Univerzita Palackého v Olomouci

Filozofická fakulta

Katedra anglistiky a amerikanistiky

Training Czech Learners of English to Perceive and Produce English Word-final Voicing

Magisterská diplomová práce

Bc. Eliška Herudková

Anglická filologie – Česká filologie

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

Olomouc 2015

Training Czech Learners of English to Perceive and Produce English Word-final Voicing

(Magisterská diplomová práce)

Autor: Bc. Eliška Herudková

Studijní obor: Anglická filologie – Česká filologie

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

Počet stran (podle čísel): 122

Počet znaků: 122375 (pouze text práce bez seznamu literatury a bez příloh)

Olomouc 2015

Prohlašuji, že jsem tuto diplomovou práci vypracovala samostatně a uvedla úplný seznam citované a použité literatury.

V Olomouci dne …………………. Podpis……………………..

I would like to thank Mgr. Václav Jonáš Podlipský, Ph.D. for giving me help, advice and guidance, without which I would not be able to complete the thesis. I would also like to thank Mgr. Šárka Šimáčková, Ph.D. for letting me work with the students of her courses. Finally, I would like to thank all the participants of my experiment for dedicating their time to help me.

Table of Contents

1. Introduction

In Czech, the phonological contrast between voiced and voiceless obstruents is neutralized in word-final positions. In English, on the other hand, the process of neutralization does not take place word-finally and the phonological contrast between voiced and voiceless obstruents is preserved in this position.

The phonetic implementation of phonological obstruent voicing in English is cued by several acoustic features. The relevance of the acoustic cues to phonological obstruent voicing varies, depending on the position the obstruent occupies in the word. The relevant cues to word-final obstruent voicing are present in the consonant itself (voicing of the closure duration, final obstruent duration) as well as in the preceding vowel (vowel duration) (Lisker 1986; Dmitrieva, Jongman, and Sereno 2010). It is believed to be a phonetic universal that vowels are longer before voiced obstruents and shorter when preceding voiceless obstruents (Chen 1970, 157). It has been suggested that in English, vowel duration outweighs the cues found in the consonant itself (Raphael 1971, 1031; Kluender, Diehl, and Wright 1987, 153; Luce and Charles-Luce 1985, 1949).

Previous research concerning other languages that apply the same phonological rule as Czech and therefore neutralize word-final obstruent voicing (e.g. Broersma 2005, 2008, 2010; Dmitrieva, Jongman, and Sereno 2010) has found out that native speakers of such languages rely on vowel duration less than native English speakers. The universal phonetic rule which states that vowel duration differs, depending on the voicing of the following obstruent, still applies to these languages. The duration of vowels varies to a much smaller degree than it does in English where this phonetic effect has phonologized to become the basis for the voicing contrast. For example in Dutch, which neutralizes the voicing contrast in word-final positions, vowels preceding a voiced intervocalic fricative are believed to be on average 40 ms longer than those preceding a voiceless intervocalic fricative (Broersma 2010, 1637). In English, however, the difference is more than three times bigger (Broersma 2010, 1637).

Several recent studies have answered the question whether there is a tendency to use different vowel duration before voiced and voiceless obstruents in Czech (Šimáčková 2003; Podlipský and Chládková 2007; Podlipský 2008). It was found out that vowel duration varies, depending on phonetic environment. The difference is, however, not as significant as it is in English.

Apart from that, Sehnalíková (2010) explored Czech native speakers' word-final voicing perception. The results suggest that neutralization is complete in Czech and that it is not likely that Czech native speakers rely on vowel duration in their lack of perception of underlying word-final obstruent voicing. The usage of vowel duration as a cue to word-final obstruent voicing, was, however, not a subject of the thesis.

The present thesis will be concerned with training Czech learners of English to be able to perceive and produce the word-final voicing contrast more efficiently. It will draw on previous research and focus on the development of the ability to use vowel duration as the main cue to word-final obstruent voicing as is typical of native English perception. In order to achieve the intended goal, high variability phonetic training using the identification task has been chosen as it has been adopted as a successful method in training the acquisition of non-native contrasting categories (e.g. Jamieson and Morosan 1986; Logan, Lively, and Pisoni 1991; Lively et al. 1994) and reliance on the correct acoustic cues (e.g. Broersma 2008; Iverson, Hazan, and Bannister 2005; Iverson and Evans 2009; Francis, Baldwin, and Nusbaum 2000).

The thesis will be divided into five chapters. The first chapter, Introduction, will open with information about the implementation of phonological voicing of obstruents in English and in Czech. Generally, it is believed that in word-final positions, phonological voicing of obstruents is neutralized in Czech, while in English, it is preserved. On the other hand, it has been suggested that neutralization of phonological voicing in word-final positions may not be complete in some languages, including Czech. Section 1.2. will provide information about results of research concerned with incomplete neutralization. In 1.3, the range of acoustic cues to word-final voicing in English will be introduced and more information about the use of vowel duration as a cue to phonological voicing of obstruents in various languages will be given.

Training methods, used for an improvement of non-native perception, will be discussed in 1.4. The potential effect the perceptual training may have on non-native production will be discussed in 1.5. The Introduction will be closed with research questions and hypotheses, formulated in 1.6.

The second chapter, Methodology, will give information about the present experiment. First, information about the participants will be provided in 2.1. The stimuli will be introduced in 2.2. The last section, 2.3. Methodology, will be devoted to the procedure employed.

The results of the experiment will be discussed in the third chapter, Results. The chapter will be divided into two sections – 3.1. Percentages of Correct Responses, and 3.2. Response Times. The Results will be followed by a Discussion, the fourth chapter, where it will be discussed, whether the hypotheses were confirmed or rejected. The thesis will end with a Conclusion, which will summarize the results of the present experiment.

1.1. Word-Final Voicing

Both English and Czech have phonological sets of voiced – voiceless stops, fricative and affricate sounds (obstruents). The implementation of the voicing contrast is, however, different in both languages.

In English, the phonologically voiced obstruents are often phonetically voiced only in certain positions. Full phonetic voicing occurs in positions between two voiced sounds (e.g. in *labour, eager, rub out*). Partial devoicing may, however, appear in intervocalic positions (Cruttenden 2001, 152) as maintaining voicing during closure cannot be achieved for long. In initial positions, voicing is marked and phonologically voiced stops and affricates become devoiced (e.g. in *bill, done, jazz*). Phonologically voiceless stops become aspirated in initial positions in order to maintain the voiced – voiceless contrast (e.g. in *pip, tip, kit*), except for when they are preceded by *s* (e.g. in *stew, sky, spoon*) (Ladefoged 2001, 57). In word-final positions, phonetic voicing is realized. The phonetically voiced obstruents have, however, only little voicing word-finally (Roach 1998, 33).

Phonologically voiceless obstruents remain phonetically voiceless in most positions in English. The exceptions include the realization of the glottal fricative /h/, which stays phonetically voiceless in word-initial positions but is voiced between voiced sounds (Cruttenden 2001, 178), and tapping of /t/ between two vowels, the second one of which is unstressed (Ladefoged 2001, 58).

In Czech, phonetic realization of phonological voicing is different. It has been suggested that unlike in English, devoicing of phonemically voiced obstruents is rare and does not appear systematically (Skarnitzl and Šturm 2014, 200). Furthermore, the voicing contrast is neutralized in certain positions. Word-finally, before a pause, all obstruents are pronounced as phonetically voiceless (e.g. in *led, mez, hod*) and no cues indicating devoicing are preserved, suggesting that devoicing takes place at a deeper level. When an obstruent is surrounded by other consonants, voicing assimilation takes place in Czech. Regressive assimilation is the most common type of assimilation and it affects phonemes in morpheme-final and word-final positions. Preceding phonemes therefore become either voiced or voiceless, depending on the voicing of the following sound or group of sounds. In consonant clusters, voicing of the whole group of obstruents is affected by voicing of the last obstruent (Skarnitzl 2011, 123). In English, the process of voicing assimilation is not as common as it is in Czech. It only affects voiceless fricatives in the production of grammatical words by some speakers (Cruttenden 2008, 299).

Czech and English, therefore, apply different rules in the implementation of the phonological voicing in the sets of voiced and voiceless obstruents. In addition, native speakers of both languages choose different strategies when using regressive assimilation of voicing. It is, therefore, believed that in the production of English by Czech learners of English, negative transfer takes place and becomes perceptible to native listeners (Skarnitzl and Šturm 2014, 201). At the same time, as the voicing contrast between obstruents is neutralized word-finally in Czech, negative transfer should take place also in isolated words.

It has been found out that native speakers of languages which do not realize the voicing opposition in word-final positions do not use the correct cues to word-final voicing in English (e.g. Luce and Charles-Luce 1985; Flege 1989; Crowther and Mann 1992; Crowther and Mann 1994; Broersma 2005, 2008, 2010). The goal of the present thesis is to train Czech learners of English to use the right cues to English word-final voicing. The nature of the cues will be explained in chapter 1.3. Cues to Word-Final Voicing.

1.2.Incomplete Neutralization

Various studies focusing on other languages than English (e.g. Polish, German, Catalan, Dutch) addressed the question whether neutralization is complete in these languages or not and if it does not take place, what cues to the voicing distinction do native speakers of such languages preserve (e.g. Slowiaczek and Dinnsen 1985, Slowiaczek and Szymanska 1989, Port and O’Dell 1985, Charles-Luce 1985, Warner et al. 2004, Dinnsen and Charles-Luce 1984). The present chapter reviews findings of these studies concerning the issue of the potentially incomplete neutralization.

1.2.1. Incomplete Neutralization

Polish is one of the languages that (like Czech) neutralize the voicing contrast between obstruents in word-final positions (Ostaszewska and Tambor 2000, 108). In 1985, Slowiaczek and Dinnsen published a paper studying the process of neutralization in Polish. Five Polish native speakers participated in their experiment. They were asked to read a carrier sentence containing words that differed in the final obstruent. The results that were obtained suggest that a slight difference in the production of words containing underlyingly voiced and voiceless[1] final obstruents is made – the mean duration of vowels preceding the underlyingly voiced obstruents was longer than the duration of vowels before underlyingly voiceless obstruents. Apart from that, for three of the subjects, voicing into closure/frication was longer for underlyingly voiced obstruents than for the underlyingly voiceless obstruents. Based on the results, the authors concluded that devoicing in Polish is non-neutralizing.

The experiment by Slowiaczek and Dinnsen (1985) was extended by Slowiaczek and Szymanska (1989), who focused on perception of word-final voicing. In their experiment, they involved both Polish and English native speakers and presented them with pairs of Polish monosyllabic words, differing in the underlying voicing of the final obstruent. According to their results, the identification of the underlying voicing was poor and the listeners in both groups tended towards the voiceless alternative. It was concluded that the neutralization in perception must be complete.

German is also considered one of the languages that neutralize the voicing contrast in word-final positions (Russ 2010, 141). Port and O'Dell (1985) analysed both production and perception of word-final obstruent voicing by German native speakers. The findings of their research suggest that German native speakers preserve underlying distinctions between words containing final voiced and voiceless obstruents in both

[1] As underlyingly voiced, words, whose orthographic representation contained a symbol for a final voiced consonant where a voiceless sound is pronounced, were presented (e.g. *karb* [karp]). The underlyingly voiceless obstruents where represented as such in writing (e.g. *karp* [karp]).

production and perception. In the subjects' production, the effect of voicing into closure, burst duration and vowel duration was reported. In the test of perception, high accuracy in the identification of the correct answer was achieved. The authors, therefore, concluded that neutralization in German is not complete.

In another paper on German, Charles-Luce (1985) took into consideration the place of articulation of the final stop, phonetic environment of the stimuli and other factors. Also according to his conclusions, neutralization in German is incomplete. The effects of the underlying voicing were, however, dependent on the place of the articulation of the final stop (dental/velar), on the lexical item itself, on the phonetic environment (vowel duration was used as a cue to underlyingly voiced and voiceless fricatives, while closure duration served as a cue to underlyingly voiced and voiceless stops) and on the position in the sentence that the word occurs in.

Another language that is said to neutralize the voicing contrast in final obstruents is Catalan (Hualde 1992, 393). Dinnsen and Charles-Luce (1984) carried out an experiment focused on the production of the word-final voicing contrast and in Catalan. Unlike in Polish and in German, in Catalan the word-final underlying voicing contrast is not preserved in orthography. Dinnsen and Charles-Luce (1984), therefore, suggested that the production of speakers of such languages that represent the voicing contrast orthographically, would be guided by the orthographical representation. The authors suggested that if the results of their experiment indicated that neutralization is incomplete in Catalan (a language with no orthographical representation of the underlying voicing contrast in word-final positions), it would truly challenge the rule of neutralization.

Five speakers of Catalan participated in their experiment. Their task was to read carrier sentences, containing Catalan monosyllabic words, differing in the underlying voicing of the final obstruent. The duration of a vowel preceding the final obstruent, closure duration and voicing into closure were measured. No effects of underlying voicing were found for the three factors for the subjects as a group. Individually, two of the participants preserved the voicing contrast, though. Each of them, however, used a different cue to the voicing contrast. The results, therefore, provided only a little evidence of the presence of incomplete neutralization in Catalan.

The presence of incomplete neutralization was also studied for Dutch (Warner et al. 2004). In Dutch, similarly as in the above mentioned languages, the phonological voicing contrast is supposedly neutralized in word-final positions. In 2004, Warner et al. published the findings that they obtained from a series of experiments on word-final obstruent devoicing in Dutch. In their research, they focused on various factors and studied perception of natural speech as well as of synthetized speech, in which either vowel duration or closure duration was manipulated. They also focused on the role of orthography. Their results suggest that: “incomplete neutralization effects are small, variable and task dependent.” (Warner et al. 2004, 273).

1.2.2. Incomplete Neutralization in Czech

It is believed that in Czech, neutralization of the voicing contrast between obstruents in word-final positions takes place – minimal pairs of words such as *plot* and *plod* are, therefore, pronounced identically (in this case, both words are pronounced as [plot]) (Palková 1994, 132). In morphologically related word forms, however, the voicing contrast is realized (as in *ploty* [ploti] and *plody* [plodi]). It has been suggested that the underlying voicing contrast in word-final positions may not be neutralized completely and that there still might be acoustic cues to the underlying voicing in Czech (Kuzla and Duběda 2008; Sehnalíková 2010).

Kuzla and Duběda (2008) carried out an experiment in order to find out whether neutralization of the word-final obstruent voicing contrast is incomplete in Czech. Twenty Czech native speakers participated in the experiment. The participants were asked to read sentences, which included the words *plod* and *plot* in diverse segmental and prosodic contexts. In the analysis of the recordings, the closure duration of the final obstruent was measured, as well as the duration of the preceding vowel, the duration of aspiration and the amount of voicing during closure.

According to the results, the distinction between the underlyingly voiced – voiceless word-final obstruents was kept. More voicing during closure was observed for /d/ and a longer closure duration was noted for /t/. Vowel duration and duration of aspiration turned out as irrelevant.

Kuzla and Duběda (2008), therefore, concluded that neutralization of the word-final voicing contrast is incomplete in Czech. However, their results were only preliminary as they only managed to analyse data for three subjects. At the same time, they only included one minimal pair of obstruents in their experiment (/t/ – /d/). It has been suggested in other studies (e.g. Charles-Luce 1985) that production is dependent on, for example, the place of articulation of the final obstruent or on the lexical item itself.

Sehnalíková (2010) also tested, whether the word-final voicing contrast is neutralized completely in Czech. In her experiment, 45 native Czech speakers took part. Five of them were recorded, while the other 40 were tested in perception. The participants of the perceptual tests were divided into two groups according to their knowledge of English – one of the groups had no experience with English, while the other one consisted of learners of English. The group of Czech native speakers with no experience with English listened to stimuli composed of Czech word pairs as well as fillers with enhanced vowel duration. The Czech learners of English listened to English stimuli (real words and fillers). The subjects of the perceptual test were given a 2-alternative forced-choice task. The group that was recorded produced Czech stimuli for the perceptual test.

According to the analysis of the results, none of the participants that were recorded produced a significant difference between words ending in voiced – voiceless obstruent. In the test of perception, Czech speakers with no knowledge of English generally did not perform above chance in categorizing voicing in coda. In general, they were more accurate in the identification of voiceless word-final obstruent than in the identification of voiced word-final obstruent. The group of Czech learners of English performed with a higher accuracy in the perceptual test. The responses were, however, also more accurate for words ending with a voiceless obstruent than with a voiced obstruent. As the listeners were unable to distinguish the word-final voicing contrast in their L1, it was concluded that neutralization is complete in Czech. It was also concluded that ability to distinguish word-final voicing contrast must be acquired in the process of L2 learning.

Neutralization of the voicing contrast may be incomplete in some languages. The speakers of these languages might differentiate between words that end in an underlyingly voiced and voiceless obstruent by using various acoustic cues to the voicing contrast. The cues to word-final voicing in English are described in chapter 1.3. Further, the chapter presents findings of various studies, which show that although neutralization might be incomplete in other languages, the cues to word-final voicing contrast might be used in a different extent by native speakers of languages other than English than by English native speakers.

1.3. Cues to Word-Final Voicing

In English, there is a number of acoustic cues leading to the recognition of the phonological voicing in obstruents. The nature of the cues depends on the phonetic context in which the obstruent is situated. There are different acoustic cues to phonological voicing in syllable-initial positions, in intervocalic and syllable-final positions. In the initial positions, obstruents are usually devoiced and aspiration, duration and a starting frequency of the first formant transition play crucial roles in differentiating them (Raphael 2005, 189–190).

In syllable-final positions, obstruents are not aspirated and aspiration, therefore, cannot be used as a reliable acoustic cue to phonological voicing in these positions. A number of cues to syllable-final voicing have been recognized. There are cues to final stop voicing as well as cues to fricative and affricate voicing. Such cues are defined with respect to the nature of the consonant.

As for stops, the nature of a burst/release (its duration, intensity, spectral properties) serves as a differentiating factor, as well as the duration of the closure and the character of the formant transitions (Raphael 2005, 192; Wang 1959, 66). The distinction between voiced and voiceless fricatives can be indicated by the duration of a fricative noise (Raphael 2005, 195).

As a universal cue to the difference between voice and voiceless obstruents – stops as well as fricatives and affricates – duration of a vowel, which precedes the final obstruent, is considered. It has been found out that English vowels are significantly longer before voiced obstruents and shorter before voiceless obstruents (Chen 1970, 129). It has been observed that vowel duration differs according to the nature of the following obstruent in other languages than in English, too (e.g. Broersma 2010, Dmitrieva, Jongman, and Sereno 2010, Crowther and Mann 1992, 1994; Podlipský and Chládková 2007).

A number of studies have attempted to challenge the notion of vowel duration being the primary cue to word-final voicing (e.g. Hillenbrand et al. 1984; Luce and Charles-Luce 1985; Wang and Wu 2001). Luce and Charles-Luce (1985) tried to examine both vowel duration and closure duration as acoustic cues to word-final obstruent voicing in a systematic way. In their experiment, they compared the two cues in a number of contexts and came to the conclusion that the duration of a closure had failed to distinguish voice categories in more than 50% of instances. Vowel duration, however, served as a reliable cue in all instances. Vowel duration is considered a reliable cue to word-final obstruent voicing and its nature will be described in the next chapter.

1.3.1. Vowel Duration as a Cue to Word-Final Voicing

1.3.1.1. Vowel Duration as a Cue to Word-Final Voicing in English

In an early paper by House and Fairbanks (1953), the influence of consonant environment on the duration of vowels in American English was studied. In their experiment, vowel duration was measured in monosyllabic words, depending on obstruent voicing, place of articulation and manner of articulation of the following consonant. Phonological voicing of obstruents was identified as the most important factor influencing the length of the preceding vowel. The mean duration of an English vowel which precedes a voiceless consonant is according to their data 174 ms. Vowels followed by a voiced obstruent are approximately 253 ms long. The mean difference in vowel duration was reported to be 79 ms.

The phenomenon was investigated further by Peterson and Lehiste (1960) who studied the influence of consonants both preceding and following a vowel on vowel duration. They came to the conclusion that in English, vowel duration is affected by the nature of the following consonant and that the syllable-initial consonant does not have any impact on the duration of the vowel. They also concluded that in English, vowels are the longest before voiced fricatives, then before voiced stops, nasals, voiceless fricatives and voiceless stops.

In addition to that, they found out that the same vowels are longer when followed by a voiced obstruent than by a voiceless obstruent in the ratio of 3:2. The mean duration of vowels preceding voiced obstruents was 197 ms. Vowels preceding voiceless obstruents lasted 297 ms on average. The mean difference in vowel duration was, therefore, 100 ms.

The two early papers presented above were both concerned with vowel duration affected by obstruent voicing with respect to production. In 1971, Lawrence J. Raphael published a paper focused on perception of the phenomenon. In the experiment described in his paper, a number of recorded and synthetized stimuli were presented to a group of American English native speakers. The tokens were changed so that the final obstruent would always be voiceless. Vowel length was preserved so that it would vary in words with previously voiced obstruents and in words with originally voiceless obstruents.

According to the results, tokens with longer vowels were perceived as voiced and those with shorter vowels were considered voiceless. Thus, the fact that vowel duration changes according to voicing of a following consonant in English was confirmed through the analysis of both perception and production.

1.3.1.2. Vowel Duration as a Cue to Word-Final Obstruent Voicing in Other Languages than English

The results of the papers by House and Fairbanks (1953), Peterson and Lehiste (1960) and Raphael (1971) were, however, applicable only to English and a question, whether the same principles operate also in other languages, arose. In his paper, Matthew Chen (1970) presented results from his observations of the use of vowel duration as a cue to word-final voicing in English, Korean, French and Russian. His goal was to find out whether vowels become longer before voiced consonants and shorter before voiceless consonants in other languages than in English and whether this principle can be regarded as a language universal or if it is only language specific to English.

According to the results of his research, all languages show a similar tendency, suggesting that vowel duration differs with respect to the phonological voicing of the obstruent following it in all four languages. The differences in vowel duration were, however, smaller in the other three languages than they were in English. While in English, vowels preceding a voiced obstruent were on average 92 ms longer than vowels followed by a voiceless obstruent, in French the mean difference was only 53 ms, in Russian it was 29 ms and in Korean 28 ms (Chen 1970, 138). It was, therefore concluded that: "(a) it is presumably a language-universal phenomenon that vowel duration varies as a function of the voicing of the following consonant, and (b) the extent, however, to which an adjacent voiced or voiceless consonant affects its preceding vowel durationwise is determined by the language-specific phonological structure." (Chen 1970, 139)

Since the publication of Chen's (1970) paper, scholars have been interested in the usage of vowel duration as an acoustic cue to the final phonological obstruent voicing in other languages. Slowiaczek and Dinnsen (1985) published a paper on neutralization of the voicing contrast in word-final positions in Polish (for more information about the experiment see section 1.2.1. Incomplete Neutralization). The analysis of the results of their experiment revealed a significant main effect of underlying voicing for 4 of their 5 subjects. The length of Polish vowels preceding underlyingly voiced obstruents was about 10% bigger than the length of those that were followed by voiceless obstruents (the mean duration of vowels preceding voiced obstruents varied from 114 ms to 155 ms; the mean duration of vowels followed by voiceless obstruents was between 102 ms and 137 ms). Although the goal of their paper was different, the results corresponded with the suggestion in Chen's (1970) paper that differentiations in vowel length are present language-universally.

At the same time, the mean durations of vowels preceding voiced obstruents produced by native Polish speakers were much shorter than those of native English subjects of the experiments by House and Fairbanks (1953), Peterson and Lehiste (1960) and Chen (1970). The mean duration of vowels followed by voiced obstruents were not longer than 155 ms in Polish, while in English they were approximately 250 ms long (see the results above).

In the 1985, Patricia Keating published an article opposing Chen's theory. She investigated vowel duration as a cue to syllable-final voicing in Polish and in Czech. In her study, she focused on production and recorded 24 native speakers of Polish (the results for Czech will be discussed separately in section 1.3.1.3 Vowel Duration as a Cue to Obstruent Voicing in Czech) reading a pair of words *rata* and *rada*. The words differed in the voicing of the medial obstruent as in Polish, the voicing contrast is preserved in this position. She came to the conclusion that: "Polish vowel duration does not vary systematically according to the voicing of the following consonant." (Keating 1985, 121) The mean difference in duration was 0.99.

As for other Slavic languages, the phenomenon has been studied for Russian, Serbian and Czech. In Russian and in Czech, the voicing contrast is neutralized in word-final positions (Cubberley 2002, 73; Palková 1994, 133). In Serbian, neutralization of the voicing contrast in word-final positions does not take place (Sussex and Cubberley 2006, 216). The effect of vowel duration in Russian was first studied by Chen (1970) in the paper introduced above. Later, Pye (1986) found out that vowel duration before voiced obstruents was 5-20 ms longer than that before voiceless obstruents in Russian. According to the results, vowel duration differed with respect to the phonetic properties of the obstruent (the difference in duration was the largest for bilabial stops and the smallest for coronal stops). The results of his research were, however, questioned by Dmitrieva, Jongman, and Sereno (2005), who claimed that Pye (1986) did not provide enough statistical evidence in his paper. Dmitrieva, Jongman, and Sereno (2005) arranged their experiment so that the results would provide a satisfying answer to the complexity of the studied phenomenon. According to their findings, Russian native speakers, regardless their proficiency in English (which however played a significant role in the duration of the pronounced vowel), generally tended to use longer vowels before voiced obstruents. The mean difference in vowel duration was 6 ms. The mean duration of vowels preceding voiceless obstruents was 147 ms. Vowels before voiced obstruents lasted approximately 153 ms. The mean duration of vowels followed by voiced obsruents was, therefore, smaller than in English.

In 2008, Mirjana Sokolovic-Perovic carried out an experiment in the use of vowel duration as a cue to final obstruent voicing in Standard Serbian (Sokolovic-Perovic 2009). The results suggested that Standard Serbian has a voicing-conditioned vowel duration contrast. The mean duration of vowels preceding voiced obstruents in final positions was 137 ms, while the mean duration of vowels followed by voiceless obstruents was 114 ms. The mean difference in vowel duration was 23 ms, which is, however, smaller than in English.

In 1994, Crowther and Mann published a paper on the use of vowel duration as a cue to final obstruent voicing in Arabic. Their hypothesis was that Arabic native speakers would utilize their native-language knowledge of the difference between short and long vowels in the use of vowel duration as a cue to English word-final obstruent voicing. The results of their experiments, however, indicate that Arabic native speakers use vocalic duration only in a small extend and that they rather rely on the F1 offset frequency when differentiating the voicing contrast in English word-final obstruents. The mean difference in vowel duration was 28 ms. The vowel was longer when preceding a voiced obstruent – the mean duration before the voiced obstruent was 152 ms, while the mean duration of a vowel preceding the voiceless obstruent was 124 ms. Apart from the native speakers of Arabic, there were also native English participants in the experiment. In comparison with the Arabic subjects' performance, the mean difference in vowel duration of the native English speakers was 75 ms, which was almost three times higher.

In Dutch, the voicing contrast between obstruents is believed to be neutralized in word-final positions and all obstruents are supposedly pronounced as voiceless word-finally (Broersma 2005, 3891). At the same time, Dutch has a phonological distinction between long and short vowels and vowel duration is used as a cue to the contrast in vocalic length, as well as to the voicing contrast in word-medial positions (Broersma 2005, 3891). It is, therefore, believed that Dutch prepares its native speakers for the use of vowel duration as a cue to English word-final obstruent voicing (Broersma 2005, 3895).

A series of perceptual experiments were carried out in order to find out in what extend native Dutch speakers utilize vowel duration as a cue to obstruent voicing word-finally (e.g. Broersma 2005, 2008, 2010). The results suggest that although native speakers of Dutch use vowel duration as a cue, they do not use it in such extend as native speakers of English do. It was, therefore, concluded that the L1 experience with vowel length did not affect the participants' performance when categorizing English word-final voicing (Broersma 2010, 1643).

The character of the cues to the voicing contrast in German have been discussed in several papers focused on the implementation of the phonological obstruent voicing contrast in word-final positions in German (e.g. Charles-Luce 1985, Port and O'Dell 1985, Smith et al. 2009). Charles-Luce (1985) found out that vowel duration was only used as a cue to word-final underlying voicing of fricatives. The mean difference in the duration of vowels preceding word-final underlyingly voiced and voiceless fricatives was approximately 10 ms. In the experiment in production by Port and O'Dell (1985), a mean difference in vowel duration of 15 ms was reported for pairs of words ending in underlyingly voiced and voiceless stops. The results of the experiment by Smith et al. (2009) also suggest that native German speakers utilize vowel duration as a cue to the underlying word-final voicing. The mean difference in vowel duration was 13 ms, which was similar as in the previous studies.

Dinnsen and Charles-Luce (1984) focused on final-obstruent devoicing in Catalan (for more information see section 1.2.1 Incomplete Neutralization). Vowel duration was one of the factors that were measured for each participant of their experiment. According to the data, one of the subjects shortened vowels before underlyingly voiceless obstruents by 20% more than before underlyingly voiced obstruents. Although the data are not as rich as for the other languages, they also suggest that vowel duration may differ before word-final voiceless and voiced obstruents in Catalan.

The use of vowel duration as a cue to word-final voicing was studied also for Mandarin Chinese speakers. In Chinese, the voicing contrast does not exist in word-final positions (Flege 1989, 1684). In 1989, Flege published a paper whose results suggest that Chinese native speakers depend on the release burst, rather than on vowel duration as a cue to English final stop voicing.

Crowther and Mann (1992), on the other hand, came to different conclusions. In their experiments, they examined the use of the cue of vocalic duration among native Japanese and Chinese learners of English. In Japanese, similarly as in Chinese, stops do not occur in word-final positions (Crowther and Mann 1992, 711). However, Japanese includes phonemically long and short vowels (Crowther and Mann 1992, 711).

According to the results, Crowther and Mann (1992) reported that although it is applied in less extend than it is in English, vowel duration was used as a cue to the voicing contrast by the subjects of their experiment. The mean difference in vowel duration made by the Japanese learners of English was 25 ms. The mean difference in vowel duration among the Chinese participants was 12 ms. Both groups, therefore, used vowel duration as a cue to word-final obstruent voicing, although in a smaller extend than native English speakers.

Research that deals with the use of vowel duration as a cue to word-final obstruent voicing in English and in other languages presents varied results. Generally, a tendency to differentiate vowel duration in accordance to voicing of the following obstruent can be observed. However, the mean differences in vowel duration are usually much smaller in other languages than in English. Due to that, vowel duration may not be used as a cue to word-final obstruent voicing by foreign learners of English, who might employ other cues, not as reliable as distinct vowel duration, to the word-final English voicing contrast.

1.3.1.3. *Vowel Duration as a Cue to Word-Final Obstruent Voicing in Czech*

Research concerning the usage of vowel duration as a cue to the voicing contrast between obstruents in Czech is quite recent. Although Patricia Keating (1985) partially focused on Czech in her paper, she only worked with recordings of three Czech native speakers. Her hypothesis was that in Czech, vowel duration would not be used as a cue because in Czech, phonemic vowel-length contrasts are used. She, therefore, assumed that vowel duration would be reserved for the phonemic length contrast (Keating 1985, 121). According to her results, vowels were shortened before voiceless consonants. The mean difference was, however, not statistically significant and the conclusion was that in Czech, vowel duration does not indicate consonant voicing.

The assumption that vowel length which is phonological in certain languages affects the use of vowel duration as a cue to English final consonant voicing in the non-native production and perception has been made in the papers studying these languages. For example Crowther and Mann (1992) assumed that their Japanese subjects performed better than the Chinese subjects because of the phonemic vowel length distinction in Japanese. Broersma (2010, 1637) mentions that the subjects of her experiment, who had been chosen among Dutch native speakers, may utilize their experience with Dutch vowels that are phonemically long and short in their perception of English.

The role of phonemic vowel length in relation to obstruent voicing has been also studied for Czech. In 2007, a paper reporting two experiments regarding the relationship between vowel length and consonant voicing in Czech was published (Podlipský and Chládková 2007). Both experiments were focused on perception and both included nonsensical words that contained phonemically either long or short vowel. According to the results, Czech listeners do not utilize the vowel length distinction when differentiating between voiceless and devoiced obstruents. When it comes to voiced and voiceless obstruent context, however, a tendency that: “quantitatively ambiguous vowels were more likely to be perceived as short before a voiceless coda obstruent than before a (phonetically) voiced coda obstruent” was shown (Podlipský and Chládková 2007, 7). Thus, even in Czech, vowel duration variation, which depends on the voicing of the following obstruent, occurs.

The claim can be supported by results gained in other studies. Although the paper by Šimáčková (2003) had different objectives, the fact that Czech speakers make differences between short and long vowels with respect to the voicing of a consonant which follows that vowel, was revealed. The subjects of her experiment were both Czech and American. Although the Czech talkers produced vowels of different duration, the length differences were not as big as among the American English speakers.

Machač and Skanitzl (2007) recorded Czech speakers producing VCV sequences and then measured durations of consonants and vowels from different perspectives. The findings of their study prove that vowels are shorter before voiceless stops and longer before voiced stops in Czech. They also suggest that the character and length of a vowel affects the duration of the stop in Czech.

Podlipský (2008) examined whether English learners of Czech use vowel duration as a cue to coda voicing in their perception of Czech. The participants in his experiment included both American English and Czech native speakers. The American participants redefined vowel duration as a cue to Czech vowel quantity. More important for the present thesis were findings about the Czech speakers. Among Czech speakers, a tendency to require a slightly longer vowel before a voiced coda than before a voiceless coda was revealed.

Sehnalíková (2010) studied the perception of word-final voicing by Czech native speakers. According to the results of her research, Czech native speakers do not use vowel duration as a cue to word-final voicing in the perception of their native language. When perceiving English word-final voicing, they were very accurate in identifying the correct words. The author of the paper, however, states that: “it is not clear if the vowel duration variation was used as a cue to final voicing.” (Sehnalíková 2010, 70) Based on the findings of the subjects’ perception of Czech words, it is very likely that Czech native speakers do not use vowel duration as a cue to word-final obstruent voicing.

The results of the above mentioned papers indicate that generally, the duration of Czech vowels is partially dependent on the voicing of the obstruent that follows. At the same time, mean differences in vowel length before voiced and voiceless consonants are statistically not large. Keating (1985) reports that the duration of vowels which precede /t/ was in average 193.7 ms, before /d/ it was 204.2 ms (Keating 1985, 122).

In the experiment focused on perception by Podlipský and Chládková (2007), vowel length was intentionally adjusted and it was observed, what the most frequent choice would be among the participants. In the first experiment, a vowel of about 172 ms was chosen in both environments. In the second experiment, the mean difference between vowels chosen in voiced and voiceless context was about 3.3 ms.

Šimáčková (2003) reported that for the Czech native participants in her experiment, /ɛ/ was 122 ms long in the voiced context and 98 ms long in the voiceless context. The vowel /æ/ lasted 151 ms in the voiced context and 128 ms in the voiceless context. The differences in duration are, however, very small compared to those produced by American English native speakers in identical words.

Machač and Skarnitzl's paper (2007) includes a table with durations of different vowels before various plosives. Although in several cases vowels that precede voiced stop are longer than those which are followed by a voiceless one, the difference in duration was generally not significant.

Podlipský (2008) found out that in the perceptual test he gave to Czech native speakers: "the difference between the mean boundary location in the voiced coda context (100.89 ms) and the mean boundary location in the voiceless-coda context (96.95 ms) was approximately 3.93 ms." (Podlipský 2008, 10535)

These data suggest that Czech speakers do not use contrastive vowel duration as a cue to consonant voicing. As in English vowel duration is the most important cue to final obstruent voicing, it is necessary that Czech learners of English get accustomed with the rule.

1.3.1.4. The Role of the Degree of Proficiency

A number of studies have analysed the fact whether vowel duration as a cue to obstruent voicing is likely to be acquired intuitively throughout the process of language learning and whether the learner's degree of proficiency in the second language has any impact on their perception and production of correct vowel duration.

For example, Dmitrieva, Jongman, and Sereno (2010) studied the extent of final devoicing among speakers of Russian. Their participants were divided into three groups – monolingual native speakers of Russian, Russian native speakers learning English, American English native speakers learning Russian. They were asked to read Russian word pairs. It was found out that Russian learners of English employ vowel duration more than monolingual speakers of Russian. The more proficient the speakers were in English, the more they employed distinctive vowel duration in their production of the stimuli. Similarly, the more proficient the American learners were in Russian, the smaller the differences in vowel duration were in their production.

For Czech, Skarnitzl and Šturm (2014) published a paper concerning how the degree of foreign (Czech) accent affects the speaker's treatment of English word-final obstruent voicing. In their experiment, three groups of Czech native speakers varying in the degree of Czech accent in their pronunciation of English were involved. In their task, the participants were asked to read a news bulletin. The obtained results proved the authors' hypothesis that the more the subjects' speech is Czech-accented, the more likely to neutralize the voicing contrast word-finally they would be.

Broersma (2010) studied the use of vocalic duration as a cue to final fricative voicing by Dutch and British English listeners. Two experiments, each involving different participants, were performed. The experiments were focused on perception. The participants were presented with words whose final consonants were modified in order to find out whether the participants would use vowel duration as a cue to their voicing. The native speakers of Dutch that took part in her experiment were all proficient in English. Although they used vowel duration as a cue, they were unable to use it in such an extent as the British English native speakers did.

In the research by Smith et al. (2009), word-final obstruent voicing was studied for German. The native German speakers who participated in their experiment had a good knowledge of English as they had been studying English for approximately 10 years and they had been living in the USA for 3.8 years on average. They were asked to read minimal pairs of similar (orthographically and phonemically) English and German words, differentiating in the voicing of the word-final obstruent. According to the results, the participants utilized vowel duration in a greater extend in their production of English than in German (the mean difference in vowel duration was 41 ms for the English words, while the mean difference in vowel duration for the German words was only 13 ms). The results suggest that the degree of proficiency played a significant role in the experiment, although it: "was not a specific focus of [the] study concerns." (Smith et al. 2009, 272)

In general, it is suggested that with rising proficiency in English, foreign learners start to perceive and produce the difference in the duration of vowels which precede voiced and voiceless obstruents. However, the differences still remain much smaller than among native English speakers. It is, therefore, obvious that the use of vowel duration as a cue to final obstruent voicing is hard to acquire and that it is necessary to find a way to teach learners of English how to use it.

1.4. Training Methods

In the previous sections, the problem that non-native speakers of English have to deal with in their production and perception of English word-final obstruent voicing was introduced. The problem of an insufficient reliance on vowel duration as a cue to final obstruent voicing has been widely studied and there is a number of papers on various languages documenting the fact that learners of English do not use the acoustic cues to final obstruent voicing in a native-like way (see chapter 1.3.1.2 Vowel Duration as a Cue to Obstruent Voicing in Other Languages than English).

In the past, it was believed that language learners lose the ability to discriminate among phonological categories that are not present in their native language in adulthood unless they are consistently exposed to them since their early development (see Pisoni et al. 1982 for the review of literature). This view has been challenged and it has been suggested and examined that under specific conditions, adult language learners are able to acquire new, non-native phonological categories in order to improve their linguistic abilities (Pisoni et al. 1982) and on the top of that, they are able to transfer the newly learnt knowledge to other contexts from those that they are trained in (McClaskey, Pisoni and Carrell 1983). Moreover, after a series of experiments by Rochet (1995) it was concluded that accurate perception of L2[2] phonological categories is crucial for production to be native-like.

The primary aim of this thesis is to use the well-established methods to train Czech learners of English to both perceive and produce English word-final obstruent

[2] L2 = second language; a person's non-native language

voicing. First, the training tasks and the perceptual training methods will be introduced. Next, the effects of the perceptual training on production will be discussed.

1.4.1. Perceptual Training Methods

A number of methods in training adult's acquisition of non-native sound patterns have been put into practice. Researchers have been focusing on training both the perception and the production of novel non-native phonetic categories. Overall, the methodologies vary in the duration of the training and its design. Within studies focused on perception, the nature of the stimuli that is presented to the subjects (the listeners) and the number of talkers that get recorded are two main differentiating factors.

The duration of the training depends on the intended goal and on the nature of the problem that the training is aimed at. Therefore, there is variability within the duration of training. Generally, training programmes are either short-term or long-term (Gordon 2008, 45). A short-term training does not exceed one session on one day (Logan and Pruitt, 1995). A long-term training extends over a longer period of time, depending on the design of the training.

Most commonly, training methods are designed so that they compose of a pretest, a training session (one or more, depending on the design) and a posttest. Sometimes, generalization tests are involved. The purpose of generalization tests is to assess the subjects' ability to adapt their newly gained knowledge to new conditions (e.g. speakers, phonetic environments).

The methodological approaches to training perception involve stimuli which either composes of naturally produced tokens or synthetized tokens. Sometimes, both types of tokens are involved in order to provide the subjects with both the natural variability as well as with focus on the phonetic property that is being acquired, which can be achieved through synthetization of natural speech.

The number of talkers that get recorded is the last differentiating factor of the training methods. Some training methods use multiple speakers, while some use only a single talker. The number of speakers varies and depends on the requirements of the specific training.

Among the training methods in perception, there are two tasks that are most commonly put in use – it is the identification task and the discrimination task (apart from that, visual feedback or a combination of visual and auditory stimuli might be used (Wang and Munro 2004, 540). In the following chapter, the identification task and the discrimination task will be described first. Next, the training methods will be accounted for. The following methods will be discussed: low variability phonetic training, high variability phonetic training and the perceptual fading technique.

1.4.1.1. Identification Task vs. Discrimination Task

The identification task and the discrimination task are two training tasks, which differ in both their design and the results, which they produce.

The identification task is based on a forced-choice principle. The subjects listen to a single stimulus and subsequently, they are presented with options. Then they have to choose from one of the alternatives that they are given, according to what they think is the right representation of the stimulus heard. During the training sessions, the choice is followed by an immediate feedback.

While in the identification task, the subjects listen to a single token, in the discrimination task, they are presented with either two or three stimuli in sequence (Wang and Munro 2004, 540). There are three types of the discrimination task: the category change paradigm, the same/different paradigm (also called AX) and the oddity paradigm (also called ABX) (Wang 2002, 23). The category change paradigm is used in infant speech perception (Wang 2002, 23) and is, therefore, not relevant to the subject of this thesis. In the AX task, the subjects listen to two tokens and then they have to decide whether the tokens are same or different. In the ABX task, the subjects listen to three tokens: A, B, X. Then they decide whether the token X is the same as the token A or B.

Several studies comparing the two tasks have been published. The first one to deal with it was a paper by Jamieson and Morosan (1986). The aim of their research was to design a method that will prove to be effective in adult's speech laboratory training. They argued against the discrimination task, claiming that: "discrimination training rarely improves the categorization of non-native speech contrasts." (Jamieson and Morosan 1986, 207). Apart from that, they believed that the discrimination task shifts the subjects' attention to the acoustic differences between phonemes. They believed that the desired outcome of a training are learners with a newly acquired ability to classify speech sounds into categories that are relevant to L2, rather than subjects trained to recognize intraphonemic differences.

Their experiment consisted of a pretest, 3 training sessions and a posttest. The pretest and the posttest consisted of both the identification task and the discrimination task. The training sessions included identification task-based tests. The subjects were divided into two groups – a control group and a training group. The control group received only the pretest and the posttest, while the training group participated in all parts of the experiment. The results suggest that the identification training was successful as the training group's performance improved from pretest to posttest, while no such improvement was noticed for the control group.

A later paper by Flege (1995) re-examined the efficiency of both methods. The goal of the research was to train native Mandarin Chinese speakers to perceive the word-final /t/ – /d/ contrast in English. In the experiment, the subjects were divided into two groups. One group received the identification task, the other group was given the same/different discrimination task. Both groups received a pretest, ten training sessions and two posttests. The stimuli was synthetized for both groups. The results suggested that both tasks were equally effective. However, according to the author: "Those who received identification training responded more positively than those who received same/different training." (Flege 1995, 439)

Recently, the two tasks were re-examined and compared by Carlet and Cebrian (2015), whose training was designed to improve the perception of English vowels among native Spanish speakers. The subjects were divided into three groups – one of them received the identification task, one of them received the discrimination task, the third one was a control group with no training. The two groups which received training participated in a pretest, training and a posttest. The results suggest that although the discrimination task yielded good results, the identification task was more effective.

Generally, the identification task has been considered to be more effective and is believed to "have dominated recent training studies." (Wang 2002, 25) It is also believed that: "identification tasks have yielded better results in [the] studies, possibly because they lead trainees to direct their attention to the specific characteristics of a speech sound that make it differ from the other member of the contrastive pair." (Wang and Munro 2004, 541).

1.4.1.2. *Low Variability Phonetic Training*

Low variability phonetic training is a method of training perception of non-native phonetic categories, which involves one talker and a single phonetic context (Wong 2012, 37). It was first practiced by Strange and Dittman (1984), whose work is considered "the groundwork for future non-native phoneme contrast training." (Bradlow 2007, 294)

In their study, Strange and Dittman (1984) aimed to train the perception of /r/ – /l/ among Japanese learners of English. Their training included a pretest, 14 – 18 training sessions and a posttest. The pretest and the posttest included naturally produced stimuli, while during the training sessions, the subjects were presented with digitalized speech. The stimuli included monosyllabic words with /r/ – /l/ in their initial positions.

The subjects were given a variety of tasks, including a minimal-pairs test, an identification task and a discrimination test in their pretest. The training sessions consisted of a same/different discrimination task with immediate feedback. The posttest used the same tasks as the pretest – the identification and discrimination tasks and the minimal-pairs test.

The results of the study indicate that the training was successful in terms of discrimination of the synthetic stimuli. The subjects were, however, unable to extend their ability to distinguish between /r/ – /l/ in naturally produced words, which they were presented with in their pretest and their posttest.

The low variability phonetic training thus uses stimuli produced by a single talker, which is synthetized for the purposes of the training. Apart from that, the stimuli do not change during the training and the same phonetic context remains.

1.4.1.3. *High Variability Phonetic Training*

High variability phonetic training was designed by Logan, Lively, and Pisoni (1991) in order to: "circumvent some of the difficulties associated with Strange and Dittman's [(1984)] study." (Logan, Lively, and Pisoni 1991, 876) According to the authors, the problematic aspects of Strange and Dittman's (1984) training procedure were above all the single phonetic environment in which the contrasting /r/ – /l/ were presented, the fact that only a single talker was recorded, as well as the use of synthetic tokens. They believed that through the inclusion of a variety of phonetic contexts and naturally produced stimuli by multiple speakers, the subjects would be capable of generalization to novel tokens. Apart from that, they believed that the training would be beneficial on a long-term basis.

In their paper, Logan et al. (1991) focused on training Japanese learners of English to identify /r/ – /l/. Their work was based on a similar assumption as that of Strange and Dittman (1984) and of Jamieson and Morosan (1986), believing that the acquisition of new phonetic categories is difficult in adulthood and that an appropriate laboratory training is necessary.

The training consisted of a pretest, the training itself and a posttest. In all parts of the training, the identification task was used as the only training procedure. The stimuli included words with /r/ – /l/ in all phonetic environments. The words were all naturally produced. There were six native English speakers recorded in order to provide variability.

According to the results, all subjects showed an improvement in their identification of the stimuli. The results, however, varied with respect to the phonetic environment (more improvement in the identification of word-final contrasts than in the identification of the word-initial contrasts), response times (response times were faster for such phonetic environments, which had not posed problems to identification from the beginning; response times for the identification in such phonetic environments, which were initially problematic, decreased during training) and the talker (stimuli produced by certain talkers were identified more accurately than stimuli produced by other talkers).

A number of studies using the high variability phonetic training followed, focusing not only on the discrimination of English /r/ – /l/ by Japanese native speakers (e.g. Yamada 1993; Lively et al. 1994; Bradlow et al. 1997, 1999; Iverson, Hazan, and Bannister 2005), but also on other areas within non-native phonetic perception and production: as mentioned in 1.4.1.1., Flege (1995), for example, trained native speakers of Mandarin Chinese to perceive English word-final /t/ – /d/ contrast; Wang et al. (1999) trained native American English speakers to perceive contrasts in Mandarin Chinese tones. Pruitt, Jenkins and Strange (2006) trained both American and Japanese native speakers in perception of Hindi dental and retroflex stops.

In summary, the high variability phonetic training uses naturally produced stimuli from multiple speakers in order to provide natural variability to the learners. It is believed to: "enhance participants' long-term recall and force them to focus on phonetic cues that underlie categorical distinctions." (Wong 2012, 37) Although the identification task is more widely used than the discrimination task, both are applicable to the high variability phonetic training.

The high variability phonetic training was first directly compared to the low variability phonetic training by Wong (2012) in order to assess the efficiency of the former over the latter. The author of the paper trained Cantonese learners of English to perceive and produce English /æ/ – /e/ contrast. The subjects were divided into three groups. One of the groups was a control group, while one of the other two received the high variability phonetic training and the last one got trained using the low variability phonetic training method. The perceptual training included a pretest, a posttest, ten training sessions and two generalization tests. The production was tested with a pretest, a posttest and a test of generalization.

The results confirmed the author’s hypotheses that the subjects trained with the high variability phonetic method would outperform the other groups in perception and that they would be able to transform their knowledge into production more effectively. Apart from that, the group trained with the high variability phonetic method showed the largest degree of generalization to new talkers and phonetic environments. It was, therefore, concluded that: “exposure to highly variable stimuli is necessary for the subjects to form robust phonetic representations by learning which acoustic cues are relevant to a specific sound.” (Wong 2012, 40)

1.4.1.4. The Perceptual Fading Technique

The perceptual fading technique is an approach which involves a use of synthetic stimuli in a specific way, which differentiates this type of training from the low variability training and the high variability training. It was first used by Jamieson and Morosan (1986).

Apart from comparing the efficiency of the identification task and the discrimination task, Jamieson and Morosan’s (1986) research introduced a perceptual fading training method, which made use of a gradual enhancement of primary acoustic cues relevant for the phonetic contrast.

Their goal was to train the English /ð/ – /θ/ contrast among Canadian Francophones. The training involved a pretest, two training sessions and a posttest. The training sessions were designed in the following manner: the subjects were first presented with maximally enhanced tokens and they gradually progressed towards less enhanced tokens. The maximally enhanced tokens were designed to signal the most important category differences – the subjects were meant to focus on the most important primary cues to the trained contrast. As they progressed through the training, the category differences began to fade away. Towards the end of the training, the tokens became ambiguous and more acoustic variability was introduced to the listeners. The results of the research were positive – the subjects were successfully trained in the identification of both natural and synthetized stimuli containing the target sounds.

The perceptual fading technique has become an effective method in training perceptual category differences and it was later used for example by Rochet (1995) for training native Mandarin Chinese speakers' identification of French voiced and voiceless stops. Another example of the use of the fading technique is a study by Pruitt (1995) on the perception of Hindi dental and retroflex stops by English and Japanese speakers.

1.5.Effects of Perceptual Training on Production

The present thesis focuses on training both the perception and the production of the word-final voice contrast. The most commonly used methods of perceptual training were introduced in chapter 3.1. The current chapter deals with the link that has been found in between speech perception and production, which suggests that if the training is well-designed, both areas of L2 acquisition can develop simultaneously.

The effects that perceptual training has on speech production and the possible correlation between the two areas were first investigated by Jamieson and Rvachew (1992, 1994), Rvachew (1994), Rochet (1995) and Bradlow et al. (1997). In the introduction to their paper, Bradlow et al. (1997) state that although there had been studies which aimed to examine the relationship between production and perception (e.g. Sheldon and Strange 1982; Yamada et al. 1994), "they [did] not provide quantitative information about how the changes in one domain (i.e. perception) affect performance in the other domain (i.e. production)." (Bradlow et al. 1997, 3)

The studies by Jamieson and Rvachew (1992, 1994) and Rvachew (1994) were focused on training children with speech problems to perceive and produce phonetic categories correctly. They believed that the methods which were used in contemporary speech pathology were based on wrong assumptions, as they did not consider speech perception to play a relevant role in the process of learning correct speech production. Jamieson and Rvachew, however, believed that perception and production are interrelated.

Apart from that, they assumed that children with speech problems can be trained to identify and produce sounds in a similar manner as L2 learners. They believed that there are three patterns of making errors in speech perception and production, which both groups make: "(1) Neither member of the contrast is present in the underlying system; (2) Both members of the phoneme pair belong to a single category in the underlying system; and (3) Both members of the contrasting pair exist as separate categories in the underlying system, but they are differentiated in terms of nonstandard cues." (Jamieson and Rvachew 1994, 1200)

In their experiment, Jamieson and Rvachew (1992) used the perceptual fading technique to train correct perception of English fricatives. The training was designed in a similar way as the one used by Jamieson and Morosan (1986) to train adult L2 learners. The subjects were children aged 5 to 7 with functional articulation disorders. The stimuli was synthetized and it was introduced to the subjects gradually, starting with the most extreme tokens, ending with less extreme versions, including naturally produced tokens. According to the results, the training was beneficial both with respect to perception and production.

Rvachew (1994) trained preschool children with phonological impairment, who had problems with the articulation of /ʃ/. The high variability phonetic training method was selected for the training – naturally produced words were, therefore, used as the training material. Apart from that, traditional speech production training was used to teach children the correct articulation of the target sound. The children were divided into three groups. Two of the groups received relevant training, while a third group was trained on words unrelated to the target sound. According to the results, the groups which received the relevant training showed an improvement in both perception and production of /ʃ/, while the third group did not. The results thus proved how beneficial perceptual training is for production.

Rochet (1995) used the perceptual fading technique to improve native Mandarin Chinese speakers' perception and production of French voiced and voiceless stops. The subjects attended six training sessions, during which they listened to a series of synthetized French /pu/ – /bu/ tokens. According to the results, both perception and production of the stops developed. On the top of that, the analysed production results showed a change in the VOT durations, which were closer to those of native French speakers.

In 1997, Bradlow et al. published a follow-up to the studies by Logan et al. (1991) and Lively et al. (1993, 1994). The goal of their research was to examine the link between perception and production "in the absence of any explicit production training, and across a wide range of phonetic contexts." (Bradlow et al. 1997, 3)

In their experiment, the high variability phonetic training with the identification task was used to train a group of native Japanese speakers to perceive and produce the /r/ – /l/ contrast. The subjects were recorded before the perceptual pretest and after the perceptual posttest. The recordings were then assessed by a group of native American English speakers in two evaluation tests.

Similarly as Rochet (1995), Bradlow et al. (1997) observed an improvement in production after the perceptual training, proving the existence of a link between perception and production. In both evaluation tests, the data obtained from the posttest in production were generally considered to be better representations of the categories than the data recorded in the production pretest.

Previous research has found out that speech production can be trained alongside speech perception. The experiments, which were carried out with children exhibiting speech impairment (Jamieson and Rvachew 1992, Rvachew 1994), as well as the experiments with adult L2 learners (Rochet 1995, Bradlow et al. 1997), all successfully improved the subjects' articulation and perception of the target sounds.

1.6.Research Questions and Hypotheses

Based on the facts that have been discussed in the sections above, the following research questions (Q) arose and the following hypotheses (H) have been made:

Q_1: Does the degree of proficiency in L2 affect the subjects' performance?

H_1: The degree of proficiency in L2 is believed to positively influence the ability to recognize non-native speech categories. Providing that the subjects of the training are fluent in English, the amount of their correct responses should be high in the experiment.

The subjects should be able to correctly identify the correct answers with high accuracy already in the pretest, which precedes the actual training.

Q_2: Will the perception of English word-final obstruent voicing improve after using the high variability phonetic training with the identification task?

H_2: The high variability phonetic training is believed to be an effective method in training the acquisition of non-native speech categories. At the same time, the use of the identification task is supposed to facilitate successful acquisition of the target characteristics of speech. Providing that the subjects undergo the high variability phonetic training using the identification task, their perception of English word-final obstruent voicing should improve. The improvement should be signalled by shorter response times, as well as a higher number of correct answers after the training.

Q_3: Will production be affected by the perceptual high variability phonetic training?

H_3: It is believed that an improvement in perception is accompanied with an advancement in production, providing that an appropriate training method is used. It is, therefore, assumed that if the subjects of the high variability training improve in their perception of English word-final obstruent voicing, their production should improve as well.

Q_4: Will training with enhanced vowel duration improve the subjects' ability to utilize vowel duration as a cue to word-final obstruent voicing in English?

H_4: In Czech, neutralization of word-final obstruent voicing is complete and vowel duration is not used as an acoustic cue to the voicing contrast, while in English vowel duration serves as the most important cue to word-final obstruent voicing. The training should, therefore, focus on the subjects' ability to utilize distinctive vowel duration as a cue to English word-final obstruent voicing. If the training involves stimuli with enhanced vowel duration, the subjects should be able to learn to use vowel duration as a cue to word-final voicing. At the same time, they should be capable of generalization to perception of naturally produced stimuli. Moreover, their ability to use vowel duration as a cue should be reflected in their production.

Q_5: Do phonetic characteristics of the word-final obstruents play a role in the identification?

H_5: The results of the experiment should reveal whether the subjects' identification differs with respect to the phonetic properties of the word-final obstruent and if it does, it should show up which obstruents are most easily recognized. Furthermore, the training should result in a more precise identification of those obstruents, whose classification was initially poor.

2. Methodology

The goal of the present thesis was to train Czech learners of English in perception and production of English word-final obstruent voicing. The present chapter will introduce the training experiment. In section 2.1., information about the participants of the training will be given. Section 2.2. will inform about the stimuli used in the training. In the last section, 2.3., the procedure will be described.

2.1. Participants

2.1.1. Perceptual Training

There were 36 participants in the perceptual training, 12 of them were male and 24 of them were female. They were all Czech native speakers. All of the participants were university students majoring in English, aged 19–30 years. During the time of the training, all of the participants were attending a seminar in phonetics. Moreover, some of them had previously received lectures in phonetics.

The participants were divided into three groups. Group 1 consisted of 6 male students and 7 female students. In Group 2, there were 4 male students and 10 female students. Group 3 comprised of 2 male students and 7 female subjects.

2.1.2. Production Testing

18 students participated in tests of the effects of perceptual training on production. The participants were selected from the groups of students which underwent the perceptual training. 6 students – three male, three female – from each of the three groups took part in the experiment.

2.2. Stimuli

The stimuli selected for the training consisted of real English monosyllabic words. The words formed 58 minimal pairs, differing in the final obstruent. The following pairs of obstruents were selected to appear in the word-final positions: /p/ – /b/, /t/ – /d/, /k/ – /g/, /f/ – /v/, /s/ – /z/, /ʧ/ – /ʤ/. All words were structured CVC. The vowels preceding the final obstruents were both monophtongs and diphthongs. For the complete list of the stimuli, see Appendix 1.

The recordings of the stimuli for the perceptual training were obtained from the online MacMillan Dictionary using Audacity. The stimuli were produced by three male and three female native British English speakers. It was considered necessary to use naturally produced stimuli and not synthetized tokens in the training as it has been believed that training with synthetic stimuli does not yield as good results as training with natural speech (e.g. Logan, Lively, and Pisoni 1991, Ingvalson, Ettlinger, and Wong 2014). The amount of speakers and their gender were supposed to provide the required variability to the training.

No fillers and no nonsense words were used neither in the perceptual training, nor in the pretest and the posttests in perception. Similarly, no fillers or nonsense words were used in the production experiment. It was not considered necessary to include fillers as the main goal of the training was to shift the subjects' attention to the relevant acoustic cues, not to assess their ability to identify the word-final obstruent voicing contrast.

The stimuli were divided into two groups, containing the same word pairs. In one of the groups, the stimuli were enhanced, while in the other group, the stimuli remained in their original version. For the enhanced stimuli, several features were adjusted in order to shift the listeners' attention to the variability of vowel duration which is used as a cue to the voicing contrast in word-final positions in English. The changes were made through scripts created in Praat (Boersma and Weenink 2014; see the scripts in Appendix 3 and Appendix 4). First, vowel duration and duration of constriction were manipulated. When adjusting the durations of vowels, the initial and final 20% of the vowels remained unchanged so that the CV and VC formant transitions would not get speeded up or slowed down. In stimuli ending with a voiced obstruent, the original vowel duration was multiplied by 1.35. Vowel duration in stimuli with a final voiceless obstruent was multiplied by 0.85.

The durations of constrictions were also manipulated. In the enhanced stimuli, the constriction duration of the final obstruents was equal to a mean of duration between the original constrictions of the voiced and voiceless obstruent in their minimal pair.

In addition, for half of the stimuli with a final voiced obstruent, low frequencies were filtered out in order to avoid using the potential voicing as a cue. In the word-final voiced fricatives, all frequencies under 500 Hz in stimuli produced by a male speaker were filtered out. In stimuli produced by a female speaker, frequencies lower than 600

Hz in word-final fricatives were erased. During the noise bursts of stops, the same procedure as for the fricatives was carried out with the frequencies. During the stop closures, the frequencies under 1100 Hz in stimuli produced by male speakers and under 1300 Hz for stimuli by female speakers were filtered out. On the top of that, noise intensity of voiced fricatives was increased by 5 dB.

In the other half of the stimuli that originally contained a final voiced obstruent, the final obstruents were made ambiguous as they were replaced by their voiceless counterparts. After the voiced obstruent was replaced, the noise intensity of the voiceless obstruent was decreased by 6 dB for fricatives and by 4 dB for stops and the constriction duration was decreased to be half way between the original duration and the duration for the voiced counterpart of the obstruent.

2.3.Procedure

The experimental design was based on a pretest – training – posttest procedure. First of all, a pretest in production was made. A week later, a pretest in perception followed. On the same day, the training was carried out, followed by a posttest in perception. In order to assess the efficiency of the training, the subjects were given another perceptual posttest two weeks after the training. On the same day, the production posttest was made.

2.3.1. Perceptual Training

There were three groups of participants in the perceptual experiment. Group 1 (Enhanced Sounds Group) was presented with the set of stimuli with enhanced vowel duration. Group 2 (Original Sounds Group) was trained using the original versions of the recordings of the stimuli, in which vowel duration was not enhanced. The last group (Control Group) was a control group. The Control Group did not receive any training, they were only given the pretest and the two posttests. The purpose of having a control group was to assess whether repeated testing and no training would shape the subjects' performance or not.

Before the training, all groups were given the same pretest. The test included all 58 minimal pairs of words in their original versions with no enhancement of vowel duration. The test itself was both created and presented to the participants in Praat (Boersma, Weenink 2014). First of all, the participants were asked to fill in a short questionnaire, asking for their initials, age, gender and the type of test they were about to take (pretest / posttest / posttest2). Next, they were provided with headphones, written English instructions and a trial test. The test used a two-alternative identification task with no feedback: a minimal pair of words was displayed on the screen and at the same time, the participants heard one of the displayed words. Their task was to decide which word they heard and press a corresponding button. The stimuli were presented to the subjects in a random order. No repetitions were made. After all the minimal pairs of words were went through, the task was finished. The completion of the test took approximately 5 minutes.

The training phase for the Original Sounds Group and the Enhanced Sounds Group immediately followed the pretest. The task had the same design for both groups. Both groups were presented with 58 minimal word pairs. What the training tasks differed in, was the nature of the stimuli. Similarly as the pretest, the training task was both created and presented to the subjects in Praat (Boersma and Weenink 2014). Firstly, the participants filled in a questionnaire with their initials, age, gender and their identification number. Then they were given headphones and instructions for the task in Czech. Before the actual training task started, they were asked to complete a practice test in order to get acquainted with the testing method.

In the training, the two-alternative identification task was also used. Similarly as in the pretest, the subjects were shown a minimal pair of words and after listening to the sound, they were supposed to choose one of the options on the screen. However, unlike in the pretest, the participants were enabled to relisten to the stimuli before choosing their answer. On the top of that, they were given an immediate feedback – their answer was highlighted either red (incorrect) or green (correct).

In order to improve the subjects' identification of the stimuli, the following procedure was employed in the training phase: if the identification was incorrect, the stimulus did not get deleted from the task. It was, on the other hand, repeated. The task, therefore, consisted of a number of rounds, in which the participants were repeatedly presented with the stimuli they were not successful in identifying. The items that were successfully identified in the first round were deleted afterwards. Once the subjects began to identify the incorrect items correctly, they had to do so three times before the item was deleted from the task. The purpose of such a design was to provide the listeners with a maximum amount of exposure to the items they find difficult to recognize. The maximum number of trials was 450. After that, the training phase ended. The training phase was completed within 15–20 minutes in the Enhanced Sounds group. The training phase for the Original Sounds group lasted approximately 10 minutes.

After the training session, the participants were asked to do two posttests. Both posttests employed a procedure identical to the pretest. All groups were presented with the original versions of the stimuli with no enhancement in vowel duration. The completion of the posttests took approximately 5 minutes for each test.

2.3.2. Production Testing

Six subjects from each group were selected to take the tests in production. 21 minimal pairs were used in the production experiment (for the list of stimuli see Appendix 2). Both the pretest and the posttest proceeded in the same way. The subjects were recorded individually in a soundproof room, using the H4n Handy Recorder. First, they were asked to provide their initials and information about their age and whether they had previously attended any phonetic courses. Next, each participant received spoken instructions and a sheet of paper with a list of stimuli. Each participant was asked whether the instructions were clear and whether they were familiar with all the words on the list.

Afterwards, the participants were provided with a pair of headphones connected to the recording device. They were recorded while reading out loud words that they were presented with on a computer screen. The presentation of the words was created through a script in Praat (Boersma and Weenink 2014). The recording lasted approximately 1 minute for each participant.

3. Results

The present chapter informs about the results of the experiment. For reasons stated at the end of this chapter, only the results of the perceptual training are included. In order to assess the effectivity of the training in perception, results from all parts of the experiment (pretest, training phase, posttest, posttest 2) had to be obtained from each participant[3]. However, there were some participants, who were not present for at least one of the parts of the training. The progress, therefore, could not have been traced for these subjects and the data that were obtained from them were not submitted to the analysis. From the Enhanced Sounds Group, only 9 out of 13 participants completed all the parts of the experiment. In the Original Sounds Group, 12 out of 14 of the results were complete. In the Control Group, 8 out of 9 participants did all three tests.

The main goal of the experiment was to train the participants to utilize vowel duration as a cue to word-final voicing. If the training was successful, the groups who received it, would have an increased amount of correct responses in the posttests in comparison with the pretest. At the same time, the duration of their response times would be shorter in the posttests, compared to the pretest. In order to evaluate the progress, percentages correct and reaction times were submitted to two separate repeated measures ANOVAs with a Treatment Group (Enhanced Sounds Group, Original Sounds Group, Control Group) as the between-subject independent variable and with Test (Pretest, Posttest, Posttest 2), Segment Type (P, T, T, F, S, respectively B, D, G, V, Z) and Voicing (Voiced, Voiceless) as the within-subject factors.

3.1. Percentages of Correct Responses

Looking at the percentages correct first, the highest mean percentage of correct responses was 85.5% and it was achieved by the Original Sounds Group (see Figure 1). The Control Group’s mean percentage correct was 83% and the Enhanced Sounds Group had 81% mean percentage correct. Although the mean percentages of correct responses are very close for all groups, there are big differences among the participants in each group. While the lowest percentage correct for the Original Sounds Group was 81%, the highest result

[3] For the Control Group, data from the pretest, posttest and posttest 2 had to be present.

was 89%. Similar differences were also found for the other groups. Despite slight differences in the amounts of percentages correct, the ANOVA did not find a significant main effect of Treatment Group overall ($p > 0.05$).

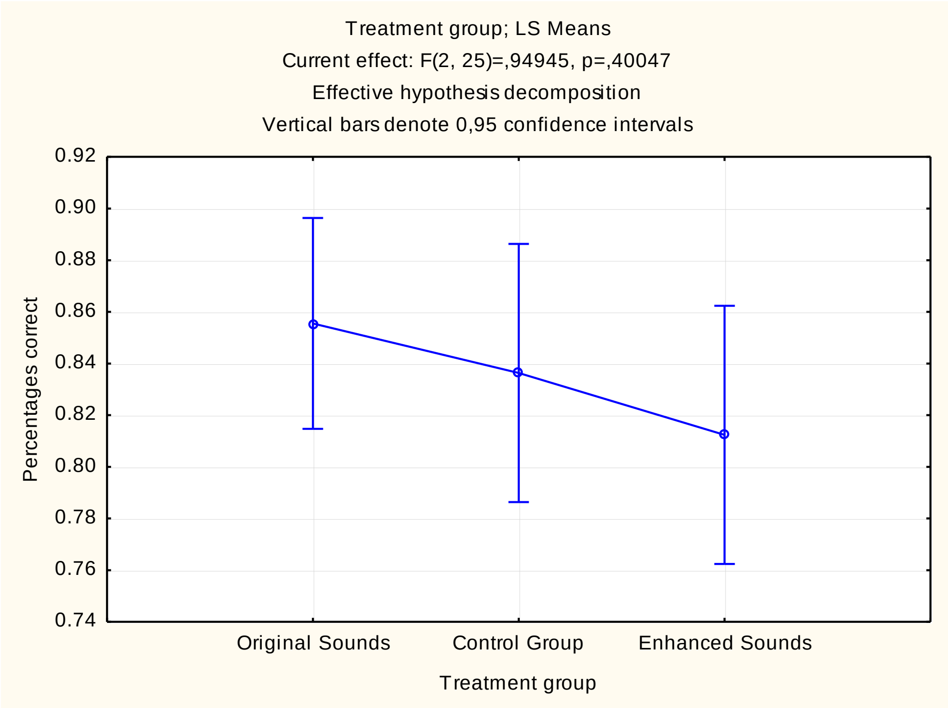

Figure 1 Percentages Correct: The Main Effect of Treatment Group

On the other hand, the main effect of Test was significant (current effect: $F(2, 50) = 5.0663$, $p = 0.00992$) (see Figure 2). According to the results obtained from the three groups together, the number of correct responses increased from the Pretest to the Posttest. The mean percentage of correct responses was 81% in the Pretest and 85% in the Posttest. Moreover, the amount of correct responses did not decrease significantly in the Posttest 2 (the mean percentage correct was 84%), suggesting that the tendency to identify the stimuli correctly does not fade away in the course of time.

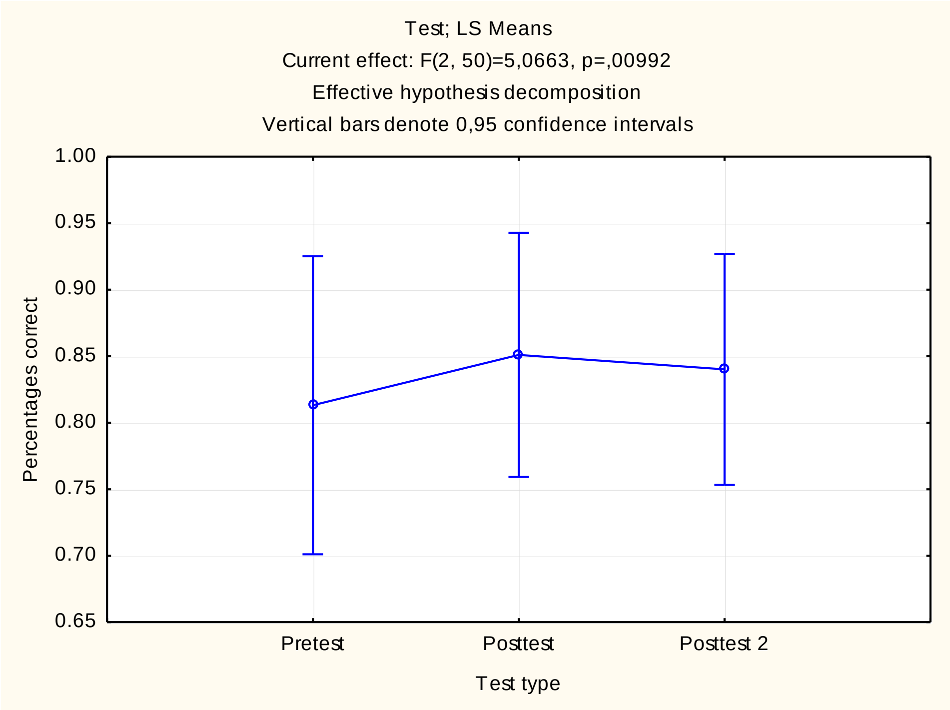

Figure 2 Percentages Correct: The Main Effect of Test

However, because the results presented in Figure 1 and Figure 2 are very general and not much can be concluded from them about how successful the training was, the interaction between the Treatment Group and the Test was measured. According to the results shown in Figure 3, there was not a significant interaction between the Treatment Group and the Test ($p > 0.05$). Although the interaction is not statistically significant, generally, small progress can be seen for the Enhanced Sounds Group (the mean percentage of correct responses increased from 79% to 83%) and for the Control Group (the mean percentage correct changed from 80% in the Pretest to 86% in the Posttest). Subsequently, the amount of correct responses did not decrease in the Posttest 2.

Figure 3 demonstrates that although the performance of the groups did not differ much in general, there were big differences among the participants within the groups – for example in the Enhanced Sounds Group, the lowest percentage of correct responses in the Pretest was 58%, while the highest percentage correct was 99%. A similar difference is within the Control Group (59% the lowest percentage correct, 100% was the highest percentage correct).

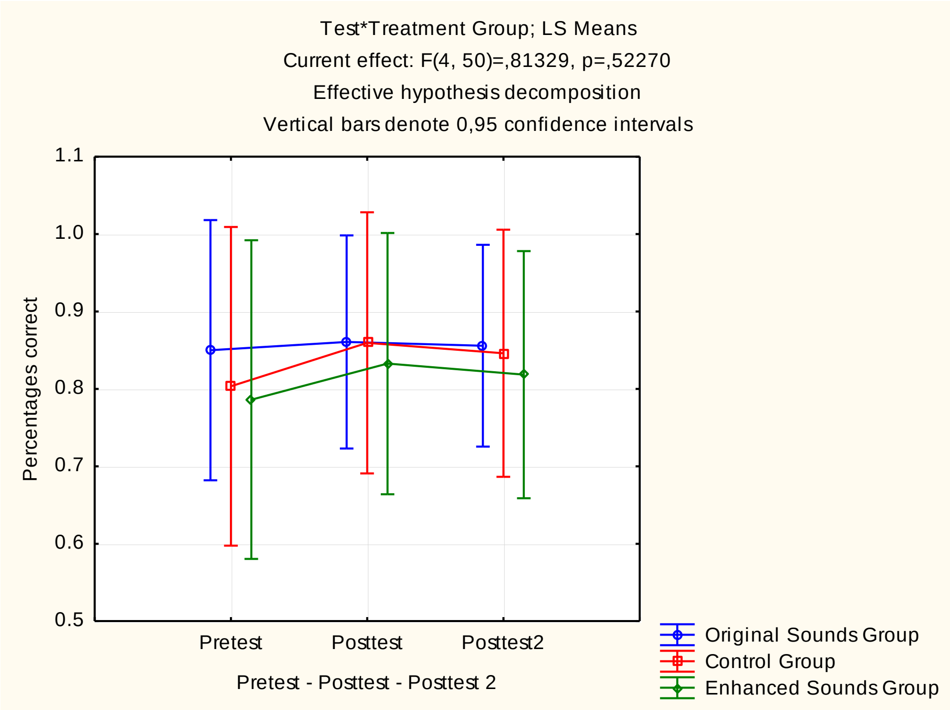

Figure 3 Percentages Correct: The Interaction Between Treatment Group and Test

Figure 4 shows the percentage of correct responses as a function of phonological voicing of the word-final obstruent. The main effect of Voicing was significant (current effect: $F(1, 25) = 6.7781$, $p = 0.01530$). It was found out that the subjects' performance was better for the word-final voiceless obstruents (the mean percentage correct was 86%) and worse for the word-final voiced obstruents (the mean percentage correct was 80%).

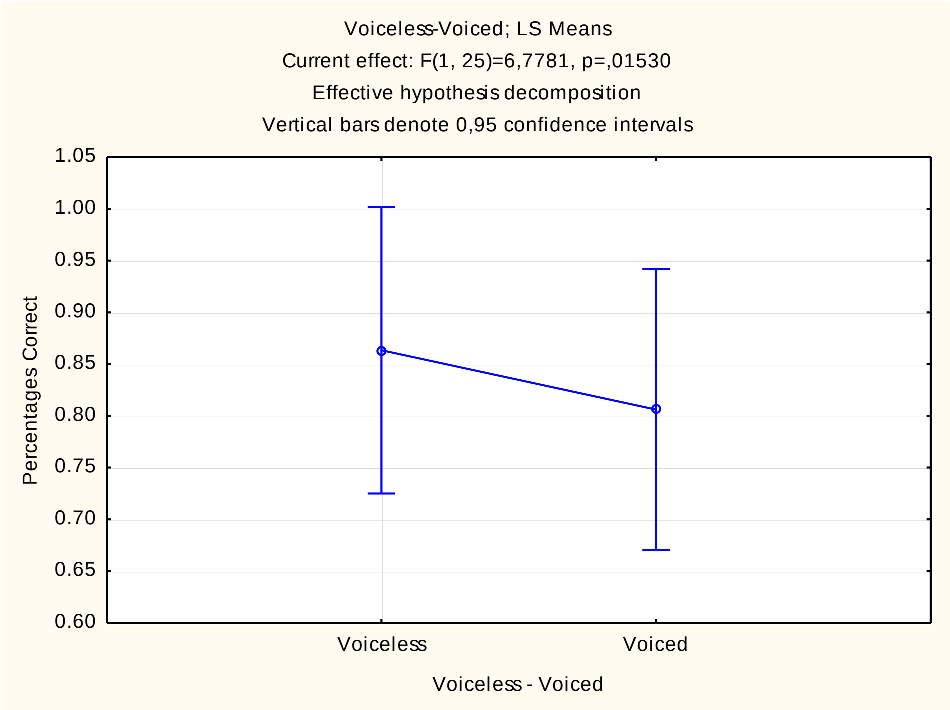

Figure 4 Percentages Correct: The Main Effect of Voicing

The main effect of Segment Type was significant (current effect: $F(4, 100) = 18.910$, $p = 0.000$) (see Figure 5). The graph shows the performance of all groups together. According to the results, the highest mean percentage of correct responses was for stimuli ending in /f/ – /v/ (87%). However, for the /f/ – /v/ contrast, the responses varied mostly significantly – the highest percentage of correct responses was 98%, while the lowest percentage was 76%. The differences among participants were the lowest for stimuli ending in /t/ – /d/ (80% was the lowest percentage correct, 90% was the highest percentage correct). The mean percentage correct for word-final /t/ – /d/ was 85%. The poorest performance was for the word-final /p/ – /b/ contrast (the mean percentage of correct responses was 75%).

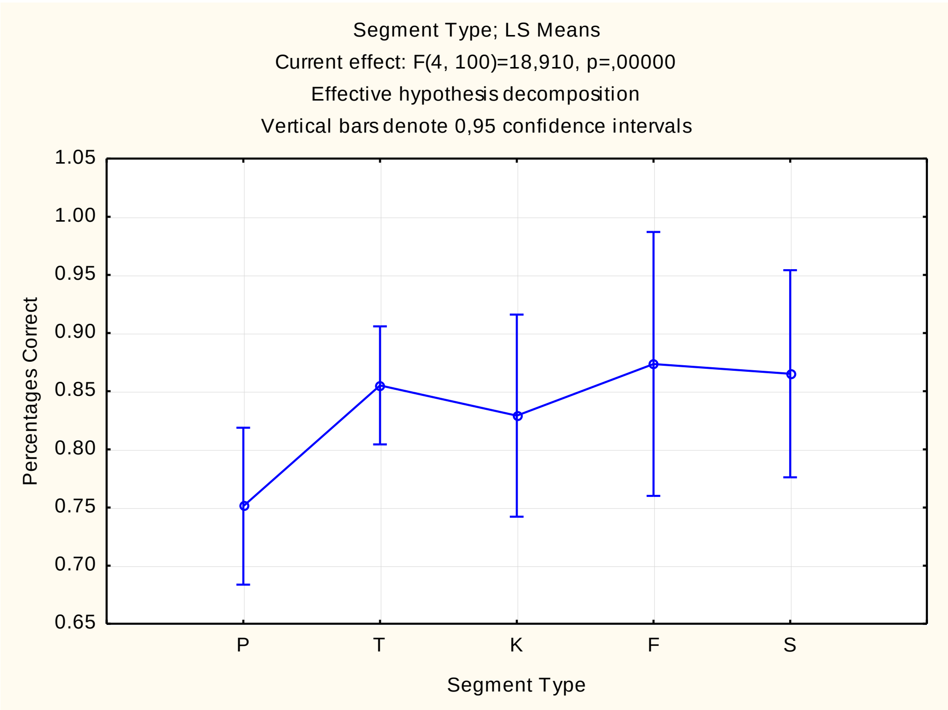

Figure 5 Percentages Correct: The Main Effect of Segment Type

In order to provide more detail, Figure 6 shows the interaction between Segment Type and Voicing. There appears to be a significant interaction between the two factors (current effect: $F(4, 100) = 3.7331$, $p = 0.00712$. Generally, the participants tended to identify the voiceless segments more accurately, which is in accordance with the results in Figure 4. There were, however, exceptions to the general tendency. First, it was the identification of the minimal pair /k/ – /g/. The mean percentage of correct responses for the voiceless word-final /k/ was 82%, while the mean percentage correct for the voiced word-final /g/ was 83%. Another minimal pair whose identification diverted from the general tendency was /s/ – /z/. The amount of correct responses for both segments was, again, very similar and none of the segments were identified with a higher accuracy than its counterpart – the mean percentage correct for /s/ was 87%, for /z/ it was 86%. The identification of /f/ – /v/ is also interesting. While the voiceless word-final /f/ was identified with a high degree of correctness (93% was the mean percentage correct for /f/, as opposed to 81% for /v/), the percentages correct for the voiced /v/ varied significantly from 66% to 96%.

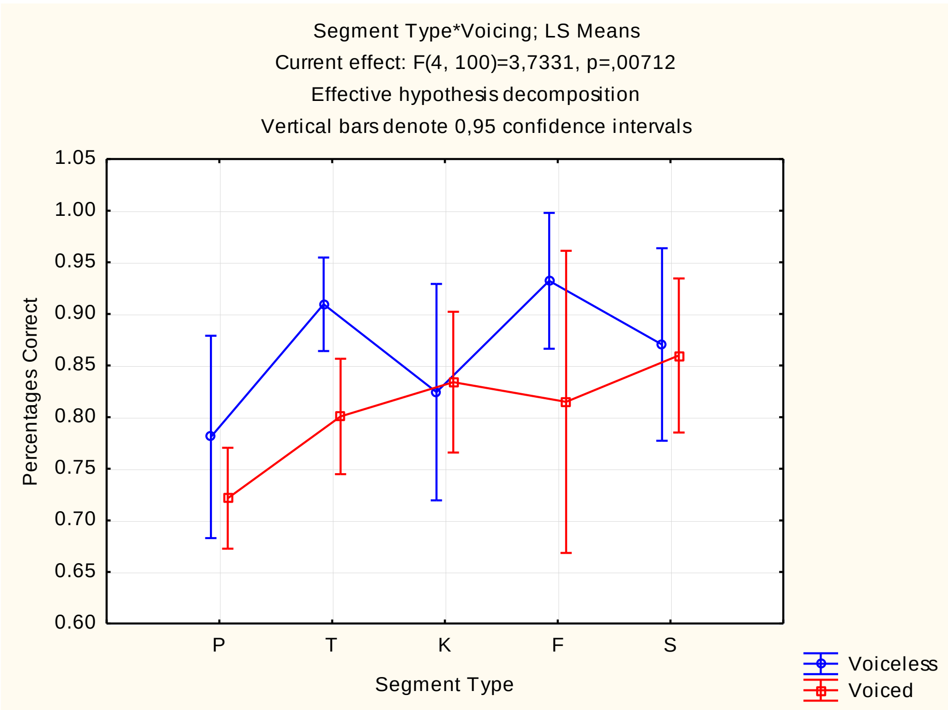

Figure 6 Percentages Correct: The Interaction Between Segment Type and Voicing

Although Figure 5 and Figure 6 both show that the Segment Type and its Voicing have a significant influence on the percentages of correct responses, they do not reflect progress of the groups throughout the training. The interaction between the Treatment Group, Segment Type and Test is shown in Figure 7. ANOVA revealed a significant effect for this interaction (current effect: $F(16, 200) = 1.9561$, $p = 0.01762$).

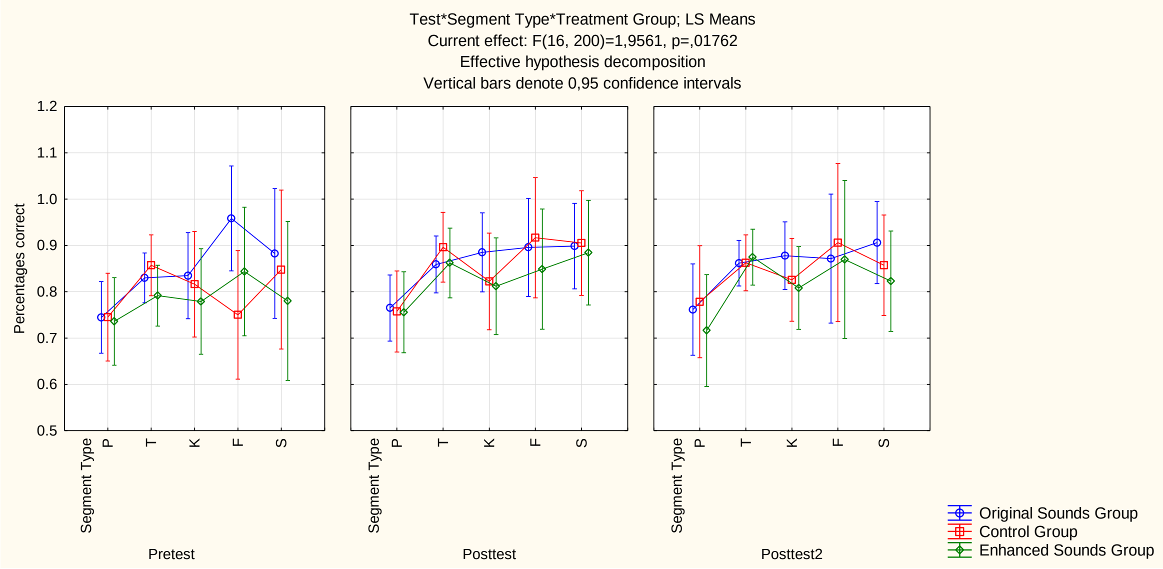

Figure 7 Percentages Correct: The Interaction Between Segment Type, Test and Treatment Group

For the minimal pair /p/ – /b/, the results were quite consistent. All groups performed similarly in its identification both in the pretest and in the first Posttest (the mean percentages correct were 74% for all groups in the Pretest and 75% in the Posttest). No progress in the identification of /p/ – /b/ can be seen for any of the groups. In the Posttest 2, the amount of correct responses of the Enhanced Sounds Group slightly decreased (to 71%).

The minimal pair /t/ – /d/ was most correctly identified by the Control Group in the Pretest and in the first Posttest. The Control Group's performance did not significantly improve from the Pretest (with the mean percentage correct being 85%) to the Posttest (the mean percentage correct was 89%), similarly as the Original Sounds Group (the mean percentages correct were 83% in the Pretest and 86% in the first Posttest). The Enhanced Sounds Group, however, performed better in the Posttest than in the Pretest. Their mean percentage of correct responses rose from 79% to 86%. Moreover, none of the groups' performance decreased in the Posttest 2.

As for the /k/ – /g/ minimal pair, a small improvement from the Pretest to the first Posttest was made by the Original Sounds Group (the mean percentage of correct responses rose from 83% to 89%) and the Enhanced Sounds Group (the increase was from 78% to 81% in the mean percentage of correct responses). The Control Group's responses got neither better, nor worse (the mean percentage correct was 81% in the Pretest and 82% in the Posttest). None of the groups performed worse in the Posttest 2.

The correct responses for the minimal pair /f/ – /v/ were varied in the Pretest. The Original Sounds Group's mean percentage of correct responses was 96%. The Enhanced Sounds Group's mean percentage of correct responses was much lower – it was 84%. The Control Group's mean percentage of correct responses was only 75%. In the first Posttest, however, the Control Group's performance improved to 92% correct. The Enhanced Sounds Group did not improve and their responses remained at similar levels as in the Pretest (85%). The Original Sounds Group performed slightly worse than in the Pretest (90%). None of the groups' results got worse in the Posttest 2.

The last minimal pair was /s/ – /z/. In its identification, the Original Sounds Group outperformed the other groups in the Pretest (88% was the mean amount of correct responses for the Enhanced Sounds Group, while the Control Group only got 78% as the mean rate and the Enhanced Sounds Group achieved 85%). The results improved after the training and in the two posttests, all groups performed similar.

3.2.Response Times

Apart from percentages correct, response times were measured during the experiment. The main effect of Treatment group was not found significant ($p > 0.05$). Figure 8 shows the response times of all the groups in all tests together.

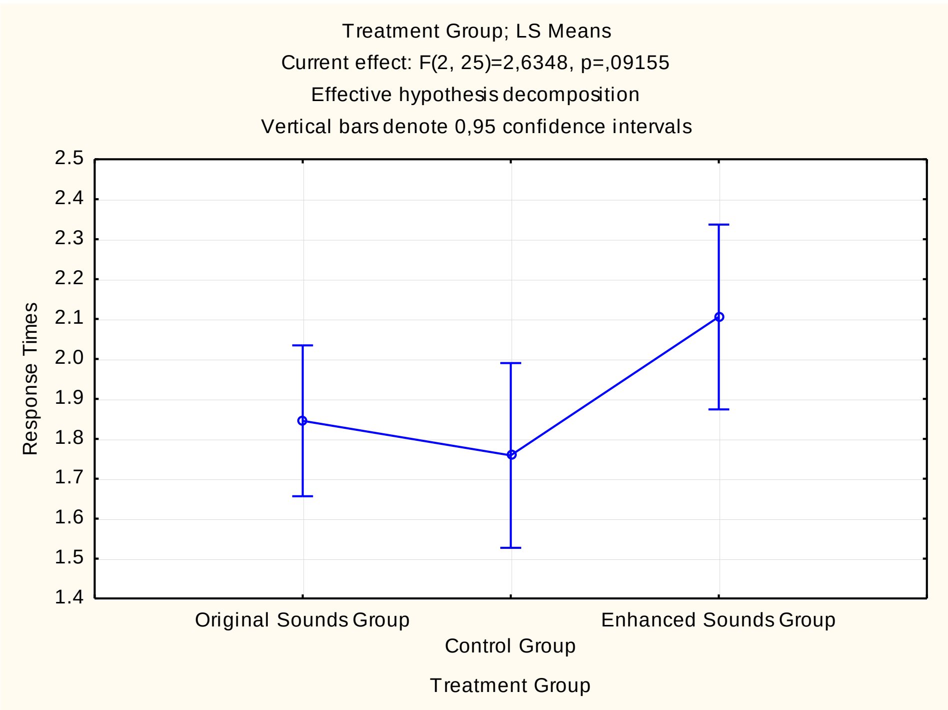

Figure 8 Response Times: The Main Effect of Treatment Group

The main effect of Test was, however, significant (current effect: $F(2, 50) = 13.405$, $p = 0.00002$). Figure 9 shows that overall, the response times were lower in the first Posttest than in the Pretest. In the Posttest 2, the response times increased. The mean response time in the Pretest was 2.18 s, while in the first Posttest, it was 1.69 s. In Posttest 2, the mean response time rose to 1.83 s.

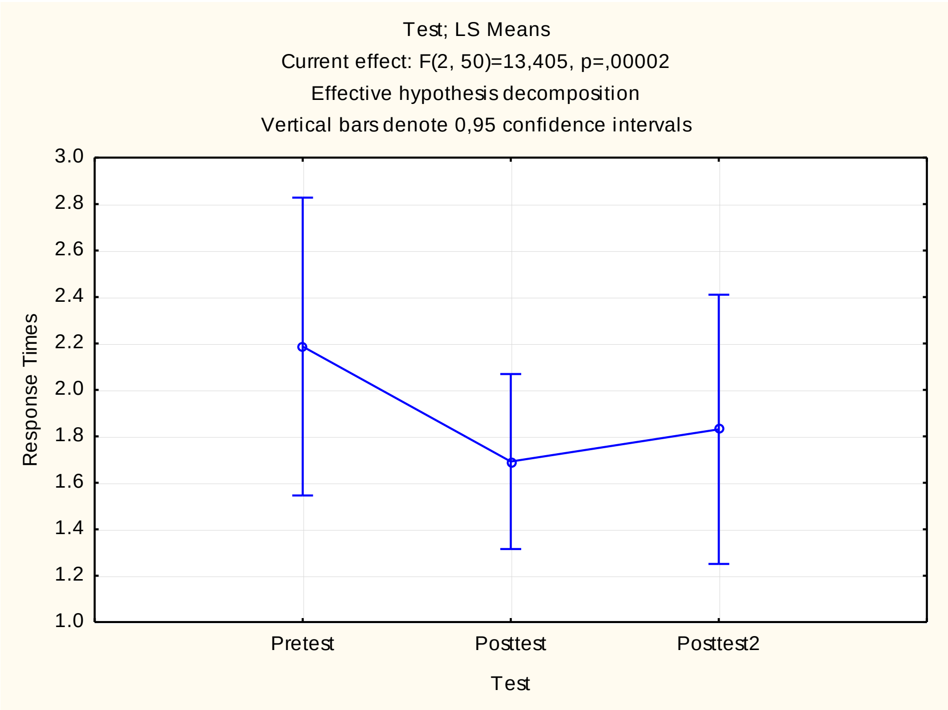

Figure 9 Response Times: The Main Effect of Test

Looking at the interaction between Treatment Group and Test, ANOVA did not find a significant effect ($p > 0.05$). Figure 10 shows response times for each group in each test, suggesting that none of the groups differed significantly from the other groups in their response times.

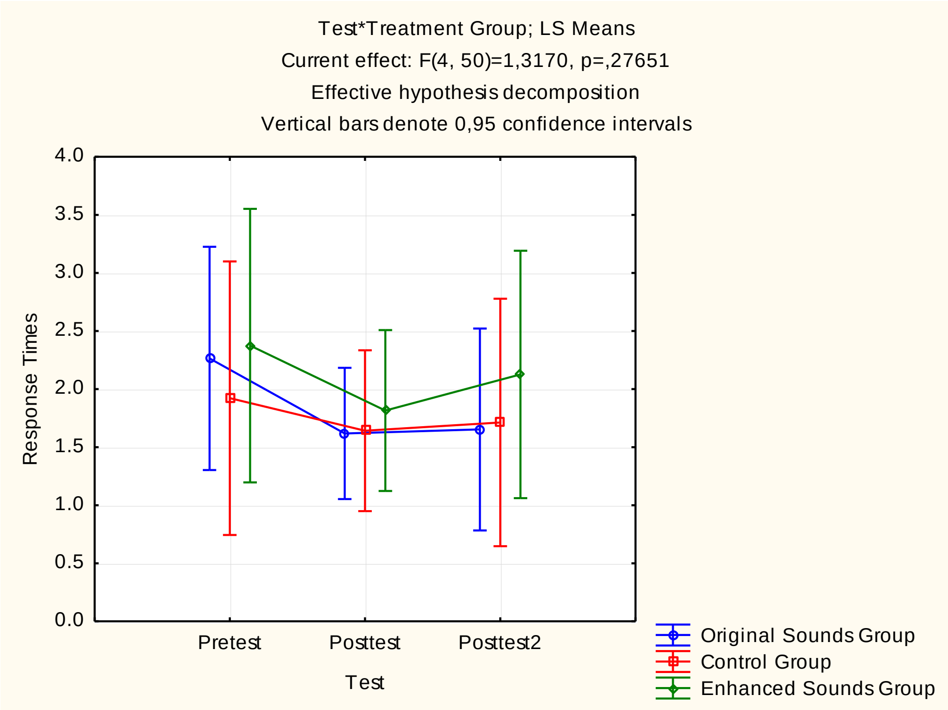

Figure 10 Response Times: The Interaction Between Treatment Group and Test

The main effect of phonological Voicing is shown in Figure 11. According to the ANOVA, the main effect was significant (current effect: $F(1, 25) = 7.3606$, $p = 0.01189$), although the effect was weak. The mean response time for stimuli with a voiceless obstruent in the final position was 1.9 s. Stimuli with a voiced obstruent in the final position had 1.86 s mean response time.

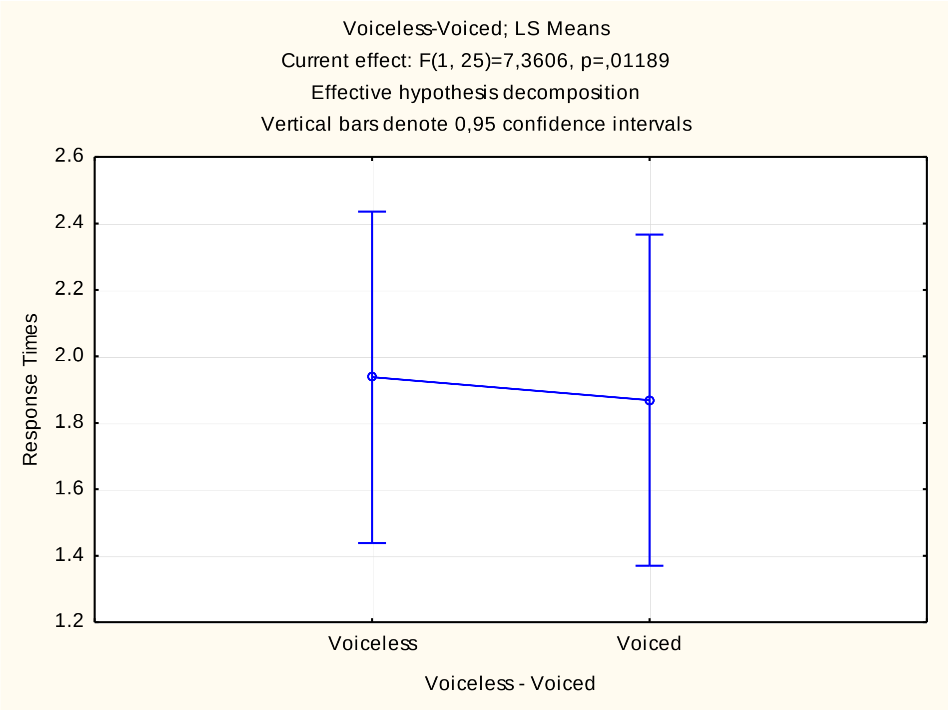

Figure 11 Response Times: The Main Effect of Voicing

Figure 12 shows a comparison of response times for the segments individually. The main effect of segment type was significant (current effect: $F(4, 100) = 11.499$, $p = 0.00000$). According to the analysis, the shortest mean response time, 1.79 s, was measured for the /k/ – /g/ minimal pair. The longest mean response time was 2.03 s and it was measured for the minimal pair /s/ – /z/. Overall, the differences in response times were large among the participants. The largest differences were measured for the /s/ – /z/ minimal pair, for which the shortest response time was 1.6 s and the longest 2.4 s. The smallest differences in response times among the participants were measured for /k/ – /g/, which was identified in the shortest time of 1.5 s and in the longest time of 2.08 s.

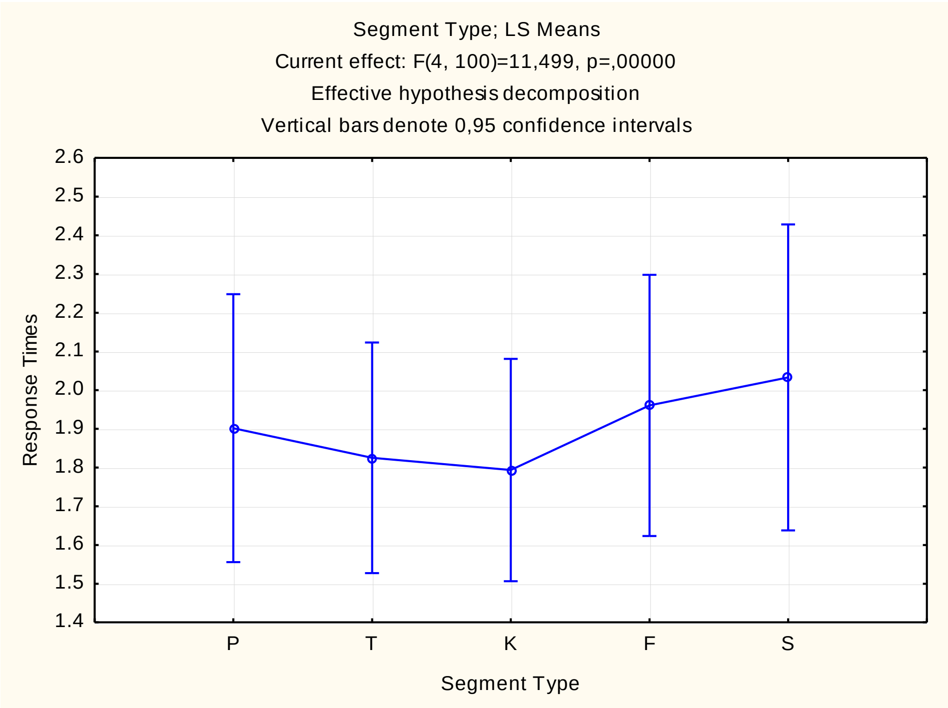

Figure 12 Response Times: The Main Effect of Segment Type

The main effect for the interaction between Treatment Group and Segment Type was significant (current effect: $F(8, 100) = 1.9753$, $p = 0.05718$) (see Figure 13). The response times vary, depending on the Treatment Group – generally, the response times are longer for the Enhanced Sounds Group than for the other two groups. The shortest response times were measured for the Control Group. The Original Sounds Group and the Control Group, however, performed very similar. Overall, the three groups had the longest response times for both pairs of fricatives. The /s/ – /z/ contrast took the participants the longest to identify (the mean response times were 1.8 s for the Control Group, 2.0 s for the Original Sounds Group and 2.3 s for the Enhanced Sounds Group).

The minimal pairs of stops were identified in shorter response times. The shortest response times were measured for /k/ – /g/ in all groups (the mean response times were 1.71 s for the Control Group and for the Original Sounds Group, 1.92 s for the Enhanced Sounds Group).

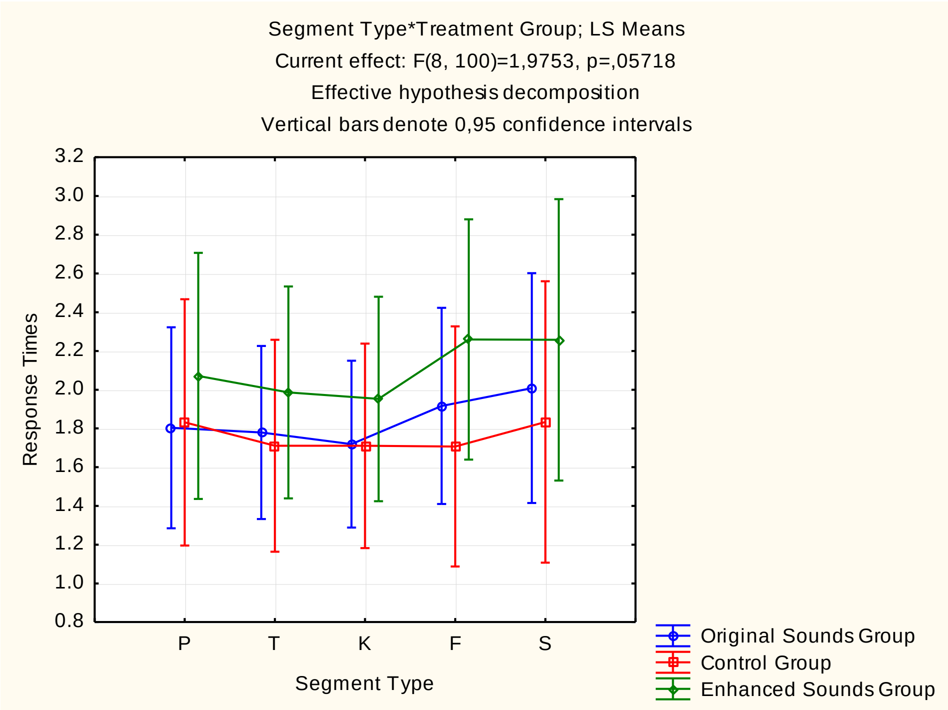

Figure 13 Response Times: The Interaction Between Segment Type and Treatment Group

Separate graphs in Figure 14 show the subjects' performance for each segment type in each test. The interaction amongst Treatment Group, Test and Segment Type was not found significant ($p > 0.05$), which suggests that none of the groups that received the training progressed after the training and neither did the Control Group.

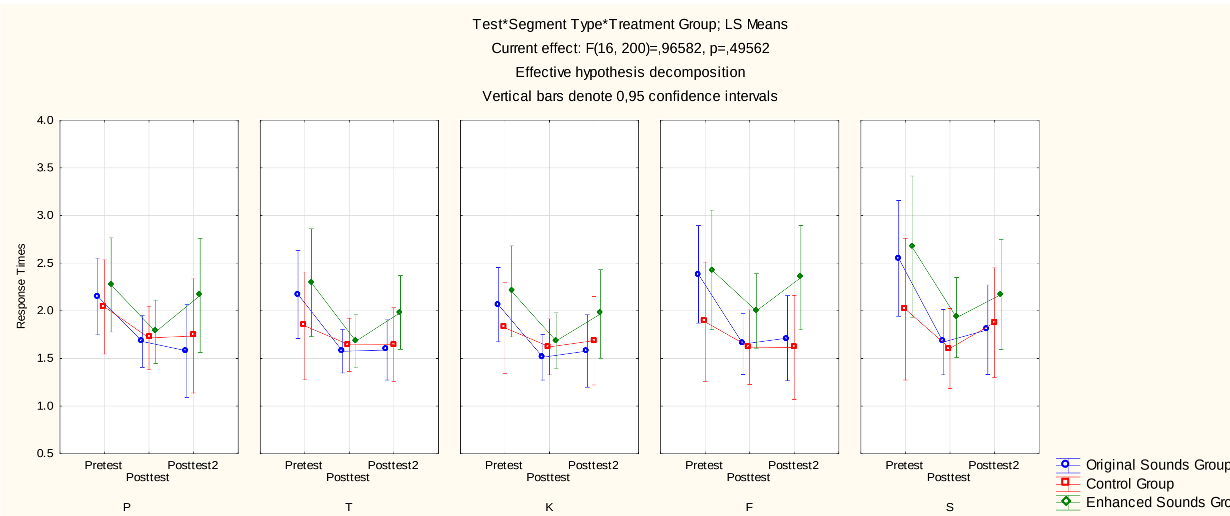

Figure 14 Response Times: The Interaction Between Test, Segment Type and Treatment Group

The present chapter introduced the results of the perceptual experiment. It was found out that neither the percentages correct, nor the response times improved after the training. Moreover, the results suggest that the Control Group outperformed the groups which received the training in several areas. The training, therefore, turned out to be unsuccessful. This result makes the hypothesis H_3 unanswerable – an improvement in production could not have taken place as no progress in perception was made. The results of the tests in production would, therefore, be insignificant and will thus not be presented in the thesis.

4. Discussion

The purpose of the present experiment was to train native Czech learners of English in the correct perception and production of English underlying word-final obstruent voicing. The results of the experiment will be discussed in the present chapter. They will be correlated with the hypotheses that were introduced in section 1.6 and they will provide answers to the research questions.

First, it was hypothesized (H_1) that because the participants were already proficient in English and they had previously attended a course in phonetics, they would already perform with high accuracy in the pretest. The hypothesis was based on results of pervious research (e.g. Dmitrieva, Jongman and Sereno 2010; Wang and Wu 2001; Skarnitzl and Šturm 2014) which suggest that the ability to perceive and produce non-native phonetic categories (or to rely on acoustic cues that are not relevant in their L1) improves with a rising degree of proficiency in L2. The progress should, therefore, take place even without explicit phonetic instruction.

The results of the present experiment indeed show that the subjects' ability to correctly identify English word-final voicing was high at the beginning of the experiment. According to the results, none of the groups' percentages of correct responses were smaller than 60%, which suggests that the participants always identified more than half of the stimuli correctly. The degree of proficiency in English is thus an important factor to be taken into consideration.

According to the second hypothesis (H_2), the use of high variability phonetic training with the identification task should have improved the subjects' performance in the experiment. As discussed in chapter 1.4.1.3 (High Variability Phonetic Training), the high variability phonetic training is believed to be an effective method for training perception in L2. At the same time, it has been suggested that the identification task is an efficient type of task for training (see chapter 1.4.1.1. Identification Task vs. Discrimination Task). The combination of the high variability phonetic training method with the identification task should have helped the subjects to improve in their L2, English.

However, the results suggest that although progress from the Pretest to the first Posttest took place (see Figure 2, Figure 9), it was only a general tendency. Figure 3 shows that when results for each Treatment Group were measured separately, no significant increase in the percentages of correct responses was seen in neither of the groups. Similarly, no significant improvement was measured for the response times (see Figure 10). Although the response times shortened generally (see Figure 9), none of the groups outperformed the others. A small decrease in the duration of response times was found for the Enhanced Sounds Group and the Original Sounds Group. The difference was, however, not significant and it cannot be concluded from the results that any of the two training programmes were successful.

Several factors may have caused the failure of the perceptual training. First of all, the experiment may have been too short. The completion of the training task took approximately 15 minutes and it did not take place repeatedly. Previous studies that used the high variability phonetic training and were successful (e.g. Logan, Lively, and Pisoni 1991; Lively et al. 1994; Flege 1995; Wong 2012; Wong 2013), employed a long-term training design, consisting of several training sessions in the course of many days / weeks. Perhaps a more intensive training, which would include more than one session, would be more effective.

Another factor that may have had an effect on the results of the experiment is the number of participants. Out of the original 36 participants, only 29 completed all parts of the training. Moreover, in the Enhanced Sounds Group, which received training with enhanced vowel duration and was, therefore, a key group of the experiment, only 9 subjects completed all parts of the training. Because the number of participants, who were trained on the enhanced stimuli, was very low, it is difficult to make generalizations about the usefulness of the training method that had been used.

Finally, generalization to natural stimuli after the training with enhanced tokens might have been difficult for the participants in the Enhanced Sounds Group. Perhaps gradual fading of the enhancement of the stimuli (see section 1.4.1.4 The Perceptual Fading Technique for more information about the method) could have been more effective. The subjects' attention would at first be drawn to a maximally enhanced vowel duration, which would gradually change towards its natural length.

It has already been suggested that the third hypothesis (H_3) cannot be proven. According to it, the participants would transfer their progress in perception to production. The fact that participants of perceptual training (both high variability phonetic training and the perceptual fading technique) transfer their progression in perception of L2 to production has been suggested by the results of several experiments (e.g. Jamieson and Rvachew 1992, 1994; Rvachew 1994; Rochet 1995; Bradlow et al. 1997). These studies propose that an improvement in perception is a prerequisite for progress in production. However, the present perceptual training was unsuccessful. The ANOVA did not find a significant interaction between the Treatment Group and the Test, neither in terms of percentages of correct responses, nor with respect to response times. As the prerequisite for the progress in production was not met, the effect of perceptual training on production could not have been measured and the experiment was, therefore, unsuccessful.

The fourth hypothesis (H_4) presumed that training with enhanced stimuli would focus the subjects' attention to the use of vowel duration as a cue to English word-final obstruent voicing. Results of previous research on Czech (e.g. Keating 1985, Sehnalíková 2010) demonstrate that native Czech speakers do not utilize different vowel duration as a cue to the voicing contrast. On the other hand, in English, it is believed to be the prominent cue to word-final obstruent voicing (e.g. Luce and Charles-Luce 1985). To teach the ability to use distinctive vowel duration before pairs of voiced – voiceless obstruents was, therefore, a key aim of the present experiment as it would make the subjects' perception and production of English more native-like. However, as suggested above, the Enhanced Sounds Group's performance did not change from the Pretest to the Posttest. These results indicate that the participants did not learn to use vowel duration as a cue to English word-final voicing and that the hypothesis H_4 has, therefore, not been proven.

The last hypothesis (H_5) concerned phonetic properties of the obstruents in word-final positions. It was assumed that the results would differ with respect to the characteristics of the word-final obstruent in the stimuli. It has indeed been found out that the subjects' performance is dependent on these properties. First of all, a significant main effect was found for phonological voicing of the final obstruent. Altogether, the mean percentage of correct responses was higher for voiceless obstruents (86%) than for voiced obstruents (80%). The results are consistent with the findings of the experiment by Sehnalíková (2010), who obtained a similar ratio between voiced and voiceless obstruents for the group of native Czech learners of English (77% of correct responses for voiceless coda and 68% for voiced coda). Apart from that, Sehnalíková (2010) tested native Czech speakers with no knowledge of English and the results she obtained suggest that only 32% of the responses for voiced coda were correct, while the percentage of correct responses for voiceless coda was 72%.

According to the results of both studies, native Czech speakers tend to identify English word-final voiceless obstruents with a relative ease, which might be due to the fact that in Czech, the voicing contrast is neutralized in this position and Czech speakers are, therefore, familiar with obstruents being voiceless word-finally. At the same time, the percentage of correct responses seems to increase with the level of proficiency in English as the groups of Czech learners of English in the present experiment and in the experiment by Sehnalíková (2010) all performed much better in identifying English word-final voiced obstruents than the monolingual Czech speakers in Sehnalíková's (2010) experiment. This partly confirms the hypothesis H_1 of the present study that the degree of proficiency in L2 affects the subjects' performance.

Secondly, a question, whether certain minimal pairs of obstruents were identified more easily than the others, arose. It was also studied, which minimal pairs posed a difficulty to the subjects. Native Czech speakers are familiar with all obstruents that were presented to them in the experiment from their native language. They should have, therefore, identified all word-final obstruents with equal accuracy. However, minimal pairs of word-final fricatives were generally identified more accurately than minimal pairs of word-final stops. Specifically, the highest percentages correct were measured for the minimal pair of labiodental fricatives /f/ – /v/. The lowest percentages of correct responses were calculated for the bilabial stops /p/ – /b/. The alveolar stops /t/ – /d/ and the alveolar fricatives /s/ – /z/ were recognized with a similar accuracy. The amount of correct responses for the minimal pair of velar stops /k/ – /g/ was similar to the number of correct responses for the alveolar obstruents.

Finally, based on the hypothesis H_5, minimal pairs of such obstruents, whose identification was the poorest in the Pretest, should have been classified with more accuracy after the training. According to the results in Figure 7, some improvement in the amount of percentages of correct responses was achieved by the groups that received the training (but also by the Control Group). The progress was reported for the minimal pairs of /t/ – /d/, /k/ – /g/ and /s/ – /z/ and it was in most times achieved by the Enhanced Sounds Group. Nevertheless, the percentages correct did not rise for the minimal pair /p/ – /b/ in any of the groups.

The hypothesis H_5 was not fully confirmed. Although the subjects' performance was dependent on phonological voicing and on the characteristics of the obstruents, no consistent improvement was made. The two groups that received the training became more accurate in the identification of only some segments. In some cases, the Control Group progressed too. It is, therefore, not possible to make generalizations from the results.

Overall, out of the five hypotheses that have been formulated in section 1.6, only the hypotheses H_1 (concerning the role of the degree of proficiency in L2) and partly H_5 (about the subjects' reactions to different segment types) have been confirmed. The other hypotheses were generally concerned with the effectiveness of the present training and they were not confirmed. The training turned out to be unsuccessful and suggestions for future research were provided.

5. Conclusion

The aim of the present thesis was to contribute to the research concerning training L2 learners in the acquisition of the correct acoustic cues to non-native contrasting categories. Specifically, the use of contrastive vowel duration as a cue to English word-final obstruent voicing by Czech learners of English was the subject of the paper.

In Czech, the voicing contrast between obstruents is neutralized (see 1.1., 1.2.) and vowel duration is not used as an acoustic cue to it in such an extent as it is in English (see 1.3.1.3.). In order to make the perception and production of Czech learners of English more native-like, a training experiment aimed at English word-final voicing was designed. Out of the perceptual training methods that have been put in practice by other researchers, the high variability phonetic training using a two-alternative identification task was chosen as it had yielded successful results before (see 1.4.). On the top of that, it was suggested that if the perceptual training is successful, progress in production takes place (see 1.5.).

Advanced Czech learners of English were chosen to participate in the experiment (see 2.1.). They were divided into three groups – two training groups (Original Sounds Group, Enhanced Sounds Group) and a control group. Each of the two groups that received training got a task with the same design, differing in the stimuli used (see 2.2.).

The results of the perceptual experiment suggest that the training was not successful. The groups that received the training performed similarly as the control group. Generally, the percentages of correct responses did not rise after the training. At the same time, the response times did not change significantly from the pretest to the posttest 1 in neither of the trained groups. The prerequisite of the progress in production was successful perceptual training, which was not achieved. It was, therefore, concluded that no change in the participants' production of English word-final voicing took place.

It was concluded that the training might not have been productive for several reasons (see chapter 4). The duration and intensity of the training were considered the most important factor that could have affected the results. Perhaps a long-term training programme could have been more efficient. The small number of participants and the choice of the training method might have also had impact on the results.

Although the training was not as successful as predicted, it has been found out that the subjects' performance differs depending on the nature of the stimuli – some obstruents turned out to be more easily recognized than others. The results also suggest that Czech learners of English tend to identify word-final voiceless obstruents better than their voiced counterparts. In addition to that, the results indicate that due to their high level of proficiency in English, the subjects generally performed well in the experiment as the amount of correct responses never dropped under 60%.

The results of the present experiment should serve as groundwork for future research. Specifically, it should be studied whether the training programme would yield different results if it was carried out in a long-term process. At the same time, it would be interesting to include several groups of ESL[4] learners with different degrees of proficiency. Lastly, perhaps a choice of more than one training method could be done in the future in order to assess their efficiency.

[4] ESL = English as a Second Language

Shrnutí

Tato práce je zaměřena na nácvik percepce a produkce konsonantické znělosti na koncích slov u českých studentů angličtiny. Fonologické inventáře angličtiny a češtiny obsahují dvojice fonémů odlišující se příznakem znělosti. V každém jazyce ovšem platí jiná pravidla pro realizaci konsonantické znělosti. Na koncích slov před pauzou v češtině dochází k neutralizaci znělostního protikladu a všechny obstruenty se realizují jako neznělé – slova jako *plot* a *plod* se vyslovují identicky jako [plot] (Palková 1994, 132). Ke znělé výslovnosti konsonantů na koncích slov dochází v případě, že po slově nenásleduje pauza, ale nové slovo s počátečním znělým konsonantem – v takových případech hovoříme o regresivní asimilaci znělosti (Skarnitzl 2011, 123).

V angličtině je znělostní asimilace spíše výjimečná a týká se pouze frikativ v gramatických slovech (Cruttenden 2008, 299). Na rozdíl od češtiny navíc v angličtině nedochází k neutralizaci znělosti na koncích slov. Znělost konsonantů v těchto pozicích není plná tak jako například mezi dvěma vokály a realizace znělostního protikladu je signalizována především délkou předchozího vokálu. Vokály předcházející neznělé konsonanty jsou v angličtině významně kratší než vokály před znělými konsonanty (např. Chen 1970).

Na základě výsledků řady studií bylo zjištěno, že délka předchozích vokálů je primárním signálem rozdílné znělosti následujících konsonantů v angličtině (např. Hillenbrand et al. 1984; Luce a Charles-Luce 1985). Výsledky dalších výzkumů naznačují, že v jiných jazycích než v angličtině se délka vokálů před konsonanty ve znělostních opozicích neliší v tak velké míře (např. Chen 1970; Broersma 2010; Slowiaczek a Dinnsen 1985; Pye 1986; Crowther a Mann 1994). K podobným zjištěním vedly studie zaměřené na češtinu (např. Keating 1985; Machač a Skarnitzl 2007; Šimáčková 2003; Podlipský 2008).

Jelikož v angličtině dochází k realizaci fonologického znělostního protikladu konsonantů na koncích slov, zatímco v češtině je znělostní protiklad v této pozici neutralizován a délka samohlásky v češtině není užívána jako primární signál znělosti, zatímco v angličtině tomu tak je, Sehnalíková (2010) provedla studii produkce a percepce anglické znělosti na koncích slov českými mluvčími. Na základě výsledků její práce dochází v produkci k neutralizaci znělosti na koncích slov. V percepčních testech došlo k přesnější identifikaci neznělých koncových konsonantů než jejich znělých protějšků, a to jak u českých mluvčích bez znalosti angličtiny, tak i u pokročilých studentů anglického jazyka.

Na základě poznatků z výše citovaných studií bylo usouzeno, že je třeba konsonantickou znělost na koncích slov české studenty angličtiny naučit správně používat. Pro nácvik výslovnosti existuje řada metod, např. *low variability phonetic training, high variability phonetic training, the perceptual fading technique*. Pro tuto práci byla vybrána metoda *high variability phonetic training*, která je specifická tím, že studenti poslouchají přirozenou (ne digitalizovanou) řeč, produkovanou více než jedním rodilým mluvčím. Pro trénink byla vybrána úloha *identification task,* v níž studenti vybírají ze dvou variant zvuk, který slyší. Užitím dané metody se má zlepšit jak percepce, tak produkce trénovaného jevu.

Efekt, jež percepční metody nácviku výslovnosti mají na percepci, byl sledován například ve studiích zabývajících se odstraněním výslovnostních problémů u dětí (Jamieson a Rvachew 1992, 1994; Rvachew 1994) nebo nácvikem výslovnosti /r/ – /l/ u japonských studentů angličtiny (Bradlow et al. 1997). Bylo zjištěno, že percepční trénink zlepšuje jak percepci, tak produkci účastníků.

Nácviku výslovnosti v této práci se zúčastnilo 36 studentů angličtiny. Všichni hovořili plynule anglicky a studovali angličtinu na univerzitě. Studenti byli rozděleni do tří skupin. První skupina (Enhanced Sounds Group) byla trénována na slovech, v nichž byla upravena délka vokálu, délka závěru ve finálním konsonantu a frekvence. Úpravou slov se chtělo docílit toho, že se pozornost studentů přesune na délku samohlásky předcházející finálnímu konsonantu a jejich percepce a produkce se tak stane bližší rodilé anglické. Druhá skupina (Original Sounds Group) byla trénována na slovech, která byla zanechána v původní podobě tak, jak byla vyslovena rodilými mluvčími. Poslední skupina (Control Group) se neúčastnila nácviku, ale pouze kontrolních testů. Tato skupina byla zařazena pro srovnání účinnosti tréninku. Z každé skupiny bylo navíc vybráno šest studentů pro testování produkce.

Slova, která účastníci výzkumu poslouchali, tvořila minimální páry skutečných anglických jednoslabičných slov, lišících se ve znělosti finálního konsonantu. Do percepčních testů bylo zahrnuto 58 minimálních párů slov, pro testování produkce bylo vybráno 21 minimálních párů.

Nácvik se skládal z následujících částí: pretest, nácvik, posttest 1, posttest 2. Pretest a oba posttesty byly identické pro všechny skupiny a účastníky a byly vytvořeny v Praatu (Boersma, Weenink 2014). Studenti při nich poslouchali digitálně neupravená slova a určovali správné odpovědi na základě *identification task*. Nácvik probíhal stejně v obou skupinách (Enhanced Sounds Group, Original Sounds Group) a byl stejně jako testy vytvořen v Praatu. Studenti vybírali odpovědi na základě *identification task*. Na rozdíl od testů ovšem byli ihned po zodpovězení upozorněni na to, zda je odpověď správná, nebo ne. Pokud byla odpověď špatná, minimální pár nebyl vymazán z tréninku, ale objevoval se opakovaně, dokud student nezačal daná slova identifikovat správně.

Výsledky percepčních testů neukázaly pokrok ani jedné ze skupin, a to ani v počtu správných odpovědí, ani ve změně reakčních časů. Nebylo tedy dokázáno, že užitá metoda nácviku byla užitečná. Problémem mohlo být, že nácvik byl příliš krátký a účastnila se jej pouze malá skupina studentů. Je tedy možné, že za jiných časových podmínek, by daná metoda mohla splnit svůj cíl.

Testování produkce proběhlo ve dvou částech, a to ve formě pretestu a posttestu. Oba testy měly identický průběh. Studenti byli individuálně zváni do zvukotěsné místnosti, kde byly pořizovány nahrávky jejich čtení seznamu testovaných slov. Jelikož smyslem celého výzkumu bylo zhodnocení účinku metody nácviku percepce, která nedosáhla svého cíle, bylo nemožné sledovat pokrok v produkci. Výsledky produkčních testů tudíž nebyly zveřejněny.

Ačkoli výzkum týkající se nácviku nebyl úspěšný, podařilo se potvrdit některé údaje, zjištěné v dřívějších studiích. Bylo tak například potvrzeno, že úroveň pokročilosti v cizím jazyce ovlivňuje studentovu percepci, a tudíž i jeho výkon v testování. Počet správných odpovědí v žádné ze skupin neklesl pod 60%, vždy tedy byla identifikována více než polovina slov správně. Mimoto výsledky prokazují, že čeští studenti lépe identifikují neznělé konsonanty na koncích slov. Toto zjištění replikuje výsledky z výzkumu Sehnalíkové (2010). Téma nácviku percepce a produkce konsonantické znělosti na koncích slov českými studenty angličtiny je důležitou oblastí výzkumu, které je třeba nadále věnovat pozornost.

References

Boersma, Paul, and David Weenink. 2014. Praat: Doing Phonetics by Computer [Computer Program]. Retrieved May 15, 2014, from htttp://www.praat.org/.

Bradlow, Ann R., David B. Pisoni, Reiko A. Yamada, and Yoh'ichi Tohkura. 1997. "Training Japanese Listeners to Identify English /r/ and /l/: IV. Some Effects of Perceptual Learning on Speech Production." *Journal of the Acoustical Society of America* 101 (4): 2299–2310.

Bradlow, Ann R., David B. Pisoni, Reiko A. Yamada, and Yoh'ichi Tohkura. 1999. "Training Japanese Listeners to Identify English /r/ and /l/: Long-term Retention of Learning in Perception and Production." *Perception & Psychophysics* 61 (5): 977–985.

Bradlow, Ann R. 2007. "Training Non-native Language Sound Patterns: Lessons from Training Japanese Adults on the English /r/-/l/ Contrast." In *Phonology and Second Language Acquisition*, edited by Jette G. Hansen Edwards, and Mary L. Zampini, 287–308. Amsterdam: John Benjamins Publishing Company.

Broersma, Mirjam. 2005. "Perception of Familiar Contrasts in Unfamiliar Position." *Journal of the Acoustical Society of America* 117: 3890–3901.

Broersma, Mirjam. 2008. "Flexible Cue Use in Nonnative Phonetic Categorization." *Journal of the Acoustical Society of America* 124: 712–715.

Broersma, Mirjam. 2010. "Perception of Final Fricative Voicing: Native and Non-native Listeners' Use of Vowel Duration." *Journal of the Acoustical Society of America* 127: 1636–1644.

Carlet, Angelica, and Juli Cebrian. 2015. "Identification vs. Discrimination Training: Learning Effects for Trained and Untrained Sounds." *Proceedings of the 18th ICPhS.*

Charles-Luce, Jan. 1985. Word-final Devoicing in German: Effects of Phonetic and Sentential Contexts." *Journal of Phonetics* 15: 187–190.

Chen, Matthew. 1970. "Vowel Length Variation as a Function of the Voicing of the Consonant Environment." *Phonetica* 22: 129–159.

Crowther, Court S., and Virginia A. Mann. 1992. "Native Language Factors Affecting Use of Vocalic Cues to Final Consonant Voicing in English." *Journal of the Acoustical Society of America* 92: 711–722.

Crowther, Court S., and Virginia A. Mann. 1994. "Use of Vocalic Cues to Consonant Voicing and Native Language Background: The Influence of Experimental Design." *Perception & Psychophysics* 55 (5): 513–525.

Cruttenden, Alan. 2001. *Gimson's Pronunciation of English*. 6th edition. London: Arnold.

Cruttenden, Alan. 2008. *Gimson's Pronunciation of English*. 7th edition. London: Hodder Education.

Cubberley, Paul. 2002. *Russian: A Linguistic Introduction*. Cambridge: Cambridge University Press.

Dinnsen, Daniel A., and Jan Charles-Luce. 1984. "Phonological Neutralization, Phonetic Implementation and Individual Differences." *Journal of Phonetics* 12: 49–60.

Dmitrieva, Olga, Allard Jongman, and Joan Sereno. 2010. "Phonological Neutralization by Native and Non-native Speakers: The Case of Russian Final Devoicing." *Journal of Phonetics* 38: 483–492.

Flege, James E. 1989. "Chinese Subjets' Perception of the Word-final English /t/-/d/ Contrast: Performance Before and After Training." *Journal of the Acoustical Society of America* 86 (5): 1684–1697.

Flege, James E. 1995. "Two Procedures for Training a Novel Second Language Phonetic Contrast." *Applied Psycholinguistics* 16: 425–442.

Francis, Alexander L., Kate Baldwin, and Howard C. Nusbaum. 2000. "Effects of Training on Attention to Acoustic Cues." *Perception & Psychophysics* 62 (8): 1668–1680.

Gordon, Leslie S. 2008. "Factors Affecting English Speakers' Perception of L2 Spanish Vowels." PhD diss., Georgetown University.

Hillenbrand, James, Dennis R. Ingrisano, Bruce L. Smith, and James E. Flege. 1984. "Perception of the Voiced–Voiceless Contrast in Syllable-final Stops." *Journal of the Acoustical Society of America* 76: 18–26.

House, Arthur S., and Grant Fairbanks. 1953. “The Influence of Consonant Environment upon the Secondary Acoustical Characteristics of Vowels.” *Journal of the Acoustical Society of America* 25 (1): 105–113.

Hualde, Jose I. 1992. *Catalan*. Oxon: Routledge.

Ingvalson, Erin M., Marc Ettlinger, and Patrick C. M. Wong. 2014. “Bilingual Speech and Learning: A Review of Recent Trends.” *International Journal of Bilingualism* 18 (1): 35–47.

Iverson Paul, Valerie Hazan, and Kerry Bannister. 2005. “Phonetic Training with Acoustic Cue Manipulations: A Comparison of Methods for Teaching English /r/-/l/ to Japanese adults.” *Journal of the Acoustical Society of America* 118: 3267–3278.

Iverson, Paul, and Bronwen G. Evans. 2009. “Learning English Vowels with Different First-language Systems II: Auditory Training for Native Spanish and German Speakers.” *Journal of the Acoustical Society of America* 126 (2): 866–877.

Jamieson, Donald G., and David E. Morosan. 1986. “Training Non-native Speech Contrasts in Adults: Acquisition of the English /ð/–/θ/ Contrast by Francophones.” *Perception & Psychophysics* 40: 205–215.

Jamieson, Donald G., and Susan Rvachew. 1992. “Remediating Speech Production Errors with Sound Identification Training.” *Journal of Speech-Language Pathology and Audiology*: 201–210.

Jamieson, Donald G., and Susan Rvachew. 1994. “Perception, Production and Training of New Consonant Contrasts in Children with Articulation Disorders.” Paper presented at the Third International Conference on Spoken Language Processing, Yokohama, Japan, September 18–22.

Keating, Patricia A. 1985. “Universal Phonetics and Organizations of Grammars.” In *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*, edited by Victoria A. Fromkin, 115–132. Orlando: Academic Press.

Kluender Keith R., Randy L. Diehl, and Beverly A. Wright. 1988. “Vowel-length Differences Before Voiced and Voiceless Consonants: An Auditory Explanation.” *Journal of Phonetics* 16: 153–169.

Kuzla, Claudia, and Tomáš Duběda. “Incomplete Neutralization and Prosodic Conditioning of Voice Assimilation in Czech.” Paper presented at the Consonant Clusters and Structural Complexity Workshop, Institute of Phonetics and Speech Processing, Ludwig-Maxmilians-University, Munich.

Ladefoged Peter. 2001. *A Course in Phonetics*. 4th edition. Boston: Heinle & Heinle.

Lisker, Leigh. 1986. “‘Voicing’ in English: A Catalogue of Acoustic Features Signaling /b/ versus /p/ in Trochees.” *Language and Speech* 29 (1): 3–11.

Lively, Scott E., David B. Pisoni, Reiko A. Yamada, Yoh’ichi Tohkura, and Tsuneo Yamada. 1994. “Training Japenese Listeners to Identify English /r/ and /l/. III. Long-term Retention of New Phonetic Categories.” *Journal of the Acoustical Society of America* 96, 2076–2087.

Logan, John S., Scott E. Lively, and David B. Pisoni. 1991. “Training Japanese Listeners to Identify English /r/ and /l/: A First Report.” *Journal of the Acoustical Society of America* 89: 874–886.

Logan, John S., and John S. Pruitt. “Methodological Issues in Training Listeners to Perceive Non-native Phonemes.” In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by Winifred Strange, 351–377. Timonium: York Press.

Luce, Paul A., and Jan Charles-Luce. 1985. “Contextual Effects on Vowel Duration, Closure Duration, and the Consonant/Vowel Ratio in Speech Production.” *Journal of the Acoustical Society of America* 78: 1949–1957.

Machač, Pavel, and Radek Skarnitzl. 2007. “Temporal Compensation in Czech?” *Proceedings of 16th ICPhS*: 537–540.

McClaskey, Cynthia L., David B. Pisoni, and Thomas D. Carrell. 1983. “Transfer of Training of a New Linguistic Contrast in Voicing.” *Perception & Psychophysics* 34 (4): 323–330.

Ostaszewska, Danuta, and Jolanta Tambor. 2000. *Fonetyka i fonologia wspolczesnego jezyka polskiego.* Warszawa: PWN.

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

Peterson, Gordon E., and Ilse Lehiste. 1960. “Duration of Syllable Nuclei in English.” *Journal of the Acoustical Society of America* 32 (6): 693–703.

Pisoni, David B., Richard N. Aslin, Alan J. Percy, and Beth L. Hennessy. “Some Effects of Laboratory Training on Identification and Discrimination of Voicing Contrasts in Stop Consonants.” *Journal of Experimental Psychology: Human Perception and Performance* 8 (2): 297–314.

Podlipský, Václav J. 2008. “Vowel Duration as a Cue for Coda Voicing and the Perception of Second-language Vowel Quantity.” *Proceedings of the Acoustics’08 Paris Conference,* 6153–6157.

Podlipský, Václav J., and Kateřina Chládková. 2007. “Vowel Duration Variation Induced by Coda Consonant Voicing and Perceptual Short/Long Vowel Categorization in Czech.” In *Speech Processing*, edited by Robert Vích, 68–74. Praha: Institute of Photonics and Electronics AV CR.

Port, Robert, and Michael O’Dell. 1985. “Neutralization and Syllable-final Voicing in German.” *Journal of Phonetics* 13: 455–471.

Pruitt, John S. 1995. “The Perception of Hindi Dental and Retroflex Stop Consonants by Native Speakers of Japanese and American English.” PhD diss., University of South Florida.

Pruitt, John S., James J. Jenkins, and Winifred Strange. 2006. “Training the Perception of Hindi Dental and Retroflex Stops by Native Speakers of American English and Japanese.” *Journal of the Acoustical Society of America* 119 (3): 1684–1696.

Pye, Shizuka. 1986. “Word-final Devoicing of Obstruents in Russian.” *Cambridge Papers in Phonetics and Experimental Linguistics* 5: 1–10.

Raphael, Lawrence J. 1971. “Preceding Vowel Duration as a Cue to the Perception of the Voicing Characteristic of Word-final consonants in American English.” *Journal of the Acoustical Society of America* 51: 1296–1303.

Raphael, Lawrence J. 2005. “Acoustic Cues to the Perception of Segmental Phonemes.” In *Handbook of Speech Perception,* edited by David B. Pisoni, and Robert Remez, 182–206. Oxford: Blackwell.

Roach, Peter. 1998. *English Phonetics and Phonology*. 2nd edition. Cambridge: Cambridge University Press.

Rochet, Bernard L. 1995. “Perception and Production of Second-language Speech Sounds by Adults.” In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by Winifred Strange, 379–410. Timonium: York Press.

Russ, Charles V. J. 2010. *The Sounds of German*. Cambridge: Cambridge University Press.

Rvachew, Susan. 1994. “Speech Perception Training Can Facilitate Sound Production Learning.” *Journal of Speech and Hearing Research* 37: 347–357.

Sehnalíková, Eva. 2010. “Perception of Word-final Voicing in English and Czech.” MA thesis, Univerzita Palackého v Olomouci.

Sheldon, Amy, and Winifred Strange. 1982. “The Acquisition of /r/ and /l/ by Japanese Learners of English: Evidence that Speech Production Can Precede Speech Perception.” *Applied Psycholinguistics* 3 (3): 243–261.

Skarnitzl, Radek. 2011. *Znělostní kontrast nejen v češtině*. Praha: Epocha.

Skarnitzl, Radek, and Pavel Šturm. 2014. “Assimilation of Voicing in Czech Speakers of English: The Effect of the Degree of Accentedness.” *Research in Language* 12: 199–208.

Slowiaczek, Louisa M., and Daniel Dinnsen. 1985. “On the Neutralizing Status of Polish Word-final Devoicing.” *Journal of Phonetics* 13: 325–341.

Slowiaczek, Louisa M., and Helena Szymanska. 1989. “Perception of Word-final Devoicing in Polish.” *Journal of Phonetics* 17: 205–212.

Smith, Bruce L., Rachel Hayes-Harb, Michael Bruss, and Amy Harker. 2009. “Production and Perception of Voicing and Devoicing in Similar German and English Word Pairs by Native Speakers of German.” *Journal of Phonetics* 37 (3): 257–275.

Sokolovic-Perovic, Mirjana. 2009. “Voicing-Conditioned Vowel Duration in Southern Serbian.” *Newcastle Working Papers in Linguistics* 15: 126–137.

Strange, Winifred, and Sibylla Dittmann. 1984. "Effects of Discrimination Training on the Perception of /r-l/ by Japanese Adults Learning English." *Perception & Psychophysics* 36 (2): 131–145.

Sussex, Roland, and Paul Cubberley. 2006. *The Slavic Languages*. Cambridge: Cambridge University Press.

Šimáčková, Šárka. 2003. "'Engela's Eshes': Cross-linguistic Perception and Production of English [æ] and [ɛ] by Czech EFL Learners Trained in Phonetics." In *Proceedings of the 15th International Congress of Phonetic Sciences*, 2293–2296. Barcelona: UAB.

Wang, William S-Y. 1959. "Transition and Release as Perceptual Cues for Final Plosives." *Journal of Speech, Language, and Hearing Research* 2 (1): 66–73.

Wang, Hsu S., and Jui-Chun Wu. 2001. "The Effect of Vowel Duration on the Perception of Postvocalic Voiced/Voiceless Consonants." *Concentric: Studies in English Literature and Lingusitics* 27 (2): 35–52.

Wang, Xinchun. 2002. "Training Mandarin and Cantonese Speakers to Identify English Vowel Contrasts: Long-term Retention and Effects on Production." PhD diss., Simon Fraser University.

Wang, Xinchun, and Munro, Murray J. Munro. 2004. "Computer-based Training for Learning English Vowel Contrasts." *System 32*: 539–552.

Warner, Natasha, Allard Jongman, Joan Sereno, and Rachel Kemps. 2004. "Incomplete Neutralization and Other Sub-phonemic Durational Differences in Production and Perception: Evidence from Dutch." *Journal of Phonetics* 32: 251–276.

Wang, Yue, Michelle M. Spence, Allard Jongman, and Joan A. Sereno. 1999. "Training American Listeners to Perceive Mandarin Tones." *Journal of the Acoustical Society of America* 106 (6): 3649–3658.

Wong, Janice W. S. 2012. "Training the Perception and Production of English /e/ and /æ/ of Cantonese ESL Learners: A Comparison of Low vs. High Variability Phonetic Training." *SST*: 37–40.

Wong, Janice W. S. 2013. "Does Proficiency Matter? Effects of High Variability Phonetic Training on the Perception and Production of English Vowels by Cantonese

ESL Learners with High and Low Proficiency Levels.” Paper presented at the Third International Conference on English Pronunciation: Issues and Practices, Murcia, Spain, May 8–10.

Yamada, Reiko A. 1993. “Effects of Extended Training on /r/ and /l/ Identification by Native Speakers of Japanese.” *Journal of the Acoustical Society of America* 93 (4): 2391.

Yamada, Reiko A., Winifred Strange, James S. Magnuson, James S. Pruitt, and William D. Clarke. 1994. “The Intelligibility of Japanese Speakers’ Production of American English /r/, /l/, and /w/ as Evaluated by Native Speakers of American English. In *Proceedings of the International Conference of Spoken Language Processing*, 2023–2026. Yokohama: Acoustical Society of Japan.

List of Figures

Appendix 1: List of Stimuli for the Perceptual Experiment

/f/ – /v/

calf	calve
duff	dove
safe	save
strife	strive

/k/ – /g/

back	bag
buck	bug
chuck	chug
clock	clog
cock	cog
dock	dog
duck	dug
hack	hag
muck	mug
peck	peg
pick	pig
sack	sag
shack	shag
stack	stag
tuck	tug

/p/ – /b/

bop	bob
cap	cab
cop	cob
cup	cub
hop	hob
nip	nib
pup	pub
sop	sob
tap	tab

/s/ – /z/

base	baize
dose	doze
face	phase
fuss	fuzz
grace	graze

/t/ – /d/

bat	bad
beet	bead
bet	bed
bit	bid
but	bud
coat	code
cot	cod
cut	cud
fat	fad
fate	fade
feet	feed
grate	grade
greet	greed
hit	hid
kit	kid
mat	mad
mate	maid
pat	pad
pate	paid
pot	pod
seat	seed
site	side
sweet	swede
tight	tide

/ʧ/ – /ʤ/

rich	ridge

Appendix 2: List of Stimuli for the Production Experiment

/f/ – /v/

calf	calve
duff	dove
safe	save

/k/ – /g/

back	bag
dock	dog
pick	pig
tuck	tug

/p/ – /b/

cap	cab
pup	pub
sop	sob
tap	tab

/s/ – /z/

dose	doze
face	phase
fuss	fuzz

/t/ – /d/

beet	bead
coat	code
fate	fade
kit	kid
pate	paid
site	side
bat	bad

Appendix 3: Script 1 Used for the Enhancement of the Stimuli

The following scripts (Appendices 3, 4, 5, 6, 7) were created using Praat (Boersma and Weenink 2014) by Mgr. Václav Jonáš Podipský. The scripts in Appendix 3 and Appendix 4 were used for changing the quality of the sounds in training. Appendix 5 contains a script that was used to create the pretest and the two posttests. The script in Appendix 6 was used to create the training session with the enhanced stimuli. In Appendix 7, there is a script that was used for training with the original sounds.

```
manipulationEdge = 0.2
enhanceDurBy = 0.35
shiftOfCdur = 1
fricIntShift = -6
stopIntShift = -4

tab = Read from file: "pairs.txt"
nPairs = Get number of rows
sound = Read from file: "soundsNEW edited.wav"
grid = Read from file: "soundsNEW_edited.TextGrid"
gridTab = Down to Table: "no", 6, "yes", "no"

for i to nPairs
	;appendInfoLine: i
	@getTimes: "vl"
	@getTimes: "vd"
	select vlwd
	@enhanceVdur: "vl"
	Save as WAV file: "stimuli/vl/enhancedVdur/'vl$'_'enhanceDurBy'.wav"
	select vdwd
	half1 = Extract part: 0, vdvcbound, "rectangular", 1, "no"
	select vdwd
	if manner$ = "f"
		select vlwd
		half2 = Extract part: vlvcbound, vlWdur, "rectangular", 1, "no"
```

```
		int = Get intensity (dB)
		Scale intensity: int + fricIntShift
		select half1
		plus half2
	else
		select vlwd
		half2 = Extract part: vlvcbound, vlWdur, "rectangular", 1, "no"
		int = Get intensity (dB)
		Scale intensity: int + stopIntShift
		select half1
		plus half2
	endif
	ambivdwd1 = Concatenate
	manambivdwd = To Manipulation: 0.01, 75, 600
	durambivdwd = Extract duration tier
	Add point: vdvcbound, 1
	Add point: vdvcbound + 0.001, (vlcdur - ((vlcdur - vdcdur) * shiftOfCdur)) / vlcdur
	Add point: vdvcbound + vlcdur - 0.001, (vlcdur - ((vlcdur - vdcdur) * shiftOfCdur)) / vlcdur
	Add point: vlctmax, 1
	select manambivdwd
	plus durambivdwd
	Replace duration tier
	select manambivdwd
	ambivdwd2 = Get resynthesis (overlap-add)
	Save as WAV file: "stimuli/vd/shifted/swappedC/'vd$'.wav"
	@enhanceVdur: "vd"
	Save as WAV file: "stimuli/vd/shifted/swappedC-enhanced-vdur/'vd$'_swap_'enhanceDurBy'.wav"
endfor


procedure getTimes: .voice$
	select tab
	'.voice$'$ = Get value: i, .voice$
```

```
		'.voice$'Sex$ = Get value: i, .voice$ + "_sex"
		manner$ = Get value: i, "manner"
		select gridTab
		row = Search column: "text", '.voice$'$
		'.voice$'Tmin = Get value: row, "tmin"
		'.voice$'Tmax = Get value: row, "tmax"
		'.voice$'Wdur = '.voice$'Tmax - '.voice$'Tmin
		select grid
		'.voice$'Grid = Extract part: '.voice$'Tmin, '.voice$'Tmax, "no"
		'.voice$'vtmin = Get start point: 2, 2
		'.voice$'vcbound = Get end point: 2, 2
		'.voice$'vdur = '.voice$'vcbound - '.voice$'vtmin
		'.voice$'ctmax = Get end point: 2, 3
		'.voice$'cdur = '.voice$'ctmax - '.voice$'vcbound
		select sound
		'.voice$'wd = Extract part: '.voice$'Tmin, '.voice$'Tmax, "rectangular", 1, "no"
		Scale intensity: 70
		Save as WAV file: "stimuli/'.voice$'/orig/" + '.voice$'$ + ".wav"
endproc


procedure enhanceVdur: .voice$
		man'.voice$'wd = To Manipulation: 0.01, 75, 600
		dur'.voice$'wd = Extract duration tier
		Add point: '.voice$'vtmin, 1
		if .voice$ = "vd"
				Add point: '.voice$'vtmin + ('.voice$'vdur * manipulationEdge), 1 + enhanceDurBy
				Add point: '.voice$'vcbound - ('.voice$'vdur * manipulationEdge), 1 + enhanceDurBy
		else
				Add point: '.voice$'vtmin + ('.voice$'vdur * manipulationEdge), 1 - enhanceDurBy
				Add point: '.voice$'vcbound - ('.voice$'vdur * manipulationEdge), 1 - enhanceDurBy
```

```
        endif
        Add point: '.voice$'vcbound, 1
        select man'.voice$'wd
        plus dur'.voice$'wd
        Replace duration tier
        select man'.voice$'wd
        enh'.voice$'wd = Get resynthesis (overlap-add)
        Scale intensity: 70
endproc
```

Appendix 4: Script 2 Used for the Enhancement of the Stimuli

```
manipulationEdge = 0.2
enhanceDurvl = 0.15
enhanceDurvd = 0.35
shiftOfCdurToVl = 0.5
mcCutoff = 1100
fcCutoff = 1300
mbCutoff = 500
fbCutoff = 600
fricIntIncrease = 5

tab = Read from file: "pairs.txt"
nPairs = Get number of rows
sound = Read from file: "soundsNEW edited.wav"
grid = Read from file: "soundsNEW_edited.TextGrid"
gridTab = Down to Table: "no", 6, "yes", "no"

for i to nPairs
	;appendInfoLine: i
	@getTimes: "vl"
	@getTimes: "vd"
	select vlwd
	@enhanceVdur: "vl"
	Save as WAV file: "stimuli/vl/enhancedVdur/'vl$'_'enhanceDurvl'.wav"
	select vdwd
	half1 = Extract part: 0, vdvcbound, "rectangular", 1, "no"
	select vdwd
	if manner$ = "f"
		half2 = Extract part: vdvcbound, vdWdur, "rectangular", 1, "no"
		filt = Filter (stop Hann band): 0, 'vdSex$'bCutoff, 80
		int = Get intensity (dB)
		Scale intensity: int + fricIntIncrease
```

```
		select half1
		plus filt
	else
		half2 = Extract part: vdvcbound, vdctmax, "rectangular", 1, "no"
		filt = Filter (stop Hann band): 0, 'vdSex$'cCutoff, 80
		select vdwd
		half3 = Extract part: vdctmax, vdWdur, "rectangular", 1, "no"
		filt2 = Filter (stop Hann band): 0, 'vdSex$'bCutoff, 80
		select half1
		plus filt
		plus filt2
	endif
	ambivdwd1 = Concatenate
	manambivdwd = To Manipulation: 0.01, 75, 600
	durambivdwd = Extract duration tier
	Add point: vdvcbound, 1
	Add point: vdvcbound + 0.001, (vdcdur - ((vdcdur - vlcdur) * shiftOfCdurToVl)) / vdcdur
	Add point: vdctmax - 0.001, (vdcdur - ((vdcdur - vlcdur) * shiftOfCdurToVl)) / vdcdur
	Add point: vdctmax, 1
	select manambivdwd
	plus durambivdwd
	Replace duration tier
	select manambivdwd
	ambivdwd2 = Get resynthesis (overlap-add)
	Save as WAV file: "stimuli/vd/shifted/ambiguous-c/'vd$'.wav"
	@enhanceVdur: "vd"
	Save as WAV file: "stimuli/vd/shifted/ambi-c-enhanced-vdur/'vd$'_filt_'enhanceDurvd'.wav"
endfor

procedure getTimes: .voice$
	select tab
	'.voice$'$ = Get value: i, .voice$
```

```
	'.voice$'Sex$ = Get value: i, .voice$ + "_sex"
	manner$ = Get value: i, "manner"
	select gridTab
	row = Search column: "text", '.voice$'$
	'.voice$'Tmin = Get value: row, "tmin"
	'.voice$'Tmax = Get value: row, "tmax"
	'.voice$'Wdur = '.voice$'Tmax - '.voice$'Tmin
	select grid
	'.voice$'Grid = Extract part: '.voice$'Tmin, '.voice$'Tmax, "no"
	'.voice$'vtmin = Get start point: 2, 2
	'.voice$'vcbound = Get end point: 2, 2
	'.voice$'vdur = '.voice$'vcbound - '.voice$'vtmin
	'.voice$'ctmax = Get end point: 2, 3
	'.voice$'cdur = '.voice$'ctmax - '.voice$'vcbound
	select sound
	'.voice$'wd = Extract part: '.voice$'Tmin, '.voice$'Tmax, "rectangular", 1, "no"
	Scale intensity: 70
	Save as WAV file: "stimuli/'.voice$'/orig/" + '.voice$'$ + ".wav"
endproc

procedure enhanceVdur: .voice$
	man'.voice$'wd = To Manipulation: 0.01, 75, 600
	dur'.voice$'wd = Extract duration tier
	Add point: '.voice$'vtmin, 1
	if .voice$ = "vd"
		Add point: '.voice$'vtmin + ('.voice$'vdur * manipulationEdge), 1 + enhanceDur'.voice$'
		Add point: '.voice$'vcbound - ('.voice$'vdur * manipulationEdge), 1 + enhanceDur'.voice$'
	else
		Add point: '.voice$'vtmin + ('.voice$'vdur * manipulationEdge), 1 - enhanceDur'.voice$'
		Add point: '.voice$'vcbound - ('.voice$'vdur * manipulationEdge), 1 - enhanceDur'.voice$'
```

```
		endif
		Add point: '.voice$'vcbound, 1
		select man'.voice$'wd
		plus dur'.voice$'wd
		Replace duration tier
		select man'.voice$'wd
		enh'.voice$'wd = Get resynthesis (overlap-add)
		Scale intensity: 70
endproc
```

Appendix 5: Script Used for Pretest and Posttest

```
pauseAfterEvry = 50
;minDistanceFromFirst = 6


pairs = Read from file: "stimuli/pairs.txt"
allStimList = Create Strings as file list: "stimList", "stimuli\*.wav"
Randomize
nAllStim = Get number of strings
playListTab = Create Table with column names: "playListTab", nAllStim, "subject task file leftButton rightButton response rw rt"
for j to nAllStim
	select allStimList
	stim$ = Get string: j
	select playListTab
	Set string value: j, "file", stim$
endfor


trainList = Create Strings as file list: "stimList", "stimuli\train\*.wav"
Randomize
nTrain = Get number of strings


beginPause: "A short questionnaire"
	comment: "Please fill out this short questionnaire."
	word: "Initials", "AB"
	word: "Age", "22"
	choice: "Sex", 1
		option: "female"
		option: "male"
	choice: "Test", 1
		option: "pre-test"
		option: "post-test"
		option: "post-test2"
```

```
	natural: "Participant number", "1"
clicked = endPause: "Continue", 1

subjCode$ = "'participant_number'_'initials$'_'age$'_" + left$ (sex$, 1)

select playListTab
Save as tab-separated file: "results\'subjCode$'_'test$'.txt"

label INTRO0
demoWindowTitle: "Listening task"
demo Black
demo Times
demo Font size: 24
demo Select inner viewport: 0, 100, 0, 100
demo Axes: 0, 100, 0, 100
demo Paint rectangle: "silver", 0, 100, 0, 100
demo Text: 50, "centre", 55, "half", "##Maximize this window now.#"
demo Text: 50, "centre", 45, "half", "Click to show instructions."
while demoWaitForInput ( )
	goto INTRO1 demoInput ("•→")
endwhile

label INTRO1
demo Erase all
demo Select inner viewport: 0, 100, 0, 100
demo Paint rectangle: "silver", 0, 100, 0, 100
demo Text: 50, "centre", 75, "half", "You will hear words recorded by different people."
demo Text: 50, "centre", 65, "half", "Decide what the word was and click on the corresponding button."
demo Text: 50, "centre", 35, "half", "Click to hear four examples."
while demoWaitForInput ( )
	goto TRAINING demoInput ("•→")
	goto INTRO0 demoInput ("←")
```

```
endwhile

label TRAINING
for i to nTrain
	select trainList
	file$ = Get string: i
	sound = Read from file: "stimuli\train\'file$'"
	soundDur = Get total duration
	@screen
endfor

label INTRO2
demo Erase all
demo Paint rectangle: "silver", 0, 100, 0, 100
demo Black
demo Font size: 24
demo Select inner viewport: 0, 100, 0, 100
demo Text: 50, "centre", 65, "half", "Is the volume fine? (To play the examples again press the spacebar.)"
demo Text: 50, "centre", 45, "half", "The task shouldn't last more than 5 minutes."
demo Text: 50, "centre", 35, "half", "You'll be given the chance to take a break."
demo Text: 50, "centre", 25, "half", "Click to run the task."
while demoWaitForInput ( )
	goto THETEST demoInput ("•→")
	goto INTRO1 demoInput ("←")
	goto TRAINING demoInput (" ")
endwhile

label THETEST
for i to nAllStim
	select playListTab
	file$ = Get value: i, "file"
	shortName$ = file$ - ".wav"
```

```
	sound = Read from file: "stimuli\'file$'"
	soundDur = Get total duration
	@screen
	select playListTab
	Set string value: i, "subject", subjCode$
	Set string value: i, "task", test$
	Set string value: i, "response", response$
	Set string value: i, "rw", rw$
	Set string value: i, "leftButton", opt1$
	Set string value: i, "rightButton", opt2$
	Set numeric value: i, "rt", rt
	Save as tab-separated file: "results\'subjCode$'_'test$'.txt"
	if i mod pauseAfterEvry = 0 and i < nAllStim
		@pause
	endif
endfor

label END1
demo Erase all
demo Black
demo Font size: 24
demo Select inner viewport: 0, 100, 0, 100
demo Paint rectangle: "silver", 0, 100, 0, 100
demo Text: 50, "centre", 60, "half", "You've finished. Your answers have been recorded."
demo Text: 50, "centre", 40, "half", "Thanks very much! You can close this window now."
select all
Remove

procedure screen
	demo Erase all
	demo Font size: 34
	demo Select inner viewport: 0, 100, 0, 100
	demo Paint rectangle: "silver", 0, 100, 0, 100
```

```
demo Black
demo Text: 50, "centre", 80, "half", "Which word did you hear?"
demo Grey
demo Line width: 3
demo Draw rounded rectangle: 32, 46, 42, 58, 3
demo Draw rounded rectangle: 54, 68, 42, 58, 3
corOpt = randomInteger(1,2)
if corOpt = 1
	other = 2
elif corOpt = 2
	other = 1
endif
opt'corOpt'$ = file$ - ".wav"
select pairs
pairRow = Search column: "vl", opt'corOpt'$
if pairRow = 0
	pairRow = Search column: "vd", opt'corOpt'$
	opt'other'$ = Get value: pairRow, "vl"
else
	opt'other'$ = Get value: pairRow, "vd"
endif
select sound
stopwatch
Play
demo Paint rounded rectangle: "cyan", 32, 46, 42, 58, 3
demo Paint rounded rectangle: "cyan", 54, 68, 42, 58, 3
demo Black
demo Text: 39, "centre", 50, "half", opt1$
demo Text: 61, "centre", 50, "half", opt2$
while demoWaitForInput ( )
	if demoClickedIn (32, 46, 42, 58)
		rawRt = stopwatch
		rt = rawRt - soundDur
```

```
				response$ = opt1$
				if response$ = opt'corOpt'$
					rw$ = "r"
				else
					rw$ = "w"
				endif
				goto NEXTTRIAL
			elsif demoClickedIn (54, 68, 42, 58)
				rawRt = stopwatch
				rt = rawRt - soundDur
				response$ = opt2$
				if response$ = opt'corOpt'$
					rw$ = "r"
				else
					rw$ = "w"
				endif
				goto NEXTTRIAL
			endif
		endwhile
		label NEXTTRIAL
endproc

procedure pause
		demo Erase all
		demo Black
		demo Font size: 24
		demo Select inner viewport: 0, 100, 0, 100
		demo Paint rectangle: "silver", 0, 100, 0, 100
		demo Text: 50, "centre", 60, "half", "Now you can take a short break if you like."
		demo Text: 50, "centre", 40, "half", "Click to resume the task."
		while demoWaitForInput ( )
			goto RESUMED demoInput ("•")
		endwhile
```

```
    label RESUMED
endproc
```

Appendix 6: Script Used for Training with Enhanced Stimuli

```
#	This script uses the Praat Demo Window to run a little experiment
#
#	Copyright (C) 2014 Jakub Bortlík and Jonáš Podlipský
#
#	jakub.bortlik@gmail.com
#
#	This program is free software: you can redistribute it and/or modify
#	it under the terms of the GNU General Public License as published by
#	the Free Software Foundation, either version 3 of the License, or
#	(at your option) any later version.
#
#	This program is distributed in the hope that it will be useful,
#	but WITHOUT ANY WARRANTY; without even the implied warranty of
#	MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.  See the
#	GNU General Public License for more details.
#
#	You should have received a copy of the GNU General Public License
#	along with this program.  If not, see <http://www.gnu.org/licenses/>.
#

	###############
	##  VARIABLES
	###############

# COLORS
background$ = "0.6, 0.8, 0.9"
buttons$ = "0, 0.6, 0.9"

# BUTTONS
nButtons = 2
x1 = 10
```

```
x2 = 60

y1 = 40

y2 = 40

xsize = 30

ysize = 15


trial = 0


# How many times do the listeners have to make a mistake for the particular stimulus to be repeated later:

repIncor = 1


# How quickly will correctly recognized items be deleted from the list:

progress = 0.75

# if "progress = 1" the Nr. of deleted items will be the same as the Nr. of incorrect items added in the previous round.

# if "progress > 1" more correct items are deleted than have incorrect ones been added.



# The round after which items can be removed from the list (PRACTICE ROUND = 1st round)

deleteAfterRound = 3


# at least "minProgress" items will be deleted (this has to be an integer):

minProgress = 50


# Delete an item when correct answers exceed incorrect answers by "train"+1:

train = 1


# skip item if correct answers > incorrect answer by a critical number

# the critical number:

maxCorrRepeat = 4


# Stop the experiment after this many trials:

lastTrial = 500
```

```
# Pause the experiment after "pauseAfter" items:
pauseAfter = 80

# number of items in the PRACTICE ROUND:
practice = 4


	####################
	##  IDENTIFICATION
	####################

beginPause: "Identifikace"
	comment: "Vyplňte prosím svoje údaje. Pořadové číslo Vám určí experimentátor."
	word: "Inicialy", "AB"
	word: "Vek", "20"
	choice: "Pohlavi", 1
		option: "zena"
		option: "muz"
	word: "Poradove cislo", "01"
	clicked = endPause: "Pokračovat", 1

pohlavi$ = left$ (pohlavi$, 1)
subjCode$ = "'poradove_cislo$'_'inicialy$'_'vek$'_'pohlavi$'"

### This part loads data and creates tables to be used during the experiment:
stimList = Create Strings as file list: "Stimuli list", "stimuli/enh/*.wav"
nStimuli = Get number of strings

for i to nStimuli
	selectObject: stimList
	stimFull$ = Get string: i
	stim$ = stimFull$ - ".wav"
```

```
	Set string: i, stim$
	incor'stim$' = 0
	cor'stim$' = 0
endfor

results = Create Table with column names: "results", 0, "stimulus opt1 opt2 correct cor incor response rw replayed rt"
pairs = Read from file: "stimuli/pairs.txt"


	##################
	##  INSTRUCTIONS
	##################
demoWindowTitle: "Identifikační experiment"
demo Erase all
demo Colour: "black"
demo Select inner viewport: 0, 100, 0, 100
demo Axes: 0, 100, 0, 100
demo Paint rectangle: "{'background$'}", 0, 100, 0, 100

demo Erase all
demo Colour: "black"
demo Select inner viewport: 0, 100, 0, 100
demo Axes: 0, 100, 0, 100
demo Paint rectangle: "{'background$'}", 0, 100, 0, 100
demo Text special: 50, "centre", 80, "half", "Helvetica", 30, "0", "Instrukce"
demo Text special: 50, "centre", 65, "half", "Helvetica", 20, "0", "B\e<hem experimentu usly\s<\i'te jednotliv\a' slova"
demo Text special: 50, "centre", 60, "half", "Helvetica", 20, "0", "vysloven\a' rodil\y'm mluv\c<\i'm angli\c<tiny."
demo Text special: 50, "centre", 50, "half", "Helvetica", 20, "0", "Pot\e' dostanete na v\y'b\e<r ze dvou mo\z<nost\i'"
demo Text special: 50, "centre", 45, "half", "Helvetica", 20, "0", "a Va\s<\i'm \u'kolem bude ur\c<it, jak\e' slovo jste sly\s<eli."
```

```
demo Text special: 50, "centre", 30, "half", "Helvetica", 20, "0", "\U'kol si  vyzkou\s<\i'te v cvi\c<n\e'm kole."

demo Text special: 50, "centre", 20, "half", "Helvetica", 20, "0", "%%Pro pokra\c<ov\a'n\i' klikn\e<te nebo stiskn\e<te mezern\i'k.%"


while demoWaitForInput ( )
	goto PRACTICE demoInput ("• ")
endwhile



	###############
	##   PRACTICE
	###############


label PRACTICE


demo Erase all
demo Colour: "black"
demo Select inner viewport: 0, 100, 0, 100
demo Axes: 0, 100, 0, 100
demo Paint rectangle: "{'background$'}", 0, 100, 0, 100
demo Text special: 50, "centre", 80, "half", "Helvetica", 30, "0", "Demo"
demo Text special: 50, "centre", 65, "half", "Helvetica", 20, "0", "U ka\z<d\e' nahr\a'vky si mus\i'te vybrat jednu z mo\z<nost\i'."
demo Text special: 50, "centre", 60, "half", "Helvetica", 20, "0", "Pokud mysl\i'te, \z<e ani jedno slovo neodpov\i'd\a' nahr\a'vce,"
demo Text special: 50, "centre", 55, "half", "Helvetica", 20, "0", "vyberte to, kter\e' je j\i' v\i'ce podobn\e'."
demo Text special: 50, "centre", 45, "half", "Helvetica", 20, "0", "Ka\z<dou nahr\a'vku si m\uo\z<ete p\r<ehr\a't je\s<t\e< jednou."
demo Text special: 50, "centre", 35, "half", "Helvetica", 20, "0", "B\e<hem dema se m\uo\z<ete na cokoliv zeptat,"
demo Text special: 50, "centre", 30, "half", "Helvetica", 20, "0", "pozd\e<ji u\z< nebude mo\z<n\e' experiment p\r<eru\s<ovat."
demo Text special: 50, "centre", 15, "half", "Helvetica", 20, "0", "%%Pro pokra\c<ov\a'n\i' klikn\e<te nebo stiskn\e<te mezern\i'k.%"
```

```
while demoWaitForInput ( )
	goto MAIN demoInput ("• ")
endwhile



	################
	##  MAIN PHASE
	################


label MAIN
```

This "for cycle" goes through the list of stimuli until either the "number of trials = lastTrial";

or all the items have been identified correctly a certain number of times.

In each round (except the PRACTICE round), incorrect answers from the previous round are repeated once more.

```
for r to 100
	# The number of stimuli in a round is stored as nStimuli'r', so that rounds can be compared
	# to find out how many incorrect items have been added from the previous round:
	selectObject: stimList
	nStimuli'r' = Get number of strings

	# The PRACTICE round is shorter and its results are deleted after it is done (see end of the main "for cycle"):
	if r = 1
		nStimuli'r' = practice
	endif

	if r > deleteAfterRound
		# After the 'deleteAfterRound' round a certain number of always correctly identified items is removed from the stimuli list:
		prevRun = r - 1
		newItems = nStimuli'r' - nStimuli'prevRun'
```

```
itemsToDelete = ceiling (newItems * progress)
if itemsToDelete < minProgress
	itemsToDelete = minProgress
endif

# This "for cycle" removes the required number of correct items, provided there are any such items:
removed = 0
for n to itemsToDelete
	candidate = 0
	# This "for cycle" identifies the correct items that could be deleted:
	for i to nStimuli'r'
		selectObject: stimList
		stim$ = Get string: i
		if incor'stim$' = 0
			candidate = candidate + 1
			position'candidate' = i
		endif
	endfor
	# This "if jump" removes randomly one correct item:
	if candidate > 0
		rand = randomInteger (1, candidate)
		Remove string: position'rand'
		removed = removed + 1
	endif
	nStimuli'r' = Get number of strings
endfor
if itemsToDelete > removed
	stillRemove = itemsToDelete - removed
	for n to stillRemove
		candidate = 0
		# This "for cycle" identifies the correct items that could be deleted:
		for i to nStimuli'r'
```

```
				selectObject: stimList
				stim$ = Get string: i
				if (incor'stim$' + train) < cor'stim$'
					candidate = candidate + 1
					position'candidate' = i
				endif
			endfor
			# This "if jump" removes randomly one correct item:
			if candidate > 0
				rand = randomInteger (1, candidate)
				Remove string: position'rand'
			endif
			nStimuli'r' = Get number of strings
		endfor
	endif
endif

selectObject: stimList
Randomize

for i to nStimuli'r'
	selectObject: stimList
	stim$ = Get string: i

	if r > deleteAfterRound
		if incor'stim$' + maxCorrRepeat < cor'stim$'
			goto SKIPITEM
		endif
	endif

	# For each item in the stimuli list randomly assign "opt1$" and "opt2$" values
(different in each round):
	random = randomInteger(1,2)
```

```
if random = 1
	other = 2
elif random = 2
	other = 1
endif
opt'random'$ = left$ (stim$, index (stim$, "_") - 1)
correct$ = opt'random'$
selectObject: pairs
pairRow = Search column: "vl", opt'random'$
if pairRow = 0
	pairRow = Search column: "vd", opt'random'$
	opt'other'$ = Get value: pairRow, "vl"
else
	opt'other'$ = Get value: pairRow, "vd"
endif
# Pause the experiment after a specified number of stimuli:
if trial > 0 and trial mod pauseAfter = 0
	demo Erase all
	demo Colour: "black"
	demo Select inner viewport: 0, 100, 0, 100
	demo Axes: 0, 100, 0, 100
	demo Paint rectangle: "{'background$'}", 0, 100, 0, 100
	demo Text special: 50, "centre", 80, "half", "Helvetica", 30, "0", "P\r<est\a'vka"
	demo Text special: 50, "centre", 65, "half", "Helvetica", 20, "0", "Te\d< si m\uo\z<ete chv\i'li odpo\c<inout."
	demo Text special: 50, "centre", 60, "half", "Helvetica", 20, "0", "A\z< budete p\r<ipraveni pokra\c<ovat,"
	demo Text special: 50, "centre", 55, "half", "Helvetica", 20, "0", "klikn\e<te nebo stiskn\e<te mezern\i'k."
	while demoWaitForInput ()
		goto ENDPAUSE demoInput ("• ")
	endwhile
endif
```

```
label ENDPAUSE

selectObject: results
Append row
row = Get number of rows
Set string value: row, "stimulus", stim$
Set string value: row, "opt1", opt1$
Set string value: row, "opt2", opt2$
Set string value: row, "correct", correct$

trial = trial + 1
replayed = 0
correct = 0
wrong = 0

played = 0
@trialScr: "'background$'", "'background$'", "'background$'"
stopwatch
@playSound
played = 1
@trialScr: "'buttons$'", "'buttons$'", "'background$'"

label REPEAT_MAIN

### This "while cycle" enables the user to click the buttons:
while demoWaitForInput ( )
	for k to nButtons
		x = x'k'
		y = y'k'
		if replayed = 0 and demoClickedIn (46, 54, 48, 55)
			replayed = 1
			@trialScr: "'background$'", "'background$'", "'background$'"
```

```
                        stopwatch
                        @playSound
                        @trialScr: "'buttons$'", "'background$'", "'background$'"
                        selectObject: results
                        Set string value: row, "replayed", "yes"
                        goto REPEAT_MAIN
                elif demoClickedIn (x, x+xsize, y, y+ysize) or demoInput ("k")
                        rt = stopwatch
                        selectObject: results
                        Set string value: row, "rt", fixed$ (rt, 3)
                        Set string value: row, "response", opt'k'$
                        if opt'k'$ = correct$
                                correct = k
                                Set string value: row, "rw", "R"
                                cor'stim$' = cor'stim$' + 1
                                # The columns "cor" and "incor" can be deleted from the "results" table.
                                # They only make the values "visible" in the table, but the script runs without them.
                                Set numeric value: row, "cor", cor'stim$'
                                Set numeric value: row, "incor", incor'stim$'
                        else
                                wrong = k
                                Set string value: row, "rw", "W"
                                incor'stim$' = incor'stim$' + 1
                                Set numeric value: row, "incor", incor'stim$'
                                Set numeric value: row, "cor", cor'stim$'
                                # This "if jump" adds incorrectly identified items to the "stimList":
                                if incor'stim$' >= repIncor and r > 1
                                        selectObject: stimList
                                        currentRows = Get number of strings
                                        Insert string:  currentRows + 1, stim$
```

```
					endif
					selectObject: results
				endif
				Save as tab-separated file: "results/'subjCode$'_results.Table"

				@trialScr: "'background$'", "'background$'", "'buttons$'"

				### This "while cycle" makes it possible to click the "new item" button:
				while demoWaitForInput ()
					goto DECIDED_MAIN demoClickedIn (46, 54, 40, 47) or demoInput (" ")
				endwhile
			endif
		endfor
	endwhile
	label DECIDED_MAIN
	if trial = lastTrial
		goto FINAL
	endif
	label SKIPITEM
endfor

# Reset values and create a new "results" table after the PRACTICE round:
if r = 1
	trial = 0
	selectObject: stimList
	nStimuli = Get number of strings
	for i to nStimuli
		stimFull$ = Get string: i
		stim$ = stimFull$ - ".wav"
		incor'stim$' = 0
```

```
cor'stim$' = 0
endfor
selectObject: results
Remove
results = Create Table with column names: "results", 0, "stimulus opt1 opt2 correct cor incor response rw replayed rt"

demo Erase all
demo Colour: "black"
demo Select inner viewport: 0, 100, 0, 100
demo Axes: 0, 100, 0, 100
demo Paint rectangle: "{'background$'}", 0, 100, 0, 100
demo Text special: 50, "centre", 70, "half", "Helvetica", 30, "0", "Konec n\a'cviku"
demo Text special: 50, "centre", 50, "half", "Helvetica", 20, "0", "Jestli m\a'te n\e<jak\e' ot\a'zky,"
demo Text special: 50, "centre", 45, "half", "Helvetica", 20, "0", "te\d< m\a'te mo\z<nost se zeptat."
demo Text special: 50, "centre", 30, "half", "Helvetica", 20, "0", "%%Kliknut\i'm nebo stisknut\i'm mezern\i'ku%"
demo Text special: 50, "centre", 25, "half", "Helvetica", 20, "0", "%%spust\i'te hlavn\i' \c<\a'st experimentu.%"

while demoWaitForInput ( )
goto START_MAIN demoInput ("• ")
endwhile
label START_MAIN
endif
endfor

#################
## FINAL SCREEN
#################
```

```
label FINAL
demo Erase all
demo Colour: "black"
demo Select inner viewport: 0, 100, 0, 100
demo Axes: 0, 100, 0, 100
demo Paint rectangle: "{'background$'}", 0, 100, 0, 100
demo Text special: 50, "centre", 55, "half", "Helvetica", 30, "0", "Konec"
demo Text special: 50, "centre", 40, "half", "Helvetica", 20, "0", "D\e<kujeme za Va\s<i \u'\c<ast
na experimentu!"

selectObject: stimList
plusObject: results
plusObject: pairs
Remove



	#################
	##  PROCEDURES
	#################

procedure trialScr: .buttons$, .replay$, .next$
	demo Erase all
	demo Colour: "black"
	demo Line width: 2
	demo Select inner viewport: 0, 100, 0, 100
	demo Axes: 0, 100, 0, 100
	demo Paint rectangle: "{'background$'}", 0, 100, 0, 100

	demo Paint rounded rectangle: "{'.replay$'}", 46, 54, 48, 55, 2
	demo Draw rounded rectangle: 46, 54, 48, 55, 2
	demo Text special: 50, "centre", 52.5, "half", "Helvetica", 15, "0", "p\r<ehr\a't"
	demo Text special: 50, "centre", 50.5, "half", "Helvetica", 15, "0", "znovu"
```

```
	demo Paint rounded rectangle: "{'.next$'}", 46, 54, 40, 47, 2
	demo Draw rounded rectangle: 46, 54, 40, 47, 2
	demo Text special: 50, "centre", 44.5, "half", "Helvetica", 15, "0", "dal\s<\i'"
	demo Text special: 50, "centre", 42.5, "half", "Helvetica", 15, "0", "slovo"

	### This "for cycle" draws the two option buttons in appropriate colours:
	for k to nButtons
		x = x'k'
		tox = x+xsize
		y = y'k'
		toy = y+ysize
		selectObject: results
		opt$ = Get value: trial, "opt'k'"
		if correct = k
			demo Paint rounded rectangle: "{0, 1, 0}", x, tox, y, toy, 3
		elif wrong = k
			demo Paint rounded rectangle: "{1, 0, 0}", x, tox, y, toy, 3
		else
			demo Paint rounded rectangle: "{'.buttons$'}", x, tox, y, toy, 3
		endif
		demo Draw rounded rectangle: x, tox, y, toy, 3

		### This "if jump" hides the button labels befor playing the sound:
		if played = 1
			demo Text special: (tox+x)/2, "centre", (toy+y)/2, "half", "Helvetica", 42, "0", "'opt$'"
		endif
	endfor
	demo Text special: 50, "centre", 90, "half", "Helvetica", 36, "0", "##Kter\e' slovo jste sly\s<eli?#"
endproc

procedure playSound
	selectObject: stimList
```

```
    sound$ = Get string: i
    Read from file: "stimuli/enh/'sound$'.wav"
    Play
    Remove
endproc
```

Appendix 7: Script Used for Training with Original Stimuli

```
#	This script uses the Praat Demo Window to run a little experiment
#
#	Copyright (C) 2014 Jakub Bortlík and Jonáš Podlipský
#
#	jakub.bortlik@gmail.com
#
#	This program is free software: you can redistribute it and/or modify
#	it under the terms of the GNU General Public License as published by
#	the Free Software Foundation, either version 3 of the License, or
#	(at your option) any later version.
#
#	This program is distributed in the hope that it will be useful,
#	but WITHOUT ANY WARRANTY; without even the implied warranty of
#	MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.  See the
#	GNU General Public License for more details.
#
#	You should have received a copy of the GNU General Public License
#	along with this program.  If not, see <http://www.gnu.org/licenses/>.
#

	###############
	##  VARIABLES
	###############

# COLORS
background$ = "0.6, 0.8, 0.9"
buttons$ = "0, 0.6, 0.9"

# BUTTONS
nButtons = 2
x1 = 10
```

```
x2 = 60

y1 = 40

y2 = 40

xsize = 30

ysize = 15


trial = 0


# How many times do the listeners have to make a mistake for the particular stimulus to be repeated later:

repIncor = 1


# How quickly will correctly recognized items be deleted from the list:

progress = 0.75

# if "progress = 1" the Nr. of deleted items will be the same as the Nr. of incorrect items added in the previous round.

# if "progress > 1" more correct items are deleted than have incorrect ones been added.



# The round after which items can be removed from the list (PRACTICE ROUND = 1st round)

deleteAfterRound = 3


# at least "minProgress" items will be deleted (this has to be an integer):

minProgress = 25


# Delete an item when correct answers exceed incorrect answers by "train"+1:

train = 1


# skip item if correct answers > incorrect answer by a critical number

# the critical number:

maxCorrRepeat = 4


# Stop the experiment after this many trials:

lastTrial = 500
```

```
# Pause the experiment after "pauseAfter" items:
pauseAfter = 80

# number of items in the PRACTICE ROUND:
practice = 4


	####################
	##  IDENTIFICATION
	####################

beginPause: "Identifikace"
	comment: "Vyplňte prosím svoje údaje. Pořadové číslo Vám určí experimentátor."
	word: "Inicialy", "AB"
	word: "Vek", "20"
	choice: "Pohlavi", 1
		option: "zena"
		option: "muz"
	word: "Poradove cislo", "01"
	clicked = endPause: "Pokračovat", 1

pohlavi$ = left$ (pohlavi$, 1)
subjCode$ = "'poradove_cislo$'_'inicialy$'_'vek$'_'pohlavi$'"

### This part loads data and creates tables to be used during the experiment:
stimList = Create Strings as file list: "Stimuli list", "stimuli/or/*.wav"
nStimuli = Get number of strings

for i to nStimuli
	selectObject: stimList
	stimFull$ = Get string: i
	stim$ = stimFull$ - ".wav"
```

```
	Set string: i, stim$
	incor'stim$' = 0
	cor'stim$' = 0
endfor

results = Create Table with column names: "results", 0, "stimulus opt1 opt2 correct cor incor response rw replayed rt"
pairs = Read from file: "stimuli/pairs.txt"



	##################
	##  INSTRUCTIONS
	##################
demoWindowTitle: "Identifikační experiment"
demo Erase all
demo Colour: "black"
demo Select inner viewport: 0, 100, 0, 100
demo Axes: 0, 100, 0, 100
demo Paint rectangle: "{'background$'}", 0, 100, 0, 100

demo Erase all
demo Colour: "black"
demo Select inner viewport: 0, 100, 0, 100
demo Axes: 0, 100, 0, 100
demo Paint rectangle: "{'background$'}", 0, 100, 0, 100
demo Text special: 50, "centre", 80, "half", "Helvetica", 30, "0", "Instrukce"
demo Text special: 50, "centre", 65, "half", "Helvetica", 20, "0", "B\e<hem experimentu usly\s<\i'te jednotliv\a' slova"
demo Text special: 50, "centre", 60, "half", "Helvetica", 20, "0", "vysloven\a' rodil\y'm mluv\c<\i'm angli\c<tiny."
demo Text special: 50, "centre", 50, "half", "Helvetica", 20, "0", "Pot\e' dostanete na v\y'b\e<r ze dvou mo\z<nost\i'"
demo Text special: 50, "centre", 45, "half", "Helvetica", 20, "0", "a Va\s<\i'm \u'kolem bude ur\c<it, jak\e' slovo jste sly\s<eli."
```

```
demo Text special: 50, "centre", 30, "half", "Helvetica", 20, "0", "\U'kol si  vyzkou\s<\i'te v cvi\c<n\e'm kole."

demo Text special: 50, "centre", 20, "half", "Helvetica", 20, "0", "%%Pro pokra\c<ov\a'n\i' klikn\e<te nebo stiskn\e<te mezern\i'k.%"


while demoWaitForInput ( )

	goto PRACTICE demoInput ("• ")

endwhile



	###############

	##   PRACTICE

	###############


label PRACTICE


demo Erase all

demo Colour: "black"

demo Select inner viewport: 0, 100, 0, 100

demo Axes: 0, 100, 0, 100

demo Paint rectangle: "{'background$'}", 0, 100, 0, 100

demo Text special: 50, "centre", 80, "half", "Helvetica", 30, "0", "Demo"

demo Text special: 50, "centre", 65, "half", "Helvetica", 20, "0", "U ka\z<d\e' nahr\a'vky si mus\i'te vybrat jednu z mo\z<nost\i'."

demo Text special: 50, "centre", 60, "half", "Helvetica", 20, "0", "Pokud mysl\i'te, \z<e ani jedno slovo neodpov\i'd\a' nahr\a'vce,"

demo Text special: 50, "centre", 55, "half", "Helvetica", 20, "0", "vyberte to, kter\e' je j\i' v\i'ce podobn\e'."

demo Text special: 50, "centre", 45, "half", "Helvetica", 20, "0", "Ka\z<dou nahr\a'vku si m\uo\z<ete p\r<ehr\a't je\s<t\e< jednou."

demo Text special: 50, "centre", 35, "half", "Helvetica", 20, "0", "B\e<hem dema se m\uo\z<ete na cokoliv zeptat,"

demo Text special: 50, "centre", 30, "half", "Helvetica", 20, "0", "pozd\e<ji u\z< nebude mo\z<n\e' experiment p\r<eru\s<ovat."

demo Text special: 50, "centre", 15, "half", "Helvetica", 20, "0", "%%Pro pokra\c<ov\a'n\i' klikn\e<te nebo stiskn\e<te mezern\i'k.%"
```

```
while demoWaitForInput ( )
	goto MAIN demoInput ("• ")
endwhile


	################
	##  MAIN PHASE
	################

label MAIN

### This "for cycle" goes through the list of stimuli until either the "number of trials = lastTrial";
### or all the items have been identified correctly a certain number of times.
### In each round (except the PRACTICE round), incorrect answers from the previous round are repeated once more.

for r to 100
	# The number of stimuli in a round is stored as nStimuli'r', so that rounds can be compared
	# to find out how many incorrect items have been added from the previous round:
	selectObject: stimList
	nStimuli'r' = Get number of strings

	# The PRACTICE round is shorter and its results are deleted after it is done (see end of the main "for cycle"):
	if r = 1
		nStimuli'r' = practice
	endif

	if r > deleteAfterRound
		# After the 'deleteAfterRound' round a certain number of always correctly identified items is removed from the stimuli list:
		prevRun = r - 1
		newItems = nStimuli'r' - nStimuli'prevRun'
```

```
		itemsToDelete = ceiling (newItems * progress)
		if itemsToDelete < minProgress
			itemsToDelete = minProgress
		endif

		# This "for cycle" removes the required number of correct items, provided there are any such items:
		removed = 0
		for n to itemsToDelete
			candidate = 0
			# This "for cycle" identifies the correct items that could be deleted:
			for i to nStimuli'r'
				selectObject: stimList
				stim$ = Get string: i
				if incor'stim$' = 0
					candidate = candidate + 1
					position'candidate' = i
				endif
			endfor
			# This "if jump" removes randomly one correct item:
			if candidate > 0
				rand = randomInteger (1, candidate)
				Remove string: position'rand'
				removed = removed + 1
			endif
			nStimuli'r' = Get number of strings
		endfor
		if itemsToDelete > removed
			stillRemove = itemsToDelete - removed
			for n to stillRemove
				candidate = 0
				# This "for cycle" identifies the correct items that could be deleted:
				for i to nStimuli'r'
```

```
					selectObject: stimList
					stim$ = Get string: i
					if (incor'stim$' + train) < cor'stim$'
						candidate = candidate + 1
						position'candidate' = i
					endif
				endfor
				# This "if jump" removes randomly one correct item:
				if candidate > 0
					rand = randomInteger (1, candidate)
					Remove string: position'rand'
				endif
				nStimuli'r' = Get number of strings
			endfor
		endif
	endif

	selectObject: stimList
	Randomize

	for i to nStimuli'r'
		selectObject: stimList
		stim$ = Get string: i

		if r > deleteAfterRound
			if incor'stim$' + maxCorrRepeat < cor'stim$'
				goto SKIPITEM
			endif
		endif

		# For each item in the stimuli list randomly assign "opt1$" and "opt2$" values
(different in each round):
		random = randomInteger(1,2)
```

```
if random = 1
	other = 2
elif random = 2
	other = 1
endif
opt'random'$ = stim$ - ".wav"
correct$ = opt'random'$
selectObject: pairs
pairRow = Search column: "vl", opt'random'$
if pairRow = 0
	pairRow = Search column: "vd", opt'random'$
	opt'other'$ = Get value: pairRow, "vl"
else
	opt'other'$ = Get value: pairRow, "vd"
endif
# Pause the experiment after a specified number of stimuli:
if trial > 0 and trial mod pauseAfter = 0
	demo Erase all
	demo Colour: "black"
	demo Select inner viewport: 0, 100, 0, 100
	demo Axes: 0, 100, 0, 100
	demo Paint rectangle: "{'background$'}", 0, 100, 0, 100
	demo Text special: 50, "centre", 80, "half", "Helvetica", 30, "0", "P\r<est\a'vka"
	demo Text special: 50, "centre", 65, "half", "Helvetica", 20, "0", "Te\d< si m\uo\z<ete chv\i'li odpo\c<inout."
	demo Text special: 50, "centre", 60, "half", "Helvetica", 20, "0", "A\z< budete p\r<ipraveni pokra\c<ovat,"
	demo Text special: 50, "centre", 55, "half", "Helvetica", 20, "0", "klikn\e<te nebo stiskn\e<te mezern\i'k."
	while demoWaitForInput ()
		goto ENDPAUSE demoInput ("• ")
	endwhile
endif
```

```
label ENDPAUSE

selectObject: results
Append row
row = Get number of rows
Set string value: row, "stimulus", stim$
Set string value: row, "opt1", opt1$
Set string value: row, "opt2", opt2$
Set string value: row, "correct", correct$

trial = trial + 1
replayed = 0
correct = 0
wrong = 0

played = 0
@trialScr: "'background$'", "'background$'", "'background$'"
stopwatch
@playSound
played = 1
@trialScr: "'buttons$'", "'buttons$'", "'background$'"

label REPEAT_MAIN

### This "while cycle" enables the user to click the buttons:
while demoWaitForInput ( )
	for k to nButtons
		x = x'k'
		y = y'k'
		if replayed = 0 and demoClickedIn (46, 54, 48, 55)
			replayed = 1
			@trialScr: "'background$'", "'background$'", "'background$'"
```

```
                stopwatch
                @playSound
                @trialScr: "'buttons$'", "'background$'", "'background$'"
                selectObject: results
                Set string value: row, "replayed", "yes"
                goto REPEAT_MAIN
            elif demoClickedIn (x, x+xsize, y, y+ysize) or demoInput ("k")
                rt = stopwatch
                selectObject: results
                Set string value: row, "rt", fixed$ (rt, 3)
                Set string value: row, "response", opt'k'$
                if opt'k'$ = correct$
                    correct = k
                    Set string value: row, "rw", "R"
                    cor'stim$' = cor'stim$' + 1
                    # The columns "cor" and "incor" can be deleted from the "results" table.
                    # They only make the values "visible" in the table, but the script runs without them.
                    Set numeric value: row, "cor", cor'stim$'
                    Set numeric value: row, "incor", incor'stim$'
                else
                    wrong = k
                    Set string value: row, "rw", "W"
                    incor'stim$' = incor'stim$' + 1
                    Set numeric value: row, "incor", incor'stim$'
                    Set numeric value: row, "cor", cor'stim$'
                    # This "if jump" adds incorrectly identified items to the "stimList":
                    if incor'stim$' >= repIncor and r > 1
                        selectObject: stimList
                        currentRows = Get number of strings
                        Insert string:  currentRows + 1, stim$
```

```
						endif
						selectObject: results
					endif
					Save as tab-separated file: "results/'subjCode$'_results.Table"

					@trialScr: "'background$'", "'background$'", "'buttons$'"

					### This "while cycle" makes it possible to click the "new item" button:
					while demoWaitForInput ()
						goto DECIDED_MAIN demoClickedIn (46, 54, 40, 47) or demoInput (" ")
					endwhile
				endif
			endfor
		endwhile
		label DECIDED_MAIN
		if trial = lastTrial
			goto FINAL
		endif
		label SKIPITEM
	endfor

	# Reset values and create a new "results" table after the PRACTICE round:
	if r = 1
		trial = 0
		selectObject: stimList
		nStimuli = Get number of strings
		for i to nStimuli
			stimFull$ = Get string: i
			stim$ = stimFull$ - ".wav"
			incor'stim$' = 0
```

```
				cor'stim$' = 0
			endfor
			selectObject: results
			Remove
			results = Create Table with column names: "results", 0, "stimulus opt1 opt2 correct cor incor response rw replayed rt"


			demo Erase all
			demo Colour: "black"
			demo Select inner viewport: 0, 100, 0, 100
			demo Axes: 0, 100, 0, 100
			demo Paint rectangle: "{'background$'}", 0, 100, 0, 100
			demo Text special: 50, "centre", 70, "half", "Helvetica", 30, "0", "Konec n\a'cviku"
			demo Text special: 50, "centre", 50, "half", "Helvetica", 20, "0", "Jestli m\a'te n\e<jak\e' ot\a'zky,"
			demo Text special: 50, "centre", 45, "half", "Helvetica", 20, "0", "te\d< m\a'te mo\z<nost se zeptat."
			demo Text special: 50, "centre", 30, "half", "Helvetica", 20, "0", "%%Kliknut\i'm nebo stisknut\i'm mezern\i'ku%"
			demo Text special: 50, "centre", 25, "half", "Helvetica", 20, "0", "%%spust\i'te hlavn\i' \c<\a'st experimentu.%"


			while demoWaitForInput ( )
				goto START_MAIN demoInput ("• ")
			endwhile
			label START_MAIN
		endif
endfor



		#################
		## FINAL SCREEN
		#################
```

```
label FINAL
demo Erase all
demo Colour: "black"
demo Select inner viewport: 0, 100, 0, 100
demo Axes: 0, 100, 0, 100
demo Paint rectangle: "{'background$'}", 0, 100, 0, 100
demo Text special: 50, "centre", 55, "half", "Helvetica", 30, "0", "Konec"
demo Text special: 50, "centre", 40, "half", "Helvetica", 20, "0", "D\e<kujeme za Va\s<i \u'\c<ast
na experimentu!"

selectObject: stimList
plusObject: results
plusObject: pairs
Remove


        #################
        ##  PROCEDURES
        #################

procedure trialScr: .buttons$, .replay$, .next$
        demo Erase all
        demo Colour: "black"
        demo Line width: 2
        demo Select inner viewport: 0, 100, 0, 100
        demo Axes: 0, 100, 0, 100
        demo Paint rectangle: "{'background$'}", 0, 100, 0, 100

        demo Paint rounded rectangle: "{'.replay$'}", 46, 54, 48, 55, 2
        demo Draw rounded rectangle: 46, 54, 48, 55, 2
        demo Text special: 50, "centre", 52.5, "half", "Helvetica", 15, "0", "p\r<ehr\a't"
        demo Text special: 50, "centre", 50.5, "half", "Helvetica", 15, "0", "znovu"
```

```
	demo Paint rounded rectangle: "{'.next$'}", 46, 54, 40, 47, 2
	demo Draw rounded rectangle: 46, 54, 40, 47, 2
	demo Text special: 50, "centre", 44.5, "half", "Helvetica", 15, "0", "dal\s<\i'"
	demo Text special: 50, "centre", 42.5, "half", "Helvetica", 15, "0", "slovo"

	### This "for cycle" draws the two option buttons in appropriate colours:
	for k to nButtons
		x = x'k'
		tox = x+xsize
		y = y'k'
		toy = y+ysize
		selectObject: results
		opt$ = Get value: trial, "opt'k'"
		if correct = k
			demo Paint rounded rectangle: "{0, 1, 0}", x, tox, y, toy, 3
		elif wrong = k
			demo Paint rounded rectangle: "{1, 0, 0}", x, tox, y, toy, 3
		else
			demo Paint rounded rectangle: "{'.buttons$'}", x, tox, y, toy, 3
		endif
		demo Draw rounded rectangle: x, tox, y, toy, 3

		### This "if jump" hides the button labels befor playing the sound:
		if played = 1
			demo Text special: (tox+x)/2, "centre", (toy+y)/2, "half", "Helvetica", 42, "0", "'opt$'"
		endif
	endfor
	demo Text special: 50, "centre", 90, "half", "Helvetica", 36, "0", "##Kter\e' slovo jste sly\s<eli?#"
endproc

procedure playSound
	selectObject: stimList
```

```
        sound$ = Get string: i
        Read from file: "stimuli/or/'sound$'.wav"
        Play
        Remove
endproc
```