of all items.

The present study investigates voicing in word-final obstruents in Czech speakers of English; however, in addition to the previous study, we are adding the degree of the speakers’ foreign accent into play as a factor. In other words, we are interested in whether the degree of the speakers’ foreign accent correlates with the way they treat English obstruents in assimilatory contexts. Specifically, we hypothesize that speakers with a strong Czech accent in English will neutralize and assimilate voicing word-finally, depending on the underlying voicing status of the neighbouring sounds, while speakers who have a near-native English will not assimilate voicing in voiceless–voiced contexts, nor neutralize voicing in voiced–voiceless contexts.

2. Method

We examined twelve Czech female speakers of English, aged 20–25, all students of English and American Studies at the Faculty of Arts in Prague. Four of the speakers had been previously (Skarnitzl et al., 2005) classified as having a near-native English accent (henceforth referred to as Group A), four as manifesting a recognizable but not strong Czech accent (Group B), and four as having a strong Czech accent (Group C). The speakers were asked to read, after sufficient time for preparation, a news bulletin from BBC World Service. The recordings were obtained at the sound-treated recording studio of the Institute of Phonetics in Prague, using a studio electret microphone IMG ECM 2000 and digitized at the sampling rate of 22,050 Hz.

The recordings were automatically segmented using the Penn Phonetics Lab Forced Aligner (P2FA; Yuan & Liberman, 2008), and the boundaries of the target speech sounds were then manually adjusted following the recommendations listed in Machač & Skarnitzl (2009). By target sounds we will mean two consecutive phones – the word-

final obstruent and the initial sound of the following word; since the voicing of the former may not be independent of that of the latter in Czech speakers of English, we were interested in both of them.

In order to assess the voicing of the target sounds, fundamental frequency (F0) information – or more specifically, the presence or absence of F0 – was extracted at eleven equidistant points throughout the final obstruent using a Praat (Boersma & Weenink, 2013) script.

In subsequent analyses, we evaluated the degree of voicing using two methods. First, we characterized each word-final obstruent using the **categorization** proposed by Smith (1997) for [z] in American English: *voiceless* when voicing was present in less than 25 % of the sound’s duration, *voiced* when voicing was present in over 90 % of the sound’s duration, and *partially voiced* in the remaining cases. Second, we applied the **voicing profile** method (Shih & Möbius, 1998; Möbius, 2004) to assess the dynamic changes of the probability of voicing in normalized time. The probability of voicing at the given point of the obstruent’s duration refers to the ratio of voiced realizations to all realizations; for instance, there is a 95% probability of voicing in the temporal midpoint of Czech [z], but only a 85% probability in the midpoint of [ʒ], and as low as a 50% probability of voicing in the middle of the fricative trill [ʀ] (Skarnitzl, 2011, p. 215).

3. Results

The results suggest that speakers with a different degree of Czech accent do differ in their realization of voicing. Disregarding sequences of voiceless obstruents, we distinguished three assimilatory contexts at word boundaries: (1) the *voiced–voiceless* context (*vd-vl*), where we predicted the loss of voicing of the first obstruent in the least proficient speakers; (2) the *voiceless–voiced* context (*vl-vd*), in which we anticipated regressive assimilation of voicing in the least proficient speakers; and (3) the *voiced–voiced* context (*vd-vd*), where two voiced obstruents meet. Here, we predicted partial voicing of both obstruents in the proficient group, but either a full voicing of both obstruents or a loss of voicing of the first obstruent for the less proficient speakers.

As regards the **voiced–voiceless** context, the “Czech” neutralization (i.e., loss) of word-final voicing was strongest in Group C and weakest in Group A. There was a monotonic relationship between the number of realizations categorized as phonetically voiceless and the degree of Czech accent (Fig. 1a). The differences between the categories and groups were highly significant: $\chi^2(4; n = 849) = 39.2, p < 0.001$. The voicing profile, which adds a dynamic dimension to the results, revealed similar tendencies: the probability of voicing of the word-final obstruent decreased most rapidly in Group C (at 40% of the sound’s duration it was already low), and slowest in Group A (see Fig. 1b).

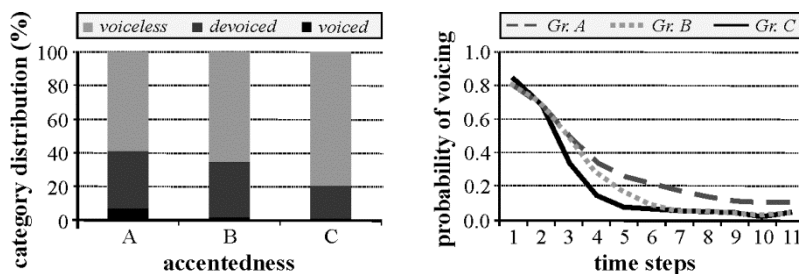


Figure 1: *Voiced–voiceless assimilatory context* for speakers with different degrees of accentedness (A = “native-like English”, C = “strong Czech accent”):
a. categorization according to Smith (1997); **b.** the voicing profile.

Unfortunately, the inverse **voiceless–voiced** context did not yield enough cases to allow drawing serious conclusions ($n = 84$). Group A behaved as expected, i.e. the proficient speakers kept the final obstruent voiceless in more cases than the other two groups, and relatively rarely assimilated voicing. However, Group B proved to be more “Czech-like” than Group C. This holds true both for the categorization according to Smith and for the voicing profile (Fig. 2a and 2b).

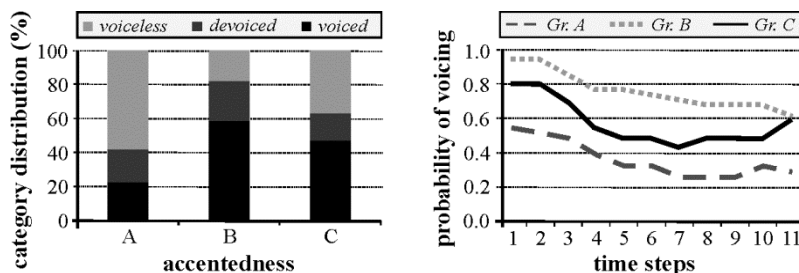


Figure 2: *Voiceless–voiced assimilatory context* for speakers with different degrees of accentedness (A = “native-like English”, C = “strong Czech accent”):
a. categorization according to Smith (1997); **b.** the voicing profile.

In sequences of **two voiced obstruents** across the word boundary, Group A manifested the greatest proportion of not fully voiced tokens, indicating the English-like process of word-final devoicing. The differences between the groups were significant: $\chi^2(4; n = 284) = 9.54, p < 0.05$. Although the probability of voicing was quite high for all groups, Group A has a drop in the profile towards the end, indicating a tendency for partial voicing (Fig. 3a and 3b).

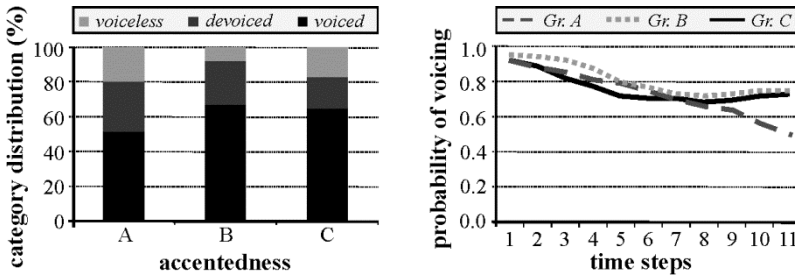


Figure 3: *Voiced–voiced assimilatory context* for speakers with different degrees of accentedness (A = “native-like English”, C = “strong Czech accent”):
a. categorization according to Smith (1997); **b.** the voicing profile.

In addition to these overall results, we analyzed the data with respect to **individual speakers** (four in each group). The first context (*vd-vl*) showed a gradual decline in the voicing profile for all Group A speakers as opposed to a sharp drop for the speakers from Group C. The individual speakers thus behaved uniformly, with the exception of a negligible two-cluster pattern in Group A.

The inverse context (*vl-vd*) revealed significant between-speaker differences. As can be seen in Fig. 4a, two speakers from Group A had a low probability of voicing, especially towards the end of the sound, while one seemed to voice all instances. However, this speaker contributed only four instances of this assimilatory context, which might explain her outlying position. Group B was more uniform, with three speakers assimilating voicing in all items and only one speaker demonstrating a lesser tendency to assimilation (Fig. 4b), comparable to speakers in Group A. Group C, with the highest degree of Czech accent, ranged from 0% probability to 100% probability of voicing; however, the two speakers in these extreme positions contributed only 2 and 3 items, respectively. The *vl-vd* assimilatory context must therefore be considered unrepresentative and the results only tentative.

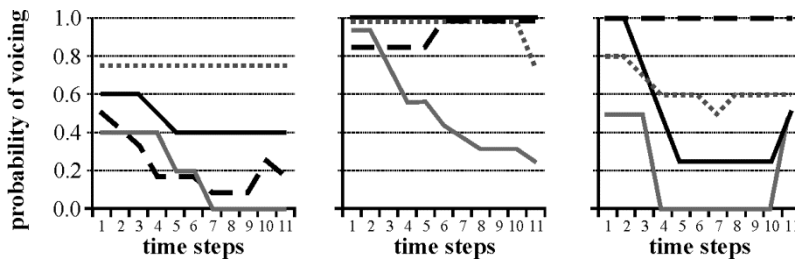


Figure 4: Voicing profiles (*voiceless–voiced assimilatory context*) for three groups of speakers (individual speakers denoted by different lines): **a.** native-like English (group A); **b.** mild Czech accent (group B) and **c.** strong Czech accent (group C).

The last context (*vd-vd*) did not reveal any significant differences between individual speakers within the groups. In Group A there was one speaker who diverged slightly, but

not markedly, in the direction towards full voicing. Group B was completely uniform, and Group C also included one subject who slightly diverged from the rest of the group towards full voicing.

Lastly, we were interested in other linguistic effects, such as the **type of word** the final obstruent appeared in. Restricting our attention to the two contexts with a sufficient number of cases (*vd-vl* and *vd-vd*), we found that the presence of voicing was stronger in synsemantic (function) than in autosemantic (content) words (see Fig. 5a and 5b). For instance, devoicing of the first consonant in *vd-vd* sequences was stronger in words like *called Barry* than *his brother*. Similarly, *world cup* in the other context was devoiced more frequently than *could face*.

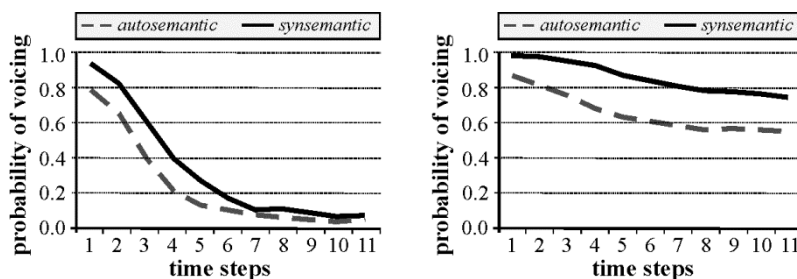


Figure 5: Voicing profiles for autosemantic vs. synsemantic words: **a.** voiced–voiceless assimilatory context and **b.** voiced–voiced assimilatory context.

The *vd-vd* context is of special interest: whereas speakers from Group A and B did not make any marked distinction between autosemantic and synsemantic words (Fig. 6a and 6b), speakers from Group C behaved quite differentially with respect to these classes (Fig. 6c). In their productions, the word-final obstruent in autosemantic words was devoiced in contrast to synsemantic words, in which it was fully voiced.

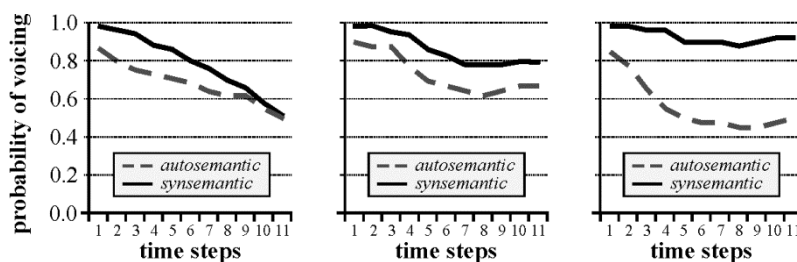


Figure 6: Voicing profiles for autosemantic vs. synsemantic words split into three groups of speakers (*voiced–voiced assimilatory context*): **a.** native-like English (group A); **b.** mild Czech accent (group B) and **c.** strong Czech accent (group C).

4. Discussion

The objective of this study was to examine the realization of voicing in Czech speakers of English with a differing degree of foreign accent. We focused on several types of sequences of two obstruents. First, we investigated the sequences of a voiced and a voiceless obstruent, such as *called Peter*. In native English, the word-final [d] tends to become partially devoiced, but does not turn into its voiceless (or, more precisely, fortis) counterpart, [kɔ:l̥d̥ pi:t̥ə]. In Czech, word-final voicing is neutralized, and we thus predicted a greater degree of devoicing in the speakers with a stronger Czech accent in their English and a lower degree of devoicing in the near-native-like group. Our results confirmed this hypothesis, with Group A producing the fewest voiceless items (i.e., with voicing present in less than 25 % of the sound's duration) and Group C the most (Figure 1). In English, some degree of devoicing of a final voiced obstruent is expected even if the following obstruent is phonologically voiced, such as *called Barry*. A similar tendency was observed in our most proficient speakers, who devoiced more than the other two groups (Figure 3). The results regarding the voiceless–voiced context (*back door*) are only tentative, given the low number of occurrences in our data, but they again show that speakers in Group A behave in accordance with expectations: most of their realizations are not fully voiced, which means that they did not produce the incorrect assimilation of voicing to the following voiced obstruent (Figure 2).

The data lend support to our division of speakers into three groups according to the degree of foreign accent, but we can still observe some idiosyncratic tendencies (Figure 4). That is to be expected, since the original evaluations are based on the listeners' overall impression in which the implementation of voicing presumably plays a marginal role.

Assimilation of voicing is typically not treated in much detail in the classic textbooks of English phonetics and phonology (see the Introduction). At the level of fine phonetic detail, however, the relationship between phonological and phonetic voicing in obstruent clusters across the word boundary may be more complex. From our cursory analyses, the phonetic voicing of the word-final obstruent does appear to partly depend, even in English, on the phonetic voicing of the following word-initial obstruent. In other words, the degree of voicing of the two sounds is not completely independent. It might therefore be useful, in our future research, to incorporate into the analyses the phonetic voicing of the following speech sound.

The more important line of our research will focus on the evaluation of wrong voicing assimilation by native listeners. While we do not expect this to be a phenomenon which should be stigmatizing, it may still play a role in the evaluation of a foreign-accented speaker (see, e.g., Gluszek & Dovidio, 2010).

Acknowledgement

The second author of this study was supported by the project GAČR P406/14-08084S, and by the Programme of Scientific Areas Development at Charles University in Prague (PRVOUK), subsection 10 – Linguistics: Social Group Variation.

References

- Boersma, P. and Weenink, D. 2013. Praat: doing phonetics by computer (Version 5.3.57). Retrieved from <http://www.praat.org>
- Butcher, A. 2004. "Fortis/lenis" revisited one more time: the aerodynamics of some oral stop contrasts in three continents. *Clinical Linguistics & Phonetics* 18: 547–557.
- Cruttenden, A. 2008. *Gimson's pronunciation of English*. London: Hodder Education.
- Flege, J. E. 1987. The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics* 15: 47–65.
- Flege, J. E. 1995. Second language speech learning: Theory, findings, and problems. In W. Strange (ed) *Speech perception and linguistic experience: issues in cross-language research*. Baltimore: York Press: 233–277.
- Gluszek, A. and Dovidio, J. F. 2010. The way they speak: A social psychological perspective on the stigma of nonnative accents in communication. *Personality and Social Psychology Review* 14: 214–237. DOI: 10.1177/1088868309359288
- Hirose, H. and Gay, T. 1972. The activity of the intrinsic laryngeal muscles in voicing control: An electromyographic study. *Phonetica* 25: 140–164.
- Jusczyk, P. W. 1993. From general to language-specific capacities: the WRAPSA model of how speech perception works. *Journal of Phonetics* 21: 3–28.
- Kohler, K. 1984. Phonetic explanation in phonology: the feature fortis/lenis. *Phonetica* 41: 150–174.
- Kuhl, P. K. and Iverson, P. 1995. Linguistic experience and the "Perceptual Magnet Effect." In W. Strange (ed) *Speech perception and linguistic experience: issues in cross-language research*. Baltimore: York Press: 121–154.
- Ladefoged, P. and Maddieson, I. 1996. *The sounds of the world's languages*. Oxford: Blackwell Publishers.
- Löfqvist, A. and Yoshioka, H. 1980. Laryngeal activity in Swedish obstruent clusters. *Journal of the Acoustical Society of America* 68: 792–801.
- Machač, P. and Skarnitzl, R. 2009. *Principles of Phonetic Segmentation*. Praha: Epocha.
- Möbius, B. 2004. Corpus-based investigations on the phonetics of consonant voicing. *Folia Linguistica* 38: 5–26. DOI: 10.1515/flin.2004.38.1-2.5
- Ogden, R. 2009. *An introduction to English phonetics*. Edinburgh: Edinburgh University Press.
- Ridouane, R., Fuchs, S. and Hoole, P. 2006. Laryngeal adjustments in the production of voiceless obstruent clusters in Berber. In J. Harrington and M. Tabain (eds) *Speech production: models, phonetic processes, and techniques*. New York: Psychology Press: 275–297.
- Roach, P. 2009. *English phonetics and phonology: A practical course*. Cambridge: Cambridge University Press.
- Shih, C. and Möbius, B. 1998. Contextual effects on voicing profiles of German and Mandarin consonants. *Proceedings of ICSLP 1998*: 3135–3138.
- Skarnitzl, R. 2002. Pronunciation practice overview in current textbooks of English. *ATE Newsletter* 13: 45–51.
- Skarnitzl, R. 2011. *Znělostní kontrast nejen v češtině*. Praha: Nakladatelství Epocha.

- Skarnitzl, R. and Poesová, K. 2008. Typology of voicing changes in Czech English. In A. Grmelová, L. Dušková, M. Farrell and R. Pípalová (eds) *Plurality and Diversity in English Studies – Proceedings from the Third Prague Conference on Linguistics and Literary Studies*. Praha: Faculty of Education, Charles University in Prague: 8–17.
- Skarnitzl, R., Volín, J. and Drenková, L. 2005. Tangibility of foreign accents in speech: The case of Czech English. In A. Grmelová, L. Dušková and M. Farrell (eds) *Proceedings from the Second Prague Conference on Linguistics and Literary Studies*. Praha: Faculty of Education, Charles University in Prague: 11–20.
- Smith, C. L. 1997. The devoicing of /z/ in American English: effects of local and prosodic context. *Journal of Phonetics* 25: 471–500.
- Yuan, J. and Liberman, M. 2008. Speaker identification on the SCOTUS corpus. *Proceedings of Acoustics '08*.