Filozofická fakulta Univerzity Palackého v Olomouci
Katedra anglistiky a amerikanistiky

Czech Learners' Implicit Knowledge of English Prosody

(Bakalářská diplomová práce)

Autor: Silvie Pavlíková
Vedoucí práce: Mgr. Václav Jonáš Podlipský, PhD.

Olomouc 2014

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

V Olomouci dne ……

……

Acknowledgements

I would like to thank my supervisor Mgr. Václav Jonáš Podlipský, PhD. for his immense help, encouragement and patient guidance he provided me throughout the process of writing this thesis. I would also like to thank all the participants who volunteered to take part in the research.

Contents

1 Introduction

Normally produced connected speech is usually not monotonous. Every language uses specific features that create the melodic and temporal dynamism of speech and these are called prosodic or suprasegmental features. These features can be defined as effects created by variations in pitch, stress, and length (Ladefoged 2011, 23). Prosodic features are different in all languages. When someone speaks in a second language (L2), it is therefore often the use of suprasegmental features from the first language (L1) that causes foreign accent.

Foreign accent in speech can be viewed as "the effect of the contact of two phonological systems, one being from a native language (L1) and the other from a non-native, second language (L2)" (Boula de Mareüil 2006). These phonological systems contain segmental as well as suprasegmental features, but while the contribution of segments to foreign accent has often been the subject of research, less attention is paid to the contribution of suprasegmentals (Boula de Mareüil 2006).

However, a foreign-accented prosody probably affects intelligibility far more than learners usually assume. Roach (1991, 120) states an example of this, when he describes the experiment of scientists who were trying to develop machines that would produce speech from a vocabulary of pre-recorded words. Speech so produced was intelligible for some very short sentences, but it was unintelligible in most cases, because the speech was very unnatural with respect to prosody.

As Volín and Skarnitzl (2010a, 271) point out, unnatural prosody, and thus accented speech, can also affect the perception of a speaker's personality, which means that a speaker with a foreign accent may also be the victim of prejudice. For that reason, acquiring the prosodic features of an L2 is an important part of a successful L2 acquisition and thus L2 learners should probably pay more attention to suprasegmentals in their speech, if they want to avoid misunderstandings and undesirable social judgements by the listeners.

Czech and English prosody differs as well. For that reason, L1 Czech speakers of L2 English should be aware of the differences and they should try to adapt their speech to the different prosodic system in order to increase the intelligibility of their speech. Tomková (2008, 130-131) observed that speakers of any language often regard their L1's

dynamic and rhythmic structure as universal and that Czech speakers of English neglect the interaction between stressed and reduced syllables which "renders their comprehension inadequate to the number of years spent studying English and their speech rather dull and monotonous." This again can lead to an undesired reception by native listeners.

Nevertheless, the acquisition of L2 prosody may be considered as one of the hardest parts of learning the second language and the complexity of prosodic system can be the reason why acquisition of suprasegmentals is seemed to be often neglected in English lessons in the Czech grammar schools. The focus of the grammar schools' teachers of English is often aimed on the acquisition of segmental features and the suprasegmentals are mostly ignored. Thus the students usually do not pay much attention to prosody and they might consider the acquisition of L2 prosody as unnecessary, for they might believe that it is sufficient for them to know the vocabulary and grammar to successfully communicate in L2. However, the facts stated above suggest that the learners who would like to achieve proper language acquisition should be acquainted with the prosodic features and stop ignoring them.

The aim of this thesis is to examine the intuitions of first year students in the Department of English and American Studies at the Palacký University for perceiving foreign accentedness caused by L2 prosody in English speech and to evaluate whether the prosodic features affect their judgement of foreign accent, before they are exposed to the explicit knowledge in their phonetic courses. The thesis deals with Czech students' awareness of the prosody and their ability to recognize foreign accent caused by the Czech prosodic features imposed on the English speech, as well as their ability to recognize the natural English prosody. This study explores the students' awareness of differences in Czech and English prosody; therefore, it could be helpful in designing courses for phonetic training. The results may suggest what the students might need to improve, regarding pronunciation and it might help to evaluate how big the focus on prosody should be in the classes.

The thesis is divided into a theoretical and a practical part. The first chapters of the thesis cover the description of English and Czech prosodic system, namely stress, rhythm and intonation, and the description of differences between Czech and English prosody. One chapter of the theoretical section comprises of the summarization of articles and research papers, which are dealing with the effect of L2 prosodic features on speech perception, which were used as an inspiration, for they deal with similar issues as this thesis.

The next chapters cover the practical research. It comprises of an explanation of the method used for the research, its assessment and interpretation of the results. There were two methods used for the research. The first method was a perceptual test based on accent recognition in manipulated recordings of English sentences. The second part consisted in a standardized interview focused on the stress placement and vowel reduction in English words, including an open-ended question about English prosody and accent in general.

2 Differences between Czech and English Prosody

Since this thesis is concerned with the learners' perception of differences between Czech and English prosody, it is important to note the differences. Hence, the following section focuses on the English prosodic system and its comparison with the Czech prosodic system. This chapter comprises general introduction to the issue of different suprasegmental features in Czech and English and it is followed by a more detailed chapters which explain the individual parts of Czech and English prosodic systems; namely stress, rhythm and intonation, and again the differences in these individual features.

As was already mentioned in the introduction, prosody is a combination of suprasegmental features which are superimposed on syllables, as variations in stress, pitch, and also length; although variations in length can affect also single segments of speech (Ladefoged 2011, 23). These features do not occur separately, but they usually combine with one another. Each one of the features can affect the accentedness of speech at various levels, depending on the difference of L1 and L2 prosodic systems. In this part, we will discuss differences of Czech and English suprasegmentals. The individual speech segments are not the focus of this research, although there are many segmental features which are different in Czech and English and which are problematic for Czech learners' acquisition and cause accentedness, e.g. different phonetic categories.

There are many authors describing the prosodic features of English, but not many who deal with the Czech suprasegmentals and the existing studies mention the topic briefly. As a consequence, there are also not many studies, which would explore the differences between Czech and English prosody. The main difference between Czech and English prosody is mostly defined by a theory that English has a stress-timed rhythm, while the rhythm of Czech speech is syllable-timed, which is mentioned for example by Ladefoged (2011, 249). However, he also points out that stress is only one of the factors

that cause different rhythm and that it is useful to divide languages into those that have variable word stress, such as English, those that have fixed word stress, such as Czech, and those that have fixed phrase stress, which is for example French. This comes from the fact that the primary stress of Czech words is fixed to the first syllable, therefore it is not used to differentiate the meaning (Palková 1994, 277), while the placement of stress varies in English, as explained below in the chapter on stress.

That might be one of the reasons why L2 English learners can find it hard to adopt English suprasegmentals properly, as English has a very complicated stress system, because the position of stress placement varies in English words and there are no specific rules how to learn the stress position. It can be said that the stress system in Czech is not as complex as in English. In fact, as Roach (1998, 88) points out, English word stress is so difficult to predict that it is better to learn the placement of stress in each word individually. Since Czech word stress is fixed to the first syllable, it is by no means as complicated as in English.

Prosodic differences between Czech and English also cause the fact that Czech speakers' pronunciation of English is often described as dull and monotonous (Tomková 2008, 130-131). This is partly because Czech does not use as strong stresses as English (Krčmová 2008), meaning that the variance between stressed and unstressed syllable is not as apparent in Czech as in English, which is why English speech with a Czech accent may sound less dynamic. As Kingdon (1959, 160) points out:

> "The difference in prominence between stressed and unstressed syllables is greater in English than in many languages. This applies equally to word stress and to sentence stress, and the latter is used more consistently than in some languages to express the speaker's meaning."

At the same time, Czech-accented English has a narrower pitch range than native English (Volín and Skarnitzl 2010a, 2010b) and, as a result, may sound more monotonous. That is the reason why some Czech speakers of L2 English can sound accented; they might apply the Czech prosodic system on their English speech and even when they put the stresses on the right position in English words, the differences between stressed and unstressed syllables in his speech are not as strong as they would be in a speech from a native English speaker.

As a consequence, many students have troubles with adopting stress and intonational patterns of English, as it sounds unnatural to them and they do not want their speech to sound affected (Tomková 2008). Generally, any learners of L2, usually apply the L1 prosody to their L2 speech, because they are familiar with the L1 prosody from an early age, and it sounds more natural for them. The students have their L1 speech prosody deeply rooted in their minds, and it is difficult to adjust to a new system (Gilbert 2008, 1). They might view English speakers' prosody as exaggerated and some speakers might never speak with a native-like prosody, even at a high proficiency level, because they want to avoid the exaggeration.

2.1 Stress

"Stress is the relative degree of force that is used by a speaker on the syllables he is uttering" (Kingdon 1965, 1). Stressed syllable is thus more prominent and produced with a greater amount of energy than unstressed syllable (Ladefoged 2011, 249). Perhaps the most obvious factor that allows listeners to identify a syllable as stressed is loudness. However, this is not the only cue for recognizing stress in a syllable. Gimson (Cruttender 2008, 236) defines the factors that work together in a syllable to create stress as loudness, pitch, quality and quantity, and he argues that stress is instead principally marked by a pitch change.

Gimson (Cruttender 2008, 235) also states that English stress is fixed in the sense that a primary stress always falls on the particular syllable of any given word, with the exceptions determined by larger rhythmical units, and it is free in the sense that it can fall on any syllable in a word. For instance, the main stress falls on the first syllable in the English words in the (a) example; on the second syllable in the words in the (b) example and on the third syllable in the (c) example.

(1) *radio, window, habit*

(2) *position, computer, banana*

(3) *cigarette, international, corporation*

This applies when these words are said in isolation, but the stress can be moved when the word is a part of a larger unit. The stress placement differs in a Czech connected speech as well, especially in the monosyllabic words and this is why stress can be marked as lexical or prosodic. Lexical stress is also called word stress, which suggests that it is placed on syllables within words and stored as a part of the lexical representation of the word.

Prosodic stress means that lexical stress can be modified when the words are part of a larger unit, for example sentence, which is why it can also be called sentence stress.

2.1.1 Word Stress

The term word stress stands for a stress placed on syllables in individual words, when the word is said in its citation form. Word in its citation form is a word pronounced in isolation in which at least one syllable is fully stressed and there is no reduction of the vowel quality (see Ladefoged 2011, 107).

As already stated, the position of word stress varies in English polysyllabic words. Also, polysyllabic words can have more than one stressed syllable. Beside the primary stress which carries the highest pitch rise, there can be also secondary stress, which is less prominent than the primary stress. English and Czech can both have primary and secondary stresses. Czech words bare a secondary stress usually on the third syllable, or generally odd syllable (Palková 1994, 287), but the secondary stress in Czech is not as prominent, and as frequent as in English.

Stressed syllables are often louder than unstressed syllables, but it is not always the case. Some syllables are more prominent than others due to the vowel quality (Cruttender 2008, 237). The reduced vowel is a sign of unstressed syllable in English, unlike in Czech. As Palková (1994, 279) describes, the quality of Czech vowels is not dependent on its position in a word and it is usually not reduced. Thus the vowel can have a strong form even in the unstressed position, which is impossible for English vowels. Moreover, English speakers usually put the stress on the longest syllable with the full vowel, while the vowel length cannot be used to signal stress in Czech, because it is used to differentiate meaning (Palková 1994, 279).

As suggested above, the stress position varies in English, while in Czech, the primary stress always falls on the first syllable, irrespective of the number of syllables in the word (Ladefoged 2011, 249); of course with the exception of certain monosyllabic functional words (prepositions, auxiliary verbs, etc.) In Czech, there is a tendency for lexical words to keep their stress, while the functional words drop it. Stress does not signal meaning in Czech words, however, it does signal boundaries between words (Palková 1994, 277). On the contrary, stress position in English words can also change the meaning of a word, for example the word *insult* is a noun when it is pronounced with the stress on the first syllable, but it is a verb when the stress falls on the second syllable.

This means, that it is possible for the Czech learners of English to make a mistake in stress placement by applying the rules of the Czech prosodic system and place the word stress always on the first syllables of English words. They might also have troubles to pronounce the English vowels correctly, as Czech vowels are rarely reduced. Hence, when the Czech learners does not learn the stress patterns of English words correctly, they might be misunderstood and they might as well have problems to understand the English words in a spoken form, thus when the words are part of a larger unit (Gilbert 2008, 14).

2.1.2 Sentence Stress

If the syllable in a word is stressed when it is said in isolation, it may become unstressed in a connected speech. As Ladefoged (2011, 116) suggests, when words are part of a larger unit, most frequently, some of the stresses are omitted. Its stress will depend upon its relative importance in the sentence, because the position of stress is determined by the meaning to be conveyed (Kingdon 1965, 8). Therefore, the pitch changes can help to show the connection between words in an utterance and the position of sentence stress depends mainly on the meaning of the utterance.

Words in a natural connected speech are hardly ever pronounced the same way, as when they are in its citation form and the words in connected speech are often distorted so much, that they would be unrecognizable without the context, especially when the speech rate is high (Gilbert 2008, 7). The word forms in a connected speech are usually reduced by various processes, for example assimilation, or elision of sounds. This is also one of the reasons why the experiment to create mechanical speech, which is mentioned in the introduction, failed. Roach (1998, 124) explains that "a significant difference in natural connected speech is the way that sounds belonging to one word can cause changes in sounds neighbouring words." This particular process is called assimilation. The changes of sounds in a connected speech create the natural speech, so when the scientists mentioned by Roach tried to develop mechanical speech, they failed, because they put together words which were pronounced in isolation, therefore the resulting speech sounded unnatural.

Although accentual patterns of connected speech are relatively free, in comparison with the accentual patterns of words, some words are predisposed by their function in a language to receive stress; hence lexical words are usually stressed, functional words unstressed (Cruttender 2008, 264). On the other hand, stress can be also added to the syllables which are normally unstressed, if they play an important part in the utterance, for

example monosyllables can take sentence stress, if they are emphasized (Kingdon 1959, 160).

As mentioned earlier, stress falls on the first syllable in Czech words. This applies when the words are said in isolation, as well as in a connected speech, because stress signals boundaries between words in Czech; e.g. *tabulka* vs. *Ta bulka* (Palková 1994, 277). Therefore, Czech speakers might have a tendency to put the stress on the first syllable in English words as well, which results in an accented speech.

However, there can be also difference between the stress in individual words, when they are said in isolation, and when they are a part of a connected speech. Monosyllabic or functional words are usually unstressed in a connected speech in English and in Czech as well, with the exception of cases when the speaker wants to emphasise the functional word; then it can bear stress.

In contrast, monosyllabic words cannot be said to have lexical stress (Kingdon 1965, 1). Monosyllables are often functional words, and they undergo the most phonetic changes in a connected speech. Words can have either strong or weak form. Strong form means that a word is in its citation form and weak form occurs when a word is in an unstressed position. Functional words occur mainly in the weak form, because they are rarely emphasized in a connected speech (Ladefoged 2011, 109). The monosyllabic words are also usually joined with the following word, so when there are more monosyllabic words following each other, the first one of them is stressed (Palková 1994, 282).

Prosodic stress is closely related to rhythm. It creates the rhythm of utterance, hence if we use stresses in an utterance incorrectly, the rhythm will be wrong as well. The following section describes the speech rhythm in more detail.

2.2 Rhythm

English is often described as a stress-timed language (see Ladefoged 2011, 249), while Czech has proven difficult to classify as either stress- or syllable-timed (Dankovičová and Dellwo 2007). This means that rhythm of an English speech is created by positioning stress in words in a way that avoids having them too close together, so that stress tend to recur at regular intervals. Hence the recursion of stresses in regular intervals is one of the processes to maintain the rhythm (Ladefoged 2011, 118). Therefore, the duration of the utterance in English is determined by a number of stresses and the vowel quality is depending on that, while in Czech, the duration is determined by a number of syllables. That may present a

difficulty during speaking for syllable-timed Czech speakers when speaking stress-timed L2 English.

In English, there seems to be a tendency to avoid succession of weak syllables in words longer than two syllables and to use the alternation of stressed and unstressed syllables with various rhythmic patterns (Cruttender 2008, 135). To maintain rhythm, stress in a word can be either omitted or moved to a different syllable than the one which is stressed in the citation form of the word, according to the context. The interval between stresses is affected by the number of syllables within the stress group, by the number and type of vowels and consonants within each syllable, and by other factors such as the variations in emphasis that are given to each word (Ladefoged 2011, 118).

However, stress is only one of the factors that cause rhythmic differences (Ladefoged 2011, 249). The occurrence of full vowels predicts the rhythm of English more usefully than stress. Syllables with a reduced vowel are shorter than syllables with full vowel and if the full-voweled syllable is immediately followed by a reduced-vowel syllable, it is shortened (Cruttender 2008, 265). Furthermore, there can be difference in the degree of rhythmicity of the speech. Sometimes we can speak very rhythmically, for example during public speaking, or we can speak arhythmically, which means without rhythm, for example when we are hesitant or nervous; for that reason, the characteristic of English as a stress-timed language is not always applicable (Roach 1991, 122-123).

On the other hand, while English vowel quality is a sign of stress placement, Czech syllables are usually not reduced and the position of stress does not depend on their quality, so the pronunciation of English vowel is yet another problematic feature, which may cause accentedness in Czech learners' English speech. In Czech, there are different devices which contribute to the maintenance of rhythm. Czech syllables tend to have the same duration, regardless of the stress placement. As has been said, Czech is a syllable-timed language and for that reason, the quality of vowels does not need to be reduced. That is why the Czech learners might use too many strong forms in their English speech and thus sound unnatural.

However, Czech stress can also be modified to create the rhythm of speech. Monosyllabic words in Czech tend to be unstressed when they are part of an utterance due to rhythm or semantic context. The monosyllabic word can be connected with the preceding or following word, for example pronouns and forms of the verb *be* (and

grammatical words in general) are almost always connected with the previous word, which carries stress at the beginning (Palková 1994, 280).

In order to achieve regular rhythm of an utterance, the vowel quality and length is reduced in unstressed syllables. Vowel quality is often difficult to describe because vowels does not have distinct boundaries, you can go smoothly from one vowel to another, unlike with consonants. The vowel is longest in an open syllable, next longest in a syllable closed by stressed consonant, and shortest in a syllable closed by a voiceless consonant (Ladefoged 2011, 98).

2.3 Intonation

Intonation, or the melody of speech, cannot be specified as precisely as other aspects of speech (Ladefoged 2011, 127). Typically, it is described with a respect to the pitch of the voice. In a normal speech, the pitch of our voice is constantly changing; it is going up or down. We speak with fixed, unvarying pitch only in very unusual situations (Roach 1991, 133). Intonation is based on a prosodic stress, thus it is the pitch pattern in a sentence (Ladefoged 2011, 24). That again proves that all the suprasegmental features are closely related and they work together to create the prosody of speech. As Kingdon (1958, xix) puts it, since intonation is based on sentence stress, and that again on word stress, it is advisable for the student who wishes to learn intonation to have a sound knowledge of stress. The difference between stress and intonation lies in that intonation is an overall melody of an utterance, while the stress marks prominence in individual words.

Pitch is especially important for so called tone languages, for example Chinese, or Vietnamese, where the pitch variations distinguish between the meanings of words. That is not the case of English, where tones do not affect the meaning of individual words (Ladefoged 2011, 260). English speakers use intonation mainly to show the difference between new and old information. The word which becomes the focus of meaning has a stressed syllable marked by a major change in pitch (Gilbert 2008, 19). According to Gimson (Cruttender 2008, 270), the pitch changes signal the primary and secondary stress in the syllables, as well as the division of utterances into intonational phrases. These phrases can also carry discoursal meaning, because the alternation in the direction of pitch change, from rising to falling, can completely change the meaning of an utterance (Kingdon 1958, xxv-xxvi). Intonational phrases most commonly correspond with clauses (Cruttender 2008, 153). They have a falling tendency, unless they are followed by another

intonational phrase. In that case, the intonational phrase which precedes another intonational phrase ends with a continuation rise (Ladefoged 2011, 120). This means that a falling pitch usually signals the end of a grammatical unit, which is typically a normal noniterrogative sentence. The last syllable in an intonational phrase is usually the tonic syllable, which is the syllable with the major pitch change, but if there the context requires an emphasis on another word it might bare the tonic accent, regardless of whether it is at the end of the phrase (Ladefoged 2011, 119).

According to the behaviour of the pitch within one syllable, we distinguish level tones and moving tones. Roach (1998, 135-145) further explains this division, and describes level tone as the one that is not moving; it is a steady tone. Falling tone is the one which descents from a higher to a lower pitch, and a rising tone moves from a higher pitch to a lower pitch. These tones can be used on one-syllable utterances, for example *yes* and *no*. Beside the rising and falling tones, there can be also fall-rise tone, or rise-fall tone. Rising tones indicate new information, while falling or level tones indicate old information. Fall-rise tone is used quite often in English, and it can have various functions, e.g. to show hesitation, limited agreement, etc. In contrast, the rise-fall tone is used to express strong feelings of approval, disapproval or surprise. However, there needs to be a term, which would refer not only to a single syllable, and this term is often called the tone-unit. The smallest tone-unit can be one syllable as well, but it is usually composed of more than one syllable. That is because in English utterance, usually only one prominent syllable carries a tone, and that syllable is called tonic syllable. This syllable carries not only the note, but also prominence, which is called tonic stress.

Furthermore, the pitch change can express the attitude of the speaker and it is used to express feelings, to ask questions or to divide utterances into syntactic units. It is not clear, if the intonational pattern is universal for all languages, but it is apparent, that many languages use the same pattern to convey similar emotional intonation (Ladefoged 2011, 254). To sum it up, Krčmová (2008) states that intonation serves mainly to signal end of an utterance, to signal interrogative utterances, to differentiate between old and new information, and to convey emotions. Roach (1998, 163) puts a sense to it, by saying that "intonation makes it easier for a listener to understand what a speaker is trying to convey."

However, as Roach (1998, 135) also points out, if a L2 English learner is not able to talk regularly with native speakers of English, or at least listen regularly to colloquial English, there is little probability that he will acquire L2 English intonation successfully.

Hence, he suggests that it is not possible to learn English intonation from a set of rules, but rather from listening and talking to English speakers and imitate their speech melody. That is due to the fact that although many languages have the same basic intonational patterns, their realization might differ in intervals between syllables, as well as in waveforms, and thus it can lead to misinterpretation of the L2 utterance (Krčmová 2008). In the research part, I am going to discuss if the students who were subjected to the experiment have the ability to perceive English intonation.

3 Earlier Research

This section deals with the review of studies concerned with the similar subject as this thesis; the perception of foreign accent caused by an unnatural prosody. These articles served as an inspiration, as we can assume that the results would be similar in many aspects, when applied on Czech and English speech.

Boula de Mareüil's (2006) paper is focused on a difference between Italian and Spanish prosody. The experiment that is described in this study comes to the conclusion that prosody is slightly more reliable than the articulation of phonemes in identifying the speaker's origin. The results show that an Italian voice with Spanish prosody is rated more Spanish-like than Italian-like and vice versa. The study also proves that prosody has an effect on the subjects' perceptual judgements and that prosody plays an important role in recognizing foreign accentedness, when other things are equal. The results of this study suggest that Czech learners might perceive English speech with Czech prosody as more Czech-like as well, however there are other factors, which might prevent this result and we cannot generalize it; one of them being that Spanish and Italian are closely related languages, while Czech and English are far more unlike.

Kamiyama's article (2004) investigates the French L1 speakers' perception of accentedness in French speech produced by Japanese learners. The stimuli used for this study were synthetized in two ways. Firstly, the quality of segments was modified, while keeping the prosody intact, secondly F0 and duration were manipulated, while keeping the segments intact. The study's results show that prosody is very important for the evaluation of naturalness of the Japanese speakers by French listeners, and that a native-like prosody could improve the evaluation, even when the speaker uses non-native-like segments. The results also show that intonation may be more important to accent ratings than duration

alone. Hence, the results again prove the importance of prosody. The method used is quite similar to the method used in this research in that they used prosody manipulation of stimuli as well.

Another paper that used similar research method came with recordings rated for foreign accent, which were digitally processed to exchange the pitch and segment durations. Melanie Pinet and Paul Iverson (2010) examined French and English speakers and their research dealt with the accent recognition in noise, and the contribution of prosody to recognizing L1 and L2 speech in noise. Therefore, the method they used was the imposition of French-accented prosody onto the recordings from English speakers. They also studied the role of the listeners' L2 experience in the accent recognition. They came to the conclusion that while:

> "English listeners were more accurate at recognizing L1 English with English prosody, the French inexperienced listeners were more accurate at recognizing French-accented speech with French prosody, and the French experienced listeners varied in the cues that they used depending on the noise level, showing more flexibility of processing" (2010, 1357).

Steven Winters and M. G. O'Brien's (2012) findings echoes the results of Pinet and Iverson (2010) for F0 contours; meaning for the stimuli without noise. This article is concerned with English and German speakers and it investigates the relative contribution of intonation and/or syllable duration to the intelligibility and perceived accentedness to both L1 and L2 speech. They come to the conclusion that intonation can increase (if it is non-native-like) or diminish (if it is native-like) the perceived accentedness of the sentence, therefore, they suggest that intonation is an important feature in the perception of foreign accent.

Volín, and Skarnitzl's paper (2010a) studies the suprasegmental acoustic cues of foreignness in Czech English. The research comprised of an evaluation of Czech speakers of English by native English speakers and by native Czech speakers, experts and naive. Their study suggests a method and cues which might be used to detect and possibly explain the Czech accent in English.

4 Research Questions

This research aims to examine the students' perception of differences between English and Czech prosody and to observe if they are able to recognize the specific prosodic features. The study's goal is to help evaluate students' knowledge in this domain and to help tutors design phonetic courses for the students, so they could learn the correct pronunciation of English suprasegmentals, as they might not be well acquainted with the English prosodic features from their previous educational background.

5 Experiment 1

There were two methods used – a perceptual test and a standardized interview. This section describes the methodology and the results of the perceptual experiment.

5.1 Methodology

The perceptual test comprised of a listening task. The participants listened to 80 pairs of English sentences, deciding which one of the pair sounded more English-like.

5.1.1 Participants

The perceptual test was done with two groups of participants; group of speakers who recorded the sentences and group of listeners who were subjects of the research. One native speaker served as a control subject.

The first group of participants, used to record the sentences, comprised of two speakers. One of them was a native English speaker and the other one was a Czech L2 English speaker. Both were male and students at the Palacký University. The native English speaker was a British postgraduate student at the Palacký University at the Department of English and American Studies. The Czech speaker was also a student at the Department, in the third year of his studies.

The second group of participants comprised of 15 listeners. They were a first year students at the Department of English and American Studies at the Palacký University who volunteered to participate in the experiment. 6 of the students were male, 9 female, all between 18-21 years of age. The participants were chosen from a linguistic introductory course and the experiment was done in October, so that the students would not be familiar with the knowledge from phonetic courses at the time of the experiment.

Part of the listening group of participants was also one native English speaker who served as a control subject. Unfortunately, I was not able to secure more native speakers, but one native speaker should be sufficient as a control for this research. However, the native speaker who participated in the test as a control was not an ideal subject for this research as he has lived in the Czech Republic for a longer period of time and he is well acquainted with Czech-accented English which might have affected the results.

5.1.2 Stimuli

Twenty sentences were used in the experiment (see Appendix 1). All sentences were pre-recorded by a native speaker of English and by a native speaker of Czech. The English recordings were chosen from a set of recorded sentences that had been used for another experiment. The Czech speaker, who could be described as having a relatively strong prototypical Czech accent, then recorded the 20 selected sentences as well. The recording took place in a sound-attenuated booth at Palacký University. Before the actual recording, the speaker was allowed to read the sentences from a sheet of paper to avoid hesitation during the recording caused by unfamiliarity with the sentences. The recording was done using Zoom H4n recorder, with 44100 Hz sampling. Praat (Boersma and Weenink, 2011) was used to manipulate the original recordings and create the actual stimulus sentences. First, both the native English (NE) and the Czech-accented English (CE) utterances were annotated marking boundaries between speech sounds and sometimes even acoustic events within speech sounds (such as the release of a stop). If a boundary between speech sounds was not clearly identifiable (e.g. in some cases the transition between a vowel and a sonorant) the portion was treated as one segment. Importantly, each boundary in one speaker's annotation corresponded to a boundary in the other speaker's annotation and so there was an equal number of segments in both. An example is given in Figure 1.

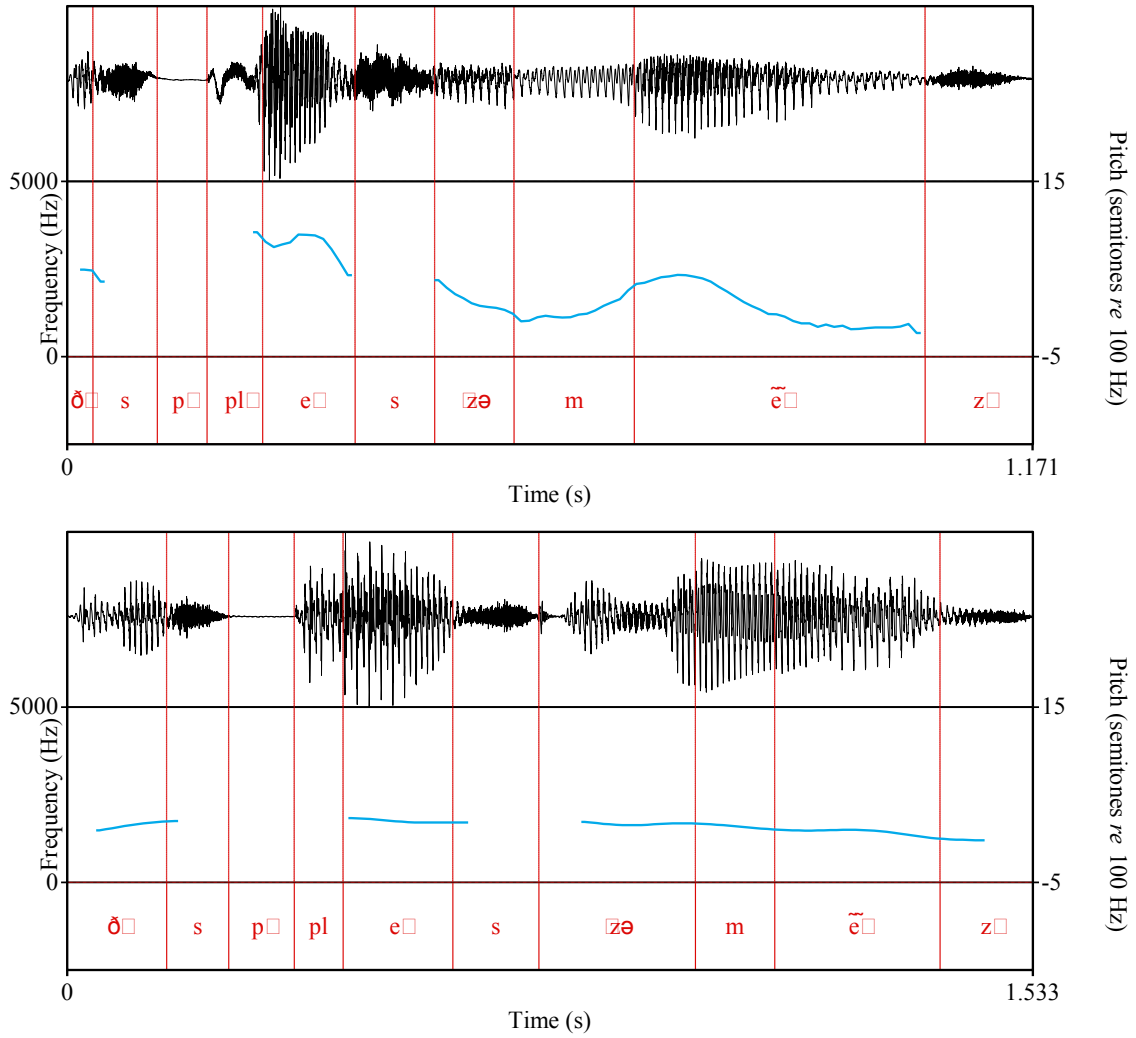

FIG. 1. Waveform and spectrogram of sentence "This place is a maze." with annotated phonetic boundaries. NE speaker's annotation of boundaries in the top panel corresponds with the CE speaker's annotation of boundaries in the bottom panel.

Then, Praat scripts were run to produce the actual warped-duration and warped-intonation stimuli for the perceptual test from each NE sentence using overlap-add resynthesis.[1] For the duration-warped stimuli, duration of each segment in a sentence were manipulated by measuring its durations in the NE and CE sentence and interpolating between these durations logarithmically using the equation

$$t' = e^{\ln(t_1)+\left\{\ln\left(\frac{t_2 s_1}{s_2}\right)-\ln(t_1)\right\}r} \quad (1)$$

[1] I would like to thank Mgr. Václav Jonáš Podlipský, PhD. for producing the manipulated stimuli and creating the test.

where *t'* is the new duration of each segment, t_1 is the duration of the NE segment, t_2 is the duration of the CE segment, s_1 is the duration of the entire NE sentence, s_2 is the duration of the entire CE sentence, and *r* is the interpolation factor, within the interval [0, 1], determining the proportion of the implanted CE temporal pattern to the original NE temporal pattern. For example the *r* of 1 means *t'* takes up as much of the total duration of the sentence as t_2 does in the CE, while the *r* of 0.5 means *t'* is exactly intermediate on a logarithmic scale between the NE segment duration and the CE segment duration scaled to the NE speech tempo (by first adjusting the total duration of the CE sentence to equal that of the NE sentence). In this way, a pair of duration-warped versions of each sentence was produced for listeners to compare: one with *r* = .15 ('native durations'), in other words one that had 85% NE and 15% CE durations, and the other with *r* = 1 ('non-native durations'), in other words one that had 100% CE durations.

F0 contours were then manipulated to produce the warped-intonation stimuli in the following way. First, temporal warping using the method described above was applied to each CE sentence with *r* = 1. That is, the new CE sentence had the total duration and all segment durations exactly equal as the NE sentence. Second, F0 tracks were extracted from this time-warped CE sentence as well as the original NE sentence. Both the F0 tracks were manually corrected for octave jumps and other artefacts. Next, the entire CE F0 contour was shifted so that its mean in semitones matched the NE mean. A new F0 contour then could be created by sampling the CE and NE contours with a 10-ms step and computing the new F0 value *i'* using the equation

$$i' = i_1 + (i_2 - i_1)\, r \tag{2}$$

where i_1 is the NE F0 value, i_2 is the CE F0 value, and *r* is the interpolation factor. In this way, a pair of F0-warped versions of each sentence was produced for listeners to compare: one with *r* = .35 ('native intonation'), in other words one that had 65% NE and 35% CE intonation, and the other with *r* = .85 ('non-native durations'), in other words one that had 15% NE and 85% CE intonation.

The reason why no original recording was used and even the 'native' durations and intonations were partially warped in the direction of non-native speech is that sometimes the resynthesized sentences did not sound perfectly natural. If listeners compared the resynthesized utterances to naturally-produced ones, their responses could reflect perception of the naturalness of speech rather than perception of English-likeness which

the experiment was intended to tap at. For the same reason, each stimulus was mixed with speech-shaped noise produced uniquely for that stimulus using Boersma's (2001) Praat script, with the signal-to-noise ratio of 7 dB. Another reason why the noise was used was to avoid a ceiling effect in listeners' responses. Finally, a 750 millisecond pause was added between the stimuli in each pair.

In sum, two pairs of stimuli were produced for each of the 20 sentences by resynthesizing the original NE recording; 'native' (85% NE, 15% CE) and 'non-native' (100% CE) warped-duration stimuli; and 'native' (65% NE, 35% CE) and 'non-native' (15% NE, 85% CE) warped-intonation stimuli; so that overall, there were 80 stimuli.

5.1.3 Procedure

The perceptual test was done in a quiet classroom at the Department of English and American Studies at the Palacký University using computers and Sanaco circumaural headphones. Control listening with a native speaker was done in a quiet environment as well, using laptop to run the test and Koss circumaural headphones. The test was completed in Praat (Boersma and Weenink, 2013). There were four buttons on the monitor in the test interface: 'first more English-like,' 'second more English-like,' 'can't decide,' and a button to replay the sentences. The students made their decision on each pair and clicked on the respective button. Before completing the actual test, the listeners received information how the test works and what they are supposed to do to finish it correctly. However, they did not know the purpose of the test, as it might have affected their responses; they might have concentrated more on the prosodic features of the stimuli and not on their overall perception. First, they were asked to complete a short demonstration (demo test) to get familiarized with the test interface, adjust the volume, or ask potential questions. The demo test contained eight pairs of the stimuli chosen from the 80 stimuli from the actual test. At the beginning of the short demo, the listeners received following information:

> "An English phrase is shown on the screen while you hear it produced by a computer speech synthesizer in two versions. (You can replay what you heard twice at most by clicking the replay button). Please decide which version of the phrase sound more English-like, the first or the second, and click the respective button. Please try not to base your decision on human-like the synthesis is but really on how close it gets to native English. None of the pairs of sentences are identical."

Because of the fact that some of the sentences sounded more synthetized than others, listeners were thus notified that they should not base their decision on that, but rather focus on the native-likeness. Each sentence was transcribed on the monitor, while being played, so that the listeners could understand the meaning better. The listeners were also told that they should not try to finish the test as soon as possible by randomly clicking on the buttons, but rather take more time and focus on their answers. Usually, the instructions were quite clear, thus the listeners proceeded immediately to the test.

The test comprised of 80 different stimuli, 40 pairs with warped-duration and 40 pairs with warped-intonation, and the position of the correct response (more native-like) was switched; the correct response was on played as the first stimulus for 40 pairs of stimuli and it was played as the second stimulus for the 40 pairs with the same stimuli, thus the listeners overall listened to 160 stimuli in 80 pairs. The listeners were given a possibility to replay each pair of the stimuli twice.

After 27 pairs of stimuli, the listeners could have a short break and proceed whenever they liked after a click. The test was done in two sessions; thus the 15 participants were divided into two groups according to their preferred time. The first session was done with 7 listeners and the remaining 8 students done the test in the following session. However, the conditions were the same for both groups. The test took about 25 minutes. After everyone finished the test, the results were extracted from each computer and saved as a text file.

5.2 Results

Overall, there were more correct than incorrect responses for both manipulated dimensions (intonation-warped stimuli and duration-warped stimuli). Means of the Czech listeners' results showed that they had a mean number of 13.07 correct responses for intonation, 4.4 incorrect responses and they could not decide on 2.53 intonation-warped stimuli, as shown in the left column in the graph below (see Figure 2). The right column shows the results for duration-warped stimuli: the listeners had a mean number of 9.93 correct responses, 4.3 incorrect responses, and 5.77 for the 'can't decide' answer, therefore the number of correct responses was higher for the stimuli with manipulated intonation.

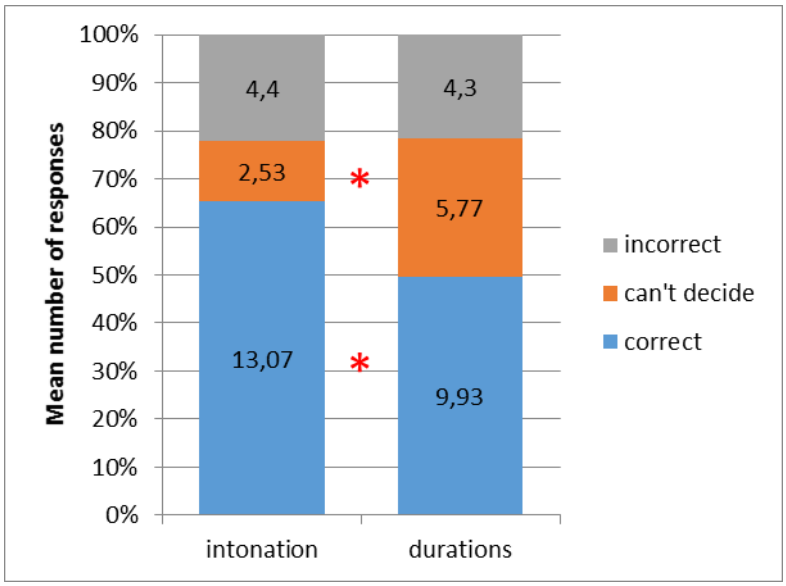

FIG. 2. Graph of the mean number of responses for stimuli with manipulated intonation and duration. Asterisks mark the significance of results (see ANOVAs results below).

5.2.1 ANOVA

Three separate repeated measures ANOVAs (Analysis of variance) were performed with the data to discover which variables were significant. The responses 'can't decide,' 'correct answer' (more native-like), and 'incorrect answer' (less native-like) were dependent variables for each measurement with manipulated dimension (intonation, duration) and position of correct phrase as the within-subject factors; there was either correct-incorrect order, or incorrect-correct order of the phrases and the correct phrase meant the more native-like stimulus.

In the first measurement, the dependent variable was the number of responses "can't decide". The results showed a significant effect ($p>.01$) for the manipulated dimension: $F(1,14)=10.01$, $p=.007$, while there was no significant main effect for the position of correct phrase: $F(1,14)=10.01$, $p=.180$ and neither was there a significant effect for the interaction of the two effects $F(1,14)=.08367$, $p=.777$, which is shown in the Figure 3.

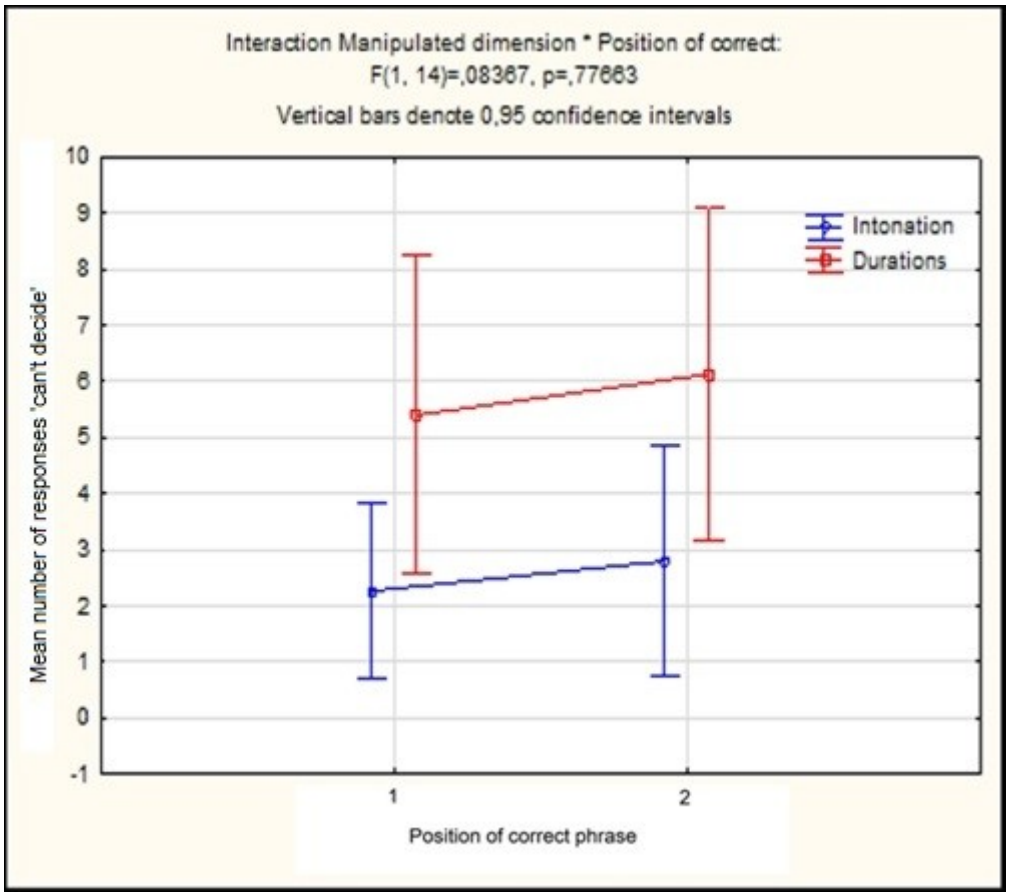

FIG. 3. ANOVA results of the interaction of position of correct phrase and manipulated dimension with dependent variable "number of responses 'can't decide'."

In the second measurement, the dependent variable was the number of correct responses. There was a significant effect for the manipulated dimension: F(1,14)=9.646, p=.008, and the position of the correct answer was not significant: F(1,14)=2.409, p=143. The interaction of the manipulated dimension and the position of correct phrase was not significant as well, F(1,14)=.000, p=1.000, see Figure 4.

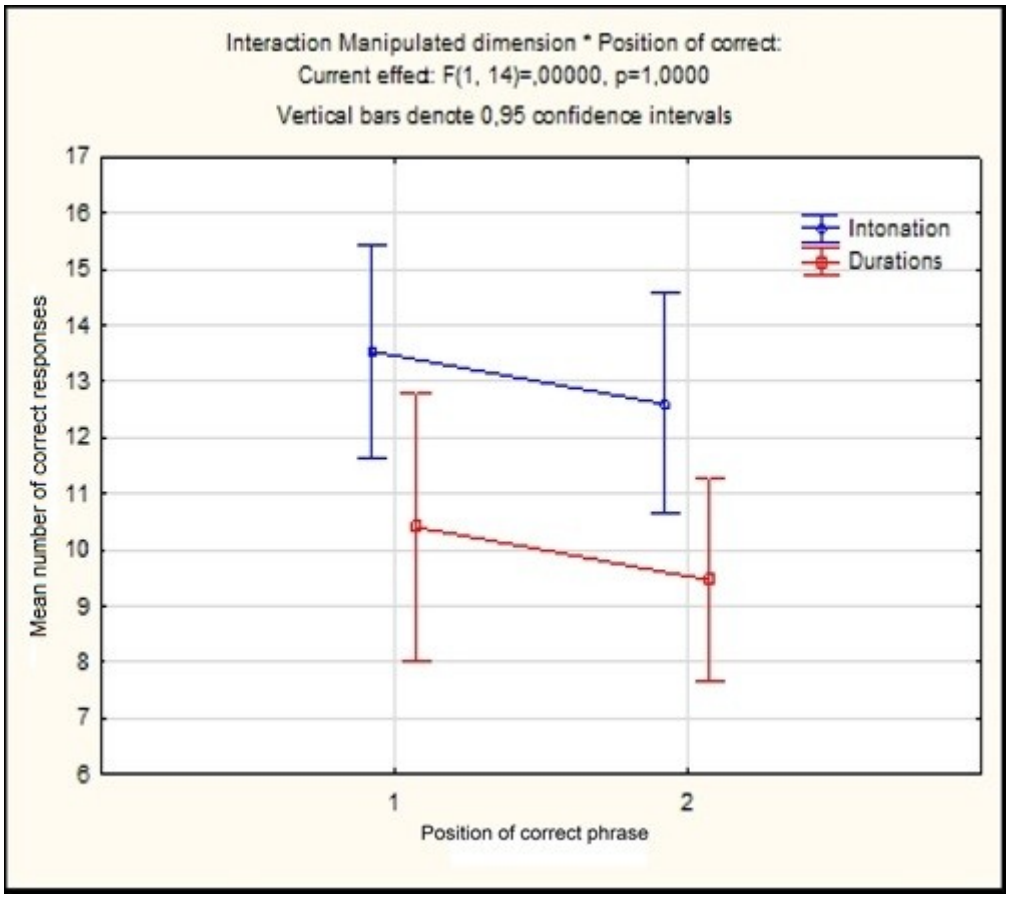

FIG. 4. ANOVA results of the interaction of position of correct phrase and manipulated dimension with dependent variable 'number of correct responses'.

In the third measurement, the dependent variable was the number of incorrect responses. The Manipulated Dimension effect was no significant: F(1,14)=.045, p=836, neither was the Position of correct phrase: F(1,14)=.259, p=619, or the interaction between them: F(1,14)=.033, p=859, as can be seen in the Figure 5 below.

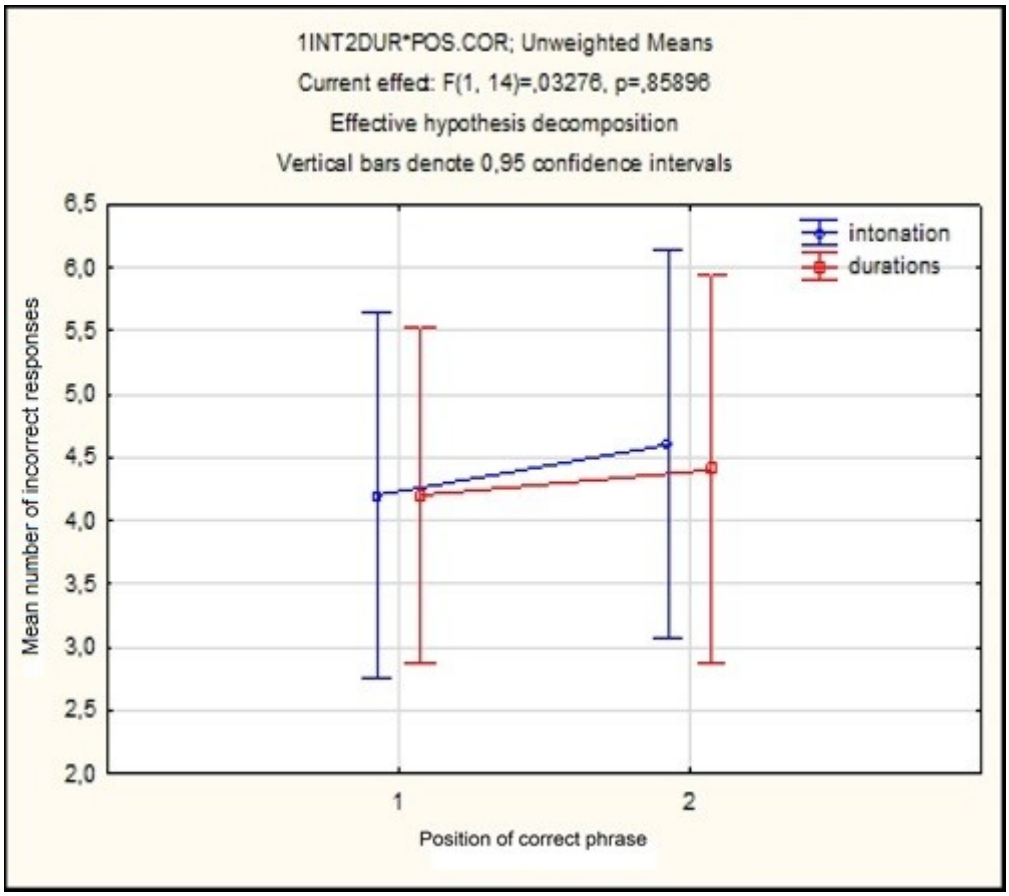

FIG. 5. ANOVA results of the interaction of position of correct phrase and manipulated dimension with dependent variable 'number of incorrect responses.'

Therefore, there was a significant effect of the Manipulated dimension for the 'correct answers' and for the answers 'can't decide' but not for the incorrect answers. Mainly, the ANOVAs results showed that there was no significant effect for the position of the correct (more native-like) answer – the position was no important in deciding, although the results with the dependent variable 'number of correct responses' showed slight preference of the first answer for the intonation-warped stimuli, but the difference was not significant and it had no effect on the results.

5.2.2 Results on individual stimuli

The results revealed that some sentences were harder to categorize as native-like or non-native-like than others. Overall, the hardest sentence to recognize with regard to the incorrect answers was the sentence '*Should I turn left now?*' for duration-warped stimuli with mean number of 3.5 correct answers, as well as for intonation-warped stimuli with mean number 6.5 correct answers, while the sentence with the lowest number of 'can't decide' responses (mean: 1.25), as well as correct answers for both duration and intonation

(mean number: 11.75), was *'Nobody had told me that Jessica was pregnant.'* The sentence which got the most answers 'can't decide' was *'How do you feel now?'* In most cases, the sentences seemed as equally hard to categorize for both duration and intonation, however there were some interesting exceptions. For instance, the sentence *'You locked yourself out?'* got the biggest number of correct responses for duration-warped stimuli, but it was in the worse section of the results for intonation.

5.2.3 Control Subject

The results of the native speaker who served as a control subject were very close to the results of the Czech learners with even less correct responses for duration-warped stimuli than the Czech listeners from the first test. The graph below shows means of the native speaker's results. There were two responses for each pair, one for the pair where the more native-like sentence appeared first, and one for the pair where the more native-like sentence appeared as second. The means were therefore as follows: 13 correct answers for intonation-warped stimuli, 4.5 incorrect and 2.5 'can't decide' answers. For duration-warped stimuli, there was mean number of 7.5 correct responses, 5.5 incorrect, and mean number of responses 'can't decide' was 7.

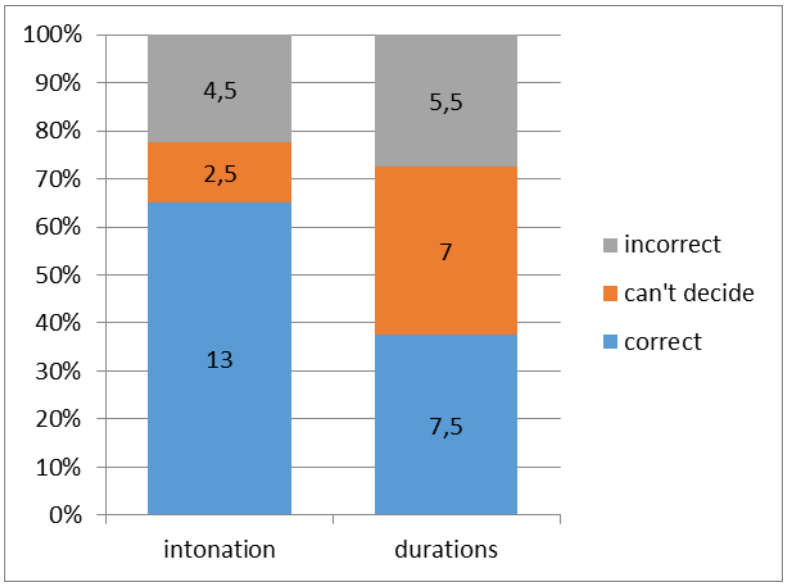

FIG. 6. Graph of the mean number of native speaker's responses.

5.3 Discussion

The results which were described in the section above suggest that the difference between native-like durations and non-native-like durations seems to be less noticeable for Czech learners than difference between native-like and non-native-like intonation. The same

suggestion applies to the native speaker. The hypothesis was that students would have a difficulty in recognizing the native-like intonation and durations and that a native speaker would be more successful in this recognition. However, the results showed that the native speaker had similar difficulty as the Czech students with recognition of the native-like prosody and his results were even worse than the results of Czech learners for duration-warped stimuli. The fact that the native speaker who participated in this experiment had this results might be caused by the fact that he is used to listen to the Czech-accented English, as he has lived in the Czech Republic for a long time and he is a professor at the Department of English and American Studies at the Palacký University. This result raises the question of the importance of an L2 experience in the perception of foreign accent in the L2. Another factor that might have influenced the results from the native speaker could have been his age, as he is about twice the age of the Czech students who participated in the perceptual test.

The findings of this study suggest that students can to some extent perceive the native English intonation or duration of speech. The results show that the listeners were not deciding randomly, because they have a rather similar number of incorrect responses for both duration and intonation, but it can be seen that they were not as sure when they were deciding on pairs with manipulated durations as when they were deciding on intonation-warped stimuli, because they have more 'can't decide' answers for stimuli with manipulated duration; they rather choose the answer 'can't decide' then randomly clicking on the buttons.

6 Experiment 2

As the second experiment, a standardized interview was used to further examine the students' intuition about prosody.

6.1 Methodology

Second experiment composed of a standardized interview. Three exercises were presented to three first year students, who kindly agreed to take part in an individual session. The exercises comprised of a set of words and the students were supposed to read the words out loud and then read them again and decide where lays the word stresses (see Appendix 2). During the interview, the participants were presented with pairs of English words differing in stress placement and whether a corresponding vowel was full or reduced. The aim was

to elicit responses about prosody without using any linguistic terminology, unless it was used by the participants themselves. Finally, they were asked an open-ended question about English prosody.

6.1.1 Participants

Three students participated in the second experiment, two male (J. and R.) and one female (M.). All of them were the first year students at the Department of English and American Studies at the Palacký University, therefore they were between 18-21 years of age as well as the first group of students who did the perceptual test. The male students also participated in the perceptual test a few weeks before this interview. The interview was done at the end of October as well so that the students would not have been familiar with a theoretical knowledge of English prosody from their courses at the Department.

6.1.2 Procedure

Student's answers were recorded in a sound-proof booth at the Palacký University at a Department of Czech Studies. The participants read the words and the pairs of words from three exercises (see Appendix 2). The words were presented in random order within each exercise and they were chosen according to their stress placement and vowel reduction. The first exercise comprised of a set of 45 words with different stress position; 8 words with the first syllable unstressed, starting with a consonant and followed by an unstressed vowel; 8 words starting with an unstressed vowel; 8 words with the first syllable stressed and some unstressed; 13 'filler' words which had no vowel reduction; and 8 words with the last syllable containing either /i/ or schwa. The second and the third exercise comprised of the pairs of words. The second exercise contained 32 pairs of words (the words within a pair had the same root) 16 of which had no shift in the stress position – 8 pairs had the first syllable stressed, and 8 had the second syllable stressed; 16 had a shift in a stress position. The third exercise were again pairs of words, but this time, they had the same form, but their syntactic function was different. Each phrase of these pairs had a stress shift according to the function. There were 16 pairs of which eight pairs had no vowel difference, just a difference in stress placement and eight had a difference in stress placement as well as a different vowel pronunciation.

The participants were asked to compare the words and possibly to tell the difference between them, or to say if there is any difference at all. After each exercise we discussed the words and their stress placement. At the end, the participants were asked

some open-ended questions about English prosody and they were asked to say some typical cues of foreign accent which can be recognized in Czech-accented English speech.

6.2 Results

As has been said, the participants were presented with three exercises which consisted of words with different stress placement. Their answers were very hesitant and it seemed that they most probably do not have any theoretical knowledge of English prosody. They were not able to guess the stress placement in most of the words, as they were not exactly sure what they should imagine under the term 'stress' or 'accent'. As J. mentioned, he is never sure where to put stresses in English words. Their answers suggest some perceptual ability, however, when they were asked to recognize stressed syllables in the words, their answers were usually incorrect.

As for the general questions about the foreign accent, all of them mentioned segmental features of a language first. J. mentioned also the stress and that English has a more prominent stresses than Czech and that Czech is 'lazier' in that case, which suggest that he might be aware of the fact that English uses stronger stresses than Czech. However, when he was supposed to recognize stresses within the words in the exercises he decided randomly and a lot of words sounded the same to him; without the change in stress placement.

R. could not decide on the stress placement as well, saying that all of the words in pairs in the second and third exercise sound the same. However, he mentioned the fact that Czech rhythm is regular, while English is not and he said that this in fact one the main factors causing foreign accented speech in Czech learners, together with different pronunciation of segments, such as /θ/ pronounced as /t/, /f/, or /s/. When describing the pair of words, he said incorrectly that verbs has a stress on the first syllable and noun mostly on the second syllable. However, he noticed that in the pairs *to record a record*, *to subject / a subject*, *to object / an object*, *to project / a project* is a vowel reduction.

M. said that Czech has a different 'system of pronunciation' and that English and Czech speakers differ in a way they 'adjust their vocal tract.' She said that intonation of speech depends on the origin of the speaker. For example, if the speaker is from the eastern part of England, he speaks differently than the speaker from the western part. As the first cue in recognition of Czech-accented English she mentioned segmentals: Czech /r/ which is not present in English and some Czech speakers use it in their English speech. She also

suggested that verbs have a rising intonation and nouns have a falling intonation is a pairs like *to increase / an increase*, *to import / an import*, *to address / an address*, but in most of the cases she regarded the pair of words in the exercises as the same in case of stress, for example even with the similar pair to the pairs which she regarded as having falling or rising intonation; *to insult / an insult.*

6.3 Discussion

The hypothesis of this research was that Czech learners are not familiar with prosodic features from their educational background (i.e. grammar school classes) and the results from the interview confirmed the hypothesis. The individual interview showed that although they are able to perceive foreign accent caused by L2 prosody imposed on L1 speech to a certain extent, as suggested by the first perceptual experiment, they are not trained to produce it and to give it importance when they are learning the L2. The students who participated in the interview were also not able to find stresses within words in most of the cases and when they were supposed to recognize differences in pairs of words with the same root, usually they were not able to do it. In most cases, they regarded the words as having the stress on the same syllable, even when it seemed that they pronounce it correctly. The open-ended questions about prosody showed that they know that there is a difference between Czech and English prosody, but usually they were not able to describe it.

7 Conclusion

To conclude, this thesis aimed to point out the importance of prosody in the perception of foreign accentedness and to examine if the first year students at the Department of English and American studies at the Palacký University are able to perceive the difference between Czech and English prosody, and also to discover if they find the prosodic features important in a perception of foreign accent.

The two methods were used to examine their perception. The first method was a perceptual test. Fifteen Czech learners and one native speaker participated in the test. The test comprised of a listening task. The participants listened to the pairs of stimuli and they were deciding which one of the pair sounds more English-like. There were two types of stimuli; stimuli with warped-intonation and stimuli with warped-duration. The sentences were manipulated with the use of native English recordings and Czech-accented English

recordings. The results showed that the students are able to perceive intonation to some extent, and even duration, but their results for perceiving duration-warped stimuli were worse than for intonation-warped stimuli. However, while the native speaker had similar results for stimuli with manipulated intonation, he had even worse results for the stimuli with manipulated duration than the Czech learners.

The second method was a standardized interview which was done with three participants. The aim of the interview was to further examine the students' perception of English prosody. There were three exercises which were used as a guide during the interview. The exercises comprised of words and pairs of words and the students were asked to determine he stress placement in these words. Then they were asked open-ended questions about English prosody and foreign accent in general. Their answers suggest that they are not familiar with the theoretical background for prosodic features and it is hard for them to recognize them.

8 Shrnutí

Tato práce je zaměřena na prosodické rysy angličtiny a češtiny, jejich rozdíly a na to, jak čeští studenti, kteří nastupují do prvního ročníku studia na Katedře anglistiky a amerikanistiky Univerzity Palackého tyto rozdíly vnímají, a jestli je dokáží rozpoznat. Práce je rozdělena na teoretickou a praktickou část. V teoretické části se zaměřuji na popsání prosodických rysů angličtiny a češtiny a jejich rozdíly, tak jak byly popsané v citované literatuře. Prosodické rysy jsou jedním z důležitých faktorů, které můžou způsobit cizí přízvuk, když mluvíme jiným než naším rodným jazykem, proto je pro studenty, kteří se chtějí naučit anglicky na výborné úrovni, klíčové mít o prosodii povědomí, obzvlášť v případě, že se jedná o tak rozdílné jazyky jako je anglický a český jazyk. Jako jeden z hlavních rozdílů mezi anglickou a českou prosodií se uvádí to, že angličtina mění pozici přízvuku ve slovech, v rámci udržení rytmu řeči, zatímco čeština má přízvuk vždy na první slabice, s výjimkou některých slov v dlouhé výpovědi, které přízvuk ztrácejí.

Výzkumná část práce se skládá ze dvou částí. První část praktického výzkumu je zaměřena na intonaci a délku trvání. Skládá se z percepčního testu, v rámci kterého posluchači poslouchali páry vět s manipulovanou intonací, nebo délkou trvání, z nichž jedna věta z páru byla blíže anglickému mluvčímu a druhá mluvčímu s českým přízvukem.

Posluchači rozhodovali v rámci každého páru, která věta zní více jako od rodilého mluvčího angličtiny. Cílem bylo zjistit, jestli tito posluchači, tedy studenti prvního ročníku na Katedře anglistky a amerikanistiky, dokáží rozpoznat anglické věty s intonací rodilého mluvčího od anglických vět s českou intonací a to samé s délkou trvání výpovědi. Percepčního testu se účastnilo 15 studentů. Výsledek testu ukázal, že posluchači dokáží rozpoznat ve většině případů rozdíl, a to zejména u vět s intonací, pro které byly výsledky o něco lepší než pro věty s pozměněnou délkou trvání. Test byl proveden také s rodilým mluvčím angličtiny, aby byly k dispozici kontrolní výsledky.

Druhá část výzkumné části se skládala ze strukturovaného rozhovoru, který měl za cíl zjistit další poznatky o vnímání anglické prosodie českými mluvčími. Tohoto testu se účastnili tři studenti. Rozhovor byl prováděn s pomocí tří cvičení se slovy a páry slov s odlišně umístěným slovním přízvukem. Ptala jsem se studentů, jestli dokáží rozpoznat přízvuk v těchto slovech, a také, co je podle nich hlavním znakem cizího přízvuku, když někdo mluví takzvanou českou angličtinou. Z těchto třech individuálních rozhovorů vyplynulo, že studenti nemají z předchozí výuky angličtiny téměř žádné znalosti o anglické prosodii, ale některé rozdíly jsou schopni vnímat.

9 References

Boersma, Paul & Weenink, David (2013). Praat: doing phonetics by computer [Computer program]. Version 5.3.56, retrieved 2 October 2013 from http://www.praat.org/

Boersma, Paul. 2001. Re: [praat-users] Speech shaped noise. Online at https://uk.groups.yahoo.com/neo/groups/praat-users/conversations/messages/59. Accessed on April 4, 2014.

Boula de Mareüil, Philippe, and Bianca Vieru-Dimulescu. 2006. "The Contribution of Prosody to the Perception of Foreign Accent." *Phonetica* 63:247–267.

Cruttenden, Alan, and A. C. Gimson. 2008. *Gimson's Pronunciation of English*. London: Hodder Education.

Dankovičová, Jana, and Volker Dellwo. 2007. "Czech speech rhythm and the rhythm class hypothesis." In *Proc. 16th ICPhS*, 1241-1244.

Gilbert, Judy B. 2008. *Teaching Pronunciation.* Cambridge: Cambridge UP.

Kamiyama, Takeki. 2004. "Perception of Foreign Accentedness in L2 Prosody and Segments: L1 Japanese Speakers Learning L2 French." *Speech Prosody 2004, Nara: Japan (2004).*

Kingdon, Roger. 1959. *The Groundwork of English Intonation.* London: Longmans.

Kingdon, Roger. 1965. *The Groundwork of English Stress.* London: Longmans.

Krčmová, Marie. 2008. *Fonetika a fonologie*. Elportál, Brno: Masarykova univerzita. <http://is.muni.cz/elportal/?id=766384>.

Ladefoged, Peter, and Keith Johnson. 2011. *A Course in Phonetics, Sixth Edition*. Boston: Wadsworth/Cengage Learning.

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

Pinet, Melanie, and Paul Iverson. 2010. "Talker-listener accent interactions in speech-in-noise recognition: Effects of prosodic manipulation as a function of language experience." *The Journal of the Acoustical Society of America* 128:1357-1365.

Roach, Peter. 1998. *English Phonetics and Phonology, Second Edition*. Cambridge: Cambridge UP.

Tomková, Kateřina. 2008. "Perception of Non-Native Pronunciation of English by Native Speakers. " Doctoral dissertation, Masaryk University Brno.

Volín, Jan, and Radek Skarnitzl. 2010a. "Suprasegmental Acoustic Cues of Foreignness in Czech English." Paper presented at the Speaker and Language Recognition Workshop: Odyssey 2010, Brno, Czech Republic, 28 June – 1 July.

Volín, Jan, and Radek Skarnitzl. 2010b. "The strength of foreign accent in Czech English under adverse listening conditions." *Speech Communication* 52:1010-1021.

Winters, Stephen, and Mary Grantham O'Brien. 2012. "Perceived accentedness and intelligibility: The relative contributions of F0 and duration." *Speech Communication* 55:486-507.

10 Appendices

10.1 Recorded Sentences

1. It was a great gig.
2. The security may not let me in.
3. Have the new brake pads arrived yet?
4. I need some simple ideas for the costume.
5. Was he still conscious when the aid finally got to him?
6. What shall I tell them to cut?
7. You must be rolling in money now, I guess.
8. What do you think about my performance?
9. I think everyone in the band played well.
10. Are you sure you didn’t put the phone in your bag somewhere?
11. We’ll never get there in time like this.
12. Nobody had told me that Jessica was pregnant.
13. Can you remember the last place you used it?
14. I think we’ve been in this street before.
15. Last time I heard from you, you were still travelling.
16. Should I turn left now?
17. I didn’t know you were back.
18. You locked yourself out?
19. This place is a maze.
20. How do you feel now?

10.2 Exercises used for the interview

10.2.1 First Exercise

1. Words beginning with unstressed syllable (consonant and unstressed vowel)

1. potato

2. photographer

3. commitment

4. position

5. connection

6. computer

7. proposal

8. banana

2. Words beginning with unstressed vowel

1. oppose

2. omission

3. opinion

4. offend

5. occasion

6. about

7. away

8. abandon

3. Words with the first syllable stressed, some unstressed

1. doctor

2. document

3. normally

4. random

5. harmony

6. painter

7. surgery

8. summary

4. Fillers – words without vowel reduction

1. Monday

2. radio

3. habit

4. window

5. baggage

6. basic

7. escape

8. easy

9. backpack

10. airline

11. roommate

12. friendly

13. duty

5. Words with last syllable containing either /ɪ/ or schwa

1. kitchen

2. private

3. secret

4. minute

5. market

6. boneless

7. chocolate

8. separate (Adj)

10.2.2 Second Exercise

Pairs of words with no stress shift

1. First syllable stressed

1. useful / usefulness

2. minimize / minimal

3. prominent / prominence

4. regulate / regular

5. understand / understanding

6. criticize / critical

7. character / characterize

8. confidence / confident

2. Second syllable stressed

1. accept / acceptance

2. develop / development

3. begin /beginner

4. collect /collection

5. confused / confusion

6. construct / construction

7. relate / relation

8. decide / decision

Words with stress shift

1. First syllable stressed

1. diplomat / diplomacy

2. company / companion

3. history / historical

4. product / producer

5. climate / climatic

6. policy / political

7. photograph / photography

8. access / accessible

2. Second syllable stressed

1. apply / application

2. inform / information

3. philosophy / philosophical

4. create / creativity

5. geology / geological

6. pornography / pornographic

7. catastrophe / catastrophic

8. connect /connectivity

10.2.3 Third Exercise

1. Group A – no V difference, just stress

1. to break up / a break-up

2. to insult / an insult

3. to increase / an increase

4. to overflow / an overflow

5. to work out / a workout

6. to turn out / a turnout

7. to import / an import

8. to draw back / a drawback

2. Group B – stress and V difference

1. to record / a record

2. to object / an object

3. to address / an address

4. to conflict / a conflict

5. to contrast / a contrast

6. to subject / a subject

7. to present / a present

8. to project / a project

11 Annotation

Title: **Czech Learners' Implicit Knowledge of English Prosody**

(Bachelor Diploma Thesis)

Study Programme: English Philology

Supervisor: Mgr. Václav Jonáš Podlipský, PhD.

Author: Silvie Pavlíková

Number of pages: 45

Number of characters: 73 191

Number of appendices: 2 + 1 CD

Number of references: 17

Olomouc 2014

Key Words: prosody, intonation, duration, rhythm, accent, perception of foreign accent

Description: The aim of this thesis is to explore intuition of the first year students at the Department of English and American Studies for recognizing English prosody, intonation and duration in particular. The research is also aimed to study how students perceive the importance of prosody for L2 acquisition and foreign accent recognition. The thesis is divided into two parts. The theoretical part describes the prosodic features of English and Czech and their differences by reviewing the relevant literature. The practical part comprises of two methods used for the research; a perceptual test which was done by 15 first year students and one native speaker and a standardized interview which was done with three first year students.

12 Anotace

Implicitní znalosti českých mluvčí o anglické prosodii

(Bakalářská diplomová práce)

Studijní program: Anglická filologie

Vedoucí práce: Mgr. Václav Jonáš Podlipský, PhD.

Autor: Silvie Pavlíková

Počet stran (podle čísel): 45

Počet znaků: 73 191

Počet příloh: 2 + 1 CD

Počet titulů použité literatury: 17

Olomouc 2014

Klíčová slova: prosodie, intonace, délka trvání, rytmus, přízvuk, percepce cizího přízvuku

Charakteristika práce: Tato práce se zabývá intuicí studentů prvního ročníku na Katedře anglistiky a amerikanistiky rozpoznávat anglickou prosodii, konkrétně intonaci a délku trvání a také má za cíl zjistit, jak anglickou prosodii vnímají, a jakou jí přikládají důležitost, co se týče osvojení si anglického jazyka a vnímání českého přízvuku v angličtině. K praktickému výzkumu byly použity dvě metody, percepční test zaměřený na rozpoznávání anglické intonace a délky trvání anglické výpovědi a standardizovaný rozhovor, který měl za cíl dále zjistit postoj českých mluvčí k anglické prosodii.