PALACKÝ UNIVERSITY IN OLOMOUC

Philosophical Faculty

Department of English and American Studies

Perception of English Lexical Stress by Czech Learners of English: *Reliance on Different Acoustic Cues*

(bachelor's thesis)

Author: Petr Tichý *(student of English Philology and Journalism)*

Supervisor: Mgr. Šárka Šimáčková, Ph.D.

Olomouc 2014

I hereby declare that this thesis represents my original work and that I have used no other sources except as noted by citations.

In Olomouc, August 18, 2014

I would like to thank my supervisor Šárka Šimáčková for her assistance and patience. Many thanks go to all the volunteers who participated in the experiment and to Keelan Mullane-Fay aka “Little Boo” for proofreading a draft of this paper. Let my gratitude be manifested through a drawing of a bunny.

Contents

1 Introduction

This paper sets out to empirically explore the nature of human perception of sound. More specifically, it focuses on the nature of perception of a single suprasegmental feature of speech— *word stress*.

English pronunciation has always been of great interest to me, and so have ways how to successfully acquire it later in life as a non-native speaker. For Czechs, pronouncing English words correctly is a challenging skill to learn, and even more so on the prosodic level. I can see this on a daily basis when teaching English or using English with other non-native speakers of Czech origin. I can see this when having to explain myself when speaking to native speakers after I have pronounced a word using the right set of phonemes to begin with, but with a stress on a wrong syllable. Good production is said to follow good perception. This is why I want to explore the topic of stress perception by non-native listeners in this paper.

I have chosen to replicate a study by Wang (2008) which shows how perception of English stress by native English speakers and speakers of Mandarin Chinese differs. Wang recorded three bisyllabic non-words and manipulated each syllable in three different acoustic dimensions: fundamental frequency, intensity and duration. Upon hearing each word, the participants were to select the accented one. While the English native speakers weighted all three cues approximately the same, the Chinese relied solely on the perceived pitch. The goal of this paper is to see which of the cues Czech speakers will rely on and discuss why. Because there are real English words mixed with the non-words in the experiment used to filter out participants who are not able to hear stress altogether, another conclusion of possible stress-deafness of Czech learners can be drawn. Simple language learning background data will be collected from the participants to see any possible correlates.

The scope of the paper is limited to lexical or *word* stress. That is stress realized on a syllable within a single word, best apparent in citation forms on words uttered in isolation. The topic of this work is not stress (or accent) imposed over higher prosodic units under the influence of suprasegmental features such as intonation or rhythm. This thesis is an empirical study of speech perception. The practical part is an experiment testing the relative importance of three different acoustic cues (duration, loudness, and pitch) of English stress for Czech learners of English. The main dimension of this work

is therefore phonetic details of second language acquisition and their possible transfer from a first language to a second language.

The paper is structured as follows: The second chapter “Literature Review” will give basic theory of stress, discuss used terms, evaluate existing studies on the topic and argue for a pair of hypotheses. The third chapter “Method” will introduce details of Wang's (2008) experiment and the present experiment, and point out all differences. The fourth chapter “Results” will state the results of the experiment. In the fifth chapter “Discussion”, I will evaluate the results and consider them in a broader context of language acquisition. In the sixth chapter “Conclusion”, I will summarize the paper, and the material cited throughout the work will be listed in the seventh chapter “References”. The end of the work is reserved for appendices.

1.1 Questions

Ultimately, this paper asks two questions. The first one is whether or not Czech learners of English can perceive English stress at all. In other words, are they 'stress-deaf', as Dupoux et al. (1997) found with French speakers? Can there be significant correlations in the basic language background information for this? The second questions is, should they indeed be able to tell a stressed syllable from an unstressed one, how do they weight the different acoustic cues that are manipulated in the experiment? Is this an example of a transfer of acoustic details of a prosodic feature from one's native language? Does the sensitivity to the different cues change over the span of learning the language or with the frequency of listening to it?

2 Literature Review

This chapter provides a degree of theoretical background for the examined phenomenon. Although the experimental part of this work is concerned solely with phonetics, in order to arrive at a set of satisfactory hypotheses, I will also briefly address the domain of phonology with its suprasegmental components. Similarly, while this paper tests and discusses how speech is perceived, production will also be discussed, since the two are invariably connected.

After a general introduction into the theory of the theme, the two languages will be systematically compared in the sub-chapters and at the end of this chapter there will

be a set of hypotheses for the experiment.

2.1 Fundamentals of Stress

After having studied the individual sounds (segments) of a language, one would usually proceed to larger features of speech. Those are ones that stretch over more than just a single segment and thus called suprasegmentals. Among them fall for example intonation, rhythm and to a certain degree also stress. While the acoustic properties of stress, which I am going to explore in this work, can in theory apply to individual vowels, they only begin to be exploited by the listeners when in context of other segments. In other words, a single sound or a syllable pronounced in isolation is always stressed, and stress therefore becomes significantly more interesting when more than one syllable are involved.

This is also where the major phonological difference between the two languages lies. While in Czech it is always the first syllable of a word that carries stress, it can be virtually any syllable of a word in English. This allows for existence of minimal pairs[1] in English, which only differ in the position of the stressed syllable. Compare *an insult* (noun) and *to insult* (verb). The noun is stressed on its first syllable /ˊinsʌlt/ whereas it is the last syllable which is stressed on the verb /inˊsʌlt/. This is not possible in Czech where placement of stress later in a word would not change its meaning and would rather create an impression of foreignness or of a regional accent.

The acoustics of stress production are a field of phonetics which is not completely understood. In 1970, Ilse Lehiste wrote that, “There is no single mechanism to which the production of stress can be attributed [...]” (Lehiste 1970, 106). It is in the nature of any complex processes that their description is difficult. Nonetheless, it seems that making a syllable stressed is a process so complicated that we still at this day lack a universal definition for either of the languages in mind.

The same is also true for the perception side of the phenomenon, which is well illustrated in Cruttenden (2008, 236) when the author says that, “any of four factors, pitch, loudness, quality and quantity may help to render a syllable more prominent than

1 Note that some authors (e.g. Bičan 2008) reject referring to pairs of words that phonologically differ only in the position of stress to be 'minimal pairs'. This is because, strictly speaking, a minimal pair would consist of word with a stressed syllable and the same word with the same syllable but unstressed. This would be a perfect opposition. In their interpretation, this definition is not able to describe a pair of words where there is not only an opposition from a stressed syllable to an unstressed syllable, but also another change (from an unstressed syllable to a stressed one) in a neighboring syllable.

its neighbours." It is only fair to mention that the author then adds that while all four cues may *help*, "it is principally pitch change which marks an accented syllable" (236). The control group in the present experiment may help give further evidence for the importance of pitch change for determining the accented syllable. Vowel quality is, however, not part of the present experiment.

What is left, it seems, to be said with certainty is that a stressed syllable is produced with more physical effort or, in other words, a greater amount of energy (Ladefoged and Johnstone 2013, 249; Lehiste 1970, 106). Even more broadly and usually from the standpoint of perception, it is possible to use the word *prominence* to describe the stressed syllables (eg., Cruttenden 2008, 235) and indeed this word is often used to refer to syllables which *stand out* without having to make judgments about the acoustics involved. The adjective *more prominent*[2] was also used in the instruction for the participants of the present experiment (see chapter Method).

It is also crucial to mention that the one partially implicit condition that allows for these definitions is that they make claims about stressed syllables *in comparison* to other syllables in the same word or a sentence. They are pronounced with *relatively* more effort or are *relatively* more prominent, not absolutely. "The absolute values are never linguistically important. But they do, of course, convey information about the speaker's age, sex, emotional state, and attitude toward the topic under discussion" (Ladefoged and Johnstone 2013, 24).

2.1.1 Choice of Terms

The two modern textbooks which I often consult here vary considerably in the terminology they use to describe the fundamental phenomena of this paper. Let me therefore present their systems and defend the choice I made for this work.

In *Gimson's Pronunciation of English* the author, "will avoid use of the term 'stress' altogether" (Cruttenden 2008, 23). They prefer using the term *prominence* for general use and *accent* for either lexical (explored here) or prosodic prominence.

Ladefoged and Johnstone (2013, 249), on the other hand, do use the term *stress*, recognizing that not every potential for stress in a citation form comes to realization in conversational speech but using the same word for both.

Some other authors (e.g., Liska 2011, 22) use *stress* for the potential of a syllable

2 For the Czech-speaking participants, the translation *výraznější* was used.

to be prominent, and *accent* for its actual realization.

Here I will use the terms according to Ladefoged and Johnstone (2013), which is also in agreement with Wang (2008), the study which is replicated here. Whether the phonological potential for prominence of a syllable or its acoustic realization is meant should be clear from context.

2.2 Particulars of English and Czech stress

Below, the phonology and acoustics of stress in English and Czech are described.

2.2.1 English

In the following two sub-chapters, a brief summary of the existing literature on the phonology of English stress and its acoustics will be provided. In the third sub-chapter, studies concerning English speakers of Czech will be mentioned.

2.2.1.1 Phonology of English Stress

English is a language with variable word stress (Ladefoged and Johnstone 2013, 249). This means that it is not always predictable where in a word the stress is placed. English stress can function as a cue for certain verb-noun opposition, compare *an 'increase* // to *to in'crease*[3]. It can also distinguish between verbs and two-word phrasal compounds such as *a 'walkout* (a noun) and *to 'walk 'out* (a verb) (112).

There is a large number of affixes that influence the position of stress in some way. They can attract stress or shift it to a different syllable in a word. "If you make a sufficiently complex set of rules, it is possible to predict the location of the stress in the majority of English words" (113).

On the word opposition of fixed—free, English stress is, "fixed, in the sense that the main accent always falls on a particular syllable of any given word, but free, in the sense that the main accent is not tied to any particular point in the chain of syllables constituting a word […]" (Cruttenden 2008, 235). There are some exception to this, however, where the stress falls on a different syllable in continuous speech, compared to where it would fall in the citation form. These are caused by larger rhythmic patterns of the context (245).

2.2.1.2 Acoustics of English Stress

As mentioned above, there is no specific, generally-accepted description of acoustic

3 The position of stress symbolized with an apostrophe (') within regular ortography.

correlates of stress in English. Fry (1955) manipulated the two-syllabic words *object*, *subject*, *digest*, *contract,* and *permit* which could be stressed on either syllable, depending on their part of speech. Duration and intensity were manipulated. Fry concluded that both duration and intensity were cues for perception of stress and that duration is a stronger cue.

Lehiste (1959, 435) concludes that "linguistically significant stress may be based on speech power, fundamental voice frequency, vowel quality, and duration […]" It is also proposed that 'correction factors' might be applied to amplitudes of vowels according to vowel quality. This is to account for their intrinsic differences. Lehiste suggests that similar 'correction factors' maybe be applied for pitch and duration. This is later expanded in Lehiste (1970).

Lieberman (1960) used noun-verb pairs similar to that of Fry (1955). Lieberman found that fundamental frequency was the most relevant acoustic correlate. Contrary to Fry (1955), however, Lieberman found that amplitude seems more important than duration. Vowel quality was not examined in the study.

Lehiste (1970, 127) comments on the role of fundamental frequency as having an 'all-or-none' effect in Fry (1958). Lehiste also states that, "additional experiments with more complex patterns of fundamental frequency change suggested that sentence intonation is an overriding factor in determining the perception of stress […]" (127). It seems that the problem of co-occurrence of stress and the intonation pattern of a tone unit makes the effects of fundamental frequency on the word level difficult to interpret.

More recent data about English stress can be drawn from studies in which native speakers of English functioned as control groups. Volín and Weingartová (2014) measured acoustic realizations of stress by Czech and English speakers. The control group used the tree cues systematically and the results were in accordance with Fry (1955). Perceptually, Wang (2008) showed that all tree cues had effects on stress judgment, the effect of fundamental frequency being stronger than those of duration and intensity.

On the other hand, Ladefoged and Johnstone (2013) write that, "the most reliable thing for a listener to detect is that a stressed syllable frequently has a longer vowel than it would have if it were unstressed" (111).

The findings of the early studies were supported by Cutler (2005) who stated

that, “over nearly 50 years, however, perceptual studies have elaborated but not fundamentally altered the early claims concerning the suprasegmental dimensions” (269).

2.2.1.3 English speakers of Czech

In a series of studies by Podlipský (2007; 2009), it was examined how native speakers of English who learned Czech later in life perceive vowel quantity in Czech. In general terms, it was studied whether non-native learners will adapt a feature that is not phonemic in their L1 as phonemic in their L2 and to what extent. This is also thought to be difficult also because vowel quantity functions as one of the cues for stress in English. While the 2007 study did not find significant differences between native and non-native listeners, in the 2009 study the non-native listeners performed only at chance in the perception experiment when there was a stress-length conflict, suggesting that there was a transfer of perceptual sensitivity to length as a cue for stress.

The question in the present paper is reversed. It is examined whether Czech learners of English will use the cue associated with stress in English in spite of its different phonological function in their L1 (short/long vowel pairs) and to what extent.

2.2.2 Czech

The following two sub-chapters will review existing literature on phonology and acoustics of Czech stress. The third sub-chapter will present studies in which learners of English of Czech origin were examined.

2.2.2.1 Phonology of Czech Stress

The phonology of stress in Czech seems to be well understood, largely because its placement is regular. Hála (1962, 298) states that there is ‘no doubt’ about the placement of stress in Czech words. It is placed on the first syllable in a word.

Palková (1994, 277) goes further and classifies Czech as a language with a fixed stress-position. In such languages, stress is not phonologically active. That means it cannot carry a discriminatory function within a given word paradigm. It does, however, fulfill a delimitative function. In other words, it can give the listener information clues of where the word boundaries lie; it cannot help the reader hear which word there is (e.g. which part of speech) but it can help mark where it starts and where it ends. This latter function is nonetheless not always taken advantage of (277). Palková then

discusses that it is often the case that while there is the potential for stress on every word in Czech, on short grammatical words it is often not realized in normal speech, creating a single prosodic foot with a neighboring lexical word (280). As the experiment in this paper is concerned with words in isolation, this is not an issue.

This is later reviewed by Šefčík (2002) who compared Czech stress to that of Vedic Sanskrit, confirming all of the above and saying that the function of stress in Czech is limited to its delimitative function where it is one of the possible clues of word boundaries (another being for instance a pause).

An attempt to re-qualify Czech as a language with free stress was made by Bičan (2008) who presented a large number of examples in which the position of stress distinguished between several possible meanings. All of them, however, were limited to a sequence of words where the presence of stress (or the lack of it) rendered either a more complex word spelled as a single unit, or a sequence of a preposition and another word. This shows that stress can mark word boundaries. While such contrasts do differentiate meaning in Czech, I consider this only a matter of the two definitions and Czech stress will be thought of as fixed throughout this paper. No possibility of having stress on a different than the first syllable in a word was provided in the paper. Moreover, the stimuli in the present experiment are single words pronounced in isolation where the possibility of them fusing with a preposition is irrelevant.

2.2.2.2 Acoustics of Czech Stress

While the placement of Czech stress is regular and highly predictable, as of now, there is no generally-accepted description for the acoustic correlates of it.

Hála (1962, 299) describes Czech stress as being essentially dynamic. More effort in stress production then often results in a higher fundamental frequency of the stressed syllable. This is described as a natural physiological process where higher intensity causes higher frequency. More effort during the production of a syllable causes the vocal chords to be more tense, and when exhaling, the rate of their vibration is therefore higher. However, there are cases in which higher fundamental frequency and higher intensity do not coincide and so, "simultaneity of [higher] tone with stress is not however a universal rule"[4] (300).

More general terms for the acoustics of stress are used by Palková (1994, 277).

4 "Paralelnost tónu s přízvukem není ovšem pravidlem." *(Czech original)*

‘Dynamic’ stress is replaced by ‘prominence’ and ‘contrast’. Making a syllable stressed is described to be a complex process which involves multiple acoustic cues. Palková revises the traditional dynamic view and talks of three potential acoustic qualities that help render a syllable more prominent: intensity; fundamental frequency and duration; the function of duration being limited due to its phonological function of contrasting short and long vowels. There are no changes of vowel quality in normal speech in Czech (279).

According to Palková (1994, 279), the most frequent correlate of stress in Czech is a change in fundamental frequency. The syllable can be marked by either having a higher frequency or a lower frequency than the following syllable (278).

Volín (2008) studied speech of two Czech Radio presenters and analyzed the differences in fundamental frequency of the first and second syllables in a prosodic foot. The results, at least for the domain of read news texts, disproved Hála’s (1962, 299) notion of natural co-occurrence of intensity and frequency peaks on stressed syllables and were in agreement with Palková’s (1994, 278) statement in the sense that the stress can be frequency-marked in either direction. Should Palková's statement however suggest that the two options of marking the stressed syllable are distributed randomly, such null hypothesis was also disproved for the given domain as there were overall only approximately 20% of feet with the first syllable being higher in frequency than the second syllable.

Intensity contours of prosodic feet in Czech were measured by Duběda (2006) who concluded that the hypothesis of positive intensity correlate of Czech stress can be refuted and, in fact, the stressed syllable, “does not coincide with any systematic intensity pattern” (Duběda 2006, 4). Similarly, Palková and Volín (2003), while studying the effect of F0 contours on determining food boundaries in Czech, found no significant effect of intensity.

As for the duration of stressed syllables from the perception point of view, the Czech control group in Podlipský (2009) performed ‘slightly but significantly better’ in a perceptual segmentation task in trials where there was a stress-quantity conflict. In other words, the participants were more likely to identify word boundaries before a short vowel rather than before a long one. This suggests that it may be that in Czech it is vowel shortening which enhances stress perception.

2.2.2.3 Czech Speakers of English

Liska (2011) examined two groups of Czech speakers (a group with a strong foreign accent and a group with less foreign features) and their realizations of English stress. They found that the group differed in their usage of duration and fundamental frequency to realize stress, with duration ratio being larger for the group with less foreign accent. Duration was overall used significantly more than F0 to mark stressed syllables. All speakers produced stressed syllables longer than their unstressed counterparts. This was hypothesized to be the case because speakers tend to "bootstrap" English stress to such phonetic cues from their L1 that either also mark stress or have a different phonological function (such as vowel length in Czech). However, some subjects from the group with less foreign accent did use F0 to a great degree. Liska suggests that this over-usage may be the speakers' attempt to sound more native-like and to match the intonation patterns of English which are vastly more varied compared to a more flat intonation contours of Czech.

Similar experiment was conducted by Volín and Weingartová (2014). In production, Czech learners of English exhibited alternations to the position of stress, as compared to the canonical position as noted by Wells (2008) and as compared to the control group. Czech learners also marked stress much less than the control group. In their production, the difference between a stressed and an unstressed syllable was much smaller compared to the control group with respect to sound pressure level and fundamental frequency. Their usage of duration, however, was comparable to that of the control group. This was speculated to be a consequence of the phonological role of vowel duration in Czech.

2.2.3 Summary of Czech and English

It is generally agreed that, as for the position of stress, Czech is a language with a fixed stress position and English is a language with a variable stress position. The acoustics of realization of stress in both languages are, however, far less understood. In English, increases in all of the three cues manipulated in the present experiment (fundamental frequency, intensity and duration of the vowels) seem to play a role in production and perception of English, although to different degrees. It is also not clear what interactions of the cues, if any, play a role in stress perception. In Czech, earlier notions of dynamic realizations were rejected and there are no unambiguous findings of acoustic correlates,

although it seems that the first vowels of stress feet in Czech have lower fundamental frequency. The duration cue does not seem to play a role because of its phonological function of distinguishing short and long vowels.

2.3 Importance of Correct Stress

Volín a Weingartová (2014) speculate that because the term *accent* refers not only to the overall impression of a pronunciation style, but in many languages also to the prominence of a syllable or a word[5], it may be that realization of prominence may be highly important for the overall impression of speech. That say that, "the term itself is motivated prosodically as if to suggest that one of the most conspicuous features of 'pronunciation other than the reference standard' is prominence distribution and prominence manifestation" (176).

And indeed, there is empirical evidence for the importance of prominence for foreign accentedness. Hahn (2004) conducted a perception experiment in which it was examined how native speakers of English comprehended, evaluated and processed non-native speech with correctly and incorrectly places stresses. With the correct placement of stressed syllables, the listeners remembered more, evaluated the speakers more favorably and, thought insignificantly, processed the information more easily. Hahn suggests that teaching suprasegmental features is important and provides evidence that, "primary stress contributes significantly to the intelligibility of nonnative discourse" (218).

Similarly, Field (2005) found that there is threat on intelligibility when the stress is not placed correctly and while the threat was not large, it is suggested that the threat would be larger in continuous speech; the material of the study were citation forms.

2.4 Hypotheses

In accordance with the review above, it was hypothesized that: **(1)** Czech listeners will be able to hear stress differences; and **(2)** they will rely on the phonologically active cue from their L1— vowel duration.

5 This is also true for Czech where *přízvuk* incorporates both meanings.

3 Method

This chapter describes the original method used by Wang (2008) with the changes that were made to it for the present experiment. First, I will talk about the stimuli, and second, I will describe the procedure of the experiment. For complete tables of results and graphs see chapter Appendices.

3.1 Stimuli

The same three nonsense words as in the original study were used: *latmab, nizdit,* and *tetsep*. The words conform to the English phonotactic structure[6]. In other words, they *could* be English words and should not be confusing to produce or perceive for native listeners or experienced learners who are familiar with English phonotactics. Each of the words contains the same vowel in both of its syllables. This is to ensure that the cues can be manipulated and explored without having to account for the intrinsic differences between the different vowels. The sound files from the original study were not available to me.

3.1.1 Recording of Stimuli

While the words were recorded by a 'trained phonetician' in the original study, in the present study a female Canadian speaker who was available to me recorded the words under my guidance. I did not interfere with the recording in the sense of pronouncing the words myself. The sought stress patterns were explained using an analogy with simple English two-syllabic words with the same stress position. The desired pronunciation was also achieved by assigning different parts of speech and meanings to the pseudo words and first pronouncing the words in a sentence and only then in isolation. This was done to ensure that there was no bias on the part of the speaker which could result in unnatural pronunciation.

The recording took place in the speaker's home using a USB microphone Samson C03U switched to the cardioid pick-up pattern. The -10 dB high pass filter on the microphone was switched off. In front of the microphone, there was a conventional pop filter set up. The words were recorded using Audacity for Windows (Audacity Team 2014) into WAVE files of 16 bit of audio depth and 44.1 kHz sampling frequency. Wang (2008) mentions that the words were recorded with two stress patterns but this is never

6 Phonotactics describes the rules of combining phonemes in a language. For example, in English no word can begin with /ŋ/ as in thing or end with /h/ as in head.

sufficiently described later in the paper in the process of manipulating the syllables. The Canadian speaker also recorded the words with both stress patters for me but only the ones with initial stress were used as the basis for the manipulations. As in the original study, the words were recorded several times and only the most clear repetition was later used as basis for the computer manipulations.

3.1.2 Manipulations of Stimuli

Manipulations of fundamental frequency (F0), intensity (int) and duration (dur) were applied to the nuclei of the syllables. There were five manipulation levels in the original study. They are illustrated in *Table 1* below.

	Levels				
Correlates	1	2	3	4	5
F0	-50	-25	0	25	50
Duration (D)	0.5	0.75	1	1.25	1.5
Intensity (I)	-9	-4.5	1	4.5	9

Table 1: the five manipulation levels (Wang, 2008)

The first step was to manipulate the two syllables in each word in such a way that they would have the same values of F0, duration and intensity. This was achieved by averaging the values of both syllables and then bringing them lower or higher in value so that their values were of the average, and therefore equal to each other. This created levels 3 of manipulation. Level 3 was used as the basis for manipulations of all the other levels. Taking fundamental frequency as an example, *Table 1* shows that level 1 for F0 is -50. This means that the first syllable of the word is 50 Hz lower than the second. This is achieved by taking the first syllable 25 Hz lower than the average (level 3) and taking the second syllable 25 Hz higher than the average, resulting in a difference of 50 Hz. Level 2 represents a difference of 25 Hz and the two syllables were thus moved higher and lower by 12.5 Hz. Levels 4 and 5 are mirror manipulations of levels 2 and 1 respectively. That means that the F0 of their first vowels was higher than those of the second.

Unfortunately, the original study is not specific enough about the units that belong to the numbers in *Table 1*. For F0, the difference of 50 Hz is mentioned in the text, but for the intensity and duration, there is no additional clarifying information.

Intensity is often defined in decibels. The present study is similar in nature to that of Fry (1955), which is also cited in the Wang (2008), who manipulated intensity and duration of English words and the values were also very similar. It is therefore very likely that the intensity values in the original study are in decibels. Level 1 of -9 dB was thus achieved by taking the first syllable 4.5 dB lower and taking the second syllable 4.5 dB higher than the average. Level 3 of the new *Table 2* (see below) for intensity now reads "0" in order to represent that there is no change in the intensity of the syllables.

The situation is however more complicated in the case duration. The values for levels 4 and 5 are not absolute negative mirrors of levels 1 and 2. The values 0.5, 0.75, 1.0, 1.25 and 1.25 are too large to be taken for seconds and too small to be in milliseconds. This suggests that the numbers are in fact duration ratios of the two syllables, which is also in accord with the Fry's (1955) study. All the numbers in the *Table 1* are given from the perspective of the first syllable. However, taking for instance levels 1 and 5 (0.5 and 1.5 respectively) from the original study, should they symbolize that in level 1 the duration of the first syllable is half of the duration of the second syllable and that in level 5 the first syllable is one and a half of the duration of the second syllable, the two levels are not mirror images. Therefore I decided to use the mean duration as a basis for the manipulations, as the mean values are also used by the two other cues. The updated manipulation levels are listen in *Table 2* below.

correlates	**levels**				
	1	2	3	4	5
F0 [Hz]	-50	-25	0	25	50
duration [proportion of mean]	-0.5	-0.25	0	0.25	0.5
intensity [dB]	-9	-4.5	0	4.5	9

Table 2: updated manipulation levels

This means that in level 1 the first syllable is shorter than the second by half of the average duration value. This is achieved by making the first syllable shorter by a quarter of the average value and making the second syllable longer by a quarter of the average value. Level 2 was achieved by making the first syllable shorter by an eighth of the average value and making the second syllable longer by an eighth too. Levels 4 and 5

are now mirror images of levels 2 and 1, respectively.

The exact values which were used for the manipulations for the word *latmab* are illustrated in *Table 3* below. Only the values in the top half *Table 3* that are not bold are entered manually, the rest is automatically calculated. For data of all three manipulated words see Appendix D.

manipulation values: ***latmab***							
orig. sound file duration [sec]		0,678571					
orig. 1st syllable nucleus (from; to) [sec]		0,094351		>	0,16471		
orig. 2nd syllable nucleus (from; to) [sec]		0,400854		>	0,45949		
orig. F0 (1st; 2nd; mean) [Hz]		175,7	170,7	>	**173,2**		
orig. intensity (1st; 2nd; mean) [dB]		74,53	63,28	>	**68,905**		
orig. duration (1st; 2nd; mean) [sec]		0,070359	0,058636	>	**0,0644975**		
after level-3 manip. nuclei (1st; 2nd) [sec]		**0,094351**	**0,1588485**		**0,3949925**	**0,45949**	
correlates		**levels**					
		1	2	3	4	5	
F0	*1st*	148,2	160,7	**173,2**	185,7	198,2	
	2nd	198,2	185,7		160,7	148,2	
	diff	-50	-25		25	50	
intensity	*1st*	64,405	66,655	**68,905**	71,155	73,405	
	2nd	73,405	71,155		66,655	64,405	
	diff	-9	-4,5		4,5	9	
	1st (rel.)	-4,5	-2,25		2,25	4,5	
	2nd (rel.)	4,5	2,25		-2,25	-4,5	
duration	*1st*	48,37	56,44	**64,50**	72,56	80,62	
	2nd	80,62	72,56		56,44	48,37	
	diff	-0,50	-0,25		0,25	0,50	
	1st (rel.)	0,75	0,875		1,125	1,25	
	2nd (rel.)	1,25	1,125		0,875	0,75	

Table 3: manipulation values of *latmab*

The first half *Table 3* shows the total duration of the sound file, starting and ending points of the nuclei of two syllables, mean values of F0, intensity and duration of the individual syllables and their averages, and the changed starting end ending points of the nuclei after the first manipulation to level 3. The “after level-3 manip nuclei” row shows the updated starting and endings points of the nuclei once the first level 3 manipulation was carried out. These were necessary to calculate for the future manipulations. Because the duration across the manipulation levels is always distributed to both syllables equally, the starting point of the first nucleus and the ending point of the second nucleus stay the same, as well as the total duration of the sound file. The ending point of the first nucleus and the starting point of the second nucleus however required updating.

The second half of *Table 3* shows the specific values of the acoustic correlates

for the individual manipulation levels. Intensity and duration are also accompanied by relative values of the correlates. This is because of the specific way the manipulations were carried out. While the values of fundamental frequency are entered in absolute numbers, intensity and duration manipulations are entered using their relative values and proportion values respectively.

3.1.2.1 Manipulation Method

All the manipulations and measurements were carried out in Praat for Windows (Boersma and Weenink 2014). The process of manipulations was the same for all the three pseudo words. The first step was to manipulate the word, which was recorded with the stress on its initial syllable, to the level 3 of manipulation. This was achieved by:

(1) creating an IntensityTier with a right angle intensity drop over one nucleus and a right angle intensity rise over the other nucleus. The values were calculated as a positive and negative difference from the mean value of intensity. No other intensity manipulations were marked into the tier. The tier was then multiplied with the Sound object of the original sound. No intensity scaling was applied to the resulting object;

(2) morphing the resulting Sound object into a Manipulation object;

(3) creating a PitchTier with only a single point with the value of the mean fundamental frequency which was obtained by averaging the two nuclei;

(4) creating a DurationTier with values of calculated ratios over both nuclei. The rest of the tier wave form is of value 1. The ratios are calculated by dividing the average duration by the actual value which results in one value smaller than 1 and one value larger than 1;

(5) selecting, one by one, the PitchTier and the Manipulation object and the DurationTier and the Manipulation object and replacing the default tiers of the Manipulation object by the newly created tiers;

(6) re-synthesizing the updated Manipulation object into a new Sound object.

This was performed for all the three pseudo words and the results were files latmab_3_3_3.wav, nizdit_3_3_3.wav and tetsep_3_3_3.wav which were then used in the scripts for all the other manipulations.

The other manipulations to all desired levels were carried out in the same way, except all 15 different tiers were created before saving the final sound files. An example of the object window of Praat for the word *nizdit* can be seen in *Figure 1* below. The

level 3 input object was renamed to “nizdit_DEFAULT” and used as a basis for the following manipulation steps.

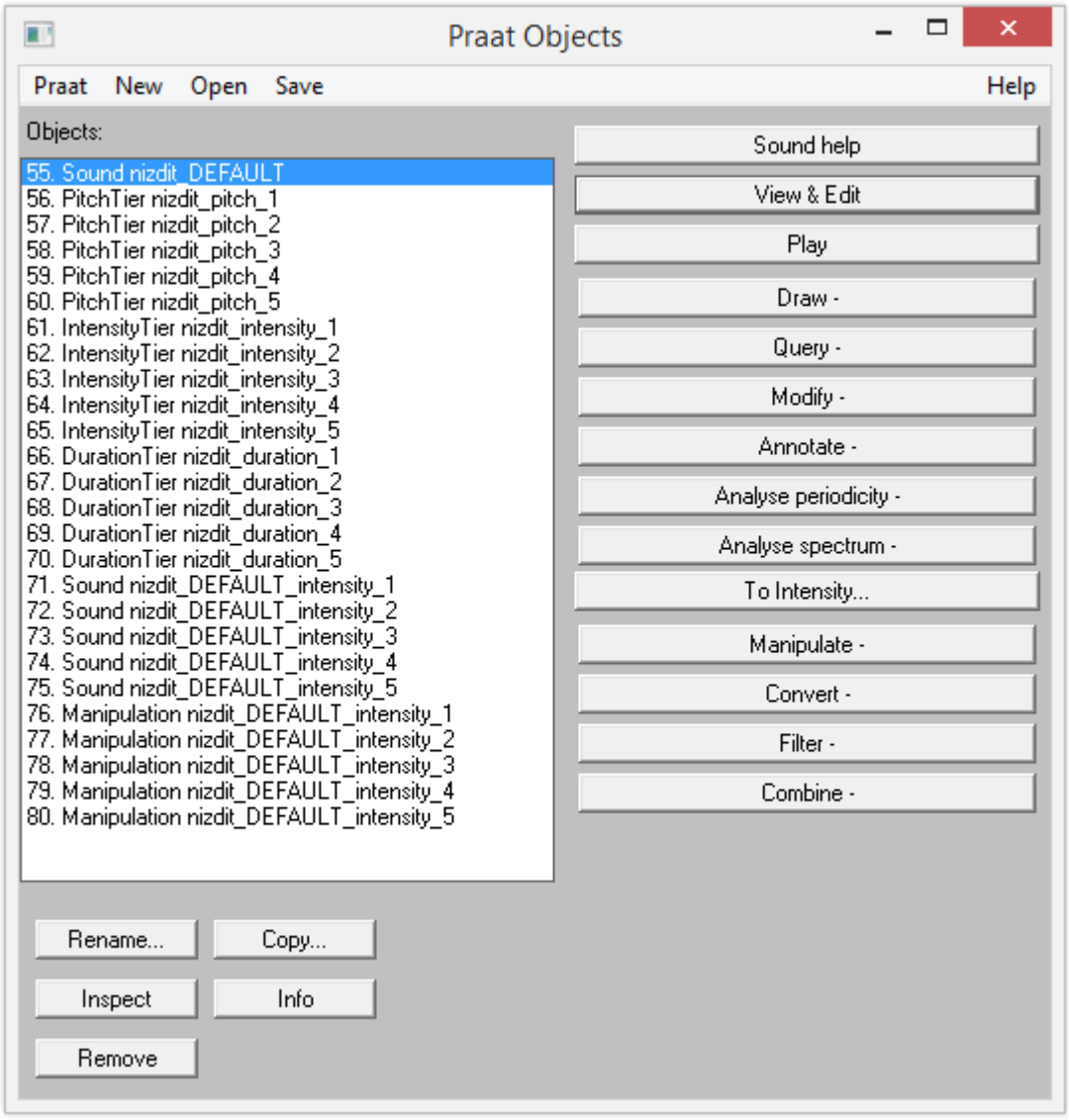


Figure 1: all manipulation objects for the word *nizdit*

The manipulation levels were also different, following the values in *Table 2*. Below you can see examples of the intensity (level 1), pitch (level 2) and duration (level 5) tiers for the word *nizdit*.

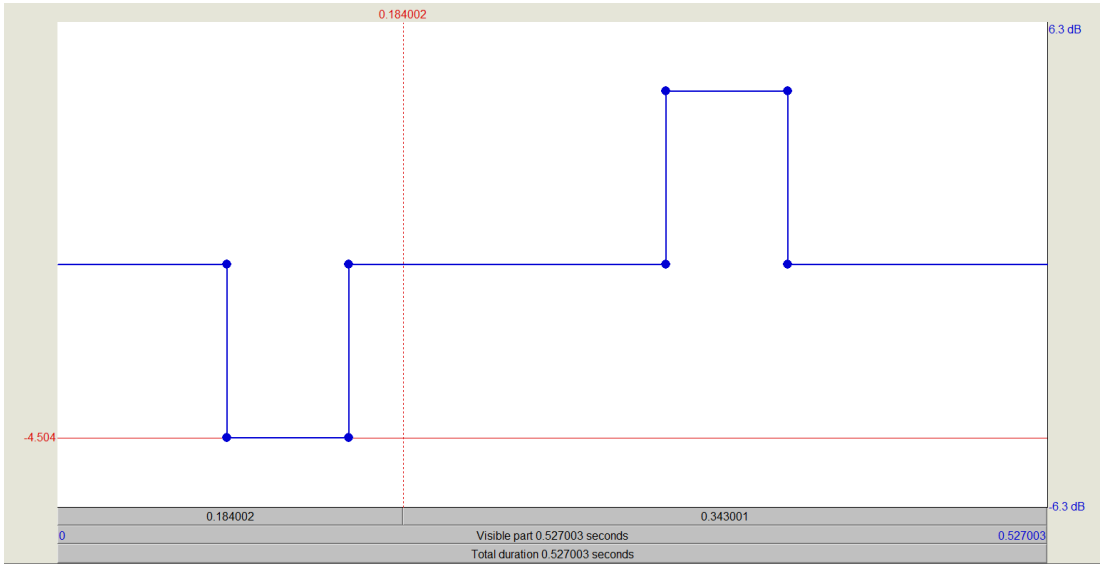


Figure 2: IntensityTier (level 1) for *nizdit*; the values are (relative) -4.5 dB for the first nucleus, +4.5 dB for the second nucleus, and the middle value of 0 dB

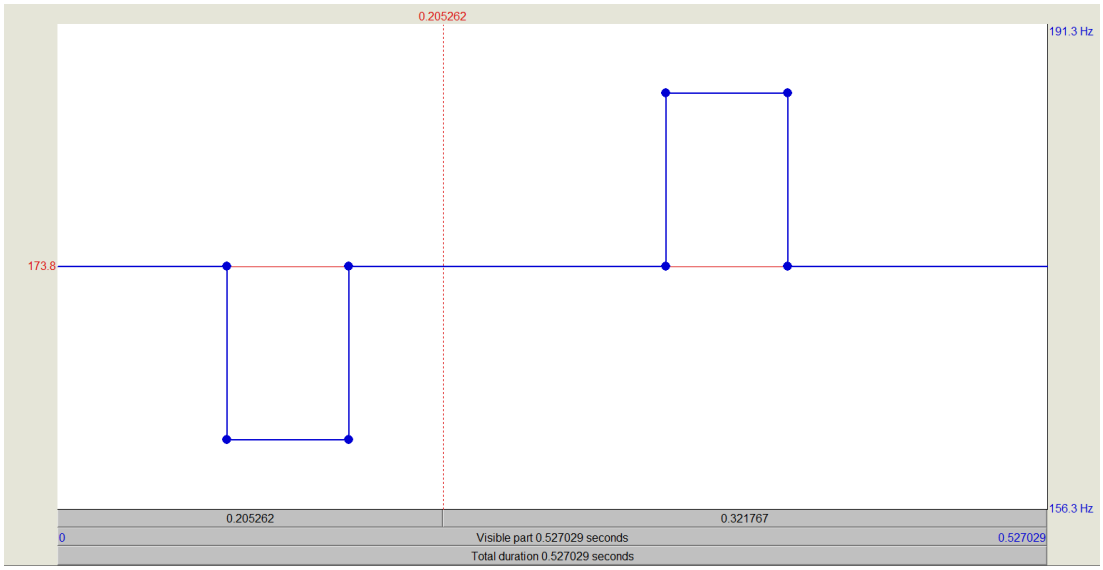


Figure 3: PitchTier (level 2) for *nizdit*; the values are (absolute) 161.27 Hz for the first nucleus, 186.27 Hz for the second nucleus, and the middle value of 173.77 Hz

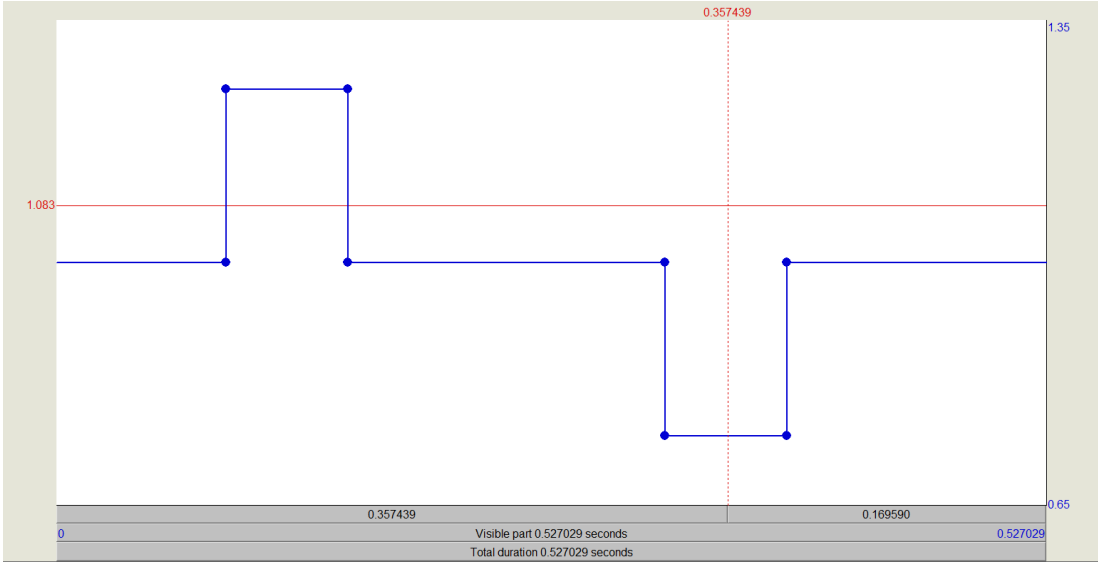


Figure 4: DurationTier (level 5) for *nizdit*; the values are (proportional) 1.25 for the first nucleus, 0.75 for the second nucleus, and the middle value of 1

The tiers for levels 3 are all only single lines; intensity level is 0 dB, pitch is the average value for the particular word (173.77 Hz for *nizdit*) and duration is 1.

It is important to realize some acoustic consequences of the manipulation methods which were used in the present experiment. When preparing the levels 3 to become the basis for the other levels, the fundamental frequency contour of all voiced parts of the sound is flattened. Any natural rises and drops that were originally present in the pitch contour are cleared. Because the changes from the middle value to the manipulated nucleus are not gradual but abrupt, this causes the resulting stimuli to sound somewhat artificial. Intensity contours of the nuclei, on the other hand, are preserved and lifted or lowered, depending on the manipulation level, as a whole. Duration is manipulated with no other changes by Praat. As far as the description in the original study is concerned, the manipulation method used in the present experiment is identical.

After having created all the necessary tiers for the particular sound file, a script (see Appendices) was executed to combine the objects and export WAVE files into a selected folder. The naming key for the sound files was "word_F0_int_dur.wav", manipulation levels being represented by numbers 1—5 (e.g. "tetsep_4_1_2.wav").

3.1.3 Control Stimuli

One hundred real English words were mixed with the pseudo words in order to mark and rule out participants who were wholly unable to hear stress. In the original study, a limit of 80% correct answers was used.

Wang (2008) does not specify which words were used. Because I wanted to minimize the effect of language experience, I sought the most frequent English words. In this way, the words should be familiar to as many of the participants as possible. The list of words in the present study is based on a free 5000 lemmas long list obtained from *Word frequency data: Corpus of Contemporary American English* (Davies 2014).

The list of a hundred words was created by selecting only bi-syllabic words from the 5000 lemmas word list, going from the most frequent, until there were 50 two-syllable long words with the stress on the initial syllable and 50 two-syllable words with stress on the final syllable. Words which were identified to bear different stress patterns in American and British English or had an alternative pronunciation with a different stress position, according to Wells (2008), were skipped. Words that are pronounced with stress on a different syllable when they represent a different part of speech were skipped, too. Also skipped were some plural forms. Among these for example "almost" (both stress patterns possible), "others" (a plural form; the list already contains "other") "research" and "record" (both can be stressed either way, depending on their part of speech). This was done to avoid confusion if a participant is familiar with the other possible stress pattern only. In *Table 4* below, you can see the first 10 words of both stress pattern groups. For the complete list of the words that were used see Appendix C.

stressed initially		
rank	word	part of speech
65	into	i
75	other	j
87	also	r
99	many	d
101	only	r
105	very	r
107	even	r
111	woman	n

stressed finally		
rank	word	part of speech
46	about	i
89	because	c
139	become	v
140	between	i
164	begin	v
180	against	i
184	again	r
213	believe	v

Table 4: first ten control stimuli words from both stress pattern groups ("rank" marks their position in the list of all 5000 lemmas)

3.1.3.1 Recording of Control Stimuli

In Wang (2008), the 100 real words were recorded by the same trained phonetician as the pseudo words. Originally, the hundred words in the present experiment were also recorded by the same speaker as the pseudo words. However, the recorded words did not seem to be of good quality for the experiment because of their rising intonation on the second syllable. This was most likely caused by wrong instruction on my part. The words were read by the speaker from a piece of paper. Unfortunately, there were not sufficient breaks between the words, causing the recording to have an intonation more typical for a sentence containing multiple short clause. Each of the clauses would rise at the end to show that there is another clause following. The rising intonation would be confusing for the participants and therefore an alternative solution was pursued.

The selected words were recorded using the LDOCE5 Viewer (Fukada 2013), an alternative dictionary viewer which uses data from the Longman Dictionary of Contemporary English (5th ed.) (Longman 2011). The recording was performed using Audacity for Windows (Audacity Team 2014) with the method of recording directly the output of the sound card of the computer. This method should not cause any drop in sound quality. The words were recorded in a sequence and then individually selected and exported as WAVE files of 16 bit of audio depth and 44.1 kHz sampling frequency.

3.2 Procedure

Altogether, there were 475 test sounds (= tokens). That is, 125 combinations of each of the tree pseudo test words (intensity level * pitch level * duration level = 5*5*5) and 100 real English words.

In Wang (2008), the presentation of the stimuli was controlled by a C++ program written specially for the experiment. The program was not available to me. The 475 stimuli were divided into 25 blocks. There were 19 words in each block. Before each block, there was a beep, and after it, there was a 30 seconds long break. There were 15 pseudo words and 4 real words within each block. The pseudo words in a block were divided into 5 sets with a real word in between every set.

In the present experiment, a MFC object type in Praat was used to present the stimuli to the participants. This is a pre-scripted way how to run multiple forced choice experiments in Praat. The choice to use the MFC in Praat was made because of the simplicity of setting up the experiment and the familiarity with Praat software as

a whole. However, three main changes had to be made to the experiment in order to fit the capabilities of the MFC Praat object type:

(1) The original experiment design was not a forced choice design. There was a 2-second long period after hearing a token in which the participants could choose the accented syllable. The following token was then presented regardless of whether a choice was or was not made, which is not possible with the MFC Praat object. Any timed-out responses were marked and discarded. 0.8% of the responses were discarded in this way in the original study. On the other hand, the present experiment was a forced choice design in the sense that the following token was not present until the choice was made. However, the response time was recorded which allows for further analysis of the time it takes to decide over particular stimuli. Responses exceeding a certain response time can be also ruled out.

(2) In the original design, the order of the tokens was random for each participant. This was not possible using the MFC Praat object while at the same time maintaining the division into blocks and sets of the pseudo words and the real words. The order of the stimuli in the present design was randomized once at the beginning and stayed the same for all the participants.

(3) While in the original design, there was a mandatory 30-second long break after each of the 25 blocks, this was not included in the same way in the present experiment. Instead, after each of the blocks, there was a text on the screen telling the participants that that may take a break, if they like. To proceed with the experiment, the participants clicked anywhere on the screen using the mouse or the touchpad buttons. The reason for this solution was to make the total duration of the experiment as short as possible for the sake of the participants' time, while at the same time keeping the possibility of taking a break between the blocks.

3.2.1 Experiment Design

From Wang (2008), it is not clear whether the first three pseudo words in each set within each of the blocks were or were not in the same order for each set or each participant. In the present experiment, they were indeed in the same order of *latmab*, *nizdit* and *tetsep* in each set, manipulation levels having been randomized once for all the participants. The first block of tokens from the experiment can be seen below in *Table 5*.

token no.	word	F0	int	dur	filename	correct answer
1	latmab	1	3	4	001-latmab-1_3_4.wav	N/A
2	nizdit	4	1	4	002-nizdit-4_1_4.wav	N/A
3	tetsep	2	3	5	003-tetsep-2_3_5.wav	N/A
4	*number*	N/A	N/A	N/A	004-number.wav	1
5	latmab	5	5	4	005-latmab-5_5_4.wav	N/A
6	nizdit	2	4	4	006-nizdit-2_4_4.wav	N/A
7	tetsep	3	4	5	007-tetsep-3_4_5.wav	N/A
8	*even*	N/A	N/A	N/A	008-even.wav	1
9	latmab	1	1	4	009-latmab-1_1_4.wav	N/A
10	nizdit	3	3	5	010-nizdit-3_3_5.wav	N/A
11	tetsep	2	4	1	011-tetsep-2_4_1.wav	N/A
12	*allow*	N/A	N/A	N/A	012-allow.wav	2
13	latmab	1	2	4	013-latmab-1_2_4.wav	N/A
14	nizdit	1	2	4	014-nizdit-1_2_4.wav	N/A
15	tetsep	5	1	3	015-tetsep-5_1_3.wav	N/A
16	*between*	N/A	N/A	N/A	016-between.wav	2
17	latmab	1	3	3	017-latmab-1_3_3.wav	N/A
18	nizdit	4	1	2	018-nizdit-4_1_2.wav	N/A
19	tetsep	1	3	3	019-tetsep-1_3_3.wav	N/A

Table 5: first block of the experiment (the column “correct answer” is only appropriate for the real words, where 1 = first syllable and 2 = second syllable)

For the complete order of the tokens in the experiment, see the script in Appendices.

The participants took the experiment one at a time, all on the same laptop computer. They all used an identical pair of AKG K 142 HD headphones. All participants were asked whether the volume level was adequate and shown how to adjust it at any further point. The testing of all the participants took place inside in different, reasonably quiet rooms.

Several things were made clear to each participant orally, on top of the written instructions on the screen. Similar language that is used here was used when talking to the participants:

(1) Some of the words in the experiment are real, simple English words but most are pseudo words that do not mean anything. However, they could all be English words and the participants were asked to think about them in such way.

(2) Participants' task was set as to identify the "more prominent" of the two syllables.[7]

(3) It was stressed that all the words have exactly two syllables.

(4) It was made clear that all the tokens in the experiment are at least a little different. The participants were told that while the pseudo words in the experiment may be spelled always the same if written, they are computer-manipulated.

(5) The participants were advised to use the buttons with the symbols "less than" and "greater than" on them in order to save time doing the experiment.

All the participants were instructed orally on top of the instructions on the screen. The native listeners were told the instructions in English but Czech was used with the group of learners. The translation "výraznější" was used for the identification task.

(6) Before the experiment round, the division into blocks with prompts for breaks after each block was also made clear.

Before the 475-token long experiment, there was a practice round. The design of the practice round was identical to the final experiment. The practice round was 8 tokens long, with stimuli that were not used in the final experiment. The original non-manipulated recordings of the three words pseudo words in both stress patterns plus two real words were used. The design of the practice round is visible in *Table 6* below.

token no.	word	filename
1	LATmab	latmab.wav
2	nizDIT	nizdit_2.wav
3	tetSEP	tetsep_2.wav
4	*MORning*	morning.wav
5	latMAB	latmab_2.wav

7 Wang (2008) does not specify what wording was used in the identification task. Since I did not want to use any terminology (such as "stress/ed" or "accent/ed") in order to eliminate possible differences due to language learning experience and/or knowledge of phonetics, the neutral and commonly used term "prominence" was used.

6	NIZdit	nizdit.wav
7	TETsep	tetsep.wav
8	*uPON*	upon.wav

Table 6: all stimuli of the practice round

The participants were prompted to ask questions before the practice round and before the final round orally and via the instructions on the screen. Below are screenshots from the experiment.

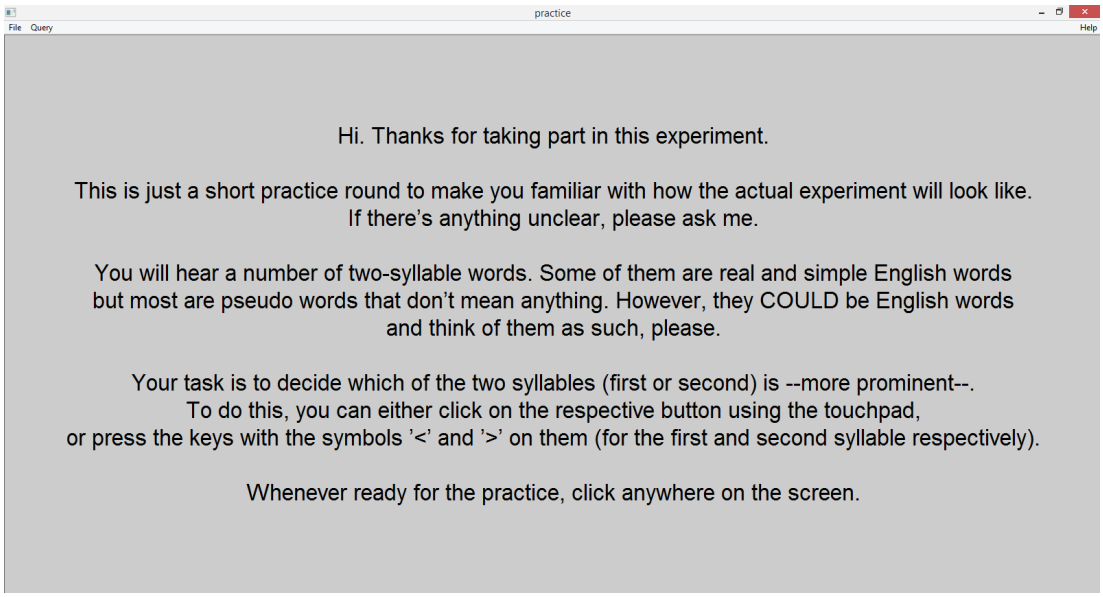


Figure 5: introduction screen of the practice round

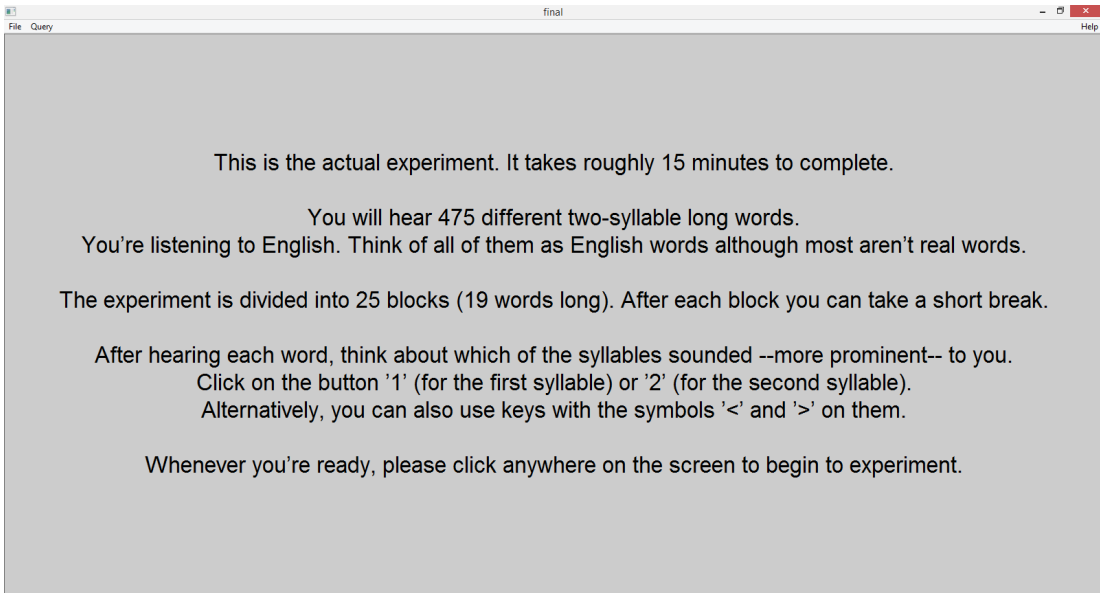


Figure 6: introduction screen of the experiment round

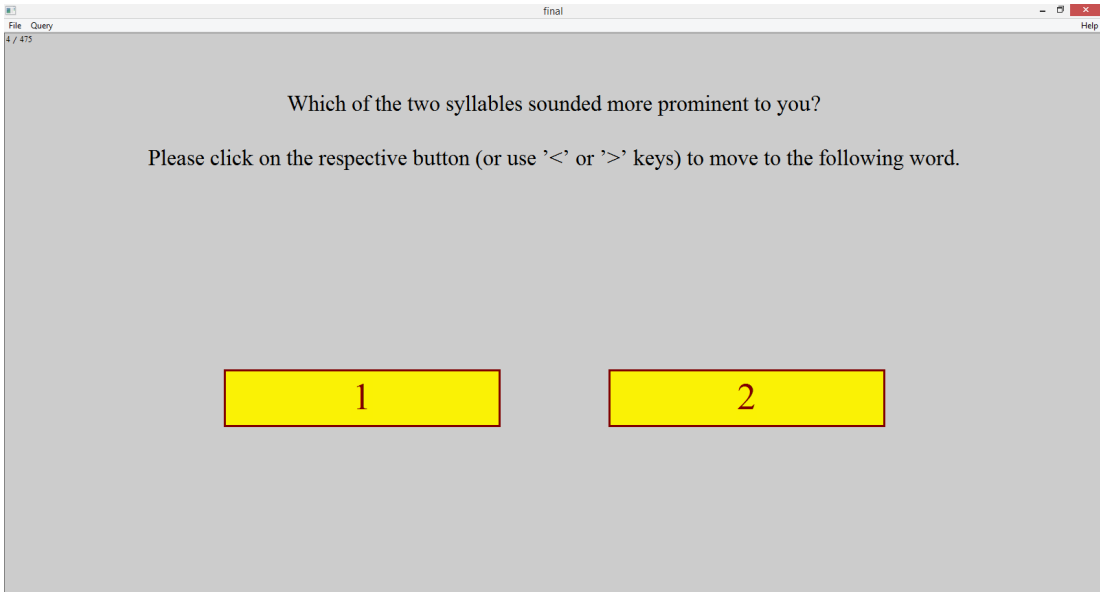


Figure 7: identification task screen

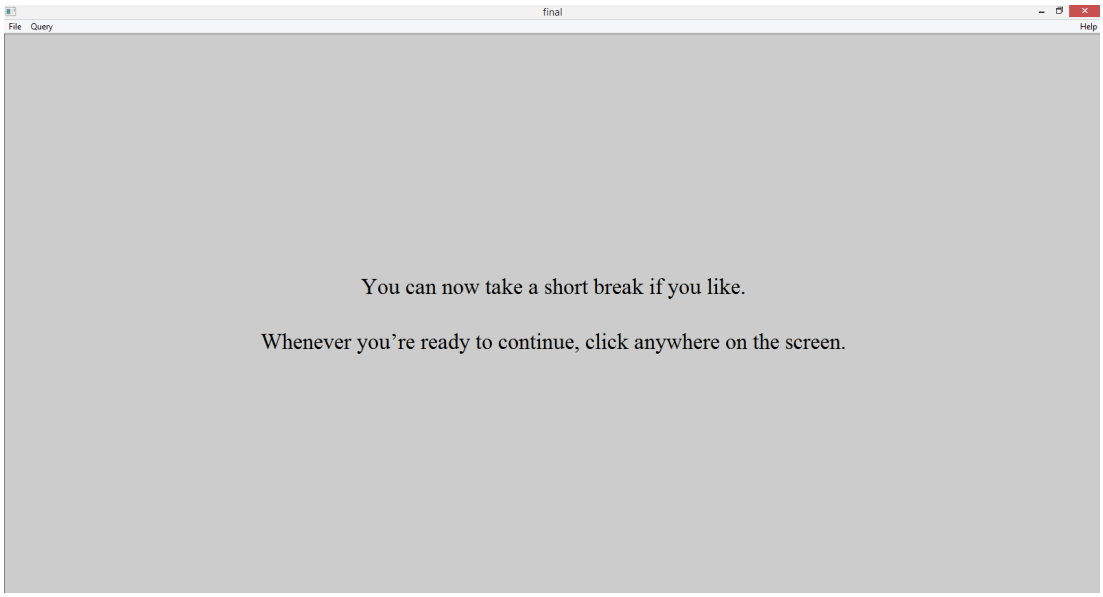


Figure 8: prompt for a break after each block

4 Results

In total, 25 subjects participated in the experiment and a total of 11875 responses were recorded (25*475). From these, 9375 (25*375) were responses to the pseudo words and 2500 (25*100) were responses to the real words. The group of 25 subjects includes 18 Czech learners of English and 7 native speakers of English, who functioned as a control group.

All participants used the keyboard (and not the touchpad) to respond to the stimuli and most did not take breaks longer than a few seconds between the blocks.

Some of the participants objected during or after the practice round that the pseudo words are the same. They were instructed that the words may indeed be spelled the same, but that they were computer-manipulated and that there are no two identical stimuli in the experiment.

4.1 Participants

Basic information and language background has been collected from the subjects. This included age, sex, possible hearing problems, formal knowledge of phonetics and knowledge of other languages apart from English, on a scale 1—5. The scale was defined from 1 being "nothing or almost nothing" to 5 being "close to a native speaker".

The subjects self-reported this information. Learners of English were also asked about their level of English proficiency (on the same scale), onset of learning, number of months spent in an English speaking country and frequency of listening to spoken English. Frequency was also defined on a scale from 1—5, based on the terms from *Table 7* below.

self-reported frequency	description
1	*once a month or less*
2	*a few times a month*
3	*weekly*
4	*a few times a week*
5	*daily*

Table 7: scale of frequency of listening to English

Finally, the native speakers were asked whether they are indeed native speakers of English and to self-define their dialect.

The average age of the participants was 26 years. 19 male participants and 7 female participants underwent the experiment. Each subject was assigned an identification number. The group of native speakers had 50 added to their number to be distinguishable even from their id only. 2 native speakers reported small hearing problems and 4 subjects (3 learners and one native speaker) reported that they have taken a phonetics course. English proficiency of the learners ranged from 1—5 with the average of 2,67.

For the full list of subjects with all the collected information see Appendix A.

4.2 Analysis

Responses to pseudo words with the reaction time of 2 or more seconds were discarded. There were 596 discarded responses which makes a percentage of 6.36% of all responses.

After the timed-out responses have been taken out, the total number of the responses to pseudo words stimuli was 8779. Out of these, 64,79% preferred the initial syllable.

4.2.1 Control Group

The level of 80% correct real words, which was set in the original study by Wang (2008), was met by 5 out of the 7 native speakers of English. The 2 subjects were therefore skipped for the following analysis. Moreover, subject 54, while having passed the 80% limit on the real words, responded to all 100% pseudo words with the initial syllable option. The subject was thus been skipped too. The control group for the analysis therefore consisted of 4 subjects (subjects 52, 53, 55 and 56). A factorial ANOVA was conducted for the control group. The dependent variable was the percentage of initial syllable responses (ISR) and there were three fixed factors: F0, intensity and duration (all three with 5 different steps). Partial Eta (η^2) squared was calculated to determine the relative importance of the different cues.

Figure 9 and *Table 8* below show the average percentage of ISR as a function of the tree different cues in all 5 levels of manipulation.

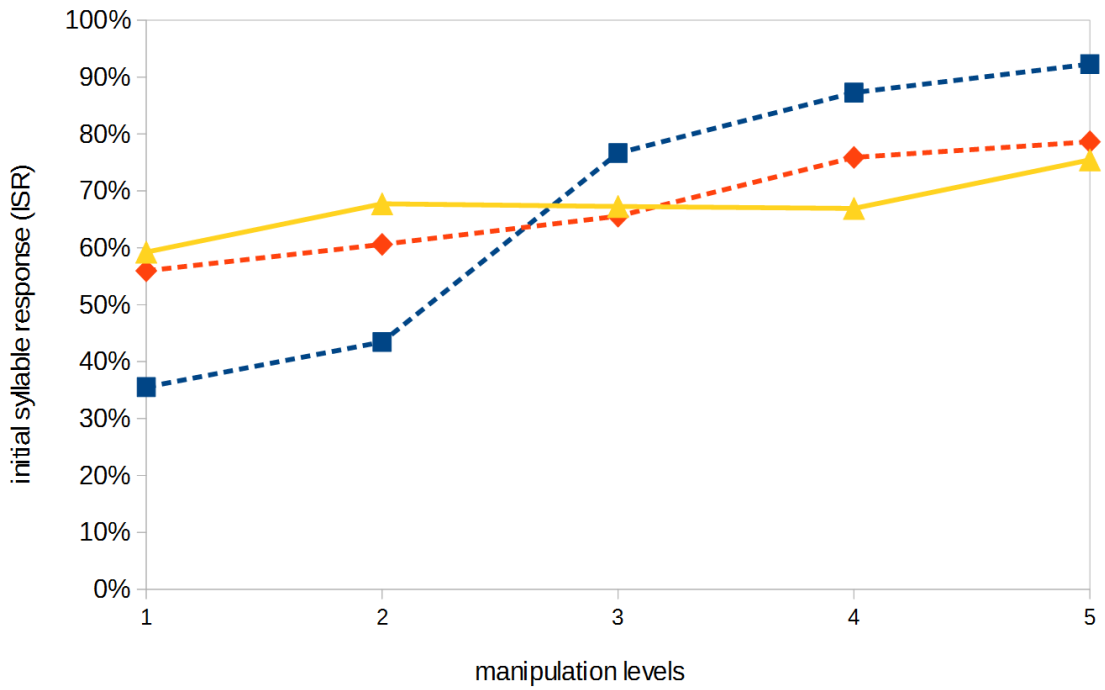


Figure 9: ISR as a function of F0, intensity and duration for all 5 levels of manipulation **(native speakers)**

	1	2	3	4	5
F0	36%	43%	77%	87%	92%
intensity	56%	61%	66%	76%	79%
duration	59%	68%	67%	67%	75%

Table 8: percentage of ISR as a function of F0, intensity and duration for all 5 levels of manipulation **(native speakers)**

Table 8 shows that for the control group of native speakers, a change in F0 from level 1 to level 5 of manipulation resulted in an increase in initial syllable responses from 36% to 92%. For intensity, the same change yielded an increase from 56% to 79%. Finally, for duration, the increase was from 59% to 75%.

The ANOVA analysis revealed that there were several significant effects on the judgment of stress by the native speakers. F0 had a significant effect on IRS, $F(4, 250) = 83.870$, $p < 0.001$, and so did intensity, $F(4, 250) = 12.009$, $p < 0.001$, and duration, $F(4, 250) = 3.3334$, $p = 0.01107$. The interaction of F0 and intensity also had a significant effect of $F(16, 250) = 1.8851$, $p = 0.02223$, and so did the interaction of F0 and duration, $F(16, 250) = 1.8888$, $p = 0.02189$. The remaining interactions had effects below the level of significance.

The calculated eta-squared for each of the significant effects or their interactions can be seen in *Table 9* below.

effect	eta-squared (η^2)
F0	0.57299933
intensity	0.161179694
duration	0.0506334076
F0*intensity	0.10765644
F0*duration	0.107844526

Table 9: eta-squared for significant effects **(native subjects)**

4.2.2 Learners Groups

Learners who did not reach 60% of correct answers from the real words were taken out from the results, and the rest were divided into two groups: a first group of learners who

scored between 80% and 100% and a second group of those who scored between 60% and 80%. 3 subjects were thus left out and there were 7 subjects in the first, more successful group and 8 subjects in the second, less successful group. For the complete results in the real-word test by all participants see Appendix B.

The same ANOVAs as for the control group were calculated for the two groups of learners. *Figure 10* and *Table 10* below show the percentage of ISR as a function of the three acoustic cues for the first group of learners and *Figure 11* and *Table 11* show the same information for the second group.

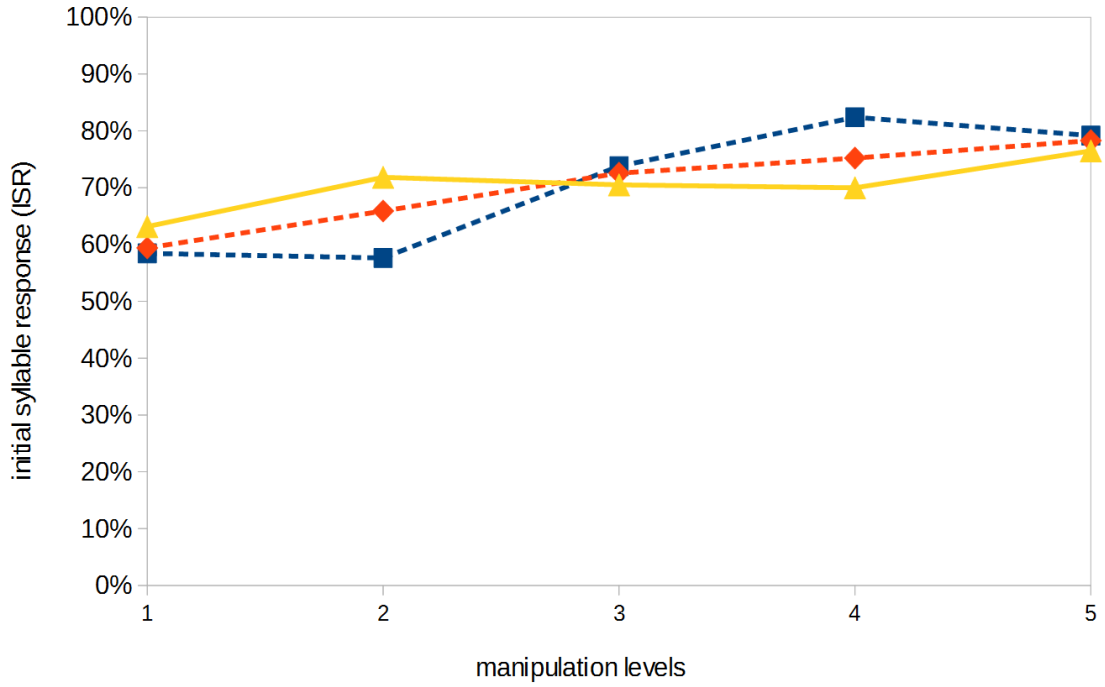


Figure 10: ISR as a function of F0, intensity and duration for all 5 levels of manipulation **(Czech speakers, group 1 = more successful)**

	1	2	3	4	5
F0	58%	58%	74%	82%	79%
intensity	59%	66%	73%	75%	78%
duration	63%	72%	71%	70%	76%

Table 10: percentage of ISR as a function of F0, intensity and duration for all 5 levels of manipulation **(Czech speakers, group 1 = more successful)**

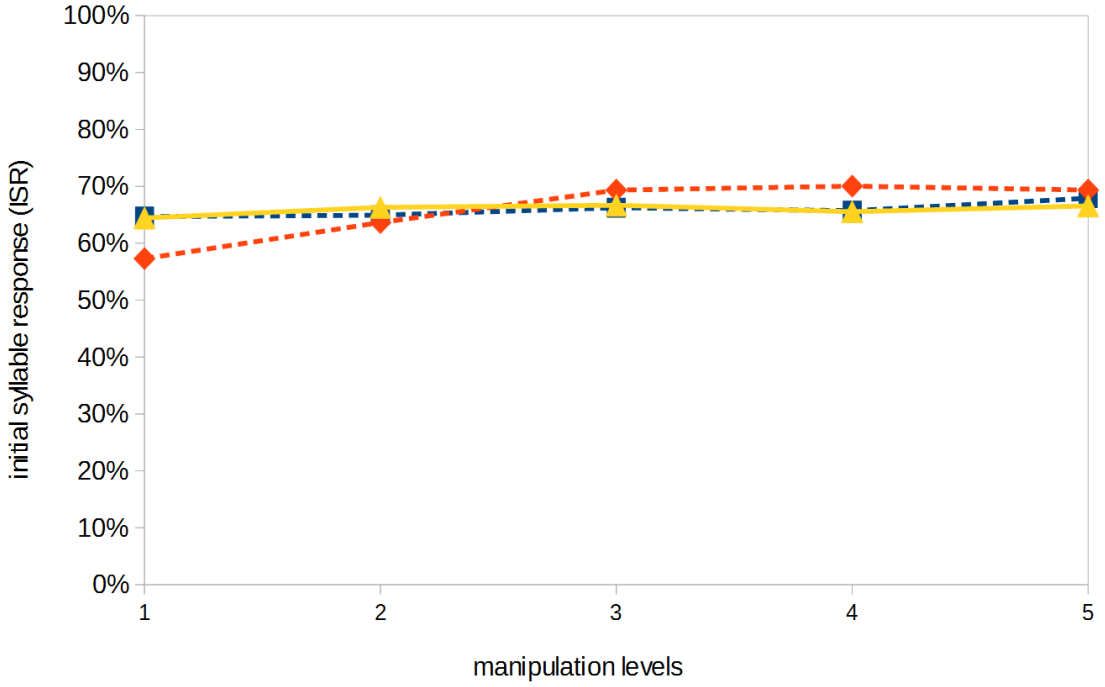


Figure 11: ISR as a function of F0, intensity and duration for all 5 levels of manipulation **(Czech speakers, group 2 = less successful)**

	1	2	3	4	5
F0	65%	65%	66%	66%	68%
intensity	57%	64%	69%	70%	69%
duration	64%	66%	67%	66%	67%

Table 11: percentage of ISR as a function of F0, intensity and duration for all 5 levels of manipulation **(Czech speakers, group 2 = less successful)**

The ANOVAs found fewer significant effects, compared to the control group. For the first group of learners (Czech 1), there were significant effects of F0, $F(4, 250) = 33.764$, $p < 0.001$, intensity, $F(4, 250) = 12.991$, $p < 0.001$, and duration, $F(4, 250) = 5.3643$, $p < 0.001$. No significant effects of interactions of the cues were found.

The partial eta-squared calculated for the three significant effects can be seen in *Table 12* below.

effect	eta-squared (η^2)
F0	0.35074389
intensity	0.172085887
duration	0.0790446203

Table 12: eta-squared for significant effects (Czech 1)

For the second group of learners (Czech 2), there was one significant effect of intensity, $F(4, 250) = 11.266$, $p < 0.001$. The eta-squared of the effect is 0.152728306.

For all the ANOVA analyses in graph form see Appendix E-G.

4.3 Language Background Correlations

A series of one-way ANOVAs were calculated to see possible correlations of certain language background information on the results in the test of 100 real words.

No significant effects of age, $F(7, 10) = 2.4293$, $p = 0.09845$, sex, $F(1, 16) = 0.60062$, $p = 0.44964$, self-reported level of English, $F(3, 14) = 1.5342$, $p = 0.24940$, onset of learning, $F(7, 10) = 1.3386$, $p = 0.32614$, a number of months spent in an English speaking country, $F(2, 15) = 1.4656$, $p = 0.26219$, or frequency of listening to spoken English, $F(4, 13) = 1.5706$, $p = 0.24040$, were found.

However, a significant effect of prior academic experience with phonetics was found, $F(1, 16) = 12.671$, $p = 0.00261$, as seen in *Figure 12* below.

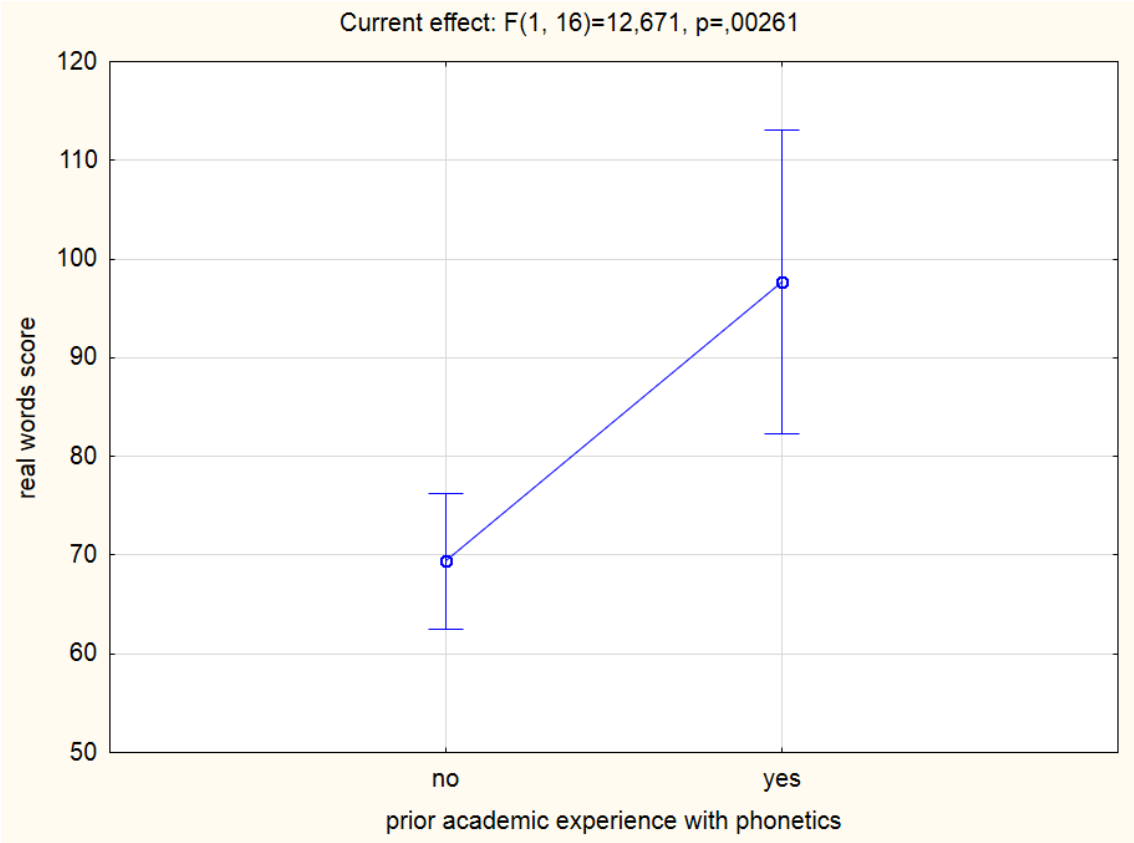


Figure 12: significant effect of phonetics experience on real words score

5 Discussion

The results of the control group show that there is a decisive effect of changes in fundamental frequency on the judgment of stress. Both duration and intensity manipulations had smaller, both also significant effects. The overall levels of initial syllable response are higher than in the original experiment by Wang (2008) and the fundamental frequency changes result in even higher percentage of initial syllable responses than in the original experiment. In other respects, however, the results are remarkably similar to Wang's. This suggests that changes in the tree manipulated cues do systematically effect native listeners' judgment. The present replication of the original study also gives support to the mentioned relative importance of the different cues. This is however limited to the used method of manipulating pseudo words in this particular way.

In the case of the learners, it can be concluded that recognizing a “more

prominent" syllable in real English words seems easy for some and difficult for others. This does not seem to be in direct correlation to the self-reported level of subjects' English, onset of learning, frequency of listening to English or the number of month spent in an English speaking country. Basic personal information such as sex and age do not correlate either. The only significant correlation was the prior experience with phonetics. Rather unsurprisingly, it seems that having the experience of making listening to sounds of a language (not English in particular, though[8]) a conscious activity does help in prominence identification tasks.

Turning to the learners and their responses to the pseudo words, it is clear that the subjects who scored better in the real words could on average exploit the manipulated acoustic cues significantly more. It could even be said that for the more successful group, the contours of the effects are in a fair way to merge with the ones of the native listeners. At this point however, the sensitivity to changes in fundamental frequency is not strong enough. Compare the strength of the effects for native speakers eta-squared F0 = 0.57, intensity = 0.16 and duration 0.05, and Czech 1 group F0 = 0.35, intensity = 0.17 and duration = 0.08.

For the less successful group, all the cues were used in approximately the same way with only the effect of intensity being of significant value. It can be speculated that the results would be different if different wording of the instructions had been used. "More prominent" may indeed be semantically closer to "louder" than to "stressed". "Stressed", however, requires some knowledge of basic linguistics. This may explain the only significant effect for the group Czech 1, which is intensity. Unfortunately, the original study does not provide this detail.

The correlation with phonetics experience can be interpreted as supportive for teaching of prominence patterns in EFL classes.

Importantly, what was empirically proven by Wang (2008) is not true for Czech speakers. While the Chinese subjects relied solely on the perceived pitch in recognizing the more prominent syllables, Czechs did not use duration to any significant degree. The notion of transfer of a phonologically important acoustic cue from L1 and using it similarly in L2 does not seem to be universally true. Czech English as an interlanguage

8 Two of the subjects (50%) who answered "yes" to the question about experience with phonetics have taken phonetics courses during their studies of German and Dutch majors, and not English. This information has unfortunately not been collected from the other two subjects.

does not seem to contain significant sensitivity to duration cues when perceiving English stress.

It may also be useful to point out that some of the manipulations in the original study by Wang (2008) were presented in a confusing way. Since it is a study of perception, it might be more helpful to use units of perception instead of acoustic qualities. This is most visible for the fundamental frequency. The relationship between frequency and perceived pitch is not that of 1 to 1 and so, in fact, the difference of 50 Hz is no twice the difference of 25 Hz, as it might seem from the description of the experiment. The difference between the two approaches for manipulations levels 1 and 2 for the pseudo word *latmab* is illustrated below in *Table 13* and *Table 14*.

latmab	**level 1**	**level 2**
syllable 1	148.2 Hz	160.7 Hz
syllable 2	198.2 Hz	185.7 Hz
difference	50 Hz	25 Hz

Table 13: differences between levels 1 and 2 for the word *latmab* defined in fundamental frequency

latmab	**level 1**	**level 2**
syllable 1	6.81	8.21
syllable 2	11.84	10.72
difference	5.03	2.06

Table 14: differences between levels 1 and 2 for the word *latmab* defined in semitones

While the ratio of the syllables defined in fundamental frequency is 2:1, in semitones the ratio comes to is 2:0.82.

5.1 Hypotheses

The hypothesis (1) that Czech listeners will indeed hear stress ***was neither proven nor rejected.*** Overall, the task was difficult for Czech listeners but it cannot be stated that they are 'stress-deaf'. **The hypothesis (2)** that Czech listeners will rely on duration as

a cue for stress for its important phonological function in their L1 ***can be rejected.*** The effect of duration on the stress judgment was the least significant out of the tree cues for group Czech 1 and not significant for group Czech 2. It seems that while Czech learners do successfully use duration to signal stress in production (Liska 2011; Volín and Weingartová 2014), this cannot be automatically taken to be true for perception. This supports the view that caution needs to taken when interpreting research results with respect to whether perception or production is tested.

5.2 Suggestions for Future Explorations

A more comprehensive experiment including vowel quality would yield valuable results. It would however have to be designed in such a way, so that it would not take too long to complete. Some participants in the present experiment said that they felt the experiment was too long. Subjectively speaking, I could see that some participants were not focused as much as I wished that they were. Ideally, the subjects' motivation could be boosted by reward or by introducing a university community service system.

I could also observe that because the three pseudo words in each set of stimuli were in the same order, some participants learned a single combination of responses which they used automatically. This could be avoided by making the order of the stimuli more or even absolutely random (disregarding the blocks and sets). As the present experiment is a replication however, the information in the original study was followed.

Also, a more comprehensive study of the extralinguistic factors influencing learners' ability to perceive stress could provide some practical solutions how to teach stress in EFL classes.

6 Conclusion

Literature on acoustics of stress and its production and perception has been reviewed. The topic is not fully understood and the present experiment contributes to the body of research with some empirical data about Czech English as an interlanguage.

Czech learners of English have been tested in a forced-choice experiment to find the “more prominent” syllable in a series of bisyllabic real English words and computer-manipulated English pseudo words. Most of the subjects performed above-chance on the real words but some only at chance. No correlation of language background

information for this was found, except that prior experience with phonetics significantly helps subjects hear stress.

The hypothesis that Czech learners will rely on vowel duration the most because it plays an important phonological role in their L1 was disproved. Instead, the more successful learners in the real word scores used the acoustic cues similarly to the native speakers, although the effect of fundamental frequency was not quite as strong.

7 References

Audacity Team. 2014. Audacity for Windows (version 2.0.5) [computer program]. Retrieved August 11, 2014, from http://audacity.sourceforge.net/.

Bičan, Aleš. 2008. "Ne-volnost přízvuku v češtině." *Varia Slavica: Sborník příspěkvů k 80. narozeninám Radoslava Večerky.* Praha: Nakladatelství Lidových novin.

Boersma, Paul, and David Weenink. 2014. Praat: doing phonetics by computer (version 5.3.82) [computer program]. Retrieved August 11, 2014, from http://www.praat.org/.

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

Cutler, Anne. 2005. "Lexical Stress." In *The Handbook of Speech Perception* edited by David B. Pisoni and Robert E. Remez, 264—289. Malden, MA: Blackwell Publishing.

Davies, Mark. 2014. *Word frequency data: Corpus of Contemporary American English*. Available online at http://www.wordfrequency.info/.

Duběda, Tomáš. 2006. "Intensity as a macroprosodic variable in Czech." *Speech Prosody 2006*:185–188.

Field, J. 2005. "Intelligibility and the listener: The role of lexical stress." *TESOL Quarterly* 39(3):399—423.

Fry D. B. 1955. "Duration and Intensity as Physical Correlates of Linguistic Stress." *The Journal of the Acoustical Society of America* 27(4):765–768.

Fry, D.B. 1958. "Experiments in the perception of stress." *Language and Speech 1:*126

—152. In: Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge (MA): MIT Press.

Fukada, Taku. 2013. LDOCE5 Viewer [computer program]. Retrieved August 11, 2014, from https://hakidame.net/ldoce5viewer/.

Hahn, L.D. 2004. "Primary Stress and Intelligibility: Research to Motivate the Teaching of Suprasegmentals" *TESOL Quarterly* 38(2):201—223.

Hála, Bohuslav. 1962. *Uvedení do fonetiky češtiny na obecně fonetickém základě.* Prague: ČSAV.

Ladefoged, Peter, and Keith Johnstone. 2013. *A Course in Phonetics*. 6th ed. Boston (MA): Wadsworth/Cengage Learning.

Lehiste, Ilse, and Gordon E. Peterson. 1959. "Vowel Amplitude and Phonemic Stress in American English." *The Journal of the Acoustical Society of America (31)*: 28—435.

Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge (MA): MIT Press.

Lieberman, Philip. 1960. "Some Acoustic Correlates of Word Stress in American English." *The Journal of the Acoustical Society of America (32):* 451—454.

Longman. 2011. *Longman dictionary of contemporary English (5th ed.)*. Harlow, England: Longman.

Palková, Zdena, and Jan Volín. 2003. "The role of F0 contours in determining foot boundaries in Czech." *Proc. 15th ICPhS Barcelona*:1783–1786.

Palková, Zdena. 1994. *Fonetika a fonologie češtiny s obecným úvodem do problematiky oboru.* Prague: Karolinum.

Podlipský, V. J. 2007. "Perception of Czech vowel quantity by English learners of Czech." *Proceedings of the 16th International Congress of Phonetic Sciences, Saarbrücken*:1709–1712.

Podlipský, V. J. 2009. "The interaction between word-boundary perception and vowel length in native and non-native speakers of Czech." *Acta Univ Palackianae Olomucensis, Fac Philos, Anglicka III, Linguistica*:89–100.

Šefčík, Ondřej. 2002. "K (ne)pohyblivosti přízvuku (typologická poznámka)." *Čeština – univerzália a specifika 4*:113–121. Praha: Nakladatelství Lidových novin.

Volín, Jan, and Lenka Weingartová. 2014. “Acoustic Correlates of Word Stress as a Cue to Accent Strength.” *Research in Language* 12(2):175–183.

Volín, Jan. 2008. “Z intonace čtených zpravodajství: výška první slabiky v taktu.” *Čeština doma a ve světě* 16(1–2):89–96.

Wang, Qian. 2008. “L2 Stress Perception: The reliance on different acoustic cues.” *Proc. of the 4th Conf. on Speech Prosody, Campinas, Brazil*:635–638.

Wells, J. C. 2008. *Longman pronunciation dictionary (3rd ed.)*. Harlow, England: Pearson Education Limited/Longman.

8 Appendices

Below you will find some more important data in table and graph forms that are presentable on paper. The complete data including all the responses, all participant information, manipulation values and all the recordings, as well as STATISTICA 64 files, Praat scripts and interactive spreadsheets can be however found in a more convenient electronic form on the attached CD or per request at *petr.tichy01@upol.cz*.

The CD contains:

(1) TICHY_BDP.doc – the present paper in a .doc format. The file was created and worked on in the open-source program LibreOffice Writer and may thus exhibit minor incompatibilities in other office software.

(2) TICHY_BDP.pdf – the present paper in a PDF format suitable for printing.

(3) DATA.ods – interactive spreadsheet from the open-source program LibreOffice Calc with all the data used in this paper. It contains several tabs:

(a) STIMULI (learners 1) – contains data of **the more successful** from the two Czech groups, after the timed-out responses were taken out. There is a graph of the relative importance of the cues of all the speakers in this tab, and below another interactive graph where results of a single participant can be viewed upon entering the desired participant id.

(b) STIMULI (learners 2) – contains data of **the less successful** from the two Czech groups, after the timed-out responses were taken out. There is a graph of the relative importance of the cues of all the speakers in this tab, and below another interactive graph where results of a single participant can be viewed upon entering the desired participant id.

(c) STIMULI (natives) – contains data of **control group**, after the timed-out responses were taken out. There is a graph of the relative importance of the cues of all the speakers in this tab, and below another interactive graph where results of a single participant can be viewed upon entering the desired participant id.

(d) 100 WORDS CHECK – an interactive tab which shows the score from the real words test upon entering all 475 response lines. On the right, the scores are kept.

(e) MANIPULATIONS – contains all values that were used for manipulating the stimuli. All the values except the basic ones are calculated.

(f) REAL WORDS – contains the complete list of real words used in the

experiment.

(g) ANOVA – contains information in a table form suitable for ANOVAs.

(h) ALL RESPONSES – contains all response lines to pseudo words after the timed-out responses have been taken out.

(4) folder “results” – all STATISTICA 64 12 files containing the stimuli responses and all relevant ANOVA and eta-squared calculations.

(5) folder “real words” – all recorded real words.

(6) folder “kwasi words” – all recorder and manipulated pseudo words, including scripts used to manipulate and export them.

(7) folder “procedure” – all experiment Praat scripts, the necessary stimuli and complete unchanged results of all participants.

APPENDIX A

participant information

no.	age	sex	hearing problems	native/learner	other languages (1-5)*	formal phonetics	*for learners only*: level of English*	onset	months in English speaking country	frequency of listening of E**	*for natives only*: dialect
1	23	m	none	l	Spanish 1	no	3	9	1	4	N/A
2	26	m	none	l	Spanish 1	no	3	9	0	5	N/A
3	24	m	none	l	-	no	2	23	0	4	N/A
4	19	f	none	l	German 3	no	2	14	0	2	N/A
5	25	m	none	l	German 2	no	2	6	0	2	N/A
6	49	m	none	l	Russian 2, German 2	no	1	never	0	1	N/A
7	24	m	none	l	Spanish 1	no	3	10	0	3	N/A
8	23	f	none	l	Spanish 1	no	3	10	0	3	N/A
9	23	f	none	l	German 1	no	3	10	0	4	N/A
10	23	m	none	l	Spanish 1	no	2	10	0	2	N/A
11	24	m	none	l	French 2	no	3	9	1	3	N/A
12	22	m	none	l	-	no	3	9	0	4	N/A
13	23	m	none	l	-	no	2	11	0	2	N/A
51	24	m	none	n	Czech 2	no	N/A	N/A	N/A	N/A	British (Black Country dialect)
14	24	m	none	l	Spanish 2, French 1	no	3	9	0	4	N/A
52	29	f	none	n	French 5, Czech 1	no	N/A	N/A	N/A	N/A	General Canadian
53	32	m	hearing better in right ear	n	Czech 2, German 1	no	N/A	N/A	N/A	N/A	British (Southern Country)
54	28	m	none	n	Czech 3, German 1	no	N/A	N/A	N/A	N/A	British (Northern English)
55	28	f	none	n	Spanish 4, Russian 3, Ukrainian 3	yes	N/A	N/A	N/A	N/A	Standard American (slightly rural)
56	30	m	none	n	French 3, Czech 1	no	N/A	N/A	N/A	N/A	American (Mid-Atlantic)
15	23	f	none	l	German 4	yes	2	10	0	2	N/A
16	33	m	none	l	German 5, Dutch 3, Russian 2	yes	3	12	3	5	N/A
57	33	m	once failed midtones hearing check	n	Italian 1, Spanish 1, Czech 1	no	N/A	N/A	N/A	N/A	American (New York)
17	23	m	none	l	-	yes	5	10	0	5	N/A
18	24	m	none	l	Russian 1	no	3	9	0	4	N/A

	*level of English
1	nothing or almost nothing
2	...
3	...
4	...
5	close to a native speaker

	**frequency of listening to English
1	once a month or less
2	a few times a month
3	weekly
4	a few times a week
5	daily

APPENDIX B

The results of the test of real words: subjects marked with an asterisk (*) were dismissed based on their low score. Native speaker 54 (**) was dismissed because they responded with the first syllable option to 100% of the stimuli.

Two groups of Czechs were created: Czech 1 with the scores between 80—100 and Czech 2 with scores between 60—80.

learners		**native speakers**	
id	score	id	score
1	88 (Czech 1)	51	75*
2	75 (Czech 2)	52	99 (control)
3	62 (Czech 2)	53	95 (control)
4	53*	54	94**
5	72 (Czech 2)	55	100 (control)
6	91 (Czech 1)	56	96 (control)
7	69 (Czech 2)	57	64*
8	88 (Czech 1)		
9	83 (Czech 1)		
10	53*		
11	50*		
12	64 (Czech 2)		
13	70 (Czech 2)		
14	60 (Czech 2)		
15	95 (Czech 1)		
16	100 (Czech 1)		
17	98 (Czech 1)		
18	63 (Czech 2)		

APPENDIX C

100 real words used in the experiment

stressed finally		
rank	word	part of speech
46	about	i
89	because	c
139	become	v
140	between	i
164	begin	v
180	against	i
184	again	r
213	believe	v
219	without	i
220	before	i
263	provide	v
265	around	i
270	away	r
271	until	c
280	among	i
293	include	v
333	create	v
340	allow	v
353	within	i
355	result	n
370	himself	p
375	enough	r
378	across	i
379	although	c
396	appear	v
406	expect	v
422	behind	i
427	remain	v
428	effect	n
431	suggest	v
433	control	n
436	perhaps	r
448	along	i
451	report	v
457	decide	v
469	police	n
473	return	v
481	explain	v
499	receive	v
515	agree	v
525	support	v
526	event	n
543	produce	v
557	itself	p
569	describe	v
641	campaign	n
657	involve	v
658	defense	n
660	increase	v
663	myself	p

stressed initially		
rank	word	part of speech
65	into	i
75	other	j
87	also	r
99	many	d
101	only	r
105	very	r
107	even	r
111	woman	n
120	after	i
124	over	i
138	never	r
143	something	p
157	student	n
166	country	n
171	problem	n
172	every	a
191	system	n
194	program	n
197	question	n
198	during	i
204	number	n
206	always	r
217	happen	v
226	under	i
230	mother	n
233	money	n
234	story	n
241	study	n
247	business	n
248	issue	n
257	little	j
264	service	n
268	father	n
272	power	n
273	hour	n
275	often	r
281	ever	r
286	member	n
296	later	r
309	minute	n
314	body	n
316	nothing	p
320	social	j
322	whether	c
326	follow	v
328	parent	n
334	public	j
339	level	n
342	office	n
346	person	n

APPENDIX D

complete manipulation values for all three pseudo words

manipulation values key					
correlates	**levels**				
	1	2	3	4	5
F0 [Hz]	-50	-25	0	25	50
intensity [dB]	-9	-4,5	0	4,5	9
duration [proportion of avg]*	0.5	0.25	0	-0.25	-0.5

**Changes have been made to this row in comaprison to the original.*

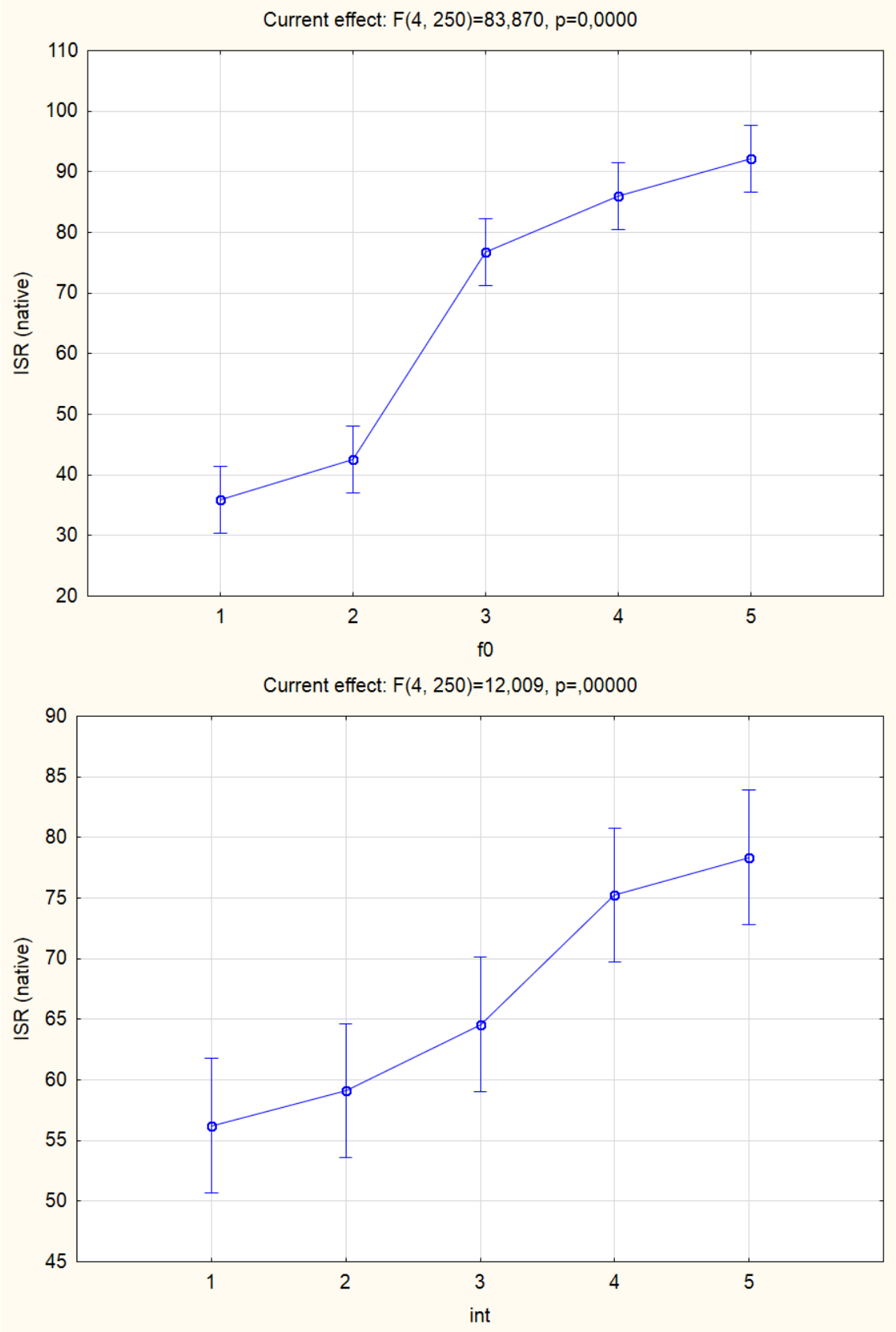
manipulation values: ***latmab***						
orig. sound file duration [sec]		0,678571				
orig. 1st syllable nucleus (from; to) [sec]		0,094351		>	0,16471	
orig. 2nd syllable nucleus (from; to) [sec]		0,400854		>	0,45949	
orig. F0 (1st; 2nd; mean) [Hz]		175,7	170,7	>	**173,2**	
orig. intensity (1st; 2nd; mean) [dB]		74,53	63,28	>	**68,905**	
orig. duration (1st; 2nd; mean) [sec]		0,070359	0,058636	>	**0,0644975**	
after level-3 manip. nuclei (1st; 2nd) [sec]		**0,094351**	**0,1588485**		**0,3949925**	**0,45949**
correlates		**levels**				
		1	2	3	4	5
F0	*1st*	148,2	160,7	**173,2**	185,7	198,2
	2nd	198,2	185,7		160,7	148,2
	diff	-50	-25		25	50
intensity	*1st*	64,405	66,655	**68,905**	71,155	73,405
	2nd	73,405	71,155		66,655	64,405
	diff	-9	-4,5		4,5	9
	1st (rel.)	-4,5	-2,25		2,25	4,5
	2nd (rel.)	4,5	2,25		-2,25	-4,5
duration	*1st*	48,37	56,44	**64,50**	72,56	80,62
	2nd	80,62	72,56		56,44	48,37
	diff	-0,50	-0,25		0,25	0,50
	1st (rel.)	0,75	0,875		1,125	1,25
	2nd (rel.)	1,25	1,125		0,875	0,75

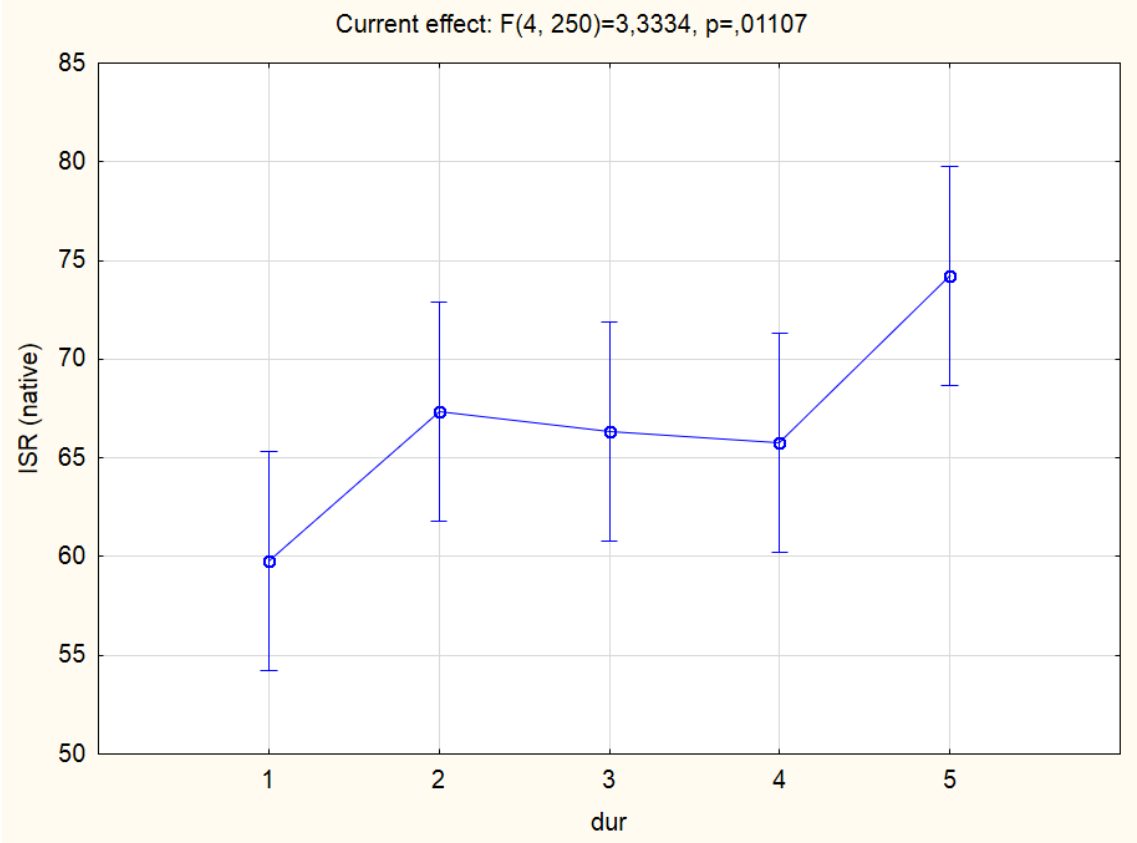
manipulation values: ***nizdit***						
orig. sound file duration [sec]		0,527029				
orig. 1st syllable nucleus (from; to) [sec]		0,090276		>	0,167693	
orig. 2nd syllable nucleus (from; to) [sec]		0,336567		>	0,388937	
orig. F0 (1st; 2nd; mean) [Hz]		184,75	162,79	>	**173,77**	
orig. intensity (1st; 2nd; mean) [dB]		74,02	70,81	>	**72,415**	
orig. duration (1st; 2nd; mean) [sec]		**0,077417**	**0,05237**	>	**0,0648935**	
after level-3 manip. nuclei (1st; 2nd) [sec]		**0,090276**	**0,1551695**		**0,3240435**	**0,388937**
correlates		**levels**				
		1	2	3	4	5
F0	*1st*	148,77	161,27	**173,77**	186,27	198,77
	2nd	198,77	186,27		161,27	148,77
	diff	-50	-25		25	50
intensity	*1st*	67,915	70,165	**72,415**	74,665	76,915
	2nd	76,915	74,665		70,165	67,915
	diff	-9	-4,5		4,5	9
	1st (rel.)	-4,5	-2,25		2,25	4,5
	2nd (rel.)	4,5	2,25		-2,25	-4,5
duration	*1st*	48,67	56,78	**64,89**	73,01	81,12
	2nd	81,12	73,01		56,78	48,67
	diff	-0,50	-0,25		0,25	0,50
	1st (rel.)	0,75	0,875		1,125	1,25
	2nd (rel.)	1,25	1,125		0,875	0,75

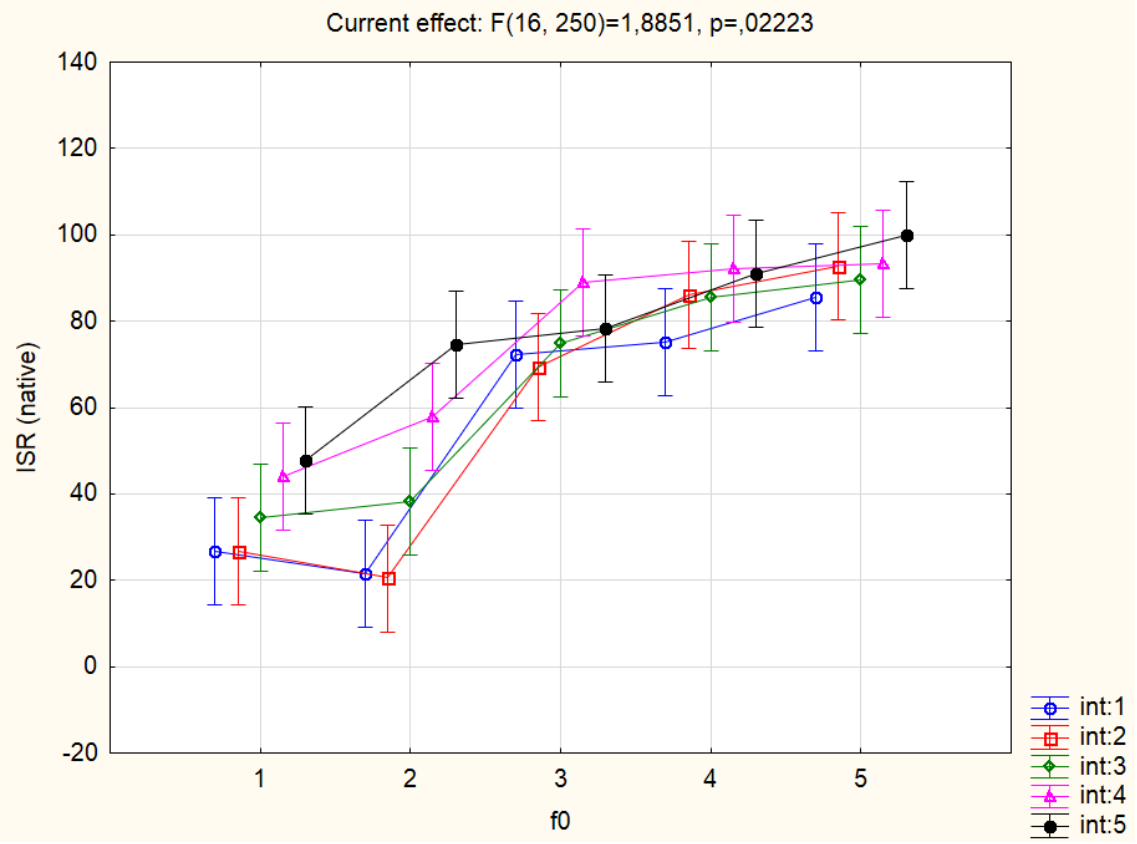
manipulation values: ***tetsep***						
orig. sound file duration [sec]		0,596644				
orig. 1st syllable nucleus (from; to) [sec]		0,059052		>	0,119514	
orig. 2nd syllable nucleus (from; to) [sec]		0,319179		>	0,363236	
orig. F0 (1st; 2nd; mean) [Hz]		186,42	163,38	>	**174,9**	
orig. intensity (1st; 2nd; mean) [dB]		74,97	68,73	>	**71,85**	
orig. duration (1st; 2nd; mean) [sec]		**0,060462**	**0,044057**	>	**0,0522595**	
after level-3 manip. nuclei (1st; 2nd) [sec]		**0,059052**	**0,1113115**		**0,3109765**	**0,363236**
correlates		**levels**				
		1	2	3	4	5
F0	*1st*	149,9	162,4	**174,9**	187,4	199,9
	2nd	199,9	187,4		162,4	149,9
	diff	-50	-25		25	50
intensity	*1st*	67,35	69,6	**71,85**	74,1	76,35
	2nd	76,35	74,1		69,6	67,35
	diff	-9	-4,5		4,5	9
	1st (rel.)	-4,5	-2,25		2,25	4,5
	2nd (rel.)	4,5	2,25		-2,25	-4,5
duration	*1st*	39,19	45,73	**52,26**	58,79	65,32
	2nd	65,32	58,79		45,73	39,19
	diff	-0,50	-0,25		0,25	0,50
	1st (rel.)	0,75	0,875		1,125	1,25
	2nd (rel.)	1,25	1,125		0,875	0,75

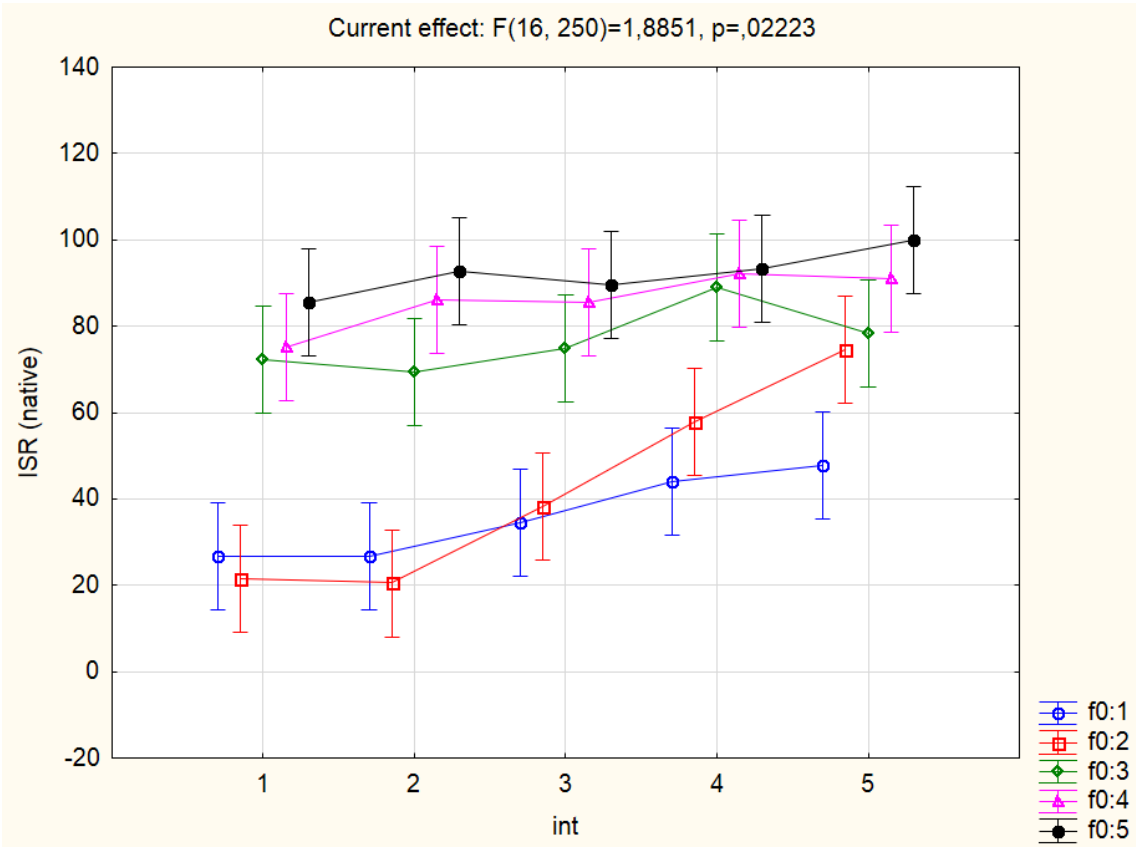
APPENDIX E

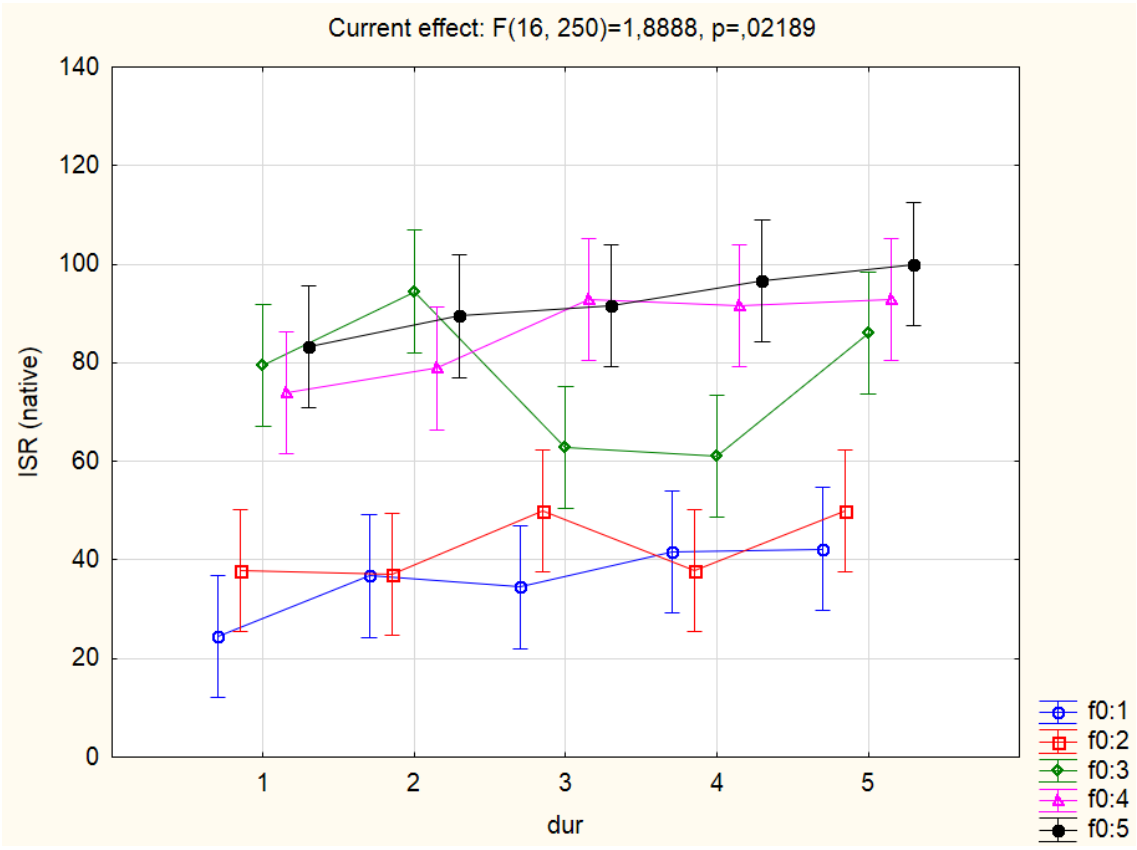
all significant ANOVA graphs for control group (simple and interactions)

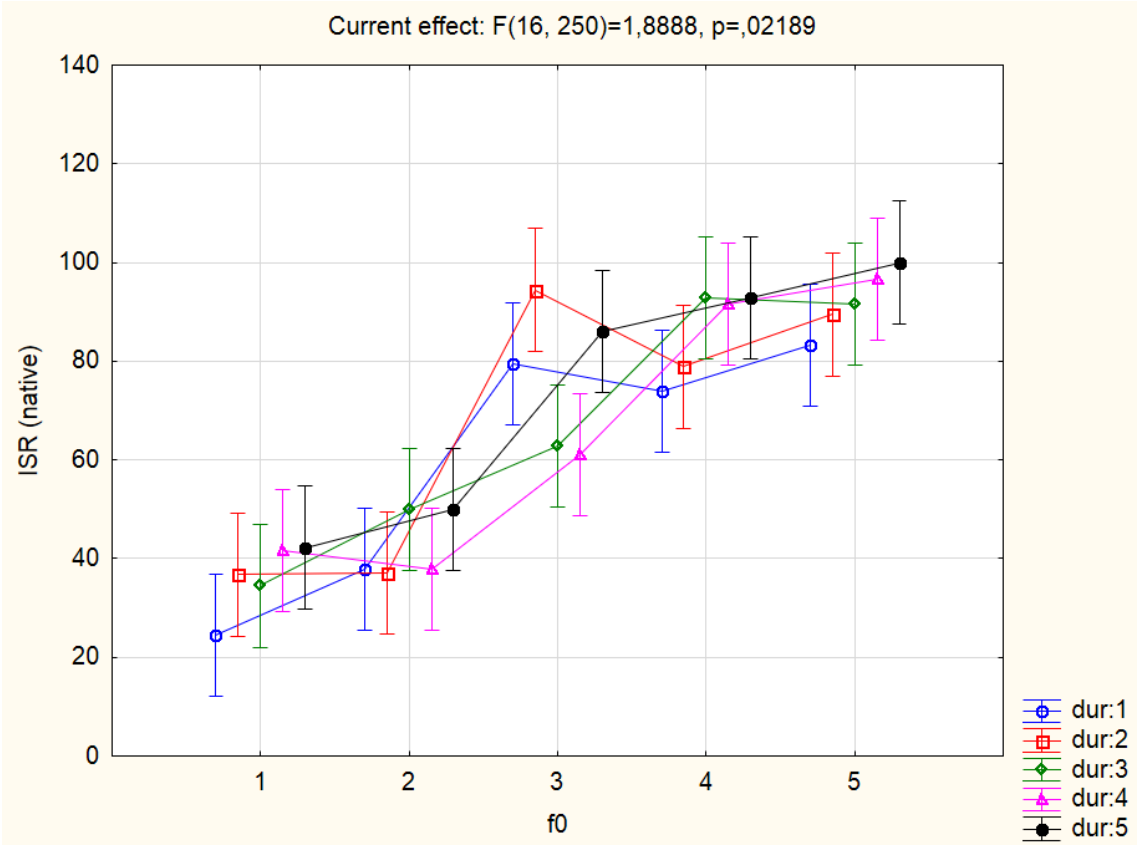



Current effect: F(4, 250)=3,3334, p=,01107
ISR (native)
85
80
75
70
65
60
55
50
1
2
3
4
5
dur


Current effect: F(16, 250)=1,8851, p=,02223
ISR (native)
140
120
100
80
60
40
20
0
-20
1
2
3
4
5
f0
int:1
int:2
int:3
int:4
int:5

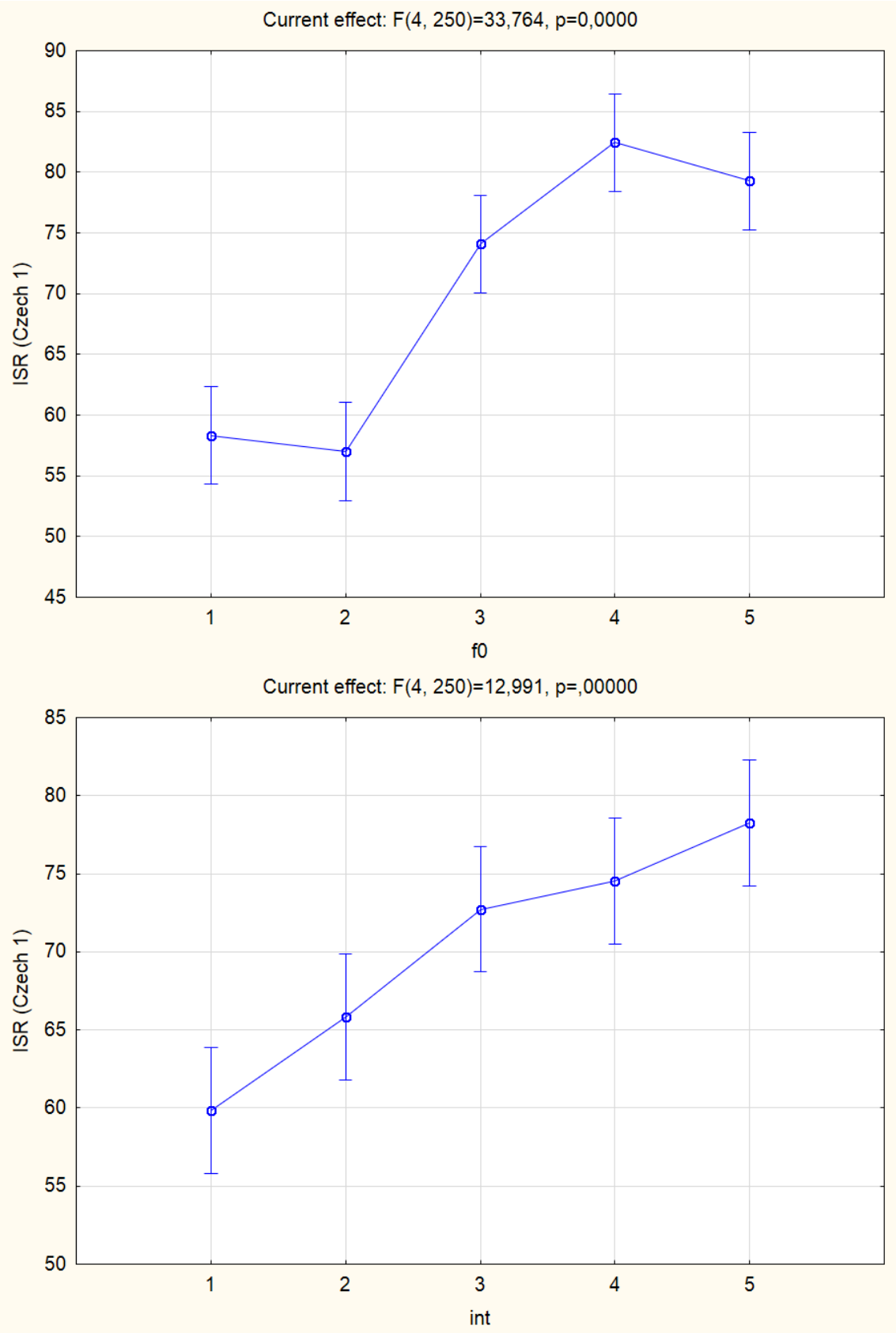

Current effect: F(16, 250)=1,8851, p=,02223
ISR (native)
int
f0:1
f0:2
f0:3
f0:4
f0:5

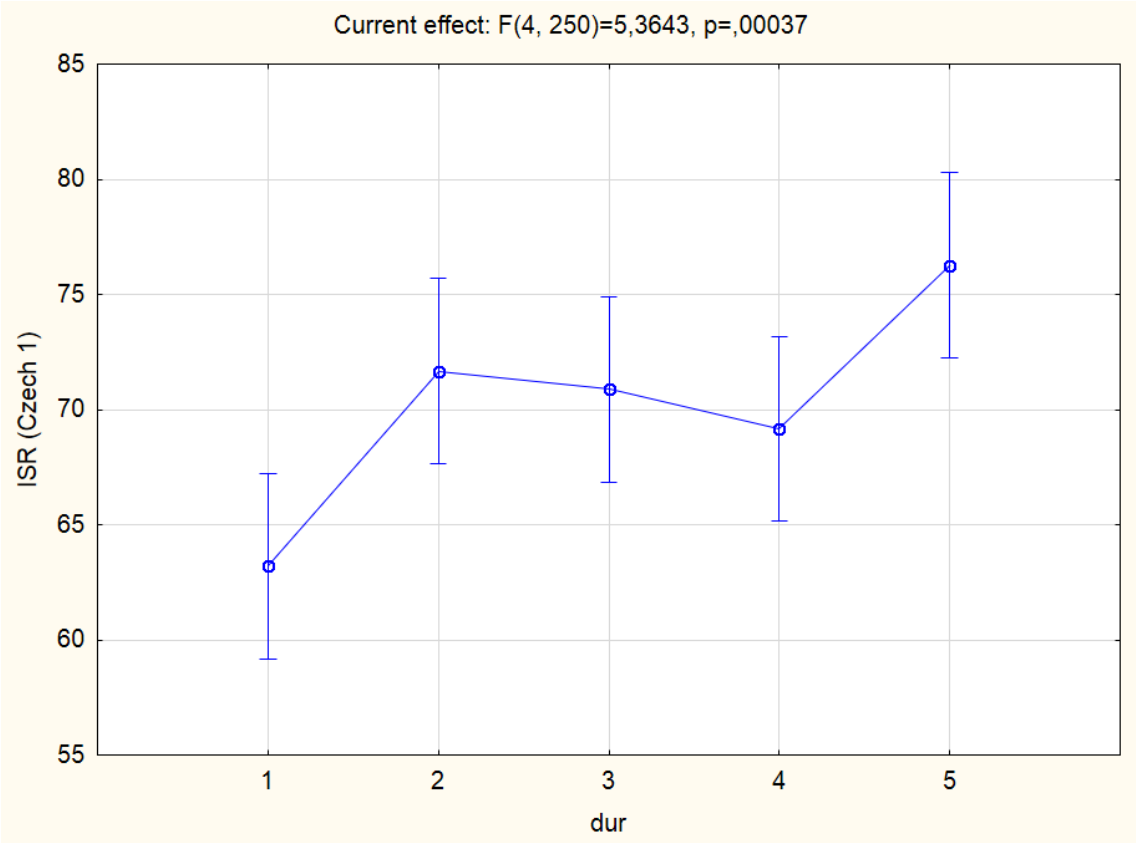

Current effect: F(16, 250)=1,8888, p=,02189
ISR (native)
dur
f0:1
f0:2
f0:3
f0:4
f0:5


Current effect: F(16, 250)=1,8888, p=,02189
ISR (native)
140
120
100
80
60
40
20
0
1
2
3
4
5
f0
dur:1
dur:2
dur:3
dur:4
dur:5

APPENDIX F

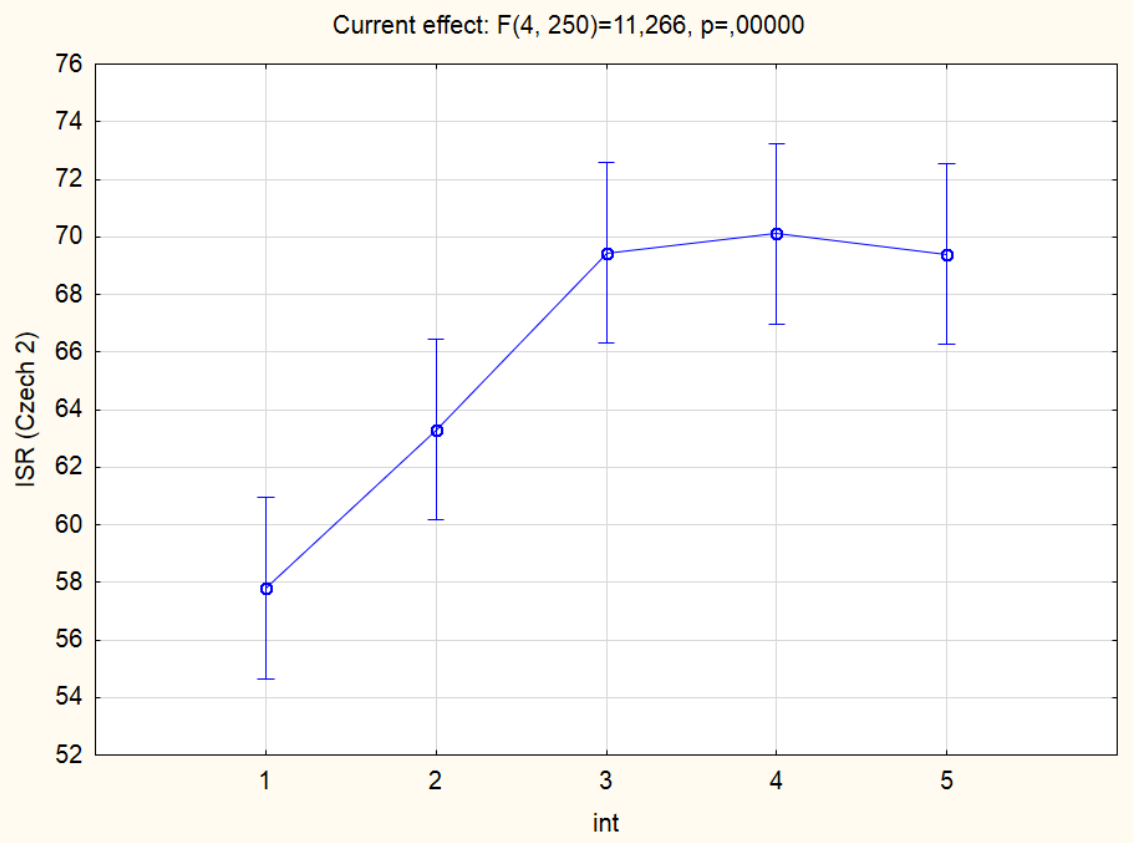
all significant ANOVA graphs for group Czech 1 (simple)




Current effect: F(4, 250)=5,3643, p=,00037
ISR (Czech 1)
85
80
75
70
65
60
55
1
2
3
4
5
dur

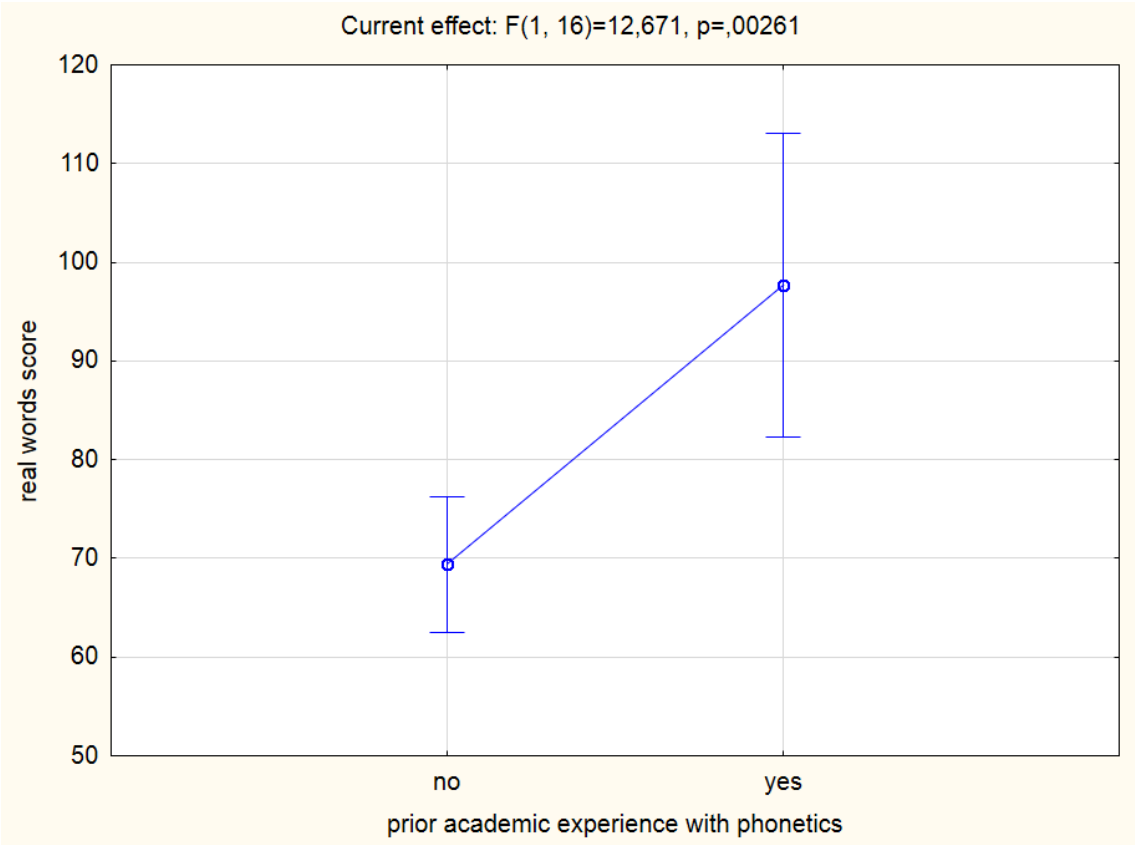
APPENDIX G

significant effect of intensity for group Czech 2 in ANOVA graph (simple)



APPENDIX G

significant effect of prior phonetics knowledge



SUMMARY

university: Palacký University, Olomouc

faculty: Philosophical Faculty

department: Department of English and American Studies

author: **Petr Tichý**

title in English: **Perception of English Lexical Stress by Czech Learners of English: Reliance on Different Acoustic Cues**

title in Czech: **Vnímání anglického slovního přízvuku českými mluvčími: důležitost jednotlivých akustických podnětů**

supervisor: Mgr. Šimáčková, Ph.D.

page count: **60** supplements: **CD (1)** references: **27**

key words in English: stress, accent, prosody, suprasegmentals, Czech, English, acoustic cues, fundamental frequency, intensity, duration

key words in Czech: přízvuk, prozódie, suprasegmentální znaky, čeština, angličtina, akustické podněty, základní frekvence, síla, délka

abstract in English: This paper is a replication of a study by Wang (2008) who manipulated English bisyllabic pseudo words in three acoustic dimensions (F0, intensity, duration) for a perception experiment in which Chinese listeners judged syllable prominence. The listeners were sensitive to pitch only and this was argued to be a result of transfer of acoustic sensitivity from a tonal L1. No such transfer of phonologically active vowel duration was found for Czechs. Listeners more successful in judging real words showed sensitivity to pseudo stimuli similar to that of native speakers'. Listeners exploited all three cues (in order of degree of reliance) F0, intensity and duration.

abstract in Czech: Tato práce je replikací studie od Wang (2008), kteří manipulovali anglická dvojslabičná pseudo slova ve třech akustických dimenzích (F0, intenzita, délka) pro percepční experiment, v němž činští posluchači posuzovali výraznost slabik. Posluchači se řídili pouze výškou tónu, což bylo považováno za výsledek přenosu akustické citlivosti z tónového L1. Žádný podobný přenos nebyl zjištěn pro fonologicky aktivní délku samohlásek pro Čechy. Posluchači, kteří byli úspěšnější při hodnocení reálných slov, předvedli podobnou citlivost na pseudo slovech jako rodilí mluvčí. Posluchači využili všech třech podnětů (v pořadí dle stupně důležitosti) F0, intenzity a délky.