Filozofická fakulta Univerzity Palackého

Katedra anglistiky a amerikanistiky

The Production and Perception of Prevoicing in Voiced Stops in Native and Non-native English

(bakalářská práce)

Autor: Erika Kopčanová

Vedoucí práce: Mgr. Václav Jonáš Podlipský, Ph.D.

Olomouc 2016

Prohlašuji, že jsem tuto bakalářskou práci vypracovala samostatně pod odborným dohledem vedoucího práce a uvedla jsem úplný seznam citované a použité literatury.

V Olomouci dne 3.5.2016 ....

Acknowledgemens

I would like to thank my supervisor Mgr. Václav Jonáš Podlipský, Ph.D. for his time, useful advice, encouragement and materials he provided me with throughout the process of writing this thesis.

I would also like to thank Mgr. Jakub Bortlík for his time and willingness to help me with processing the data.

My thanks go also to Kamila Zukalová, Filip Ouroda, Sabrina Heinrich and Matthew Robinson for their kindness and willingness to participate in my experiment.

Table of Contents

1. Introduction

Many linguists have examined the perception and production of prevoicing–the presence of vocal cord vibration during the time of closure–in various languages. Starting from the earliest research studies, Lisker and Abramson (1964) investigated voice onset time (VOT) in eleven different languages. They referred to the term voiced onset time as to the time interval between the release of a stop closure and the onset of voicing (1967, 1). In terms of voicing, linguists call some phonemic categories *voiced* and other *voiceless*. It needs to be mentioned that there is a difference between phonetic and phonological voicing. Phonetically, the contrast between the terms *voiced* and *voiceless* consists in the presence or absence of vocal fold vibration during the consonantal closure interval (Hayward 2000, 178). As regards the phonological voicing, the two terms may denote "opposed members of a phonological contrast," but sometimes they may be pronounced differently from what their abstract phonological features propose (ibid 178). Czech and English differ in terms of realization of voiced plosives. They both distinguish phonologically voiced and voiceless stops. However, English voiced stops are often produced without prevoicing (Skarnitzl 2011, 71).

The purpose of this study is to examine the influence several factors may have on production of Czech and English voiced stops in word-initial position and to design a research proposal for perception experiment.

The pilot production experiment focuses on investigating the factors that may have an impact on production of VOT in word-initial plosives: language (Czech vs. English), the speaker's sex, lexical status of the word (real word vs. nonsense word), place of articulation (labial, alveolar, velar), speech rate at which single words are pronounced one after another (unspeeded vs. speeded), incidence of the words (frequently used vs. rarely used) and vocalic environment (high vowel vs. low vowel following).

The factors can be divided in two groups. The speaker's sex, place of articulation and vocalic environment depend on the physical structure of the vocal tract.

Previous research studies suggest that there may be a difference in VOT between males and females (e.g. Ryalls et al. 1997; Morris et al. 2008). However, the results vary. The study by Ryalls et al. (1997) demonstrated shorter duration of voiced plosives produced by women, while Morris et al. (2008) found no significant differences in VOT based on speakers' sex.

In several studies (e.g. Lisker and Abramson 1964, van Alphen and Smits 2004) prevoicing was influenced by the place of articulation in a way that the velar plosives were produced with shorter prevoicing.

Concerning the vocalic environment, Morris et al. (2008) and Klatt (1975) detected the longest VOTs in consonant-vowel (CV) syllables with the high vowels /u/ and /i/.

The other factors i.e. speech rate, lexical status, stress and lexical frequency are temporal. Kessinger and Blumstein (1997) proposed that a word loses its VOT with rising speed at which the word is pronounced (157). Based on this assumption, it is reasonable to propose that words pronounced faster will have shorter prevoicing than words pronounced at normal speed.

As regards stress and lexical frequency, previous studies showed that unstressed syllables and frequently used words are pronounced faster than stressed syllables and less frequent words (Ladefoged and Johnson 2011, 252; Lam and Watson 2014).

Although the study by van Alphen and Smits (2004) did not find any significant impact of the lexical status of the word on prevoicing, the nonsense words, produced for the first time, are expected to be articulated more carefully than the real words that speaker pronounced many times before. Based on this assumption, the non-words are expected to be hyper-articulated and therefore produced with longer VOT.

Czech and English men and women will be examined in order to investigate the impact of mentioned factors on prevoicing. Based on the existing findings, the author of this thesis assumes that all of the factors will influence the production of prevoicing in word initial plosives in isolated words.

The main goal of the perception experiment is to examine whether the Czech bilingual listeners are more likely to accept an English plosive as voiced when it is pronounced without prevoicing and it is produced by a female than when it is produced by a male, in other words, is the difference in the frequency of occurrence of prevoicing in males and females due to different vocal tract anatomy reflected in listeners' perception?

Another aspect to consider is whether the absence of prevocing in Czech is tolerated the same way by Czech monolingual and Czech bilingual speakers or whether the L2 of bilingual speakers influences their perception of L1.

Last but not least, the perception experiment would investigate whether the Czech bilingual speakers have it learned that in English, a voiced stop consonant without prevoicing can still be voiced or they do not tolerate the absence of prevoicing in English at all.

This introductory section provides a brief overview of the purpose of the study as well as the definition of the key terms together with the overview of the previous studies concerning similar matter. The purpose of the following section of this paper is to review the literature and to provide the theoretical background of the topic. Then it goes on the next chapter, which describes and discusses the pilot perception experiment and the research proposal for perception experiment.

2. Phonetic Background

This chapter focuses on explaining the basic terminology regarding the research topic. It includes consonantal categorisation and characteristics of English and Czech plosives.

2.1 Consonantal Categorization

Consonants can be classified in terms of many factors. Ladefoged and Johnson (2011, 17) distinguished five of them according to which consonants can be described. For the present study, four of the five factors are relevant: state of the vocal folds, place of articulation, the position of soft palate (raised to form oral sounds or lowered to create nasal sounds) and manner of articulatory action.

As regards state of the vocal folds, consonants can be either **voiced** or **voiceless**. The voiced consonants are characterized "by the presence of glottal buzz during the interval of articulatory closure", while in case of the voiceless plosives, the glottal buzz is absent (Lisker and Abramson 1964, 384). As already mentioned in the introduction to this thesis, the terms *voiced* and *voiceless* refer to abstract phonological features, which may or may not be phonetically realized as their phonological description suggests (Hayward 2000, 178). The contrast between Czech and English stop consonants can serve as an example for demonstrating the difference between the two phenomena. In Czech, the plosives /p, t, k/ are pronounced as voiceless unaspirated and the stop consonants /b, d, g/ are pronounced fully voiced, in which case the phonetic and phonological voicing are identical. In English, however, the plosives /p, t, k/ are pronounced as voiceless aspirated, whereas the consonants /b, d, g/ often lose their voicing (Skarnitzl 2011, 71).

Another factor for classifying consonants is place of articulation. Vocal tract consist of articulators that can form sounds. "The articulators that form the lower surface of the vocal tract are highly mobile. They make the gestures required for speech by moving toward the articulators that form the upper surface" (Ladefoged and Johnson 2011, 8). The categorization of English and Czech stop consonants is described in subchapters *2.2.1* and *2.2.2* below.

The involvement of <u>the position of soft palate</u> in the production of consonants divides them into **oral** and **nasal**, according to through which cavity the air flows at the moment of creating the particular consonant (ibid 8). This study concerns only the oral sounds.

The last but not least, there are three types of consonants divided according to <u>the manner of articulatory action</u>: **stops**, **fricatives** and **approximants**. In case of stop consonants, the articulators create an obstruction in mouth so no air can go through. To create a fricative, articulators narrow the space through which the air flows and for the approximants, the articulators "simply modify the shape by approaching the others" (ibid 14).

2.2 Characteristic of Plosives

Plosives, or oral stop consonants, are created by a complete articulatory closure in the mouth together with raised soft palate, preventing the airstream from escaping through the nasal cavity. While creating a voiceless plosive, the vocal folds are not vibrating, which causes a period of silence. During the closure, the air pressure is growing and is released once the articulators come apart, which creates a burst of sound (Hayward 2000, 177). The burst is a clear indicator of plosives, for that reason it is relatively easy to recognize stop consonants and at the same time the end of the closure interval in a spectrogram. However, it is not easy to identify the beginning of voicing in voiced plosives. Machač and Skarnitzl (2009) discussed the difficulties of boundary placements. They declared that in many cases there seemed to be either more possible signs for different boundary placements or no cues for it at all (13).

Based on VOT, plosives can be divided in up to four categories: voiced aspirated, voiced unaspirated, voiceless aspirated and voiceless unaspirated plosives (Van Alphen, Smits 2004, 388). According to the number of these categories a language can differentiate, Lisker and Abramson (1964) distinguished two, three and four-category languages (388). Czech belongs to two-category languages. It recognizes voiceless and voiced, both unaspirated plosives. Although, according to Lisker and Abramson (1964), English also belongs to two-category languages, because it primarily distinguishes voiced from voiceless plosives, but moreover, it contrasts voiceless aspirated and unaspirated plosives (Skarnitzl 2011, 71).

One of the main differences between Czech and English plosives is the way of their realization. Phonologically, both Czech and English distinguish voiced from voiceless stops. However, phonetically, Czech voiced plosives are almost always produced with prevoicing, while English voiced plosives can be produced without prevoicing (ibid 71).

Figures 1 and 2 demonstrate the contrast between voiced unaspirated [b] and voiceless aspirated [p^h]. The closure interval for voiced stops contains low-frequency harmonic components, while the closure of voiceless stops is completely blank (Lisker, Abramson 1964, 184).

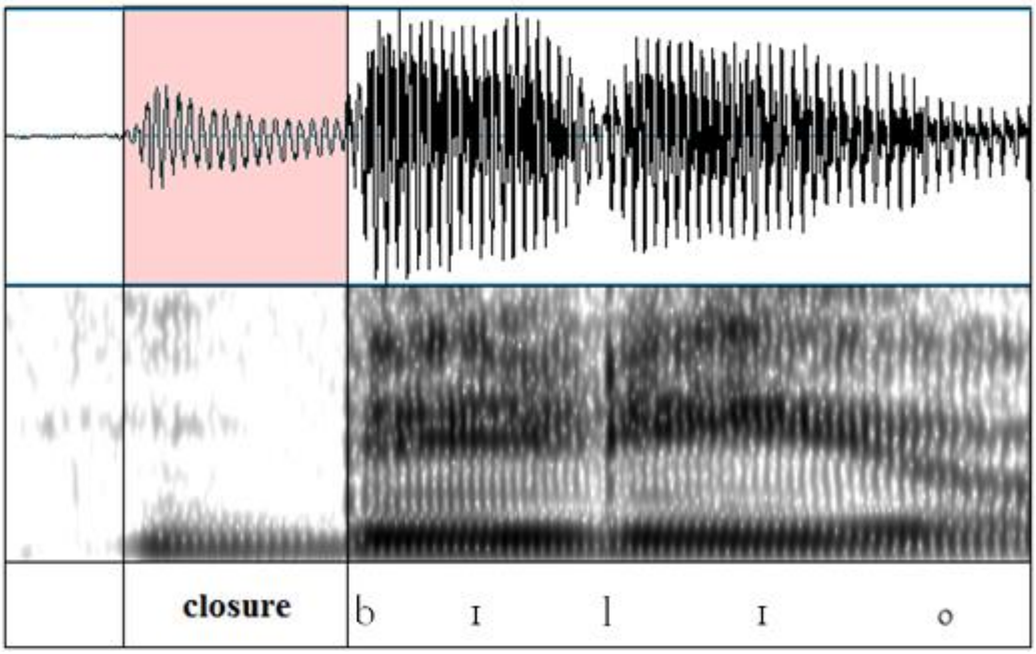

Figure 1 Waveform and spectrogram of the beginning of Czech word *billion* as pronounced by a female native speaker of Czech.

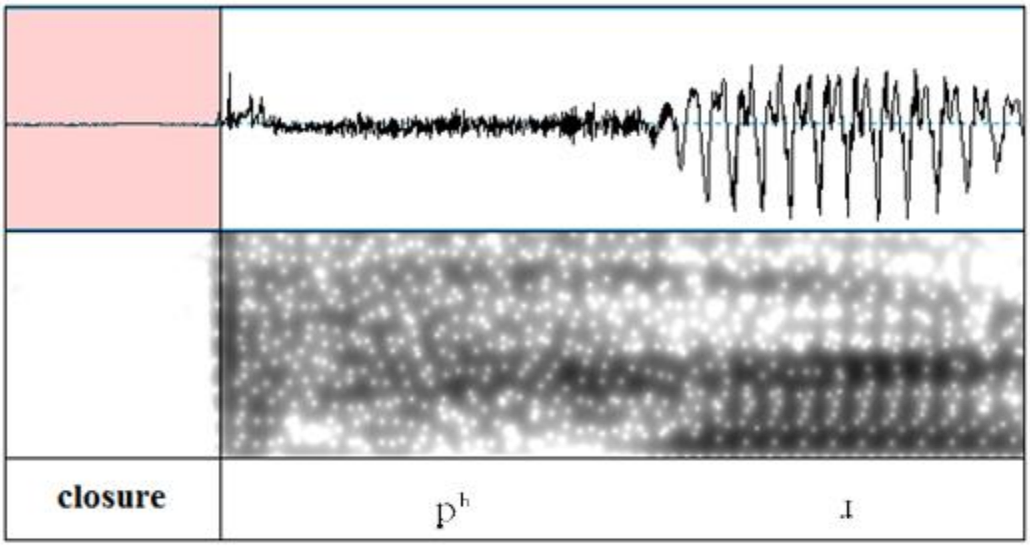

Figure 2 Waveform and spectrogram of the beginning of English word *percent* as pronounced by a female native speaker of American English.

In both cases the right line of the closure interval represents the release of the stop and the left line indicates the beginning of the closure, however, in Figure 2, it is not possible to decide where the closure began because both the closure of voiceless plosives and silence (no speech) is represented in the spectrogram with a blank space.

2.2.1 English Plosives

English distinguishes 3 pairs of plosives depending on the place of articulation–**bilabial** /p, b/, **alveolar** /t, d/ and **velar** /k, g/. These three pairs of plosives are distinguished according to where the obstacle takes place in the mouth. For bilabial plosives, both upper and lower lips are involved so that they create an obstacle by pressing against each other. In case of alveolar and velar stop consonants, the obstruction is created with the tongue pushed against either alveolar ridge for alveolar plosives or velum for creating velar plosives (Ladefoged and Johnson 2011, 11-13).

2.2.2 Czech Plosives

Apart from the bilabial /p, b/, alveolar /t, d/ and velar /k, g/, Czech contrasts one more pair of plosives–**palatal** /c, ɟ/ (Palková 1997, 225). This paper aims at directly comparing English and Czech and for this reason only the voiced plosives from the first three pairs are relevant.

3. Known Cues to Voicing in Stops in Perception and Production

The purpose of this chapter is to describe the factors that suggest voicing in stop consonants as regards perception and production. Moreover, the aim of this section is to show that prevoicing may not be as important for Czech listeners as it may seem and in its absence the stop could be perceived as phonologically voiced.

The occurrence of vocal cord vibration is in general considered to be the main cue for perceptual distinguishing between phonetically voiced and voiceless stop consonants. There are, however, other factors that may have a significant impact on voicing distinction, for example extent of F1 transition in the following vowel, intensity of burst, duration of closure, duration of the preceding (and following) vowel, or the onset F0 of the following vowel. These factors are described in following subchapters.

3.1 Duration of VOT and Aspiration

In a cross-language study carried out by Lisker and Abramson (1964), voice onset time (VOT) was defined as a time from the point of stop release to the beginning of vocal fold vibration. They divided stop consonants into three categories depending on the position of voicing in relation to release of the stop. The voicing can take place either before (for **voiced** unaspirated consonants), shortly after (for **voiceless unaspirated** consonants) or it can lag considerably behind the release (for **voiceless aspirated** consonants) (ibid 389).

Based on the range of VOT, they distinguished **negative** (-125 to -75 msec.), **short-lag** (0 to +25 msec.) and **long-lag** (+60 to +100 msec.) plosives (403). The negative prevoicing stands for voiced unaspirated stops, the short-lag VOT represents voiceless unaspirated stops and the long-lag VOT indicated the voiceless aspirated plosives (ibid 389). However, not all languages distinguish all these categories. Languages such as Hungarian and Dutch contrast two categories–voiced and voiceless unaspirated. Korean, Eastern Armenian and Thai belong to three-category languages–they differentiate voiced, voiceless unaspirated and voiceless aspirated stop consonants. As mentioned in the introduction to this paper, there are also four-category languages, for example Hindi and Marathi that contrast voiced aspirated, voiced unaspirated, voiceless aspirated and voiceless unaspirated plosives (ibid 388). As regards English and Czech, they both contrast primarily voiced /b, d, g/ and voiceless /p, t, k/ plosives (ibid 194; Skarnitzl 2011, 71). Except of these two features, English distinguishes voiceless unaspirated from voiceless aspirated plosives (Ladefoged and Johnson 2011, 57), which are not distinguished in Czech.

The length of prevoicing differs with the place of articulation. Several studies detected longer prevoicing in labial plosives and shorter in velar plosives (Lisker and Abramson 1964, van Alphen and Smits 2004). The possible reason is that the trans-glottal pressure equalizes quicker when producing /g/, since the volume of the vocal tract above the glottis is much smaller than when producing the labial plosive /b/ (Skarnitzl 2011, 56).

Another phonetic phenomenon that can help distinguish voiced from voiceless plosives is **aspiration**. The term is defined as "a period of voicelessness after the stop articulation and before the start of the voicing for the vowel" (ibid 57). As regards English, the presence of aspiration can be used as a reliable cue to voicing distinction. If aspiration is present, the word initial stop is voiceless. When there is no aspiration, the stop is voiced (Liberman, Delattre and Cooper 1958, 153).

Liberman, Delattre and Cooper (1958) believed that in English, the presence or absence of vocal cord vibration is not crucial for distinguishing between voiced and voiceless consonants as regards perception, because the distinction between these two categories is "quite perceptible in whispered speech" (153). However, van Alphen and Smits (2004), who investigated word-initial plosives in Dutch, found out that the presence or absence of prevoicing is the most reliable cue for listeners to decide whether a plosive is voiced or voiceless, yet "it is not reliably produced by speakers" (455).

3.2 Onset Frequency and Extent of F1 Transition

Another cue for the underlying (phonemic) voicing distinction is behaviour of the first formant at the beginning of the vowel following the stop. Liberman, Delattre and Cooper (1958) conducted an experiment, in which they examined "the elimination of a portion of the first-formant transition–as a cue for the distinction between voiced and voiceless stops in initial position" in English (154). This study showed an importance of presence or absence of cutback in the first formant for voicing distinction in word-initial plosives. All of the plosives /b, d, g/ changed from voiced to voiceless at different degrees of cutback in the first formant (ibid 160). The greatest amount of first-formant cutback for a voiced plosive to be perceived as voiceless was needed for alveolar plosive /d/ and little less for velar /g/. The labial plosive /b/ was perceived as voiceless at only a small amount of cutback in the first formant (ibid 157).

3.3 Intensity and Duration of Burst

According to studies carried out by Slis and Cohen (1969) and van Alphen and Smits (2004), who investigated differences between voiced and voiceless Dutch plosives, voiced plosives have shorter noise burst duration than their voiceless counterparts. Besides, they mentioned that the burst was stronger for alveolar plosives than for labials (471). Cho and Ladefoged (1999) have proposed that it might be caused by the difference in area of contact at obstruction for voiced and voiceless plosives. For the voiced plosives, the airflow is weakened when coming through vocal folds to the constriction in the mouth. Therefore, there is less energy needed for holding the obstruction and at the moment of the burst the pressure of the airflow is not strong enough to create a powerful burst. For the voiceless plosives, however, the airflow coming to the mouth is passing freely through the vocal folds, what creates bigger pressure on the constriction and subsequently a much stronger and longer bust. Moreover, "the vocal fold opening is considered to be fixed", so longer closure duration causes shorter aspiration and vice versa (213).

Repp (1979) investigated whether there is any sign for voicing distinction in the amplitude relationship between the aspiration noise and the following vocalic segment in syllable-initial stops. He proved that in English, aspiration amplitude significantly influences voicing distinction in a way that when the aspiration noise is strong, it is very likely to be perceived as voiceless. Van Alphen and Smits (2004), who investigated Dutch initial plosives, reached the same results in their study.

3.4 Fundamental Frequency (F0)

As mentioned in the introduction to this chapter, fundamental frequency, or F0, is one of the cues that help to distinguish between voiced and voiceless stops. The previous research studies of English showed higher F0 for vowels following a voiceless plosive rather than voiced plosives (e.g. Kingston and Diehl 1994; Umeda 1981, as cited in van Alphen and Smits 2004). In the study by van Alphen and Smits (2004), absolute F0 and F0 difference in Dutch initial plosives was one of the investigated factors. The results of their investigation showed that for the words started with voiced plosives, the mean F0 difference was positive with a rising F0, while for the tokens beginning with voiceless plosive the mean F0 difference was negative with a falling F0 (53). They suggested that the difference in F0 may be caused by "the lowering of the larynx during the production of voiced plosives in order to obtain sufficient trans-glottal pressure to produce vocal fold vibration" (469).

Based on previous research studies, van Alphen and Smits (2004) suggested that F0 might change depending on place of articulation. The results of their study demonstrated that in Dutch, for bilabial plosives the F0 was larger than for alveolar stop consonants (474-475).

In the case of perception, van Alphen (2004) discovered that "[t]he perception of voicing in labial plosives was influenced most strongly by the F0 difference from the burst of the plosive into the vowel" in a way that tokens with a higher F0 difference tended to be perceived as voiced (53-54).

Skarnitzl (2011) investigated the presence of F0 in plosives and fricatives in VCV position in Czech in read and spontaneous speech. His findings showed that voiced plosives in VCV position are 93.5% voiced (207). Based on previous research from different languages, he expected the biggest portion of devoicing–partial absence of F0–to occur with velar /g/. However, the results indicated that not the velar, but the alveolar plosive /d/ was devoiced the most. A possible explanation can be that in Czech, /d/ in CVC position is often pronounced as an alveolar tap (ibid 208). As regards the difference between read and spontaneous speech, the results of this study demonstrated that more devoicing in plosives in CVC position occurred in spontaneous speech (211).

4. Factors Affecting VOT

This chapter describes factors, which may influence the production and perception of prevoicing in word initial syllables. Some of these factors are driven by similar mechanism: place of articulation, speaker’s sex and vocalic environment depend on size of supra-glottal cavity, while the other factors i.e. speech rate, lexical status of the word, stress and lexical frequency are temporal.

4.1 Place of articulation

As already mentioned in chapter 2, there are three types of stops based on their place of articulation in English–labial, alveolar and velar. However, not every language distinguishes the same number of these categories. In Czech, there is one more type of stops as regards the place of articulation–palatal.

It is well documented that place of articulation has an effect on production of plosives. Lisker and Abramson (1964), who measured numerous languages including English, noticed that velar stop consonants have higher values of VOT than other plosives (399). In a study by Van Alphen and Smits (2004), the place of articulation proved to have an important influence on production of plosives in word-initial position in Dutch. According to this study, the words beginning with the voiced labial plosive were more frequently prevoiced than the words beginning with an alveolar plosive (464).

In terms of other languages, Helgason and Ringen (2008) investigated voicing in Swedish stop consonants. They detected longer prevoicing duration for labial and dental stop consonants than for velar stops (623). The reason for this is that the more anterior constriction the more space in the vocal tract and hence the more time for preserving voicing while pulmonic and supra-glottal pressure equalize. For velar plosives, the volume of the vocal tract between glottis and the obstruction is much smaller than for alveolar or labial plosives. For labials, the volume is the largest, because the tongue is not involved in maintaining the closure and also the cheeks can expand, which makes the volume between the glottis and the construction even larger and thus it is easier to produce and maintain voicing (van Alphen and Smits 2004; Helgason and Ringen 2008).

4.2 Influence of Speaker's Sex

The influence of speaker's sex proved to have a significant impact on prevoicing in several studies (e.g. Oh, 2011; Helgason and Ringen 2008; Whiteside et al. 2004). As mentioned in the introduction to this paper, the male-female differences in VOT are caused by different size of vocal tract. Men have larger vocal tract than women, which causes that the supraglottal pressure increases slower and it is easier to produce prevoicing (van Alphen, Smits 2004, 464-465). Essentially, it is the same principle that causes the place-of-articulation effect.

Morris et al. (2008) examined the influence of gender on production of voiced plosives in English. They recorded forty males and forty females producing CV syllables. However, they did not detect any significant impact of speaker's gender on production of VOT. Van Alphen and Smits (2004), who were concerned with Dutch, found out that prevoicing was more often detected in words produced by male speakers than in those produced by females. The difference in length of VOT between men (109ms) and women (89ms) was not significant (463). The reason for it might be the fact that they did not have a sufficient number of participants.

4.3 Vocalic Environment

Together with place of articulation and speaker's sex, vocalic environment depends on size of supra-glottal cavity. Therefore it is presumed that voiced stops in CV positions before low vowels will have more prevoicing than the ones before high vowels.

There are many research papers showing the influence of following phoneme on word initial plosives. In the study by van Alphen and Smits (2004), the Dutch plosives, which were followed by a vowel, were more frequently prevoiced than the ones followed by a consonant (464).

Other linguists examined the influence of high and low vowel on prevoicing in English consonant-vowel (CV) syllables. For example, Morris et al. (2008) and Klatt (1975) detected much longer VOT in English for voiceless plosives in CV syllables with /i/ or /u/ than for plosives in CV syllables with /ɑ/. Two factors might explain the difference between /ɑ/ and the rest of the vowels: (1) "lower peak oral airflows for the high vowels", which causes that the pressure decrease in the supra-glottal cavity to produce the trans-glottal pressure difference for creating the voicing will happen slowly and therefore the VOT will be maintained longer and (2) "an anterior, vertical pull on the vocal folds that would increase glottal resistance prior to high-vowel production", in which case "the phonation threshold pressure" will be more powerful and there would be more subglottal pressure needed to create the voicing (Higgins et al. 1998, as cited in Morris et al. 2008).

However, in a study by Lisker and Abramson (1967), who investigated English stops in word-initial position, the vocalic environment is labelled as "non-factor", because their data did not show any correlation between voice onset time and possible influence of surrounding vowels (15). The reason might be that their four native speakers of American English were also students of linguistics, which might lead to inaccurate results due to the speakers' knowledge in the field of study. Even though they recorded other ten speakers, who were "linguistically naive", the results appeared to be almost the same as with the four previous speakers (11).

4.4 Speech Rate

One of many factors that may influence VOT, a temporal characteristic, is the rate at which the words are pronounced. Lisker and Abramson (1967) labelled the utterance tempo as a "non-factor", because even though they had enough data, they did not find any connection "between VOT values and rate of syllable production" (15).

Kessinger and Blumstein (1997) examined the effect of speech rate on VOT in word initial labial and alveolar stop consonants in Thai, French and English. Long-lag voicing category in Thai and English together with prevoiced category in Thai and French proved to be significantly influenced by speech rate in a way that the faster the rate the shorter the VOT. As regards short-lag voicing category, no significant effect of speech tempo was detected in any of the three languages (157).

Klatt (1975) suggested two possible reasons why the speech rate influences the length of VOT. Either the duration of the consonantal closure in fast speech is too short for the glottis to fully open or it comes to “a reduction in the muscular effort expended to abduct the vocal folds” (694).

4.5 Lexical Status of the Word

Another factor influencing prevoicing is the lexical status of the word. It is reasonable to assume that nonsense words will be produced with greater prevoicing, since the speakers have never said or heard any of these words before and so they will articulate these words carefully. However, the study by van Alphen and Smits (2004) did not show any significant effect of this factor on prevoicing.

4.6 Effect of Stress

Lisker and Abramson (1967) examined the effect of stress on English stops in isolated words and sentences. They mentioned that in phonetic descriptions of English, voiceless stops are always aspirated when they occur at the beginning of a stressed syllable and always unaspirated at the beginning of an unstressed syllable in other than word-initial position. According to their findings, the highest VOT values for voiceless stops /p, t, k/ were detected “in stressed position in isolated words”, the lowest values were found in “unstressed position in sentences” and there was no significant difference in VOT values in “unstressed position in words and stressed position in sentences” (24). The author of this thesis expects stress to influence prevoicing significantly, because in a stress-timed language, such as English, stressed syllables last longer, therefore the VOT is longer (Ladefoged and Johnson 2011, 252).

In Czech, almost all words have stress on the first syllable (Palková, 1994). For that reason, the effect of stress on VOT will not concern Czech language.

4.7 Lexical Frequency

The same word is often pronounced multiple times during a discourse. A recent study by Lam and Watson (2014) examined sixteen native speakers of American English. The purpose of their study was to investigate whether "repetition reduction" is caused by "repetition of lexical items or repeated mention of referents". Even without the referent repetition, the repeated words tended to lose their duration and intensity, while the repeated referents seemed to lose their intensity only (ibid 8).

As mentioned above, the study by Kessinger and Blumstein (1997), who examined VOT in Thai, French and English, suggested that the faster the word is pronounced, the less VOT it has (157) and since the repeated lexical items tends to lose their duration, it is reasonable to believe that they lose their VOT as well.

Besides repeated mention within a discourse, the duration of prevoicing may be affected by the lexical frequency of the word, that is, how often the speaker uses the particular word in his everyday speech or encounters it in conversation. It is reasonable to suppose for words used more frequently to be more easily recognizable allowing speakers to produce voiced stops as less clearly distinct from their voiceless counterpart, i.e. with shorter prevoicing.

Lexical frequency as a possible factor influencing prevoicing will be examined in this study.

4.8 Other possible factors

This subchapter describes some additional factors that may influence the duration of prevoicing and its importance as a perceptual cue to voicing. It includes the influence of L2, the number of syllables of a word, whether the words are spoken in citation form or in connected speech and possible influence of position of the plosive in a word.

Production and prevoicing of word-initial stops in one language may be easily influenced by speaker's/listener's knowledge of other languages. The results from a study carried out by Kim (2011), who examined speakers of Standard Chinese, French, Hindi, Japanese and Korean, show that "there is a close relationship between L1 influence and proficiency in L2 and that L1 influences on L2 diminish as L2 proficiency grows" (97).

Van Alphen and Smits (2004) were wondering why their Dutch speakers did not produce prevoicing more reliably. One of their suggestions was that the speakers were all fluent in English with 6 years of learning English at school. In addition to this, all the speakers were in a daily contact with English on television and radio (487).

Concerning the number of syllables, the more syllables a word has the shorter each syllable and hence shorter VOT (Ladefoged and Johnson 2011, 101). Klatt (1975) examined the effect of the number of syllables of a word on voice onset time in English. The results of his study showed that voiceless consonants were shorter in two-syllable words than in words containing only one syllable and so was the closure interval and the duration of syllable nucleus (691).

As regards citation form versus connected speech, the pattern is similar to the one for speech rate–a word loses its prevoicing when it is pronounced quickly (Kessinger and Blumstein 1997, 157). In general, words produced in connected speech are pronounced faster than the ones in citation form.

There may be other factors, which may affect prevoicing. However, these factors are not taken into account in the present study.

5. Experiments

So far this paper has focused on theoretical background of the study. The following section will discuss the practical part. It is divided into pilot production experiment and research proposal for perception experiment.

5.1 Pilot Production of VOT

The purpose of this experiment is to obtain a portion of recordings for perception experiment described in section 5.2. The following factors should be taken into consideration as possible factors influencing the production of VOT: place of articulation, gender of a speaker, speech tempo, lexical status of a word, vocalic environment, stress and lexical frequency.

Based on previous research, it can be stated that place of articulation certainly affects production of prevoicing. Shorter negative VOT was detected in velar English plosives (Lisker, Abramson 1964; Helgason, Ringen 2008). Therefore, the author of this thesis expects velar plosives to be accompanied by shorter prevoicing than labial stops.

In terms of gender of the speakers, it is presumed that males will produce longer prevoicing than females. This presumption is based on the fact that men have larger vocal tract than women, what makes it easier for them to produce prevoicing (van Alphen, Smits 2004, 464-465).

As regards speech tempo, Kessinger and Blumstein (1997), who examined Thai, French and English initial labial and alveolar plosives, found out that words pronounced faster tend to lose their VOT (157). Based on this finding, the author of this thesis suggests that higher speech rate will lead to shortening of prevoicing in English and Czech word-initial plosives.

Concerning the lexical status of the word, even though the study by van Alpen and Smits (2004) did not show any significant influence of this factor on the production of VOT, it is presumed that nonsense words will be created with longer prevoicing than the real words.

In many research studies, word initial stop plosives in CV syllables followed by a high vowel were accompanied by longer prevoicing than stops followed by a low vowel (e.g. Morris et al. 2008; Klatt 1975). It is likely that the present study will show the same results.

Based on the study by Lisker and Abramson (1967), the factor of stress is expected to significantly influence the prevoicing in the way that stressed syllables will have significantly longer VOT than the unstressed ones.

Regarding the lexical frequency, it is expected that this factor will have an impact on duration of VOT similar to the results from the study carried out by Lam and Watson (2014).

To summarize this section, the author of this thesis expects all of the factors mentioned above to have an impact on production of VOT.

5.1.1 Method

Participants

Four people participated in this research–two Czech and two American native speakers. The gender of the speakers was equally represented in both languages–one male and one female of each nationality. The Czech participants were students of Palacky University in Olomouc; the American speakers were on an exchange program in Olomouc. The age of the speakers ranged from 21-31. None of the Czech speakers have lived in an English speaking country. Participants did not report any hearing or language impairment.

Materials

The author of this thesis prepared a list of words, which were read by the participants in a professional recording studio. It contains real words, nonsense words and fillers, which were included in order to prevent the speakers from revealing the purpose of the experiment, which might lead them to intentionally putting more stress on the first syllables of the words than they would normally do. All the word lists together with fillers are included in appendix.

Czech and English non-words are divided into six categories according to the place of articulation and the type of the following vowel. All of these words have stress on the first syllable.

English real words are divided into 18 categories. Except of the place of articulation and the type of the following vowel, English real words are divided according to their frequency and stressed syllable. The type of the following vowel was not taken into consideration as regards words with stress on the second syllable.

In reference to the frequency of the words, an online version of Macmillan Dictionary was used, where the words are rated with stars according to their frequency of usage in writing and speaking–three stars for the most frequent words, zero stars for infrequent words. Almost all of the high-frequency words from the list, which contains English real words, are rated with three stars. Words that were not rated with any star are included in the category of low-frequency words. Every category includes five items, except of high-frequency words beginning with velar stop with stress on the second syllable.

In the same manner as English real words, Czech real words are divided into categories based on place of articulation, type of following vowel and frequency. Czech words, naturally, lack the division of items according to stress.

As regards the frequency, a corpus of Czech language was used. Based on i.p.m. (instances per million positions) number, the border for high frequency is above 50 i.p.m. and the border for low frequency is under 10 i.p.m. This does not apply to words beginning with velar stop consonant, where the border is above 11 i.p.m. for high frequency and under 9 i.p.m. for low frequency. Although the border was lowered for items beginning with a velar stop, no Czech word qualifies as high-frequency with high vowel following the plosive in initial position and only three items were found to fulfil the conditions for high-frequency words with low vowel after the initial plosive. The reason for this might be the fact that most of the words beginning with “g” in Czech are loaned words.

Procedure

Four speakers, one after another, were seated in a professional recording studio and were asked to read words from Table 1-4, which were presented one by one on a computer screen. Subjects were instructed to read the words loudly and clearly in front of a microphone. Each speaker has read four lists containing words from their native language in this sequence: (1) unspeeded real words; (2) speeded real words; (3) unspeeded non-words; (4) speeded non-words Fillers were included in each of the lists such that half of the words were target words and the other half fillers. The productions were recorded using a Zoom H4n digital recorder with 41 kHz sampling rate and 16-bit quantization.

Measurements

Duration of prevoicing in every word starting with a voiced plosive was measured in Praat textgrid. The point in the time at which vocal fold vibration could be detected was considered the beginning of prevoicing and the end of prevoicing was defined as the beginning of the burst noise, which could be detected as a sudden peak in the waveform. Praat script[1] was used to measure the duration of prevoicing and to put the data in tables for better analysis. Figure 3 demonstrates an example of Czech word *dýka* starting with prevoicing.

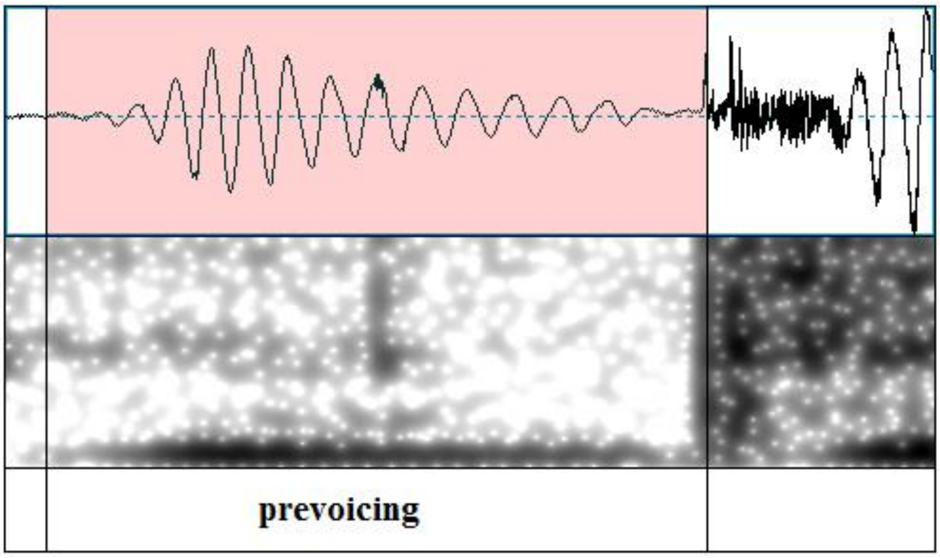

Figure 3 Waveform and spectrogram of the beginning of Czech word dýka /di:ka/ as pronounced by a Czech female.

[1] Acknowledgements to Mgr. Jakub Bortlík for writing the script.

A one-way and a two-way (factorial) ANOVA were used to analysed the data with duration of VOT as always the dependent variable.

5.1.2 Results

Not all words from the word list were taken into consideration in investigating each factor. For examining the influence of high and low vowel following the word-initial voiced plosive, only the words with the stress on the first syllable were included, since vowels in non-stressed syllables are usually pronounced as the mid-vowel /ə/.

Only the real English words were included in investigating the factor of stress, because as mentioned in chapter 4.6 of this paper, in Czech, almost all words have stress on the first syllable (Palková 1997) and including these words would lead to inaccurate results.

In examining the influence of lexical frequency of a word on the production of VOT, naturally, only real words were included.

The reason for excluding some categories of words in investigating the factor of vocalic environment, stress and lexical frequency of a word is that including all of the words from the word list would lead to inaccurate findings.

Figure 4 represents the number of prevoiced items in English and Czech. Out of 468 items, American speakers pronounced only 17.52% with prevoicing. In Czech, only one token (0.3%) was produced without prevoicing, the rest (99.7%) was prevoiced.

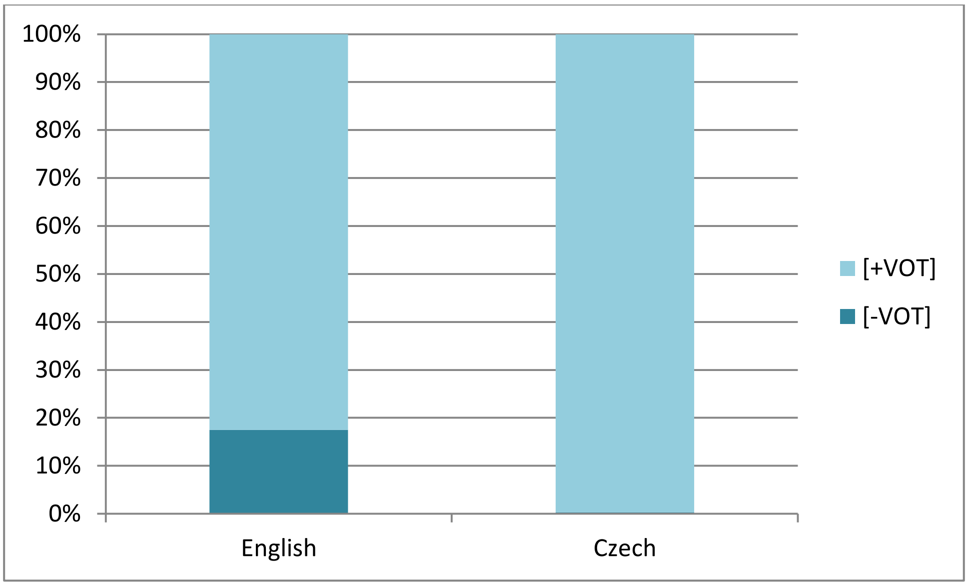

Figure 4 Percentage of prevoiced items in English and Czech.

A one-way and factorial ANOVAs (analyses of variance) were used to analyse the data, with duration of VOT always as a dependent variable. From all the examined factors, the most significant difference was found between Czech and English VOTs ($F[1, 875] = 677.33$, $p<.001$) in the expected direction with the mean -0.005 sec. for Czech language and -0.078 sec. for English language as presented in Figure 5. It can be seen that in English, the mean duration of VOT is close to zero, while in Czech it almost reaches the value -0.08 sec.

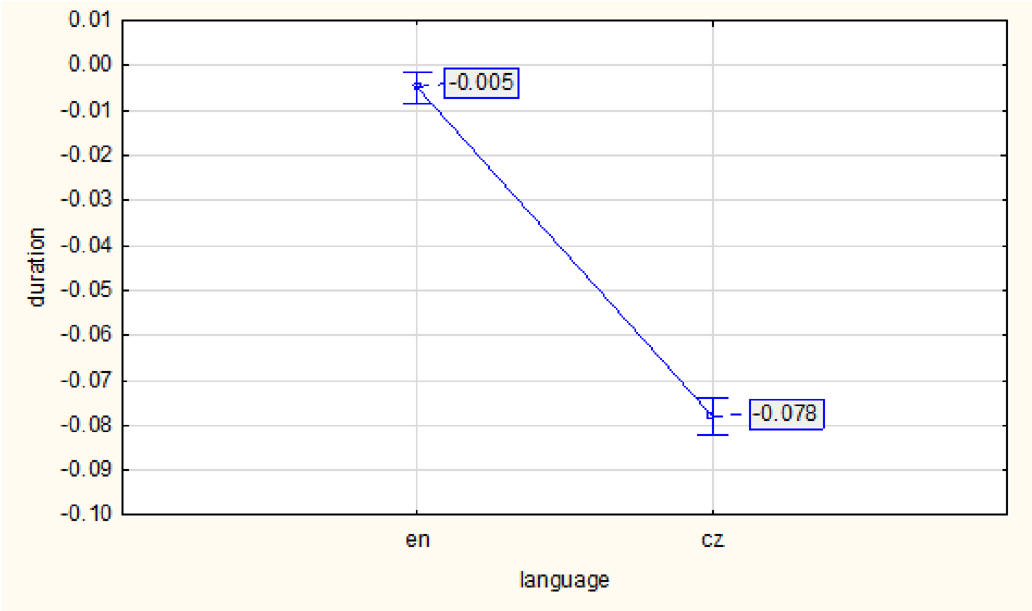

Figure 5 VOT means for Czech and English language in seconds.

A one-way ANOVA on VOT found a significant effect of gender in the expected direction with $F[1, 785] = 8.6244$, $p<.01$, presented in the Figure 6. As presumed, females produced shorter prevoicing (-0.03 sec.) than men (-0.04 sec.).

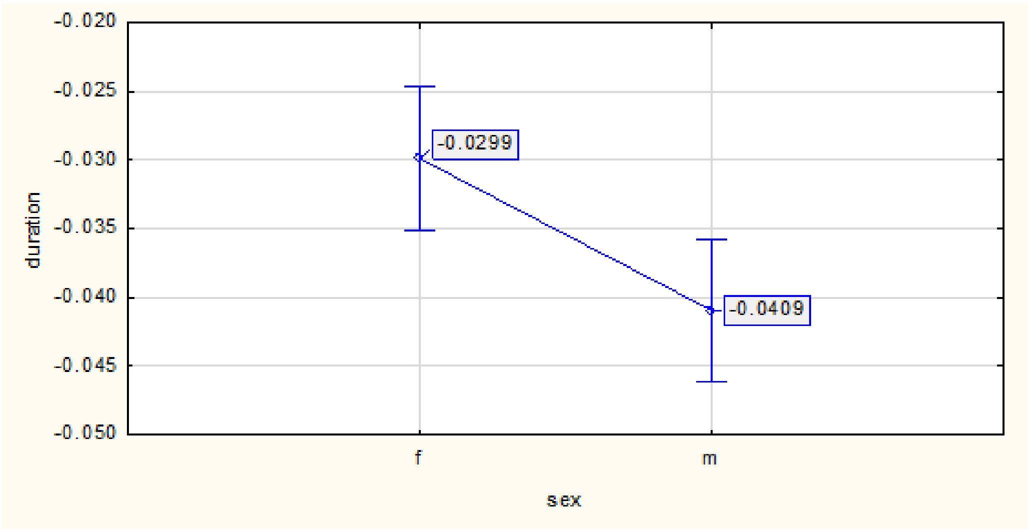

Figure 6 VOT means for speakers' sex in seconds.

Similarly to the gender, lexical status presented in Figure 7 with $F[1, 785] = 10.014$, $p<.01$ showed a significant effect on the production of VOT. As expected, the nonsense words were produced with longer VOT with mean duration -0.04 sec. than the real words with the mean duration -0.03 sec.

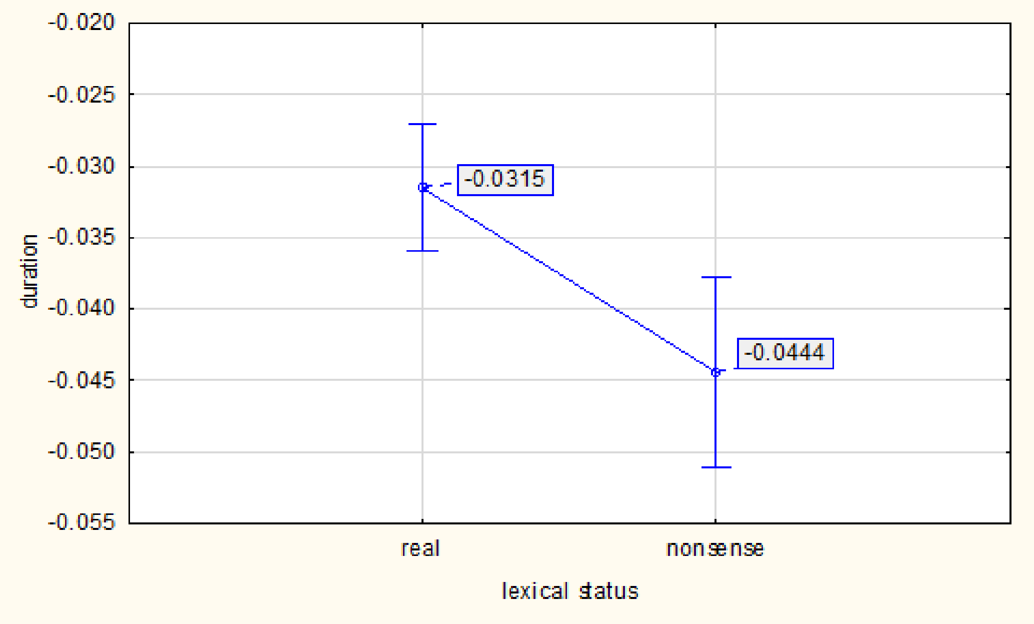

Figure 7 VOT means for lexical status of the words in seconds.

A one-way ANOVA found a significant impact of speech rate on production of prevoicing in the expected direction with $F[1, 785] = 55.8$, $p<.001$. As can be seen in Figure 8, the mean duration of VOT in unspeeded tokens is -0.05 sec., which is much longer than in speeded words, where the mean duration is -0.02 sec.

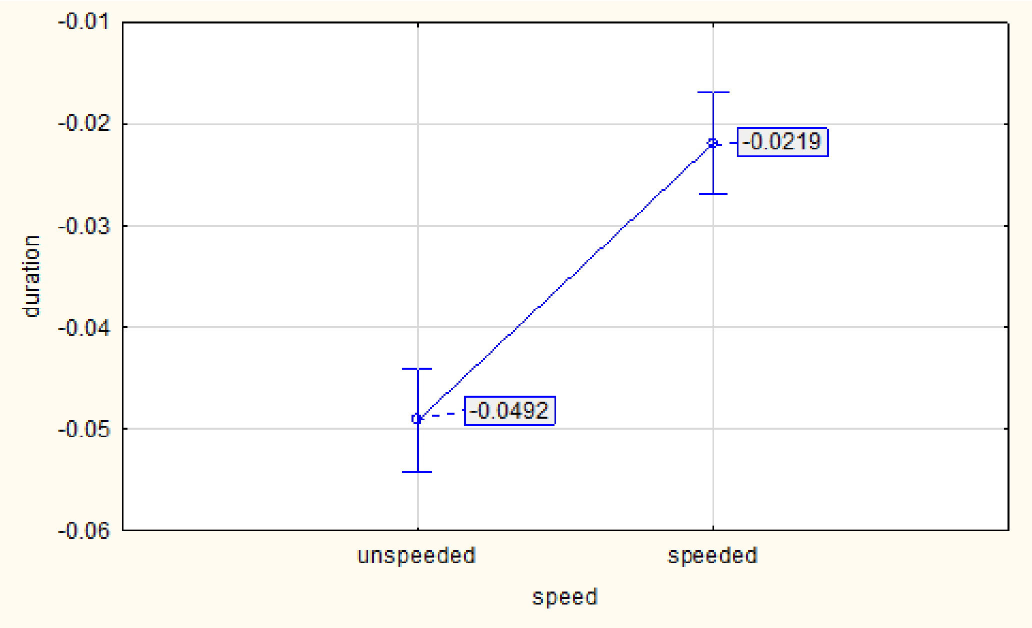

Figure 8 VOT means for speeded and unspeeded words in seconds.

Figure 9 demonstrates the VOT means for place of articulation. The longest prevoicing was produced in tokens beginning with labial plosives (-0.04 sec.) and the shortest negative VOT (-0.027 sec.) was produced in tokens beginning with velar plosives. Besides, a post-hoc LSD test found a significant difference between labial and velar stops ($p<.001$). Less significant difference was found between labials and alveolars ($p<.05$) and no significant difference was found between alveolar and velar plosives ($p<.12$).

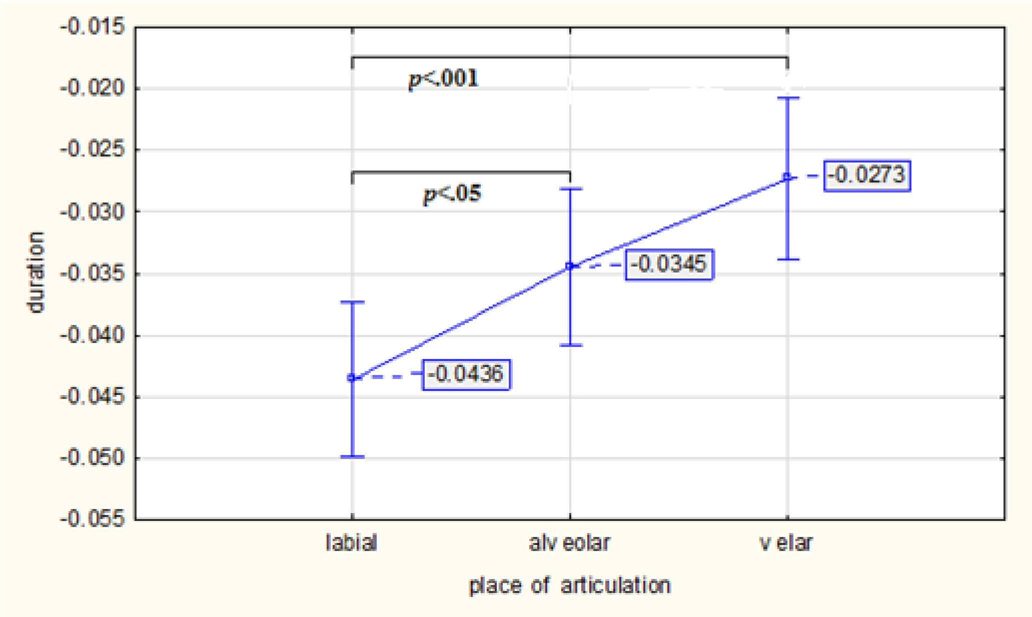

Figure 9 VOT means for place of articulation in seconds.

Another examined factor that may influence the production of VOT is lexical frequency. Figure 10 shows a significant effect of this factor on production of VOT in the expected direction with $F[1, 546] = 6.2020$, $p<.02$. The low-frequency words were produced with longer prevoicing with the mean duration -0.04 sec. than the more frequently used words with the mean duration -0.03 sec.

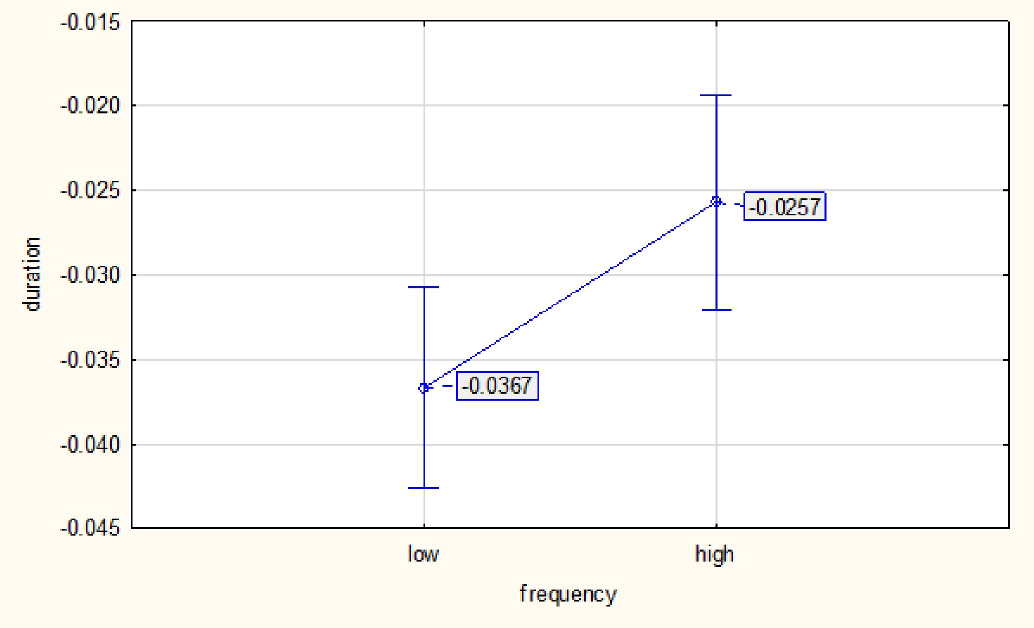

Figure 10 VOT means for lexical frequency in seconds.

Although the effect of stress influenced the production of VOT in the expected direction with $F[1, 337] = 1.4770$, $p=.225$, the difference between the mean duration of VOT for plosives in stressed (-0.0066 sec.) and unstressed syllables (-0.0002 sec.) was not significant.

The same finding goes with the effect of vocalic environment. The following vowel influenced the production of VOT in the expected direction with $F[1, 673] = .64175$, $p=.42$, though the difference between the mean duration of VOT for plosives followed by a low vowel (-0.0397 sec.) and plosives followed by a high vowel (-0.043 sec.) was not significant.

Presumably, the reason why the effects of stress and vocalic environment were not significant even though they influenced the production of VOT in the expected direction may be a lacking amount of data.

As already mentioned above, the one-way ANOVAs on VOT found a significant effect of language and speech rate. Figure 11 demonstrates the percentage of words starting with voiced plosive produced with positive and negative VOT. In Czech, 100% of the voiced plosives were prevoiced in unspeeded words and almost all (99.4%) plosives were produced with negative VOT in speeded words. In English, however, only 30.77% of plosives were produced with negative VOT in unspeeded words and 4.27% in speeded words.

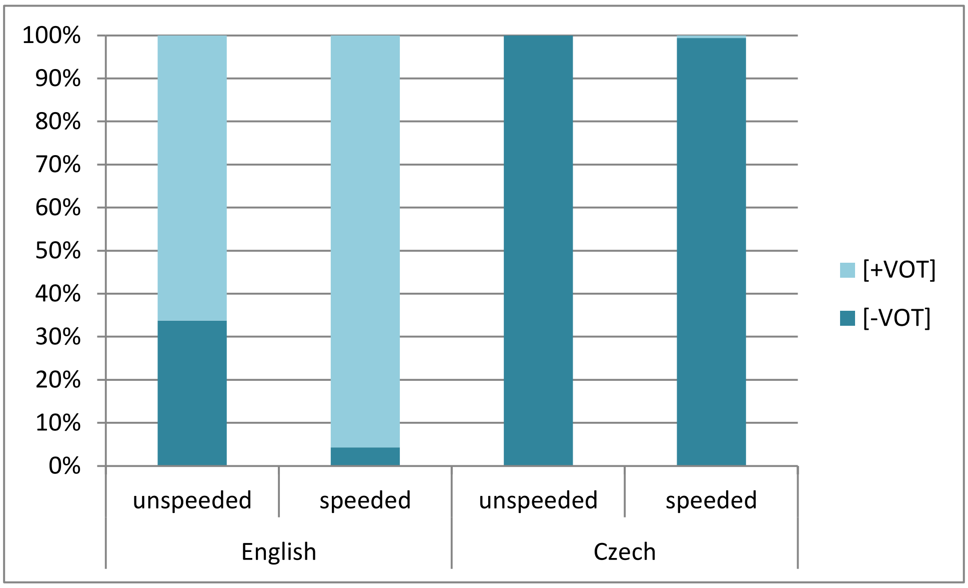

Figure 11 Percentage of tokens beginning with voiced plosives produced with and without prevoicing in speeded and unspeeded condition in Czech and English.

Separate factorial ANOVAs were used to test potential interaction between language and the other factors, i.e. place of articulation, speakers' sex, speech rate, lexical status, vocalic environment, stress and lexical frequency, respectively. Unsurprisingly, no interaction between language*place of articulation, language*sex, language*lexical status, language*vocalic environment, language*stress or language*lexical frequency was found. However, a factorial ANOVA on VOT with language and speed as factors found a significant interaction between the two factors. This interaction is represented in Figure 12. In Czech, the difference in the duration of VOT in tokens pronounced fast is not so big. The red line in the diagram represents the duration of VOT for speeded words. For Czech, the mean duration of VOT is -0.07 sec., while in English it is above zero (0.01 sec.). The duration of VOT for unspeeded words is represented by the blue line in the diagram. For Czech, the mean duration of VOT is -0.09 sec. and for English the mean duration is only -0.2 sec. As can be seen, the difference between the duration of VOT for speeded and unspeeded words in Czech is not so big, while it is more visible for English.

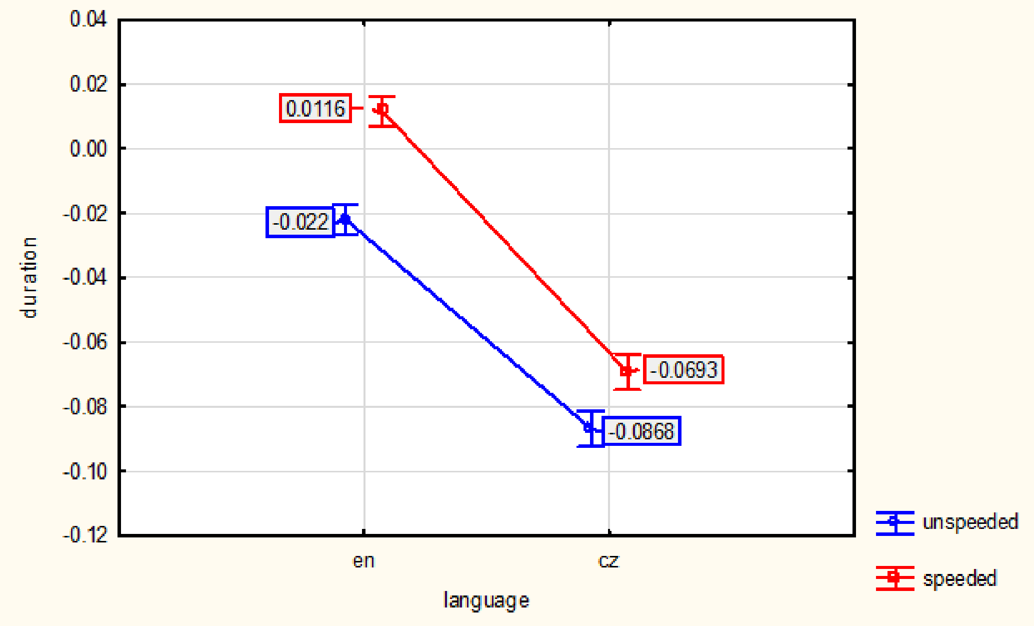

Figure 12 Factorial ANOVA showing an interaction between language and speed. VOT duration is in seconds.

5.1.3 Discussion

As expected, all of the examined factors i.e. place of articulation, speakers' sex, vocalic environment, speech rate, lexical status of a word, stress and lexical frequency had an impact on production of VOT in word-initial plosives in Czech and English, only the effect of stress and lexical frequency was not significant. All of the factors influenced the production of VOT in the expected direction.

The results also pointed out a significant difference in production of prevoicing between the two examined languages. American speakers produced only 17.52% of tokens beginning with a voiced plosive with prevoicing, while in Czech, 99.7% of the items beginning with a voiced plosive were prevoiced. Moreover, the factorial ANOVA with language and speed as factors found an interaction between the two examined languages. As figure 11 shows, Czech plosives were produced with 100% of prevoicing in unspeeded words and 99.4% in speeded words, while in English only 30.77% of voiced stops were produced with negative VOT in unspeeded words and almost no prevoicing (4.27%) was produced in speeded words. This finding demonstrates the difference between the phonetic and phonological voicing in Czech and English. In Czech, the phonological abstract feature "voiced" is almost always phonetically realized as voiced, while in English, the realization of phonologically voiced plosives does not always correspond with its classification and therefore it can be pronounced without prevoicing (Skarnitzl 2011, 71).

The expected significant influence of place of articulation of the production of VOT was detected. The longest VOT was found in the words beginning with labial plosives, shorter with alveolar plosives and the shortest prevoicing was found in words beginning with velar stops. Similar results were found in Dutch (van Alphen and Smits 2004) and Swedish (Helgason and Ringen 2008), where the tokens starting with labial or dental plosive were more frequently prevoiced and had longer negative VOT than the words beginning with velar stops. Also the cross-language study by Lisker and Abramson (1964) showed higher negative VOT values for velar plosives. This finding suggests that the smaller volume of vocal tract above glottis the shorter prevoicing.

While there was no significant effect of speaker's sex on VOT found in neither English (Morris et al. 2008) nor Dutch (van Alphen and Smits, 2004), the results of the present study showed a significant effect of this factor on VOT. As expected, women produced shorter prevoicing with the mean duration -0.03 sec. than men with the mean duration -0.04 sec. Like place of articulation and vocalic environment, speaker's sex is based on the physiological structure of the vocal tract–the less space in the oral cavity, the shorter prevoicing (van Alphen, Smits 2004, 464-465).

Another examined factor was the vocalic environment. Similarly to the results of the studies by Morris et al. (2008) and Klatt (1975) in English, the present study detected longer negative VOT before high vowels than before low vowels in both of the investigated languages. However, this influence was not significant and even though the results of the present study are together for Czech and English, the author of this thesis suggests that the prevoicing produced in Czech language could not influence the finding, since the factorial ANOVA did not show any significant interaction between language*vocalic environment.

The temporal factors i.e. speech rate, lexical status, stress and lexical frequency all influenced prevoicing of VOT in the expected direction. The results demonstrated that the faster a word is pronounced, the shorter it is and hence the prevoicing is shorter. However, stress did not influence the production of VOT significantly.

Although the study by van Alphen and Smits (2004) did not find any significant difference of speech rate on production of prevoicing, the results of this study showed that speech rate significantly influenced the duration of prevoicing in a way that unspeeded tokens were produced with more prevoicing (-0.05 sec.) than the speeded ones (-0.02 sec.).

Lisker and Abramson (1967) found more prevoicing in stressed syllables, while the word-initial voiced plosives in words with stress on the second syllables had shorter prevoicing. The results of the present study demonstrated that although the effect of stress was in the expected direction, it was not significant probably due to lacking amount of data.

While van Alphen and Smits (2004) did not find any significant effect of lexical status of a word on the production of VOT, the results of the present study detected shorter prevoicing in real words than in nonsense words, which means that nonsense words are hyper-articulated with trustier prevoicing.

Also the last investigated factor–lexical frequency–influenced the production of VOT significantly in the expected direction. The words with high lexical frequency were produced more carefully with longer duration of prevoiving than the low-frequency words.

To sum up, each of the investigated factors influenced the production of prevoicing in the expected direction. However, probably due to lacking amount of data, stress and vocalic environment did not affect the production of negative VOT significantly.

5.2 Research Proposal: Perception Experiment

This chapter focuses on proposing the perception experiment which would investigate whether Czech bilingual listeners–learners of English–can tolerate the absence of prevoicing in English voiced stops.

The pilot production experiment demonstrated that in English, prevoicing is produced less frequently than in Czech. The results also showed that velar plosives, plosives followed by a low vowel and plosives produced by females had the shortest prevoicing, while the negative VOT in labial plosives, plosives followed by a high vowel and plosives produced by males had the longest duration. Therefore, the perception experiment would examine, whether in order to hear a voiced plosive as voiced, the listeners need more prevoicing in labial stops followed by a high vowel produced by a male than in velar plosives followed by a low vowel produced by a female. Furthermore, would shortening or absence of prevoicing in these stops lead to voicing distinction by listeners?

Similar experiment was conducted in a study carried out by van Alphen and Smits (2004), who investigated perception of prevoicing in Dutch initial plosives. They researched several acoustic properties, which are produced by speakers and which may serve as a cue for voicing distinction in perception. The results of their study showed that the strongest cue for distinguishing between voiced and voiceless plosive for listeners is the presence or absence of prevoicing. All of the voiced stops that included prevoicing were perceived as voiced while the plosives without prevoicing were perceived either as voiced or as voiceless.

In order to investigate the possible influence of listeners' L2 on their perception of voiced stops in English, it is necessary to examine to what extent the Czech monolingual listeners tolerate the absence of prevoicing in native language.

The perception experiment consists of two parts. The first part examines the perception of voiced plosives in word-initial position in Czech by monolingual Czech listeners and the second part investigates whether there is any influence of listeners' L2 on their perception in English.

5.2.1 Part 1

This part of the perception experiment aims at researching whether Czech native speakers can tolerate the absence of prevoicing in their native language and if yes, whether the absence of prevoicing is more tolerated in words produced by a male than in words produced by a female.

Participants

This part of the experiment is meant for Czech monolingual speakers. No one with hearing or language impairment can participate.

Materials

Burton et al. (1989) conducted an experiment, which was aimed at perception of nonsense and real words. The stimuli were two continua, each ranged between a nonsense word and a real word. The first one ranged from the real word *duke* to the nonsense word *tuke* and the second continuum ranged from the nonsense word *doot* to the real word *toot*. The results showed that when the listeners were supposed to decide whether they hear voiced or voiceless plosive at the beginning of the word, they were inclined to preferably choose the real word rather than the nonsense word. For this reason, the word list for the perception experiment includes only nonsense words, which were invented by the author of this thesis in a way that none of the words resembles any real word. This part of the research consists only of Czech nonsense words.

Except of the target words beginning with voiced plosives used in this experiment, the word list includes the same amount of fillers starting with voiceless plosives to prevent the listeners from thinking that a token is voiceless only because so far all of them were voiced and surely there must be some voiceless tokens since they are included in the options to choose from. The whole word list is included in appendix.

To present the stimuli, Praat script will be made so that with the onset of each word, seven boxes will appear on the screen, from which the participants would choose only one. They will also have a chance to replay each token one more time. The question would be: *Which sound do you hear at the very* *beginning of the word?* The options to choose from would be: *p, t, k, b, d, g* and *none of these*.

Procedure

The stimuli should be presented over headphones, possibly in a sound-proof room. Together with the heard stimuli, a visual target containing seven different options will be presented on a computer screen. Participants will be instructed to first report the level of their knowledge of English. After that they will be told to listen to the word and then immediately decide which sound they heard at the beginning of the word and click on one of the seven possibilities–*p, t, k, b, d, g* or *none of these*. They will have a chance to replay each token one more time.

The target words should be presented three times each produced by a male and three times each produces by a female, each of the three times with a different amount of prevoicing. The stimuli will be produced in Praat in a way that the negative VOT would be (1) fully truncated, (2) truncated to -7.5ms and (3) truncated to -15ms as shown in table 1, which demonstrates an example of six different stimuli made from the Czech nonsense word *bapwick*. The same conditions will apply for each word used in this experiment, except of the fillers, which will be presented each six times–three times by a male and three times by a female–only without any change in the duration of prevoicing, so that there is the same amount of fillers and target words.

bapwick		
	Produced by a male	Produced by a female
(1)	VOT = 0	VOT = 0
(2)	VOT = -7.5ms	VOT = -7.5ms
(3)	VOT = -15ms	VOT = -15ms

Table 1 An example of different stimuli made from one word. As can be seen, the Czech nonsense word *bapwick* will be heard six times, every time under a different condition: (1) with fully discarded negative VOT, (2) with the duration of VOT -7.5ms and (3) with the duration of VOT -15ms produced by a male and a female.

Predictions

The author of this thesis presumes that since in the pilot production experiment the prevoicing was reliably produced in Czech by both males and females, even in speeded words, there will be no or very little tolerance of the absence of prevoicing.

5.2.2 Part 2

This part of the experiment focuses on examining whether Czech bilingual speakers perceive the voiced plosives the same way as the Czech monolinguals do in their native language or their L2 has an influence on their perception so that they learned to tolerate the absence of prevoicing and perceive the stops with zero or shortened negative VOT as voiced. Moreover, the aim is to investigate to what extent the listeners tolerate the absence of prevoicing. Like in the first part of the perception experiment, all of the target words will have three prevoicing values: 0, -7.5ms and -15ms.

As already mentioned, the pilot perception experiment detected the longest prevoicing in words beginning with the labial plosive, in words where a high vowel followed the initial plosive and in words produced by males. On the contrary, the shortest prevoicing was in words beginning with the velar plosive, in words where the plosive was followed by a low vowel and in words produced by females. The question rising from this finding is whether the portion of negative VOT needed for the plosive to be perceived as voiced is smaller in the words, which were produced with the shortest negative VOT (starting with the velar plosive /g/; followed by a low vowel; produced by a female) and bigger in words, which were produced with the longest negative VOT (beginning with the labial plosive /b/; followed by a high vowel; produced by a male).

Participants

This second part of the perception experiment is intended for Czech learners of English with no hearing or language impairment.

Materials and Procedure

The materials and procedure are identical with the first part of the perception experiment, only English stimuli will be added together with the same amount of fillers. The whole list of words used in the experiment is included in appendix.

Predictions

The author of this thesis assumes that an influence of L2 on L1 of the listeners will be found and that the better knowledge of English, the more tolerated absence of prevoicing. Furthermore, it is expected that bigger portion of missing prevoicing will be tolerated in words beginning with /g/ followed by a low vowel produced by a female, while the absence of prevoicing will be tolerated only minimally in words beginning with /b/ followed by a high vowel produced by a male.

7 Appendix

7.1 Part 1

Part 1 includes Czech and English real words and fillers used in the pilot production experiment.

		b	d	g
High Frequency	High Vowel	budeš bílý budova bůh buňka	důvod duše duben dům důkaz	
	Low Vowel	babička bavit banka barva barevný	další dávno dálnice dávat daně	galaxie garáž galerie
Low Frequency	High Vowel	bizon bilion budík bitky buchty	dýka dukát duha duby duny	gymnasta guma guláš gigant gusto
	Low Vowel	bagr báje balón bacil balady	dary davy datel dárce darebák	Gabriel garance galanterie gauč gama

Table 2 represents Czech real words used in the pilot production experiment.

altán anděl celek čekání cizinec	rádce ročenka lůžko lupič mezera	
hadice lágr láhev obal okno	můstek nabídka neděle nečas rodáci	meč medvěd labuť
postel polštář ponožka medovina stůl	noviny chytit jestli nejistý plamen	svíčková cena patnáct plechovka inzerát
noha ruka hlava vlasy palivo	ocel cíle ihned závislost muset	vytáhnout míle auto někde naprosto

Table 3 represents Czech fillers used in the pilot production experiment.

Stress on the **First** Syllable	High Frequency	High Vowel	book	difference	guilty
			beer	deep	goods
			beach	detailed	good
			beat	dealer	gift
			business	dear	give
		Low Vowel	bank	dance	guard
			bath	dark	gun
			battle	double	garden
			bottom	document	gather
			band	doctor	gas
	Low Frequency	High Vowel	beacon	demon	goose
			boomer	doom	geese
			boozer	digitize	goggles
			boomerang	dignify	gild
			beaver	dim	gibbon
		Low Vowel	body	dove	gulp
			balm	dangle	gulf
			baron	docusoap	gutter
			bondage	dodge	garbage
			barge	domino	gangster
Stress on the **Second** Syllable	High Frequency		because	decision	guitar 8059
			beginning	decline	goodbye 5528
			become	debate	guerrilla 224
			believe	deny	goodwill
			behind	defend	
	Low Frequency		bewitch	deactivate	gardenia 202
			bewilder	deceitful	galactic 1373
			bewail	decelerate	gazebo 454
			bilabial	decode	gazelle
			bazooka	decorum	gazette

Table 4 represents English real words used in the pilot production experiment.

always	chain	channel
allomorph	author	cherry
animal	asleep	conflict
aspect	anyway	control
chicken	anonym	corn
centre	answer	empty
cinema	zombie	engine
circle	yellow	excited
episode	wizard	fellow
epigram	winning	fish
expensive	violent	flash
ecstasy	uncovered	hack
escalator	unappealing	happy
figure	ultimate	maze
flower	sympathy	idle
formal	sunscreen	impact
frustrate	strategy	job
freedom	spellbound	knock
homework	sneaker	length
horrible	slender	link
habituated	seventy	logic
handicapped	second	marry
hangover	reverse	medal
heating	relative	milk
honesty	rebound	
justice	occasion	percent
learning	observe	ready
labial	outflow	recover
labour	obedient	region
lateral	minibar	remove

Table 5 represents English fillers used in the pilot production experiment.

7.2 Part 2

This part of appendix includes Czech and English nonsense words used in the pilot production experiment. The same stimuli would be used for the perception experiment. Besides the target words, fillers used in the production experiment are included in this part of appendix.

	b	d	g
High Vowel	buliš bůspir buvéř bikěř bikýř	duněp durpůř dyřůn dypěř dyvěl	gimýř gislůj gultýř gumůř gulim
Low Vowel	bašek bátel baptán bantář balův	damiž damíř dašen dašin dalán	gadál gamón gábol gašin gamál

Table 6 represents Czech nonsense words used in the pilot production experiment. These words would be also used for the perception experiment.

	b	d	g
High Vowel	bipwick boonidge boonesc boonpridge birinch	doopfidge dinwick doosridge disraph doowidge	guimick geaswick gooridge goomwick goonidge
Low Vowel	baswick bapridge bapwick bascridge bapwidge	dawidge dasridge datwick dariph daringe	gackorph gongridge gabanph galwick galridge

Table 7 represents English nonsense words used in the pilot production experiment. These words would be also used for the perception experiment.

muvál sinýř lápuř amýř řinák	chvápř sáfař pimelš fenižář mabél	jémeř máviša šapór pášik šápeř
pajtel natmék pévin hvádě kápteř	němpář vádeř kémuř léhif nimáp	unéfěr héšmin láběř náfoř slékiř

Table 8 represents Czech nonsense fillers used in the pilot production experiment.

miscoon oskwidge lanwick onwidge fillscott	niscoon uskwidge ranwick unwidge villscott	loomidge yoonidge rooscidge wiscidge mepridge
rabidge fludge fedge samidge falidge	labidge ludge medge hamidge validge	nepridge lorminck nedge fidge seanidge

Table 9 represents English nonsense fillers used in the pilot production experiment.

7.3 Part 3

This part of appendix includes Czech and English fillers that would be used in the perception experiment.

pontář	tamíř	kornál
poliša	tošina	kábola
pášek	tokůl	kadál
paltán	tožina	kolán
plůkiř	tokůn	kamóna
pedéř	tripůř	kornál
puskán	tlápeř	kamíř
pámeř	telůn	kempůř
piřkán	tůvík	kléskin
plékiř	třemán	kůpeř

Table 10 represents Czech fillers that would be used in the perception experiment.

pokridge	torbidge	congridge
postook	taskwick	candroop
porwick	torfidge	calidge
pantwick	tandridge	camidge
paswick	tabanph	cormnick
ponoosck	toslamm	coomidge
proockidge	tapoos	coonwidge
pansrick	torgoon	crootisc
ploocker	tarood	creringe
pemoor	timoon	croophisck

Table 11 represents English fillers that would be used in the perception experiment.

7.4 Part 4 - DVD

The DVD contains the recordings produced in the pilot production experiment together with the Praat scripts. Part of these recordings would be used for the perception experiment.

9 Works Cited

Burton, Martha W., Shari R. Baum and Sheila E. Blumstein. 1989. "Lexical Effects on the Phonetic Categorization of Speech: The Role of Acoustic Structure." *Journal of Experimental Psychology: Human Perception and Performance* vol. 15, no. 3:567-575.

Cho, Taehong and Peter Ladefoged. 1999. "Variation and Universals in VOT: Evidence from 18 Languages." *Journal of Phonetics* 27:207-229.

Hayward, Catherine. 2000. *Experimental Phonetics.* London: Longman.

Helgason, Pétur and Catherine Ringen. 2008. "Voicing and Aspiration in Swedish stops." *Journal of Phonetics* 36:607-628.

Kessinger, Rachel H. and Sheila E. Blumstein. 1997. "Effects of Speaking Rate on Voice-Onset Time in Thai, French, and English." *Journal of Phonetics* 25:143-168.

Kim, Mi-Ryoung. 2011. "The Relationship between Cross Language Phonetic Influences and L2 Proficiency in terms of VOT." *말소리와 음성과학 제3권 제3호*, 3-10.

Kingston, J. and Diehl, R. L. 1994. "Phonetic Knowledge." *Language* 70:419-454.

Klatt, Dennis. 1975. "Voice Onset Time, Frication, and Aspiration in Word-Initial Consonants Clusters." *J Speech Hear Res* 18:686-706, accessed October 4, 2013, http://jslhr.asha.org/cgi/content/abstract/18/4/686.

Ladefoged, Peter and Keith Johnson. 2011. *A Course in Phonetics.* 6th ed. Boston: Wadsworth, Cengage Learning.

Lam, Tuan Q. and Duane G. Watson. 2014. "Repetition Reduction: Lexical Repetition in the Absence of Referent Repetition." *J Exp Psychol Learn Mem Cogn.* 40(3):829-843.

Libermann, A. M., P. C. Delattre and F. S. Cooper. 1958. "Some Cues for the Distinction between Voiced and Voiceless Stops in Initial Position." *Language and Speech* 3:153-167.

Lisker, Leigh, and Arthur S. Abramson. 1964. "A Cross-language Study of Voicing in Initial Stops: Acoustical Measurements." *Word* 20:384-422, accessed May 9, 2015, http://www.researchgate.net/profile/Arthur_Abramson/publication/208032881_A_Cross-Language_Study_of_Voicing_in_Initial_Stops_Acoustical_Measurements/links/0912f50d3b9032b04f000000.pdf.

Lisker, Leigh, and Arthur S. Abramson. 1967. "Some Effects of Context on Voice Onset Time in English Stops." *Language and Speech* 1:1-28.

Machač, Pavel and Radek Skarnizl. 2009. *Fonetická segmentace hlásek*. Praha: Epocha.

Morris, Richard, Christopher R. McCrea, and Kaileen D. Herring. 2008. "Voice Onset Time Differences between Adult Males and Females: Isolated Syllables." *Journal of Phonetics* 36:308-317.

Oh, Eunjin. 2011. "Effects of Speaker Gender on Voice Onset Time in Korean Stops." *Journal of Phonetics* 39:59-67.

Palková, Zdena. 1994. *Fonetika a fonologie češtiny.* Praha: Karolinum.

Repp, Bruno H. 1979. "Relative Amplitude of Aspiration Noise as a Voicing Cue for Syllable-initial Stop Consonants." *Language and Speech* 22, no. 2:173-189.

Ryalls, John, Allison Zipprer, and Penelope Baldauff. 1997. "A Preliminary Investigation of the Effects of Gender and Race on Voice Onset Time." *Journal of Speech, Language and Hearing Research* 40:642-645.

Skarnitzl, Radek. 2001. *Znělostní kontrast nejen v Češtině.* Praha: Epocha.

Slis, I. H. and Cohen, A. 1969. "On Complex Regulating Voiced-Voiceless Distinction. I." *Language and Speech* 12:80-102.

Umeda, N. 1981. "Influence of Segmental Factors n Fundamental-frequency in Fluent Speech." *Journal of the Acoustical Society of America* 70:350-355.

van Alphen, Petra M. and Roel Smits. 2004. "Acoustical and Perceptual Analysis of the Voicing Distinction in Dutch Initial Plosives: the Role of Prevoicing." *Journal of Phonetics* 32:455-491.

van Alphen, Petra. 2004. "Perceptual Relevance of Prevoicing in Dutch." Thesis, Radboud University Nijmegen, Nijmegen.

Whiteside, Sandra P., Luisa Henry and Rachel Dobbin. 2004. "Sex Differences in Voice Onset Time: A Developmental Study of Phonetic Context Effects in British English." *Acoustical Society of America* Vol. 116, No. 2:1179-1183.

10 Anotace

Autor:	Erika Kopčanová
Katedra:	Katedra anglistiky a amerikanistiky FF UPOL
Studijní obor:	Anglická filologie - Německá filologie
Název Česky:	Produkce a percepce znělosti plozív během závěru u rodilých a nerodilých mluvčích angličtiny
Název Anglicky:	The Production and Perception of Prevoicing in Voiced Stops in Native and Non-native English
Vedoucí práce:	Mgr. Václav Jonáš Podlipský, Ph.D.
Počet stran:	55
Počet příloh:	3+DVD
Klíčová slova v ČJ:	znělost, nástup hlasivkového tónu, znělé explozívy, produkce a percepce znělosti
Klíčová slova v AJ:	prevoicing, VOT, word-initial plosive, voiced plosive, production and perception of prevoicing

Anotace v ČJ: Hlavním cílem této práce je zjistit, jestli je nějaký rozdíl v percepci a produkci znělosti v anglických a českých explozívach. Práce je rozdělená na dvě části. V první části jsou shrnuty teoretické poznatky včetně popisu signálů, které naznačují znělost a také faktory, které můžou ovlivnit nástup hlasivkového tónu (VOT). V další části je podrobně popsán produkční experiment, který zkoumá vliv určitých faktorů na znělost. Všechny z těchto faktorů ovlivnily produkci negativního VOT v očekávaným směru, čím se potvrdily hypotézy. Nahrávky získany během produkčního experimentu budou sloužit k provedení percepčního experimentu, který je navržen v třetí části téhle studie a který má za úkol zjistit, jestli se český bilingvní mluvčí naučil tolerovat absenci negativního VOT v angličtině.

Anotace v AJ: The purpose of the study is to examine, whether there is some difference in production and perception of prevoicing in Czech and English voiced stops. This paper is divided primarily into two parts. The first part concerns the theoretical background, provides descriptions of the known cues to voicing and describes factors that may influence the production of prevoicing. The next part moves on to describing the production experiment, which investigates the influence of several factors on VOT. All of the examined factors influenced prevoicing in the expected direction, which confirms the hypotheses. The recordings gained from the production experiment will serve for conducting a perception experiment, which will investigate whether a Czech bilingual spaker have learned to tolerate the absence of prevoicing in English.