Filozofická fakulta Univerzity Palackého

The effect of pronunciation instruction on the fluency of advanced Czech learners of English

(Diplomová práca)

2021

Bc. Monika Vlčková

Filozofická fakulta Univerzity Palackého

Katedra anglistiky a amerikanistiky

The effect of pronunciation instruction on the fluency of advanced Czech learners of English

(Diplomová práca)

Autor: **Bc. Monika Vlčková**

Študijný obor: **Anglická filologie**

Vedúci práce: **Mgr. Šárka Šimáčková, PhD.**

Počet strán (znaky/1800): 74,4

Počet strán: 81

Počet znakov (bez príloh): 133 863

Abstract:

This master's thesis is focussed on addressing the effects of explicit pronunciation instruction on oral reading fluency of Czech learners of English. The participants of the study are all university students at the department of English and American studies and have a good command of the English language. However, as various studies suggest (among them Derwing and Munro 2005, Piske et al. 2001, and Flege et al. 1995), even the most advanced speakers of an L2 rarely achieve native-like speech. Thus, methodology of this study is based on the comparison of the oral reading fluency, namely on the features of pitch and duration of utterances, before and after the explicit pronunciation instruction and learner's performance is compared to the performance of four native speakers to see whether the instruction can bring the L2 production to the native-like level.

In the review of literature, I explored the various definitions of fluency and its key features, addressed the differences between spontaneous speech prosody and reading prosody and explained how the text's punctuation can help or hinder fluent reading. In the second chapter of the review of literature, I examined the general tendencies of pronunciation instruction and the effectiveness of suprasegmental and segmental instruction in enhancing reading fluency. Furthermore, I examined the role of the teacher and corrective feedback, as well as the objective and subjective reading fluency assessment methods.

Within the practical part of the thesis, I analysed the instruction process that, due to the outbreak of the pandemic, took the form of distance education with delayed feedback. In the analysis of the recordings, the focus was on pitch span and pitch level and the method of the analysis was based on Patterson (2000). The second analysed feature of fluent reading was the utterance duration. The first hypothesis of this thesis is that the pitch range before the pronunciation instruction will be narrower than that of the native speaker. The second hypothesis is that the instruction will bring the non-native speakers' production to the native model. The third hypothesis is that the pitch level will not change as a result of the instruction. And the fourth hypothesis is that the duration of the utterance will be slower after the pronunciation instruction than before the instruction.

Despite the fact that the results were not unanimous for all speaker, they indicate that the L2 production before the pronunciation instruction was indeed narrower in terms of pitch span than that of the native speaker. The move towards the native model was, however, not observed and the second hypothesis had to be rejected. As was expected, the

results indicated that the pitch level did not change as a result of the instruction. And lastly, the duration of the utterance moved closer to the native model as a result of the instruction. Thus, three of the hypotheses has been supported by the results of the analysis and one has been refuted.

Keywords: L2 fluency, reading fluency, pitch range, utterance duration, L2 English, Czech learners of English, advanced learners

Abstrakt:

Táto diplomová práca je zameraná na skúmanie vplyvu výuky výslovnosti na plynulosť čítaného prejavu českých študentov angličtiny. Všetci účastníci na tomto výskume sú študentmi Univerzity Palackého v Olomouci na katedre Anglistiky a Amerikanistiky a anglický jazyk ovládajú dobre. Avšak ako rôzne štúdie dokazujú (medzi nimi napríklad Derwing and Munro 2005, Piske et al. 2001, and Flege et al. 1995), aj tí najpokročilejší hovoriaci v cudzom jazyku málokedy dosiahnu úroveň výslovnosti rodeného hovoriaceho. Preto je metodológia tejto práce zameraná na porovnávaní plynulosti čítaného prejavu, hlavne vlastností ako je intonačné rozpätie a dĺžka prejavu, pred a po výuke výslovnosti. Aby som zistila, či dokáže výuka ovplyvniť plynulosť reči a priblížiť výkon nerodených hovoriacich k tým rodeným, je výkon jednotlivých študentov ďalej porovnávaný k výkonu štyroch rodených hovoriacich.

V teoretickej časti sa zaoberám rôznymi definíciami plynulosti reči a jej základnými znakmi, ďalej popisujem rozdiely medzi prozodickými vlastnosťami spontánnej reči a čítaného prejavu a vysvetľujem aký vplyv má interpunkcia textu na plynulý čítaný prejav. V druhej kapitole teoretickej časti sa zaoberám všeobecnými tendenciami v oblasti výuky výslovnosti a vplyvom segmentálnej a suprasegmentálnej výuky na plynulosť čítaného prejavu. Ďalej vysvetľujem rolu učiteľa a jeho spätnej väzby a popisujem subjektívny a objektívny druh hodnotenia plynulosti čítaného prejavu.

V praktickej časti práce analyzujem samotnú výuku, ktorá vzhľadom na pandémiu koronavírusu prebiehala v distančnej forme s oneskorenou spätnou väzbou. V analýze jednotlivých nahrávok som sa zamerala na intonačné rozpätie a úroveň hlasu, čo bolo založené na príklade diplomovej práce Pattersona (2000). Druhá vlastnosť plynulého čítaného prejavu, na ktorú som sa zamerala je dĺžka prejavu. Prvá hypotéza tejto práce je, že intonačné rozpätie v nahrávke s pred výuky bude užšie v porovnaní s rodenými hovoriacimi. Druhá hypotéza hovorí, že výukou výslovnosti dokážeme priblížiť intonačné rozpätie nerodených hovoriacich k modelu rodeného hovoriaceho. Tretia hypotéza hovorí, že úroveň hlasu sa vplyvom výuky nezmení a štvrtá hypotéza, že čítaný prejav bude pomalší v nahrávkach po výuke než bol pred výukou.

Aj napriek tomu, že neboli jednotné pre každého hovoriaceho, štatistické výsledky naznačujú, že intonačné rozpätie nerodených hovoriacich bolo v nahrávke s pred výuky užšie než rozpätie rodených hovoriacich. Priblíženie sa intonačným rozpätím k modelu rodených hovoriacich sa však nepotvrdilo, tým pádom bola druhá hypotéza vyvrátená.

V súlade s očakávaním, výsledky analýzy ukazujú, že úroveň hlasu sa vplyvom výuky výslovnosti nezmenila. A nakoniec, doba trvania čítaného prejavu sa vplyvom výuky priblížila k modelu rodených hovoriacich. Na záver to znamená, že výsledky dokazujú pravdivosť troch hypotéz a vyvracajú jednu hypotézu.

Kľúčové slová: plynulosť reči v L2, plynulosť čítaného prejavu, intonačné rozpätie, dĺžka prejavu, angličtina ako L2, českí študenti angličtiny, pokročilí študenti

Prehlasujem, že som túto diplomovú prácu vypracovala samostatne a uviedla v nej všetky použité zdroje a literatúru.

V Olomouci dňa 30.4.2021 Bc. Monika Vlčková

Acknowledgements

I would like to express my sincere gratitude to my supervisor Mgr. Šárka Šimáčková, PhD. for her patience and guidance which immensely helped me during the process of writing this thesis. Moreover, I would like to express my thanks to all the students who willingly participated in the research and provided research material despite the unexpected complications caused by the pandemic.

— *Bc. Monika Vlčková*

Content

1. Introduction

Fluency is a broad topic and a lot of research has been already conducted in this area. There are two points of view from which we can study fluency. Extensive research has been conducted on the topic of fluency in spontaneous (extemporaneous) speech, as well as in the area of reading fluency. This thesis will posit itself within the second group of research. The main aim is to analyse the effects of explicit classroom-based pronunciation instruction on oral reading fluency of advanced Czech learners of English. The previous research conducted on reading fluency focussed mostly on native speakers and on improving reading fluency of young children, therefore, the analysis of reading fluency of non-native speakers of English will bring new ideas to the field.

Based on the specific area of fluency under study, there are various definitions of the phenomenon of fluency, however, as the oral reading fluency is the focus of this thesis, the analysis will be governed by a definition provided by Kuhn et al. (2010), which is essentially a summary of all the previous partial definitions that focussed only on certain features of fluency and lacked the overall outlook on the phenomenon. Kuhn et al.'s definition says

> [f]luency combines accuracy, automaticity, and oral reading prosody, which, taken together, facilitate the reader's construction of meaning. It is demonstrated during oral reading through ease of word recognition, appropriate pacing, phrasing, and intonation. It is a factor in both oral and silent reading that can limit or support comprehension. (2010, 242)

Under their definition, all the important features are combined and a meaningful relationship between them is created. Prosodic reading, previously overlooked in research, is now considered an essential feature of reading fluency. Dowhower (1991, cited in Kuhn and Stahl 2003) explains that there are six acoustic features that are considered to contribute to good prosodic reading. They include appropriate placement of various types of pauses, segmentation of larger text units into phrases, the length of the said phrases, but also such features as phrase-final lengthening, intonation contour and stress placement. In the methodology of this thesis, I will analyse two features of fluent reading, namely intonation contour and duration of utterances.

In the review of literature, I will first explore the definitions of reading fluency and its key features. Then, I will explore further the concept of prosodic reading, explain the differences between spontaneous speech prosody and reading prosody, and how the text's punctuation can help or hinder fluent reading. The different roles and functions of reading

prosody will close the first chapter of the review of literature. The second chapter will focus on the pronunciation instruction and the assessment of oral reading which, as will be explained further, can be either done objectively through the precise measurement of temporal features of speech or subjectively through teacher's impressions. Furthermore, various approaches to second language (L2) teaching, as well as their effectiveness on improving reading fluency will be explored.

In the methodology, I will analyse the pronunciation instruction process and the range of activities the students took part in. As the pronunciation instruction took place during the pandemic situation, the students were asked to work alone from home and were given delayed feedback by the instructor. The main research question of this thesis is, therefore, whether the limited pronunciation instruction could have had at least some effect on reading fluency of the advanced Czech learners of English. The analysis of the effects of pronunciation instruction will be carried out on the comparison of the students' performance before and after the instruction and further comparison with the performance of native speakers. To ensure the best possible outcome of the analysis, the students, as well as the native speakers, were asked to read the same text – the children's story *The Tiger Who Came to Tea.*

Numerous studies of non-native speakers showed that people who started learning a foreign language after early childhood are unlikely to achieve native-like pronunciation and fluency of speech. Although the amount of non-native features of speech may vary based on the speaker's L1, age at which the learning began, and experience with L2, the researchers seem to agree that even the most advanced speakers of an L2 rarely achieve native-like speech (Derwing and Munro 2005; Piske et al. 2001; Flege et al. 1995). However, there are studies which claim that when the right type of instruction is provided to the learners who seemed to have fossilized certain non-native patterns in their speech, some change in their pronunciation towards the native-like level is possible (Derwing et al. 1997, cited in Thomson and Derwing 2015). Based on these studies, the hypothesis is that explicit pronunciation instruction, although with delayed feedback, should have at least minor influence on the reading fluency of L2 speakers and should move towards the native-like patterns.

2. Fluency

In this chapter, the terms fluency, cognitive fluency, oral reading fluency, and reading fluency will be defined and I will consider the prosodic features (or suprasegmental features) of the English language, which constitute an important part of oral reading fluency. Furthermore, the relationship between (oral) reading fluency and reading comprehension will be briefly discussed, because it forms an extensive part of previous research in the field.

2.1. Definitions of Fluency

Generally, there are two areas in which the fluency research is being conducted. One of them deals with the notion of oral fluency in extemporaneous (spontaneous) production and the other deals with fluency in reading, where we can further distinguish silent reading fluency and oral reading fluency. In this thesis, the primary focus will be on the area of study that deals with the oral reading fluency.

2.1.1. Fluency in spontaneous speech

The term fluency does not have a general definition and has been assigned various meanings in previous research based mostly on the narrow focus of various studies. However, used as a lay term, there is a tendency to equate fluency with a general language proficiency and accuracy, be it in a native language or a foreign language (Galante and Thomson 2016, 117; Gürbüz 2017, 1854), and, therefore, it constitutes the ultimate goal of language acquisition and learning (Lems 2003, 2). Typically, fluency in this context means nativelike knowledge of grammar, pronunciation, and vocabulary of a language, and it includes the capacity to produce smooth utterances effortlessly (Rossiter et al. 2010, 584; Thomson 2015, 3). Moreover, in the second language acquisition field, the term can also mean nativelike reading and writing ability (Thomson 2015, 3), the capacity to express ideas in L2 in a similar way that would be done in L1, or even the production of unaccented speech (Segalowitz 2010, 4).

Researchers, such as Schmidt (1992) and Skehan and Foster (1999) as cited in Galante and Thomson (2016, 117) focus on the oral language performance and define fluency as "the capacity to use language in real time" instead of the underlying knowledge of how a language should be used. Galante and Thomson (2016) also mention researchers who highlight the association between fluency and speech rate, length of utterances, distribution of filled or unfilled pauses and hesitations (Galante and Thomson 2016; Rossiter et al. 2010, 585-6), and an automatic retrieval of language forms from the speaker's mind

(Gürbüz 2017, 1854). Fillmore, for example, identifies four types of fluency in L2 speech. The first one is the ability to speak long utterance without undue pauses and hesitations. The second type is the ability to produce "semantically dense sentences" without extensive use of discourse fillers (such as *you know, the thing is that* etc.). The third type is the capacity of a speaker to use appropriate forms of expression in various social contexts, which is something not mentioned in other research. The last type of fluency is the ability of a speaker to "use the language creatively and imaginatively by expressing ideas in new ways," in other words, it is the ability to make puns and metaphors in nativelike fashion (Fillmore 1979, cited in Segalowitz 2010, 4). Generally speaking, most of the researchers define fluency in terms of the fluidity of speech and focus on measurable temporal features, such as pauses and speech rate (Segalowitz 2010, 5).

Viewed from the psycholinguistic perspective, below all the outer layer of oral fluency described above, there is an underlying layer of a cognitive skill that makes a fluent production of language possible (Galante and Thomson 2016, 117). This view is explored in detail in Segalowitz (2010), where he distinguishes three types of fluency – cognitive fluency, utterance fluency, and perceived fluency. By cognitive fluency, he means the ability of a speaker to "efficiently mobilize and integrate the underlying cognitive processes responsible for producing utterances" (Segalowitz 2010, 48). This cognitive process involves planning of an utterance, which includes the language choice, then the grammatical encoding, morpho-phonological encoding, phonetic encoding, and lastly the articulation of the utterance.

From the articulation of the utterance we get to Segalowitz's second type of fluency, which is the utterance fluency. By this term, Segalowitz means the temporal characteristics of an utterance, similar to those described by other researches above, such as pauses, hesitations, length of utterance, etc. The third and last type of fluency is the perceived fluency, which refers to a listener's judgement or impression of a speaker based on the speaker's speech sample (Segalowitz, 2010, 48). With this third type of fluency, it is important to mention that although listener's judgements are often involved in the research measuring oral fluency, they are used to reflect back on the underlying cognitive fluency of a speaker. As opposed to the research in accentedness, comprehensibility, and intelligibility, where the primary focus is on the listener's perception of the speech, the research on oral fluency is more focused on the speaker's L2 speech production system (Thomson 2015, 5).

Based on Segalowitz's (2010) definition of cognitive fluency, we can understand why some researchers, or even teachers, may describe lower proficiency learners of an L2

as fluent, despite their limited knowledge of grammar, pronunciation, and vocabulary. In such contexts, fluency may refer to the ease of retrieval of the knowledge the students have about a language from their minds and it may also refer to the way they produce the L2 without undue hesitations and pauses. As opposed to the definition of fluency as a general L2 proficiency, which considers only the advanced L2 speakers as fluent, Segalowitz's definition concentrates on the fluent production of an L2, however small the speaker's knowledge of an L2 may be (Thomson 2015, 3).

2.1.2. Reading fluency

As well as for the concept of oral fluency, there is no single definition of reading fluency, but researchers seem to have reached consensus on what the major components of the concept are. Previously, (native) reading fluency was variously defined as accurate and automatic reading, but in the last couple of decades, the concept of reading prosody (which refers to the ability to use intonation, pauses, and stress in reading as if the speech was produced spontaneously) gained importance and is now considered to be one of the key features of reading fluency. This means that the communicative aim of reading and its impact on the listener contribute to the construct of reading fluency and are as important as speed and accuracy of speech (Godde et al. 2019). All these key features have been summarized in a definition by National Reading Panel (2000, cited in Swain et al. 2017, 105), which states that reading fluency is the "ability to read text quickly, accurately and with proper expression," which resembles speech. It is also important to mention that the term reading fluency encompasses both reading fluency and oral reading fluency, which may, at first, seem like the same concept, but there is a distinction between them. Lems defines reading fluency as the competency to read automatically while extracting meaning from a text, and the oral reading fluency as a "measurable performance" of the underlying competency (2003, 9).

Automaticity and accuracy, as two of the key features of reading fluency, refer to the correct and automatic decoding of individual words that comprise a text. The necessary properties of automatic reading are "speed, effortlessness, autonomy, and lack of conscious awareness," without which we cannot really speak of automatic fluent reading (Kuhn et al. 2010, 233). There is, however, one more important feature that has not been included in the above mentioned definitions and that is the concept of reading comprehension, by which we mean extracting meaning from a text. An individual can successfully comprehend the meaning of a text only once he mastered the automaticity and accuracy in decoding, because

the reader has a limited amount of attention available for the two simultaneous tasks that necessarily comprise successful fluent reading – recognition of the words and the meaning behind the words. This means that while too much attention is paid to the decoding of individual words, the less attention is left for the extraction of meaning (Kuhn and Stahl 2003, 5).

The relationship between comprehension and reading prosody is, however, not so straightforward. We cannot accurately decipher, whether prosodic reading comes before full and adequate comprehension of a text or vice versa. On the one hand, there are studies (such as Kuhn and Stahl 2003) claiming that we cannot adequately understand the meaning of a text without prosodic reading, which enables segmentation of a text into major syntactic-semantic units, hence improving comprehension. On the other hand, other studies (mentioned in Schwanenflugel et al. (2004, 1)) claim that comprehension in reading can be achieved when decoding becomes automatic, which means that prosodic reading is an addition to fluent reading, not a key factor influencing reading comprehension.

To summarize what has been mentioned so far, let us look at the study by Kuhn et al. (2010). They review the previous definitions of reading fluency and conceptualize them into four major views – fluency as accuracy and automaticity, fluency as prosody, fluency as skilled reading, and fluency as a bridge to comprehension. The first definition focuses on accuracy and automaticity, while ignoring the importance of other factors, such as phrasing, stress, or emphasis. These factors become important in the second definition of reading fluency but, according to some critics, this definition also has a major flaw in that it does not consider the importance of comprehension. In the definition of fluency as skilled reading, word decoding, prosodic reading, and text comprehension all play an important role, however, this definition also met with some criticism claiming that the definition is too broad, and that comprehension actually results from fluent reading. The critics essentially put comprehension outside the definition of fluency. From this we get to the last view, fluency as a bridge to comprehension, according to which fluency both contributes to text comprehension, but also results from it (240-242). As a way of acknowledging the previous views along with their respective criticism, Kuhn et al. proposed their own definition, which, from now on, will be used as a working definition of reading fluency for this thesis:

> Fluency combines accuracy, automaticity, and oral reading prosody, which, taken together, facilitate the reader's construction of meaning. It is demonstrated during oral reading through ease of word recognition, appropriate pacing,

> phrasing, and intonation. It is a factor in both oral and silent reading that can limit or support comprehension. (2010, 242)

2.2. Prosody in Reading Fluency

Majority of the definitions above agree that fluent reading includes the ability to read with expression, or prosodic reading, which is meant to be understood as representing the "tonal and rhythmic aspects of language" (Kuhn and Stahl 2003, 5), such as variations in pitch or intonation, stress placement on syllables, words, and phrases, and distribution of pauses (Schwanenflugel and Benjamin 2017, 3). In addition to these features, reading with expression entails chunking of words into meaningful phrases based on syntactic and semantic structure of a text (Kuhn and Stahl 2003). According to Pival (1968), reading *without* expression means monotonous reading, along with reading ′word by word′. Both are signs that an individual is unable to "apply the intonation characteristics of his spoken language to the medium of the printed page" (458) and are, therefore, understood as markers of disfluency in reading. In contrast, a good, fluent reader

> read[s] like he speaks, with appropriate rhythm and intonation, to allow the listener to easily process the speech and understand the content. He is also supposed to add expressivity to his reading: emphasize, add focus, convey emotions… (Godde et al. 2019, 2)

According to Dowhower (1991, cited in Kuhn and Stahl 2003, 5), there are six acoustic features that comprise good prosodic reading: appropriate placement of pauses, segmentation into appropriate phrases, length of the phrases, phrase-final lengthening, terminal intonation contour (falling pitch at the end of declarative sentences), and stress. The individual features of prosodic reading will be addressed further in the next subsection of the chapter. The discussion of phrasing, pauses, terminal intonation contour, and stress draws on the literature review in Godde et al. (2019).

2.2.1. Prosodic features

One of the prosodic features of fluent reading is phrasing or chunking of sentences into meaningful phrases and clause units, which facilitates comprehension (Godde et al. 2019, 3). Academics termed this feature differently in the previous research, but they tend to describe the same concept. For example, Cowie et al. (2002) termed it rhythmic organization, while Erekson (2010) uses the term syntactic prosody and Lems (2003) writes about parsing. This feature is connected to the structure of a text and because the structure

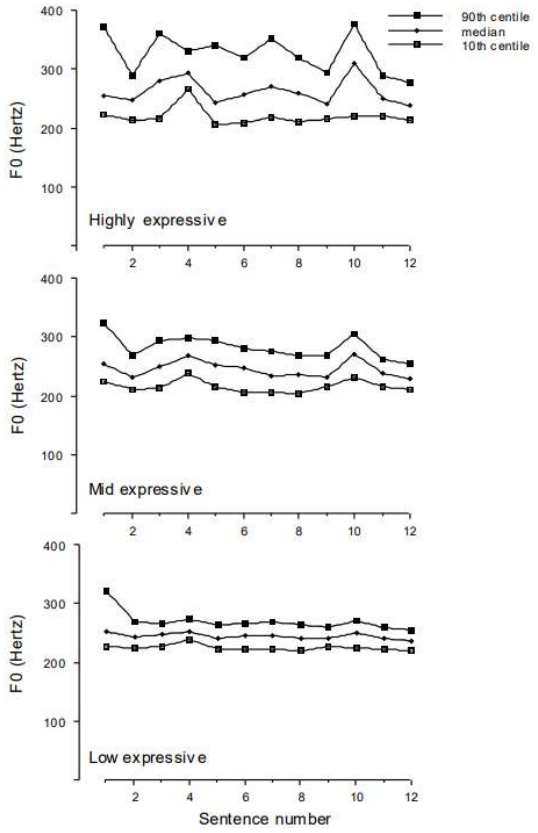
of a text is usually quite explicit in its marking of syntactic boundaries, it is presumed to be the easiest to acquire among the prosodic features of fluent reading (Godde et al. 2019, 3).

Closely linked to phrasing is the appropriate placement of pauses. We can distinguish three types of pauses: breath (respiratory), syntactic, and hesitation pauses. The first type represents the pauses produced when the reader needs to breath in and they can be accompanied by audible noises (Bailly and Gouvernayre 2012, cited in Godde et al. 2019). Frequent respiratory pauses are typical of younger readers with limited abilities of fluent reading and lungs capacity, but once they acquire automaticity in reading, their coordination of breathing and phrasing of a text becomes more stable. Among expert readers, there is a tendency to place respiratory pauses at major punctuation marks, which close a single intonation unit (Godde et al. 2019, 3). The second type of pauses are syntactic pauses, which correlate with phrasing of a text and highlight syntactic units. According to Bailly and Gouvernayre (2012), we can further distinguish between sentence-internal and sentence-final pauses, as well as paragraph-final pauses. The sentence-internal pauses tend to be shorter than sentence-final pauses, and they are generally associated with intrasentential punctuation marks, such as commas and colons. In addition to that, pauses tend to be longer before and after syntactically complex phrases and/or when the information load is too high, which means that the processing of the information may be hindered (Kuhn et al. 2010, 237). The last type, the hesitation pauses, are linked to the cognitive ability of readers and they represent problems in decoding, typical of young readers, and are considered to be ungrammatical (Godde et al. 2019, 3).

Intonation contour is another prosodic feature of great importance in reading fluency. It is closely connected to the phenomenon of fundamental frequency (F_0), which can be defined as the approximate frequency of the repetitive structure of a speech signal and is expressed in Hz. Listener's perception of the fundamental frequency is pitch (Bäckström 2020). This feature is very speaker specific because it depends on the voice quality of the speaker as well his/her age and sex. Pitch contour, or intonation, is connected to the structure of a text and is indicated by punctuation. There is a crucial difference between the pitch in declarative sentences and questions. While declarative sentences are usually indicated by initial rising and then falling pitch (declination), the yes-no questions typically end with the rising pitch. Wh-questions, on the other hand, with their sentence-final declination resemble more the declarative sentences. Such pitch patterns are, however, not obligatory for all sentences of a given type because there are also other factors, pragmatic factors to be precise, which can affect the intonation contour (Kuhn et al. 2010,

236). What is interesting is that the notion of expressivity in reading (defined above by Godde et al. (2019) as added emphasis or emotion to reading) is closely connected to this feature of prosodic reading (understood by Godde et al. (2019) as appropriate intonation and rhythm), although they are two independent variables. According to Godde et al. (2019), it is possible that a reader can read with the appropriate intonation and phrasing but without expressivity, which will result in a monotonous and boring oral reading. A nice visual representation of the difference between expressive and inexpressive readers can be seen in the Figure 2.1 below taken from Cowie et al. (2002, 28). Here, we can see that even the inexpressive readers have a tendency to vary their intonation contour depending on the

Figure 2.1 The difference between expressive and inexpressive readers according to Cowie et al. (2002). The 90th centile represents the highest point of the pitch range in each sentence, while the 10th centile represents the lowest point. The pitch changed according to the sentence type, whether it was a declarative sentence or yes-no/Wh-question.



sentence type but their upper and lower bound (90th centile and 10th centile) pitch range remains close to each other, which results in a very monotonous speech.

Another very important prosodic feature in reading is the placement of stress. This is a skill, in which we can talk about automaticity, because the knowledge of the correct stress placement in individual words seems to be retrieved without the conscious awareness of the speaker. In English, we distinguish two types of words, functional and lexical, which differ in terms of stress placement. While the functional words are usually unstressed, the lexical words always carry at least one (primary) stress. The placement of stress in lexical words can be used to distinguish different grammatical forms, such as verbs and nouns, which have the same spelling. An example of this would be the word *permit*, which can be used both as a noun and as a verb. The general tendency, in such cases, is that the nouns have the primary stress on the first syllable, whereas verbs have it on the following syllable (Kuhn et al. 2010, 236). However, it is important to say that different languages follow different rhythmic patterns, which can cause problems when learning a second language. In stress-timed languages (a group of languages to which English belongs), the stressed syllables are perceived as if they occurred at approximately the same intervals. This means that, depending on the placement within a sentence, the stress on lexical words can be dropped in order to avoid stress clashes (example *a*), the functional words, which would not normally be stressed, can carry the stress (example *b*), and/or the stress can be shifted to a different position within a word that would, if it were pronounced in isolation, carry the primary stress on the first syllable (example *c*) (Conlen 2016, 1-3).

a) *The 'big brown 'bear ate 'ten white 'mice* (Conlen 2016, 2).

b) *I 'gave it 'to the 'postman* (Kuhn et al. 2010, 236).

c) *She´s 'only six'teen* (Conlen 2016, 3).

Closely connected to the stress placement is the duration. The duration of vowels in words that carry the stress is longer than in unstressed words, and the duration is even longer when the stressed word appears in phrase-final position. The length of individual vowels depends on the speaking or reading rate of a particular speaker. Naturally, speakers who read fast will produce shorter vowels in stressed syllables than those who read at a slower rate but at the same time, the duration of stressed syllables has a tendency to become shorter when the speaker goes through longer sentences (Kuhn et al. 2010, 236).

The very last prosodic feature that will be discussed here is the phrase-final lengthening, which can be understood to mark the intonational phrase boundary by

lengthening the syllable that immediately precedes it. The duration of such syllable is longer than if the same syllable appeared in a medial position within a phrase. Such durational differences may either concern the whole phrase-final syllable or may be concentrated on either the vowel or consonant(s) in that syllable (Gósy and Krepsz 2018).

2.2.2. Difference between spontaneous speech prosody and reading prosody

Even though one of the goals of developing reading fluency is to read as one speaks, there is a difference between prosody in spontaneous speech and in reading. As various studies mentioned in Godde et al. (2019, 2) suggest, one of the major differences between these two speech productions is in the placement and duration of pauses. Those studies suggest that in reading, the pauses are shorter and occur less frequently than in spontaneous speech and, when they do occur, they are usually placed at major syntactic boundaries. Other studies (such as Bailly and Gouvernayre 2012, cited in Godde et al. 2019) claim that there is a strong connection between the occurrences of pauses and punctuation and between pauses and the structure of paragraphs, which are not available in spontaneous speech production.

Other features, such as articulation rate (measured without pauses) and intonation contour, are more language dependent, which can be supported by the findings of the following studies cited in Godde et al. (2019). While in French (Goldman et al. 2009) and in English (Hirschberg 2000) the articulation rate is higher in reading, Beinum (1991) suggests that it is not the case in Dutch. In terms of the intonation contour, Goldman et al. (2009) reported that there is a flatter contour in spontaneous speech of French speakers than there is for read utterances, but Cowie et al. (2002) suggested that, in English, such a contrast is highly dependent on context and emotional state of the speaker.

All in all, the differences between spontaneous speech and reading prosody are great enough for the listeners to easily differentiate between them (Godde et al. 2019).

2.2.3. Punctuation and prosodic reading

As has been hinted at a few times already, the structure of punctuation in a text is closely connected to the prosodic features of reading. However, as Chafe (1988, cited in Schwanenflugel et al. 2004) suggests, punctuation does not always provide cues for the appropriate prosodic reading. The use of commas is particularly tricky because in sentences like *Lesley came, she saw, and she conquered*, the commas represent pause placement, but in other sentences, such as in *Lesley wanted the one with red, white, and blue sprinkles*, they do not represent the correct placement of pauses because they were inserted there for grammatical reasons, not prosodic ones. Question marks are also not a reliable cue, because

the reader needs to have a deeper knowledge of different types of questions and their typical intonation contours – there is a different intonation contour signalled by the question mark in yes-no questions (*Did Robin go?*) and Wh-questions (*Where did Robin go?*) (Schwanenflugel et al. 2004, 3).

Historically, the punctuation units in a text coincided with the intonation units for most of the nineteenth century. This, however, changed when various punctuation rules were introduced; the rules were motivated more by grammatical constructions rather than prosody. That is the reason why, nowadays, texts necessarily contain punctuation marks which do not provide adequate cues for prosodic reading (Chafe 1988, 4-5), as the example above shows.

In terms of the duration of individual pauses signalled by different kinds of punctuation marks, Robert Lowth (the author of *A Short Introduction to English Grammar*, 1762) believed that there was a difference in duration for each of the punctuation marks in the following sequence, which equals double the amount of time attributed to the previous one. So, the pause signalled by the semicolon was twice as long as the one signalled by comma, and so on (Chafe 1988, 5):

> *Coma* (having the shortest duration of pause) → *semicolon* → *colon* → *period* (signalling the longest pause)

Chafe (1988), however, points to the fact that Lowth's belief was not accordingly supported by precise measurements of pauses and was probably based only on the popular opinion.

2.2.4. Syntactic and emphatic prosody

As Chafe (1988) suggests, every writer, when writing a text, has an image in his head of what the prosody of the text would be if it were read aloud. This may or may not be marked adequately by punctuation, which leaves the reader to interpret the text and assign such prosody to it that (s)he deems appropriate. This leads us to the difference between syntactic prosody, or the phrasing of a text, and emphatic prosody, which signals the reader's interpretation of a text (Erekson 2010).

Here, we get to the difference between fluent and expressive reading again. As some researchers (for example Cowie et al. 2002) suggest, the appropriate syntactic prosody (phrasing) makes the reading fluent, but it is the emphatic prosody (reader's interpretation of a text) that gives the expressivity to reading.

Syntactic prosody is mainly used to serve the grammatical function, that is to organize words into meaningful phrases, assign typical intonation contours to various kinds

of sentences, and make pauses in the appropriate places to divide the individual phrases. Emphatic prosody, on the other hand, goes beyond syntactic prosody and assigns the expressivity to a text based on intention, meaning, and feelings. With the use of emphatic prosody, a reader can signal irony, pleasure, irritation, curiosity, and many more pragmatic functions, which could otherwise be unclear to the listener, especially when these pragmatic functions are not specifically stated in the text (Erekson 2010).

2.2.5. The role and functions of reading prosody

There are three essential linguistic functions of prosody. The first is its ability to provide necessary information about boundary markers in speech, even reading speech, for the listener to break up the information and to understand what is being said. This is a very important linguistic function of prosody because if the prosodic boundary markers in reading are used in incorrect places, it is very difficult for the listener to understand the message. Another linguistic function of prosody is to help the listener remember what was said, to keep the information in the storage of the working memory. According to Goldman et al. (2006, cited in Kuhn et al. 2010), it is easier for people to remember poetry and poetic texts which have enhanced prosody. There is also the possibility that good prosodic rendering of a text may enhance comprehension. Prosody also serves the function of providing cues for understanding syntactically and semantically ambiguous sentences (Kuhn et al. 2010, 237; Godde et al. 2019, 1).

The reading prosody also has a paralinguistic function of carrying various discourse information. By manipulating the prosody of a certain utterance, the reader can focus the listener's attention to a desired word or phrase, to indicate a topic shift in the reading, or even to imply sarcasm and irony (Kuhn et al. 2010, 237; Godde et al. 2019, 1).

The last function of prosody is to provide emotional information about the speaker. With different prosodic rendering, the listener can distinguish between happiness and sadness in the speaker's voice, or even uncertainty. For example, happiness is typically characterised by higher pitch and faster speech rate, unlike sadness, which is characterised by slower rate and lower pitch. And the speaker can signal uncertainty by steadily raising the pitch in a sentence (Kuhn et al. 2010, 237).

Before I move to the discussion of cross-linguistic differences in intonation, let me summarise what has been mentioned so far. Prosody, as has been shown, is a key feature of reading fluency and is understood to include six features (phrasing, intonation, pauses, stress, phrase-final lengthening, duration). The distinction has been made between

spontaneous speech prosody and reading prosody and between syntactic and emphatic prosody. I also addressed the role of punctuation and how it can either aid or hinder reading fluency and, lastly, the different roles and functions of prosody in reading has been summarised.

2.2.6. English and Czech intonation

There are numerous studies focusing on cross-linguistic comparison of intonation patterns and the influence of a speaker's L1 on his/her intonation in L2 (which, in most of the cases, is English), some of which are cited in Mennen (2007, 55). Based on the comparison of the results of these studies in Mennen (2007), certain error similarities arise. One of these typical errors in the production of L2 English is narrower pitch range, which will be analysed in detail in the practical part of this thesis. To use the terminology of Ladd (1996), I will analyse the pitch *level* of the speakers (their overall height of the pitch) and pitch *span* (the range of frequencies within the speaker's utterances) and compare it to the pitch level and span of native speakers.

As the focus of this thesis is the analysis of L2 English produced by Czech speakers, it would be useful to summarise the findings of some cross-linguistic studies concerned with these two languages. Volín, Poesová, and Weingartová (2015) is one of these studies and it focuses on cross-linguistic comparison of the pitch level and span between the reading intonation of native English, native Czech, and Czech English speakers. Their analysis is based on distributional measures using various methods for the analysis, such as 80-percentile range, Variation range, Standard deviation, and Quartile range, which has been previously criticised by Patterson (2000) for their lack of perceptual validity. The results of this study suggest that there is a significant difference between the mean pitch level of native Czech and native English speakers, with the Czech speakers having lower mean F0 than the British English speakers, as can be seen in figure 2.2 below. In addition to that, the English speakers deviate more from their average F0, in both directions, than the Czech speakers. In other words, the native English speakers make use of higher as well as lower frequencies more than the native Czech speakers. Their results of pitch span using various methods of measurements suggest that the intonation range of Czech speakers is narrower than that of English speakers (for visual representation see table 2.1 below).

Figure 2.2: Mean F0 (expressed in Hz) for native Czech and British English speakers, with respect to gender differences (Volín et al. 2015, 111).

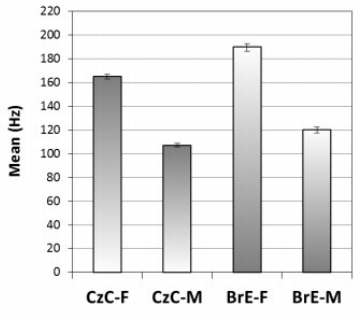
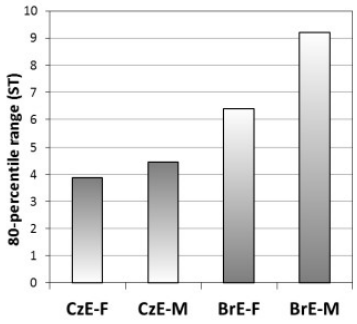


Table 2.1: Pitch span values for native Czech and native English speakers (with respect to gender), using four methods for measurements (Variation range, 80-percentile range, Quartile range, and Standard deviation) expressed in either Semitones (ST) or Hertz (Hz) (Volín et al. 2015, 115).

Speaker Group	Variation Range (ST)	80-perc. Range (ST)	Quartile Range (ST)	SD (Hz)
CzC-F	8.8	5.2	2.8	19.4
CzC-M	10.8	6.1	3.2	15.1
BrE-F	12.2	7.1	3.8	30.6
BrE-M	14	8.1	4.2	22.1

The results for the comparison of native English and Czech English production suggest that the mean F0 (pitch level) for the two groups is approximately the same, but the measurements for pitch span (represented visually only for 80-percentile range) show a significant difference between Czech English and native English production (see figure 2.4). The results also suggest that the pitch span values for Czech English are lower than both native Czech and native English values. This last result is interesting mainly because the logical expectation, also exhibited by Volín et al. (2015), would be that the Czech English values would be positioned between the native Czech and native English values. The researchers attributed the narrow pitch range of Czech English to the potential involvement of uncertainty and anxiety of speaking a foreign language.

Figure 2.4: Mean values of 80-percentile pitch span (expressed in Semitones) for Czech English and native English speakers, with respect to gender differences (Volín et al. 2015, 118).



3. Pronunciation Instruction

Explicit pronunciation instruction, as one of the aspects and methods of language teaching, is usually neglected in second or foreign language classrooms and, according to Pourhossein Gilakjani (2017), it is one of the least favourite areas of a language for the teachers to teach. Among the most often stated reasons for the lack of pronunciation instruction in classrooms, we can find the lack of phonetic knowledge at the part of the teachers, its relative irrelevance in the overall language skill, lack of time allotted to pronunciation in the syllabi, and its ineffectiveness in improving the language skills (Pourhossein Gilakjani 2017; Atli and Bergil 2012; Gordon and Darcy 2016; Derwing and Munro 2005). However, as numerous studies cited in Pourhossein Gilakjani (2017) (Derwing et al. 1998 being one of them) show, the effectiveness of pronunciation instruction in improving the L2 production is immense.

One of the reasons why explicit pronunciation instruction can lead to improved L2 production can be explained through Flege's Speech Learning Model (SLM), which says that accurate production of an L2 is preceded by the accurate speech perception, and various studies provided evidence for this claim in both segmental and suprasegmental field of language. Flege's SLM is based on the assumption that speaker's need to be aware of the subtle phonetic differences between their L1 and L2 in order to be able to acquire the L2 properly (Kissling 2014), and as numerous studies showed, an incorrect perception of an L2 is a major cause of production problems (Derwing and Munro 2005).

From the historical perspective, the pronunciation as a separate scientific field of study developed as late as in the 19th century, when the International Phonetic Association was created. Since then, the outlook on the pronunciation instruction and its importance varied greatly based on the prevalent approaches to L2 teaching at a given time. Today, pronunciation instruction should be an integral part of the current approaches to language teaching because pronunciation is considered to be a very important aspect of overall language competence and the incorrect pronunciation is the main cause of communication breakdowns (Pourhossein Gilakjani 2017).

Originally, the pronunciation instruction was governed by the nativeness principle, which had as its goal the native-like pronunciation of non-native speakers. This, however, changed and now the goal of instruction is governed by the principles of intelligibility (how well the speaker can be understood by a listener) and comprehensibility (how difficult it is for the listener to understand what the speaker says) (Atli and Bergil 2012). That might be

the reason why the majority of studies that deal with pronunciation instruction is focussed on improving intelligibility and comprehensibility.

3.1. The Role of Pronunciation in Various Approaches to L2 Teaching

Grammar translation method, as one of earliest approaches to language teaching, evolved during the 19th century and was primarily focussed on the teaching of reading and writing skills, which was based on the assumption that "literary language was superior to spoken language" (Pourhossein Gilakjani 2017, 1250). As oral communication was not the primary goal of this approach to language teaching, speaking, and listening skills were ignored and the main focus was on the teaching of grammar and expansion of vocabulary. Pronunciation, therefore, was not considered important in this approach (Pourhossein Gilakjani 2017).

In the early 20th century, the Direct Method was the prevalent approach to language teaching (Jam and Adibpour 2014). In this approach, the oral aspect of language was strongly emphasized and the focus on pronunciation instruction was from the very beginning of the language teaching process. The main method of pronunciation teaching was through intuition and imitation, that is, the students listened to a native speech and were asked to imitate it until they mastered the correct pronunciation (Pourhossein Gilakjani 2017).

During the 1940s and 1950s, the Audio-Lingual method or the Oral Approach, as it was called in the UK, also emphasized the importance of oral language skills, but the difference from the previous method was that these skilled were taught explicitly through pronunciation instruction, not only through intuitive imitation of native speech (Jam and Adibpour 2014). In this approach to language teaching, both the teachers as well as the learners should be trained in phonetics and extensive phonetic transcription was one of the methods of pronunciation instruction (Pourhossein Gilakjani 2017).

The situation changed considerably during the 1960s, when the Cognitive approach to language teaching, with its new ideas about unattainability of native-like pronunciation, gained prominence. In this approach, grammar and vocabulary teaching were considered superior to pronunciation instruction (Jam and Adibpour 2014). This, however, did not last and during the next decade pronunciation was again emphasized. In the Silent Way approach, the emphasis was on the accurate production of the target language sounds and structures, but not through the teaching of phonetic alphabet, which was a change from the earlier approach that emphasized pronunciation instruction (Pourhossein Gilakjani 2017).

Later on, the Community Language Learning approach emerged, and it concentrated again on the intuitive-imitation pronunciation learning, but the overall learning process was controlled by the learners themselves. This is similar to the two Naturalistic Methods, Natural Approach and the Total Physical Response, which emerged during the 1980s. In these approaches, the learners were not pressured to produce the language unless they managed to internalize the target language sounds and were comfortable to start speaking. During the early stages of speaking, students made errors which were tolerated by the teachers (Pourhossein Gilakjani 2017; Jam and Adibpour 2014).

One of the most recent methods of language teaching that emphasised the importance of pronunciation instruction is the Communicative Approach, the ultimate goal of which is successful communication. In this approach, the accurate pronunciation is taught through listening and imitation as well as phonetic training (Jam and Adibpour 2014).

There are essentially three main approaches to language teaching: the intuitive-imitative approach, the analytic-linguistic approach, and the integrative approach. The intuitive-imitative approach is based on the assumption that the language learner is capable of acquiring accurate pronunciation from just listening and imitating native speech, without the need for explicit pronunciation instruction whatsoever. The analytic-linguistic approach, on the other hand, emphasises the role of explicit pronunciation instruction and uses such techniques of instruction as to give the learners detailed information about how the speech is produced in the vocal tract. The last approach, the integrative one, views pronunciation as an essential part of the communication process, therefore, in this approach the learners learn the language through meaningful communicative activities (Jam and Adibpour 2014). Based on this classification of approaches to language teaching, I can say that the pronunciation instruction, on which this thesis focuses in the methodology, was a combination of intuitive-imitative and analytic-linguistic approach, because the students were provided with explicit information about the phonetic and phonological features of speech, but were also asked to imitate native speech at various points during the instruction.

3.2. Effectiveness of Pronunciation Instruction

Previous research on the effectiveness of second/foreign language pronunciation instruction showed mixed results but the overall outcome of the studies is that pronunciation instruction is effective in improving the production of L2 speakers. Thomson and Derwing (2015) reviewed 75 studies conducted on the effectiveness of L2 pronunciation instruction, 53% of which were on segmental instruction, 23% on suprasegmental, and 24% on both, and their

results showed that 82% of these studies claimed to have reached a significant improvement in L2 production. The assessment of these studies was varied, some employed reading tasks, while others focussed on extemporaneously produced speech or sentence imitation. The studies varied greatly in the time allotted to the pronunciation instruction, some sessions lasted less than an hour of a single day, while others were divided between various session over a period of one year.

Derwing et al. (1998) focussed their study on both segmental and suprasegmental instruction with the assessment based on reading and spontaneous speech tasks. Native English speakers were asked to evaluate the effectiveness of the instruction, which was divided into three categories – "segmental accuracy; general speaking habits and prosodic factors; and no specific pronunciation instruction" (393). The overall instruction lasted for 12 weeks and the data for analysis were collected before and after the instruction took place. The goal of this study, as many others done before and after, was to improve the speaker's comprehensibility, accentedness, and fluency. The results of this study showed that while both of the groups that received pronunciation instruction, either segmental or suprasegmental, improved in accentedness and comprehensibility in the reading task, only the group instructed in suprasegmental features improved in comprehensibility and fluency in the spontaneous speech task.

A very similar study was conducted by Derwing and Rossiter (2003), who also focussed on ESL speakers and the 12-week segmental and suprasegmental instruction with one group of learners who received no specific pronunciation instruction. They studied the phonological accuracy, fluency, and syntactic complexity of the extemporaneously produced speech. As was expected based on the study conducted by Skehan and Foster (1997, cited in Derwing and Rossiter 2003), which claimed that the speakers are unable to focus on all three aspects of speech under question at the same time, the results of Derwing and Rossiter's (2003) study showed that the group receiving segmental instruction showed improvement in phonological accuracy, while the rates of fluency and syntactic complexity remained unchanged. The other group, which received suprasegmental instruction, showed significant improvement in fluency and complexity of speech, while their phonological accuracy did not change significantly.

The results of the previous two studies suggest that it is the pronunciation instruction focussed on suprasegmental features of languages (intonation, rhythm, stress etc.) that can lead to significant improvement in fluency, not the segmental focussed instruction. That is why the instruction under analysis of this thesis was primarily focussed on suprasegmental

features, and based on the results from the previous studies, the fluency of the participants of this study should show improvement.

3.3. What does the Effective Instruction Entail?

In this section, I will summarise the points made in previous research about what the effective pronunciation instructions entails in English as native language (L1), English as a second language (ESL), and English as a foreign language (EFL) context. Furthermore, the role of the teacher and of his/her corrective feedback will be described. The last topic in this section to be discussed is the variety of exercises that can be implemented in order to enhance reading fluency.

An extensive number of studies concerning the effective instruction to enhance reading fluency, cited in Rasinski (2006), has focused primarily on young native speakers, who struggled either with reading rate or overall comprehension of the reading material. According to Rasinski (2006), the effective instruction in such contexts should focus on accuracy, automaticity, as well as prosody in reading at the same time (as opposed to the view promoted in Hudson et al. (2005, cited in Rasinski 2006), which holds that these individual components of reading fluency should be addressed separately) and the primary goal of such instruction should not be increased reading rate but enhanced comprehension.

Although there are studies that focus on the effectiveness of pronunciation instruction and fluency training on non-native speakers, such studies usually aim at enhancing the learner's comprehension of the reading material (Shen 2003). The findings of several studies cited in Shen (2003) suggest that the knowledge of syntactic structure can enhance comprehension of a text, in both ESL and EFL context, therefore the effective instruction in such contexts should include the explicit instruction on the L2 syntax. The difference between ESL and EFL is solely based on the exposure to English as a native language. While in ESL, the learners usually live in environments where English is the dominant language and they can engage in daily life communication with native speakers, the learners in EFL context learn English in classrooms and have very little access to native English (Ghorbani et al. 2016).

What is striking in the area of pronunciation instruction focused on non-native speakers is the lack of interest in enhancing fluency in reading in itself, not in connection with reading comprehension. This thesis aims at filling this scarcely occupied area in research by analysing the effectiveness of pronunciation instruction in EFL context on the improvement of fluency in reading without taking into account the aspect of comprehension.

Given the fact that the subjects of this research are all advanced learners of English, their comprehension of the reading material is assumed to be impeccable.

3.3.1. The role of teacher and corrective feedback

According to Rasinski (2005), despite the number of articles that concentrate on the various methods and procedures for the instruction of reading fluency, there is a very limited number of studies and articles, if any, that focus precisely on the role of the teacher in such instruction. Rasinski (2005) further claims that, based on his own experience with students, the direct involvement of the teacher in the instruction can make a tremendous difference for the student's improvement in reading fluency. In this article, Timothy Rasinski characterises the main roles of the teacher in the reading fluency instruction, which I will briefly summarise in this section.

One of the most important roles of the teacher is to make his or her students aware of the phenomenon of reading fluency, which can be done through the oral reading to students and the subsequent discussion of it. This, paired with practice of reading fluency, can provide the students with clear understanding of the concept of reading fluency and give them a specific aim they should strive to achieve in their own oral reading (Rasinski 2005).

Another very important role of the teacher is to give formative feedback to students. Without some form of a corrective feedback from the teacher, the students may practice their reading with errors without being aware of it, or they may strive for a wrong goal, such as increased rate of reading without paying enough attention to the expressivity in reading. The formative feedback may be in a form of a praise for well-read passages of a text, but also in a form of a guidance in passages that were not read properly by a student (Rasinski 2005).

According to Karimi and Esfandiari (2016), corrective feedback is an essential part of second language and foreign language learning and, given the fact that errors in learners' production are inevitable during the process of learning a language, feedback provided by the teacher during the process can facilitate learning. Although the study by Karimi and Esfandiari (2016) focussed on the corrective feedback provided during a spontaneous speech, we can assume that corrective feedback provided during reading instruction may have a similar, or even the same, effect.

The last role of the teacher, which will be mentioned here, is to monitor the progress of the students, which can be done through various kinds of assessments. This, however, will be explored further in section 3.4 below.

3.3.2. Exercises to improve reading fluency

Although studies such as Rasinski (2006), Abadiano and Turner (2005), and Swain et al. (2017) focus explicitly on improving reading fluency of native speakers, especially at a very young age, they describe a variety of instructional approaches and exercises used to enhance reading fluency, which may be implemented to foster the same development in non-native reading fluency instruction.

Previous research findings suggest that various approaches to fluency instruction contribute to enhancing reading fluency for students with, as well as without, reading disabilities. Among the most effective and influential instructional approaches are repeated oral reading combined with pronunciation instruction and feedback, and listening passage preview, or modelling by a teacher or other fluent reader (Swain et al. 2017). The same view is promoted by Hudson et al. (2005, cited in Rasinski 2006).

The Partnership for Reading (2001, cited in Abadiano and Turner 2005, 51-52) suggests five instructional approaches to students' repeated reading, which are as follows:

1) **Student-adult reading**, which is essentially a teacher or other adult's modelling of fluent reading followed by the student's oral reading of the same text and this process is repeated until the student's performance is fluent.
2) **Choral reading** also entails the model reading of an adult, but this is followed by a group reading of the text, as opposed to a single student reading. As in the previous approach, the students are encouraged to reread the text until they are fluent in its reading.
3) **Tape-assisted reading** approach has the students read along as they listen to the fluent reading of a text on an audio-tape.
4) **Partner reading** has the students paired and taking parts in oral reading to each other. In this particular approach, the paired students may not be at the same level of reading fluency.
5) **Readers Theatre** is the last approach to reading fluency instruction. In this approach, the students are encouraged to rehearse a play, which they will later perform/read in front of their peers and parents.

The Readers Theatre, or other performance-based instructional approaches, are quite popular among researchers that focus on young native speakers and they have proved to be effective in enhancing reading fluency. In terms of the material that can be used for such instruction and subsequent performance are rhythmical and interactive texts, such "poetry,

song lyrics, chants, rhymes, plays […], monologues, dialogues, and letters," which facilitate the learner's development in fluent reading, that is in the development of accuracy, speed, prosodic rendering of a text, and comprehension (Rasinski 2006, 705).

In their study of non-native speakers, Galante and Thomson (2016) analyse the efficiency of using drama and theatre to enhance the EFL learners' fluency of oral speech. Despite the fact that this study, as well as various other studies cited in their review of literature, are focused on extemporaneously produced L2 speech, the results show that the use of theatre-based instruction can be as effective in EFL as it has proved to be in English as a native language.

3.4. Assessment of Reading Fluency

In this section, I will address the different assessment methods that were used in the previous research to measure the degree of oral reading fluency. Generally, the assessment can be classified into two categories – the subjective assessment based on the listener's perception of reading fluency, and the objective assessment based on precise analysis of the visual representation of speech (spectrogram) through the use of computer programs.

The subjective assessment was primarily meant for in-class assessment of students done by a teacher, which means that they are very listener-specific and not very useful for research purposes (Godde et al. 2019). Two of the most popular assessment scales are the "Integrated Reading Performance Record" developed by National Assessment of Education Progress (Pinnell et al. 1995, cited in Lems 2003) and the Multidimensional Fluency Scale developed by Zutell and Rasinski (1991, cited in Lems 2003). These assessment scales were among the first to assess the prosody and expressivity in addition to reading rate of the (native speakers) students.

The National Assessment of Education Progress (NAEP) scale is a unidimensional 4-point scale, which assesses the student's reading development on 4 levels, which can be roughly described as reading word-by-word (level 1), reading with occasional two-word phrases (level 2), or three/four-word phrases with appropriate syntax (level 3), and reading in meaningful large phrases with preserved syntax (level 4) (Pinnell et al. 1995, cited in Lems 2003, 37).

Zutell and Rasinski (1991, cited in Lems 2003, 38), on the other hand, devised a Multidimensional Fluency Scale on which the individual features of fluent reading (phrasing, smoothness, and pace) are assessed separately on the four levels of proficiency. The levels for phrasing range from monotonous, word-by-word reading at the first level to

well-phrased units with appropriate expression at the fourth level. Similarly, levels for smoothness range from frequent pauses and disfluencies (level 1) to smooth reading (level 4) and levels for pace range from slow and laborious speech at the first level to conversational speech at the fourth level.

This Multidimensional Fluency Scale was further updated by Rasinski (2004, cited in Godde et al. 2019) and Rasinski et al. (2009, cited in Godde et al. 2019), who included one more feature of fluent reading into the scale. This fourth feature is expression, which is assessed on a 4-point scale ("1 – non-existent, 2 – poor, 3 – mostly correct, 4 – adapted interpretation") (Godde et al. 2019, 5).

Apart from the fact that these scales are strongly subjective based on the perception of the assessor, each one of them has some serious drawbacks. The drawback of the NAEP scale is the unidimensionality, as a result of which it may prove to be difficult for the assessor to place the students on one of the four levels of the scale. The drawback of the Zutell-Rasinski Multidimensional Fluency Scale is the lack of the feature of expressivity in the scale and, once this feature of fluent reading has been included, the drawback of the scale lies in the lack of clear definition of expressivity in reading, which leaves room for various interpretations of it.

The objective assessment of reading fluency based on the analysis of spectrograms has been made possible by the advancement in technology. One of the software programs that can be used to analyse various suprasegmental aspects of speech is Praat (Boersma and Weenink, 2016). This program can be used to analyse duration and placement of pauses, phrasing, pitch contour, stress and loudness of speech (Binder et al. 2013; Schwanenflugel et al. 2004).

In this research, I will use Praat (Boersma and Weenink, 2016) to assess the speech samples and I will focus specifically on the analysis of changes in pitch, which is indicated as the Fundamental frequency F_0 in spectrograms, and also on the change in duration of the analysed utterances.

3.5. The Hypothesis

The cross-linguistic research focused on Czech and English conducted by Volín et al. (2015) shows that there is a significant difference between Czech and English pitch range. Based on their findings, the first hypothesis is that the recordings collected before pronunciation instruction will have narrower pitch range than the native English speakers.

Furthermore, it is clear from the review of literature that pronunciation instruction is an effective way of improving reading fluency, either of native or non-native speakers. The pronunciation instruction that the participants of this analysis were subjected to lasted one semester and was focused on suprasegmental features of English because, as the results of Derwing and Rossiter (2003) and Derwing et al. (1998) show, it is precisely such instruction that can best aid the improvement of fluency.

Based on the assumption that some fossilized non-native patterns of speech can be corrected through the right type of instruction, which is supported by Derwing et al. (1997) and Thomson and Derwing (2015), the second hypothesis of my thesis is that the explicit pronunciation instruction, although with delayed feedback, will have at least minor influence on the reading fluency of L2 speakers and the pitch range in the recordings collected after the pronunciation instruction will move towards the native-like pattern.

As the same group of people is being analysed within a relatively short period of time and all of the NNSs are adults, we can assume that their voice quality will not change as a result of the pronunciation instruction. The third hypothesis is that the pronunciation instruction will not have an effect on the pitch level of individual speakers.

In terms of the duration of utterance, the research conducted in reading fluency indicates that one of the aims of instruction in fluency is to produce faster speech, which would, nevertheless, be expressive. Based on observation of adult native speakers' tendencies, however, we can assume that fluent expressive reading to small children is slower than fluent expressive reading to older kids or adults. The research question of this thesis is whether the Czech learners of English will resemble the tendencies of adult native speakers in their fluent oral reading despite the fact that the pronunciation instruction was focused on enhancement of speech rate. The fourth hypothesis of this thesis, therefore, is that the duration of utterance in the post-instruction recordings will be slower than in the pre-instruction recordings and it should resemble the native speakers' speech.

4. Method

The practical part of this thesis is based on the analysis of the recordings collected from advanced Czech learners of English focusing on their oral reading fluency. This chapter summarises in detail the instruction process, but also provides information about the participants and the materials used in this research, as well as the process of the analysis. All the participants were aware of being recorded for the purposes of this thesis and willingly participated in the course.

4.1. Participants

The participants included 16 students of English and American studies from Palacký University in Olomouc. All of the participants are advanced Czech learners of English, but their L2 production is accented. 9 of them were undergraduate students and 7 were students of the master's program. There will be no differentiation based on the degree of study in the analysis of their recordings. Among the participants, there were 11 female students and 5 male students. For the sake of retaining anonymity of the participants, I will not address them by their names in the analysis, instead, I will use the following codes: M1-5 (01/02) for each of the five male speakers (the number in parenthesis indicates the recording – 01 stands for the recording before instruction and 02 for the one immediately following the instruction) and F1-11 (01/02) for each of the eleven female speakers.

As the (elective) pronunciation instruction course was targeted at students with Czech accent in English and all the participants who enrolled in the course acknowledged their limitations in their L2 production, the main incentive to participate in the pronunciation instruction has been their aspiration to improve their English, in addition to receiving credits for the course.

4.2. The instruction process

The whole instruction took place within a period of 12 weeks, which is a standard duration of a semester course. The first four weeks of the course took place in a computer classroom, where each student had a set of headphones with a microphone available for active participation. The duration of each lesson was 90 minutes. Unfortunately, from week five, the instruction took the form of distance education due to the outbreak of the pandemic. There were no online lessons for this course, but the students were given enough material for self-study. In terms of feedback from the instructor, after each self-study session, the

students sent their recordings of individual exercises to the instructor, who gave each of them a delayed feedback on their performance.

During the course, the students actively participated in the exercises prepared by the instructor and supervisor of this thesis. In addition to that, each student was asked to prepare a joined presentation on topics ranging from segmental features (pronunciation of certain vowels and consonants, silent letters in common words, or the difference between General American (GA) and Received Pronunciation (RP) pronunciation) to suprasegmental features (sentence rhythm and intonation, stress placement, speaking rate, and the use of sarcasm and how to spot it).

The detailed outline of the topics and exercises discussed for individual instruction weeks can be found in the Appendix.

4.3. Material

In terms of material for the reading fluency analysis, we asked the participants to read a well-known, short children's story called *The Tiger Who Came to Tea* by Judith Kerr, which can be seen in the Appendix. They were instructed to read the story with proper expression as if they were reading the story to a small child.

Their oral reading of the story was recorded before the instruction and then immediately following it. Prior to the oral reading of the story, the participants had time to read through the text and acquaint themselves with it. We recorded their first oral reading (before instruction) in the Recording Studio at the University. As the pronunciation instruction was limited by the outbreak of the pandemic and the entire instruction took place online without access to university facilities, the students were asked to record their second oral reading of the story (after instruction) at their homes. They were instructed to read the story in a quiet place without echo and background noise and they could use any appliance suitable for the task. Such solution of the problem was not ideal, but it was the only option available for the collection of data immediately following the instruction.

Two of the students did not turn around the paper on which the story was printed during the first reading of the story, therefore, only a portion of the story was recorded (which we noticed too late for them to record the story again). For the purposes of the analysis, the recordings of the two participants will be analysed partially. In addition to that, some of the recording included false starts, repetitions of certain phrases, and throat-clearing noises, which were ignored during the analysis if it did not interfere with the analysis itself. The analysis of pitch range was be carried out on 7 declarative sentences uttered as direct

speech and, in the second part of the analysis, the duration of the sentences was also analysed. The chosen sentences for the analysis are highlighted in the story, which can be seen in the appendix.

In addition to 16 non-native speakers, I searched the internet for recordings of native speakers reading the children's story *The Tiger Who Came to Tea* and I chose two recordings read by female speakers and two by male speakers. No specific information about the speakers was provided (such as age, their occupation or region where they live) but from the recordings I assume that the speakers were relatively young adults speaking British English. For the purposes of the analysis, the code for the two female native speakers is NSf1 and NSf2, and for the two male native speakers, it is NSm1 and NSm2.

4.4. Analysis

In the analysis of the recordings of NSs and NNSs, the primary focus is placed on pitch range and the comparison of pitch range values before and after pronunciation instruction. There are various ways or methods used in the previous research to analyse pitch range. Some researchers focus on the analysis of maximum f0 and minimum f0 values, others believe this approach is prone to error and focus on 90^{th}-10^{th} percentile values or standard deviation from the mean f0 of a sentence. The method that will be used for the analysis of the recordings in this thesis is based on David Patterson's PhD. dissertation (2000).

For the analysis of pitch range, it is useful to adopt the terms developed by Ladd (1996, cited in Patterson 2000, 33), by which he distinguished between two partially independent dimensions of pitch range which sufficiently establish a speaker's pitch range. The two dimensions are *pitch span* and *pitch level.* By pitch span we understand the width or range of frequencies used in an utterance, and by pitch level the height of a speaker's range. The value of pitch level indicates whether a speaker has a high or low voice and the value of pitch span indicates whether a speaker covers a wide or narrow range of frequencies in his/her speech.

Patterson (2000) believes that the best way to analyse a speaker's pitch range is to focus on linguistically motivated high and low tones in an utterance. Furthermore, the best way to measure pitch range is to use the linear scale (expressed in Hz) to measure the speaker's pitch level and the logarithmic scale (expressed in semitones (St)) to measure the speaker's pitch span. This is given by the fact that the choice of linear scale to measure pitch span would give us incorrect results. Patterson (2000, 42) provides an example that shows why the choice of a logarithmic scale is better in the case of span. The example compares

the pitch span of 100 Hz (between the maximum f0 value 200 Hz and minimum value of 100 Hz) for a male speaker and pitch span of 200 Hz (with maximum f0 of 400 Hz and minimum f0 of 200 Hz) for a female speaker. If we used the linear scale expressed in Hz, the results would indicate that the pitch span of the male speaker is half of that of the female speaker. This is, however, not the case. That is why we use the logarithmic scale expressed in St, which indicates that the two speakers have exactly the same pitch span. The logarithmic scale would, however, not be suitable to measure the pitch level because the scale represents the difference between two tones and in the case of pitch level, there is only one f0 value for an utterance.

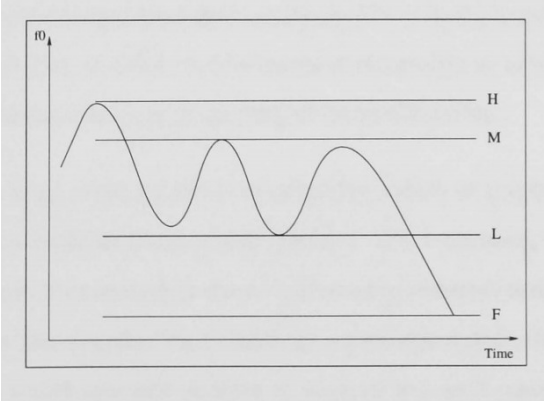
The criteria for the selection of measurement points for the analysis were based on Experiment 2 in Patterson (2000), where he came to the conclusion that the best measurement points for pitch span are the non-initial accent peak (marked in Figure 4.1 as M) and post-accentual valley (marked as L) and the best measurement point for pitch level is the sentence-final low (marked as F). It is important to mention that in the analysis of the recordings in this thesis, the measurement points are marked differently from Patterson's markings. In this thesis, the non-initial accent peak is marked M, the post-accentual valley as V, and the sentence-final low as L%. These marks are based on Mennen (2007).

The analysis of the recordings was carried out in Praat (Boersma and Weenink 2016) and, as I mentioned already, there were 7 direct speech sentences selected for each speaker and they were marked by a number from 1-7:

1. "We'd better open the door and see."
2. "…but I'm very hungry."
3. "Thank you for my nice tea."
4. "I think I'd better go now."
5. "I don't know what to do."
6. "… the tiger has eaten it all."
7. "I've got a very good idea."

As the pitch contour of the individual recordings did not resemble the one illustrated in 4.1 above and the accent peaks were distributed less neatly (the overall pitch contour was not gradually going down as can be seen in the illustration above but, in most of the cases,

Figure 4.1: Measurement points for pitch span and pitch level on an idealised pitch contour, where H indicated sentence-initial accent peak, M the non-initial accent peak, L indicates the post-accentual valley and F the sentence-final low (Patterson 2000, 94).

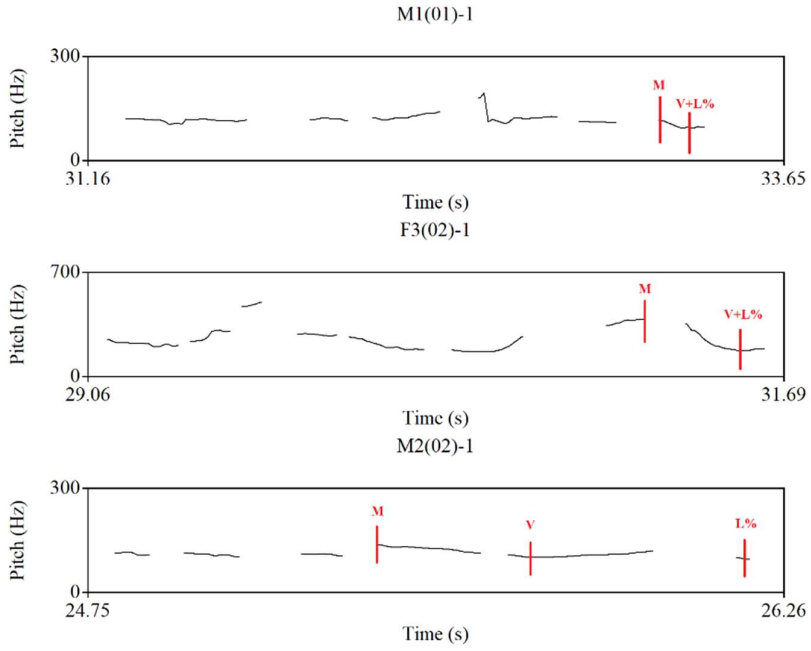


the highest accent peak appeared at the end of the utterance), the choice of the measurement points was not so straightforward. Generally, I tried to avoid locating the measurement points within voiceless segments (there was no pitch contour visible) or within segments that contained unnatural drops or rises in pitch contour.

In the following paragraphs, I will analyse individual sentences and describe the process of positioning the measurement points, as well as some of the problems that occurred in them.

For sentence 1, the non-native speakers generally placed the accent peak on one of the words in the phrase OPEN THE DOOR or on the very last word of the sentence, SEE. The decision was based on my perception of the position of the emphasis within the sentence. When the primary accent peak was located on one of the three words OPEN THE DOOR, the post-accentual valley appeared before the minor accent peak for the last word. If, however, the accent peak was located on the last word (SEE), the post-accentual valley and sentence-final low were marked at the same position. As can be seen in figure 4.2 below, the accent peak for speaker M1(01) was located on the last word (SEE), but there is one more visible accent peak in the previous part of the sentence. That is due to the insertion of an unnecessary word to the sentence ("open UP the door") on which the accent was located. Naturally, I ignored this accent peak as it was not supposed to be there. For speaker F3(02) from the same figure below, the accent peak was located on the last word as there is no visible accent peak before it, only the initial one, which is ignored in Patterson (2000) and therefore, also ignored here. The last speaker from the figure, M2(02), has the primary

Figure 4.2: Pitch contour of the first sentence for speakers M1(01), F3(02), and M2(02) with the measurement points indicated by M, V, L%.



accent peak on OPEN and a secondary peak on SEE. That is the reason why I positioned the non-initial accent peak where it is indicated in the figure below. If I compared the position of the measurement points of the non-native speakers to the position of the four native speakers, there are clear similarities between the two groups. Two of the native speakers have the accent peak located on the last word, while another two have it on one of the words of the phrase OPEN THE DOOR.

In Figure 4.3 below, you can see the problem I mentioned earlier. The software sees the last word of the first sentence (SEE) as voiceless, which means that there is no pitch contour visible for it. Because of that, there is no sentence-final low value for this speaker (F1(02)) and the mean pitch level was calculated without this value. The measurement points for calculating pitch span were positioned in the first part of the sentence, specifically on OPEN (M) and DOOR (V).

For the second sentence, the speakers placed the accent peak on either VERY, in which case the post-accentual valley was located between this word and the following one, or on HUNGRY and in that case the post-accentual valley was located at the same position with sentence-final low. The decision for the position of the measurement points in this sentence was also based on my perception; it was usually quite straightforward whether the speaker placed the emphasis on the word VERY or on HUNGRY. In figure 4.4 you can see

Figure 4.3: Waveform and spectrogram with pitch contour for speaker F1(02) indicating the lack of sentence-final pitch value on the voiceless segment.

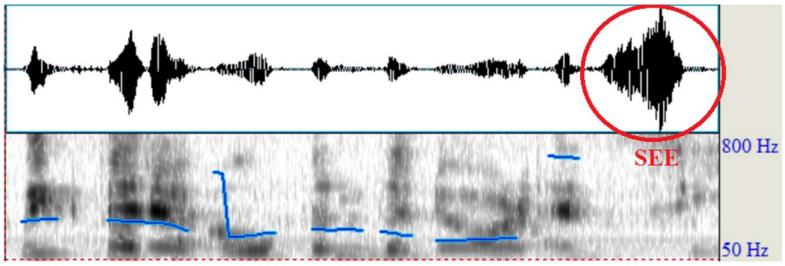
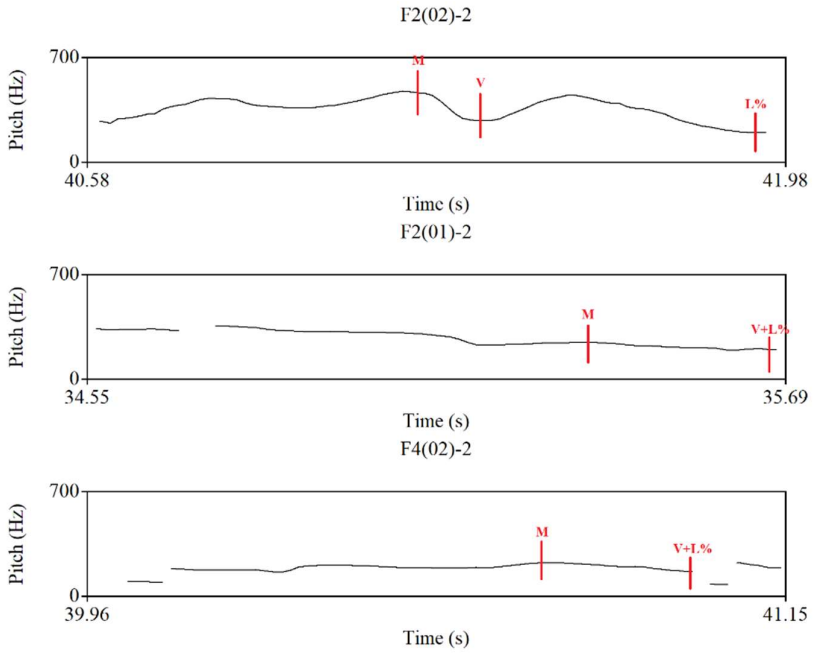


Figure 4.4: Pitch contour of the second sentence for speakers F2(02), F2(01), and F4(02) with the measurement points indicated by M, V, L%.



both cases. Speaker F2(02) has accent peaks for both of the words but the primary emphasis is located on the word VERY. What is interesting here is the fact that the same speaker placed the emphasis on the word HUNGRY in her first recording (before instruction), see middle pitch contour in the same figure. In this case, the pitch contour has no visible peak for the word VERY; the contour gradually goes down from the initial accent peak and then goes up for the word HUNGRY. The last speaker's pitch contour illustrated in the figure below indicates the position of the accent peak on HUNGRY, but the location of the post-accentual valley and sentence-final low is not at the very end of the contour. The reason for it is that the last part of the contour contains a creaky segment. The native speakers were

unanimous in the position of the accent peak for this sentence, they all emphasised the word VERY.

For sentence three, the position of the accent peak was either located on the vowel in the word NICE or on the initial consonant sound of the word TEA. In both cases, the post-accentual valley and sentence-final low were positioned on the same spot, which can be seen in figure 12 in the appendix. What is interesting to note here is the fact that the accent peak positioned on the word NICE was more common in the first recording (before instruction), while in the second recording, those speakers who placed the accent peak on NICE before instruction, usually shifted the accent peak to the word TEA after instruction. Speaker M1(02), whose pitch contour can also be seen in figure 1 in the appendix, emphasised the word MY, therefore, I placed the measurement point on that word. The four native speakers were not unanimous in the position of the accent peak in this sentence. Three of them emphasised the word NICE and one positioned the accent peak on the word TEA. The position of the valley and low were the same as that of the non-native speakers.

In sentence four, all the speakers (non-native and native alike) have the same position of the non-initial accent peak on the word GO. The position of the post-accentual valley, however, is not the same. There were speakers in whose recordings the position of the measurement point for the post-accentual valley was the same as for the sentence-final low (example M1(01) in figure 13 in the appendix), but there were also recordings in which there was a minor accent peak visible for the word NOW, therefore, the post-accentual valley point was marked before the final peak (example M5(01) in the same figure). The position of the measurement points for this sentence was also based on perception because in some recordings there was a visible accent peak for the word BETTER, but this word did not carry the main emphasis and that is why I placed the measurement point on GO (example F3(01) in the same figure).

In sentence five, all the speakers (native and non-native alike) with one exception agreed on the position of the accent peak on the sentence final word, DO. As this was the last word of the sentence, the post-accentual valley and sentence-final low measurement points were marked together in the same position. This can be seen in the example of speaker F3(02) in figure 14 in the appendix. One speaker (F4(01)), whose pitch contour can also be seen in figure 14, positioned the accent peak on the word KNOW and, as there was also a peak for the sentence-final word, the position of the post-accentual valley was marked before that last peak.

For the sixth sentence, there were two possible positions of the accent peak (for native and non-native speakers alike). The speakers either emphasised the word EATEN or the sentence-final word ALL. When the non-initial accent peak was located on the word EATEN and there was a minor accent peak on the last word, the post-accentual valley was located between the words IT and ALL. This is the case of speaker F3(01), whose pitch contour can be seen in figure 15 in the appendix. When, however, there was no accent peak on the last word and the main emphasis was placed on EATEN, the post-accentual valley and sentence-final low were marked together (speaker F6(01) in figure 15). And as can be seen in the last example in figure 15 (speaker F11(02)), there were also cases with the accent peak on ALL and, as it was the last word of the sentence, the measurement points for valley and low were marked together.

And for the seventh sentence, there were four possibilities for the position of the measurements points. The majority of non-native speakers, and all four native speakers, emphasised the word VERY and, as there was a smaller accent peak for the word IDEA, the post-accentual valley was marked before it. This can be seen in the pitch contour of speaker M1(02) in figure 16 in the appendix. Another possible position of the non-initial accent peak was on the word GOOD with another smaller accent peak for idea. As can be seen in the example of speaker M3(01) in figure 16, the post-accentual valley was marked before the peak. There were, however, cases with no accent peak for the word IDEA and in such cases the post-accentual valley and sentence-final low were marked together. This is the example of speaker M5(02) in figure 16. And there was also one speaker (F2(02) in figure 16), who placed the emphasis on the word IDEA and as this was the final word of the sentence, the post-accentual valley and sentence-final low were marked together.

For each speaker, the value of the measurement points (expressed in Hz) was manually extracted from Praat (Boersma and Weenink 2016) and recorded in an Excel table created for each speaker. From the recorded values in Hz, the mean value of all 7 M, V, and L% points was calculated for each recording and these values were then used for the calculation of a speaker's pitch span and pitch level. As was already mentioned, the mean pitch level was expressed in Hz and the pitch span in St. For the calculation of pitch span in St, the following formula was used in Excel:

St=12*IMLOG2(f0maximum value in Hz/f0 minimum value in Hz)

In the second part of the analysis, the focus was on the duration of individual sentences for each recording. This analysis was also carried out in Praat (Boersma and Weenink 2016). The measurement points were placed at the precise beginning and end of each of the 7 sentences and the duration in milliseconds was copied into an Excel table. As three speakers (M1, M3, F7) produced one or two sentences of the first recording with disfluencies, the mean values for both recordings of those speakers were therefore calculated without these sentences. However, for the individual speaker analyses, the values of all sentences were used and that is why, in some cases, the mean values for these speakers may not coincide.

5. Results

For the statistical analysis of the results, I conducted simple t-tests in Statistica (TIBCO Software Inc 2018), using either single sample t-tests against a native speaker reference or dependent samples t-tests when comparing pre-test and post-test values of the same group of speakers or individual speakers. The tests were conducted in three focus areas (pitch span, pitch level, and utterance duration) and the results are organised into subchapters according to those areas.

5.1. Pitch Span Results

For the pitch span, I ran a dependent-samples t-test to compare the mean value of semitones for the 16 non-native speakers before and after the pronunciation instruction. There was no differentiation needed for female and male speakers because of the chosen scale (logarithmic) as I already mentioned before. The results indicate that there was a nonsignificant difference between the pre-test (M = 6.17, SD = 1.60) and post-test (M = 7.68, SD = 3.37) values, $t(15) = -1.65$, $p = .12$. This result is given by the larger standard deviation from the mean value in the post-test group, which is illustrated in figure 5.1. When compared to the mean NS reference value (10.5), the results for the pre-test and post-test data indicate that the difference is statistically significant in both sample, $t(15) = -10.85$, $p < .001$ and $t(15) = -3.35$, $p = .004$ respectively.

Greater variation of the results was observed in the dependent-samples t-tests for individual speakers. The mean values were calculated from the seven sentences (4 sentences in the case of speaker M2 and F9) and these values were further analysed against a reference constant value, which represented the mean value of four native speakers (10,5), in the single-sample t-test. The illustration of the mean span values for individual speakers against the NS reference is shown in figure 5.2. The precise data of the NNSs as well as the four NSs can be seen in figure 18 in the Appendix.

For speaker **F1**, the results of the pre-test (M = 7.05, SD = 2.74) and post-test (M = 5.79, SD = 2.76) indicate a statistically significant change in pitch span, $t(6) = 2.76$, $p = .03$. In both cases (pre-test and post-test), the results indicate a statistically significant difference from the mean native speaker reference value (10,5), $t(6) = -3.33$, $p = .02$ and $t(6) = -4.52$, $p = .004$ respectively.

Figure 5.1: Box and whisker-plots graph for the pre-test vs. post-test pitch span values of 16 NNSs.

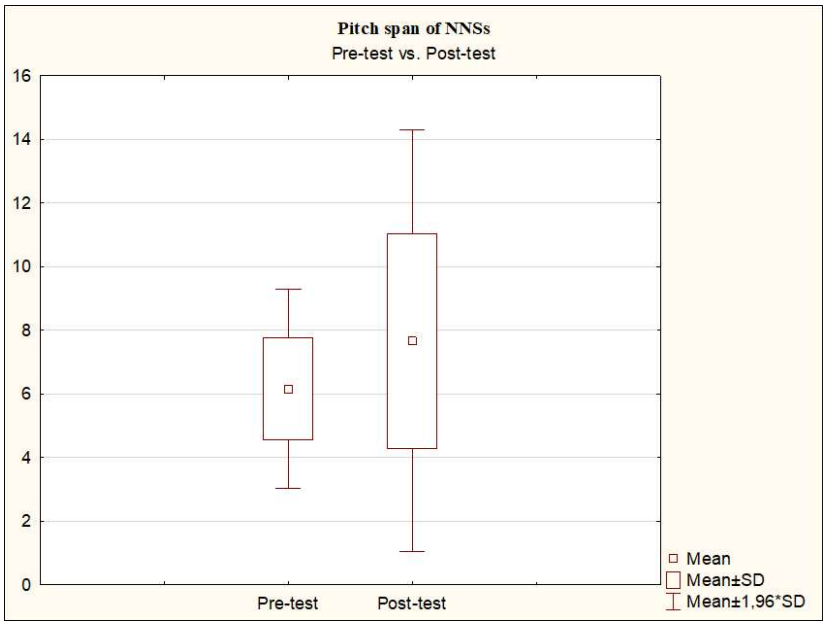
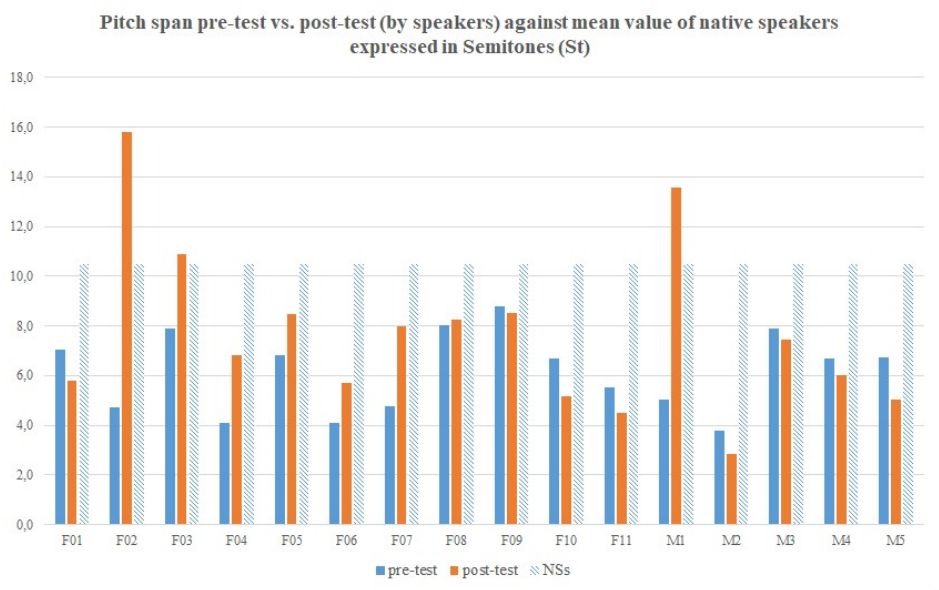


Figure 5.2: Mean pre-test and post-test pitch span values for individual speakers against a NS reference.



For speaker **F2**, the difference between pre-test (M = 4.72, SD = 3.11) and post-test (M = 15.81, SD = 3.11) values shows statistical significance, $t(6)$ = -6.79, $p < .001$. The comparison of the pre-test values to the native speaker reference (10,5) shows statistical significance, $t(6)$ = -4.91, p = .003. And even though the change in pitch span for this speaker was huge, the results of the comparison of the post-test data to the NS reference

also show statistical significance, $t(6) = 4.52$, $p = .004$. The results indicate that this speaker's pitch span was far away from the natural production of the four native speakers.

There was a statistically nonsignificant difference between the pre-test ($M = 7.89$, $SD = 2.12$) and post-test ($M = 10.9$, $SD = 3.98$) values of speaker **F3** which is given by the larger standard deviation from the mean value in the post-test recording, $t(6) = -1.57$, $p = .17$. However, the comparison to the NS reference (10,5) better indicates the improvement. The results for the pre-test pitch span values are statistically significant, $t(6) = -3.26$, $p = .02$. The results for the post-test values, on the other hand, are statistically nonsignificant, $t(6) = 0.27$, $p = .80$. The results show that this speaker is not statistically different from the four native speakers.

For speaker **F4**, the difference between pre-test ($M = 4.10$, $SD = 2.32$) and post-test ($M = 6.81$, $SD = 2.42$) values was statistically significant, $t(6) = -3.50$, $p = .01$. But, despite this significant improvement in pitch span values in the post-test recording, the results indicate that in both cases (pre-test and post-test), the comparison to the native speaker reference value (10,5) shows statistically significant difference, $t(6) = -7.29$, $p < .001$ and $t(6) = -4.03$, $p = .007$ respectively.

Similar to speaker F3, there was a statistically nonsignificant change in the pre-test ($M = 6.82$, $SD = 2.19$) and post-test ($M = 8.49$, $SD = 4.19$) values for speaker **F5**, which was due to the standard deviation in the post-test, $t(6) = -1.26$, $p = .26$. However, the improvement in pitch span of this speaker is more obvious in the statistical analysis against a native speaker reference value. In the pre-test, there was a statistically significant difference between the production of speaker F5 and the mean NS reference, $t(6) = -4.44$, $p = .004$. The post-test, on the other hand, shows that speaker F5 improved her production and the difference is nonsignificant, $t(6) = -1.27$, $p = .25$.

For speaker **F6**, there was a statistically nonsignificant change in the pre-test ($M = 4.11$, $SD = 1.60$) and post-test ($M = 5.70$, $SD = 3.78$) pitch span values, $t(6) = -1.19$, $p = .28$. In both cases, there is a statistically significant difference from the native speaker reference value (10,5), $t(6) = -10.55$, $p < .001$ and $t(6) = -3.36$, $p = .02$ respectively.

For speaker **F7**, there was a statistically near significant difference between pre-test ($M = 4.79$, $SD = 1.88$) and post-test ($M = 8.00$, $SD = 3.11$) pitch span values, $t(6) = -2.29$, $p = .06$. In the pre-test, there was a statistically significant difference from the NS reference value, $t(6) = -8.04$, $p < .001$. And for the post-test, there was also a near significant difference from the NS reference value, $t(6) = -2.13$, $p = .08$.

Speaker **F8** is interesting in that there is a statistically nonsignificant improvement in the pre-test (M = 8.04, SD = 5.11) and post-test (M = 8.24, SD = 5.80) pitch span values, $t(6) = -0.07$, $p = .95$. But in both cases, the speaker show statistically nonsignificant difference from the NS reference value (10,5), which means that the production of this speaker is comparable to that of the four native speakers, $t(6) = -1.27$, $p = .25$ and $t(6) = -1.03$, $p = .34$ respectively.

Speaker **F9** showed statistically nonsignificant improvement in her pre-test (M = 8.78, SD = 7.64) and post-test (M = 8.53, SD = 2.90) pitch span values, $t(3) = 0.08$, $p = .94$. But, similarly to the previous speaker, the production was comparable to the four native speakers in both the pre-test and post-test samples. The results in the two samples were nonsignificant, $t(3) = -0.45$, $p = .68$ and $t(3) = -1.36$, $p = .27$ respectively.

Speaker **F10** showed nonsignificant change between pre-test (M = 6.69, SD = 3.24) and post-test (M = 5.24, SD = 3.13) values towards a narrower pitch span, $t(6) = 0.93$, $p = .39$. However, in both cases, the production was not comparable to the NS reference (10,5) and showed statistically significant results, $t(6) = -3.11$, $p = .02$ and $t(6) = -4.44$, $p = .004$ respectively.

For speaker **F11**, there was a nonsignificant difference between pre-test (M = 5.53, SD = 2.46) and post-test (M = 4.50, SD = 2.33) pitch span values and this speaker also moved towards the narrower production, $t(6) = 1.53$, $p = .18$. In both cases, the production of this speaker was significantly different from the NS reference value (10,5), $t(6) = -5.35$, $p = .002$ and $t(6) = -6.80$, $p < .001$ respectively.

There was a statistically significant difference between the pre-test (M = 5.02, SD = 2.27) and post-test (M = 13.56, SD = 8.82) pitch span values of speaker **M1**, $t(6) = -2.56$, $p = .04$. In the pre-test sample, there was a statistically significant difference from the NS reference value (10,5), $t(6) = -6.38$, $p < .001$. And despite the fact that this speaker's production was wider than any of the native speakers, the results were statistically nonsignificant, $t(6) = 0.92$, $p = .39$.

For speaker **M2**, there was a statistically nonsignificant difference between pre-test (M = 3.78, SD = 3.33) and post-test (M = 2.85, SD = 2.71) and it was also towards the narrower production, $t(3) = 0.34$, $p = .76$. As the pronunciation was rather narrow even in the first recording, in both cases there is a statistically significant difference from the NS reference value (10,5), $t(3) = -4.04$, $p = .03$ and $t(3) = -5.64$, $p = .01$ respectively.

Speaker **M3** showed a statistically nonsignificant change between the pre-test (M = 7.89, SD = 4.16) and post-test (M = 7.44, SD = 5.29) pitch span values towards the narrower

production, $t(6) = 0.27$, $p = .80$. In both cases, the results showed nonsignificant difference from the NS reference value (10,5), $t(6) = -1.66$, $p = .15$ and $t(6) = -1.53$, $p = .18$.

There was also a minor and statistically nonsignificant difference between the pre-test ($M = 6.69$, $SD = 2.68$) and post-test ($M = 6.01$, $SD = 4.58$) pitch span values of speaker **M4** towards the narrower production, $t(6) = 0.30$, $p = .77$. And in both cases, the results showed statistically significant difference between the production of this speaker and the NS reference value, $t(6) = -3.76$, $p = .01$ and $t(6) = -2.59$, $p = .04$ respectively.

Speaker **M5** also showed statistically nonsignificant difference between pre-test ($M = 6.74$, $SD = 4.40$) and post-test ($M = 5.01$, $SD = 5.26$) values towards the narrower production, $t(6) = 0.82$, $p = .44$. In the case of the pre-test sample, there was a near significant difference between this speaker and the NS reference value, which means that this speaker's production before the pronunciation instruction was still comparable to that of the NSs, $t(6) = -2.26$, $p = .06$. In the post-test sample, the comparison to the NS reference value showed statistically significant difference, $t(6) = -2.76$, $p = .03$.

5.2. Pitch Level Results

For the second dimension, pitch level, I ran a dependent-samples t-test to compare the mean values (expressed in Hz) of the 16 NNSs before and after the pronunciation instruction. In this analysis, the use of the linear scale expressed in Hertz required the differentiation of the samples based on gender. Therefore, the group results are expressed separately for male and female speakers. The results indicate that there was a nonsignificant difference in the pre-test ($M = 112.68$, $SD = 28.51$) and post-test ($M = 115.88$, $SD = 23.11$) data of the male speakers, which can be seen in figure 5.3 below, $t(4) = -0.42$, $p = .69$. It is important to consider the fact that there were only five male speakers in this analysis, therefore, the results are limited, and a precise conclusion cannot be drawn from them. The results for the female speakers, on the other hand, show that the difference between pre-test ($M = 169.47$, $SD = 27.94$) and post-test ($M = 188.36$, $SD = 28.31$) data was statistically significant, $t(10) = -2.56$, $p = .03$. The illustration of the results for female speakers can be seen in figure 5.4.

The mean pre-test and post-test data of all NNSs as compared to the values of two female and two male NSs are illustrated in figure 5.5. The precise mean data for the NNSs as well as the mean values of the four native speakers can be seen in figure 19 in the Appendix. As it was with the pitch span data, the data for pitch level vary according to the speaker. I ran dependent-samples t-tests to compare the difference between the pre-test and post-test pitch level values for each non-native speaker.

Figure 5.3: Box and whisker-plots graph for the pre-test vs. post-test pitch level values of 5 male NNSs.

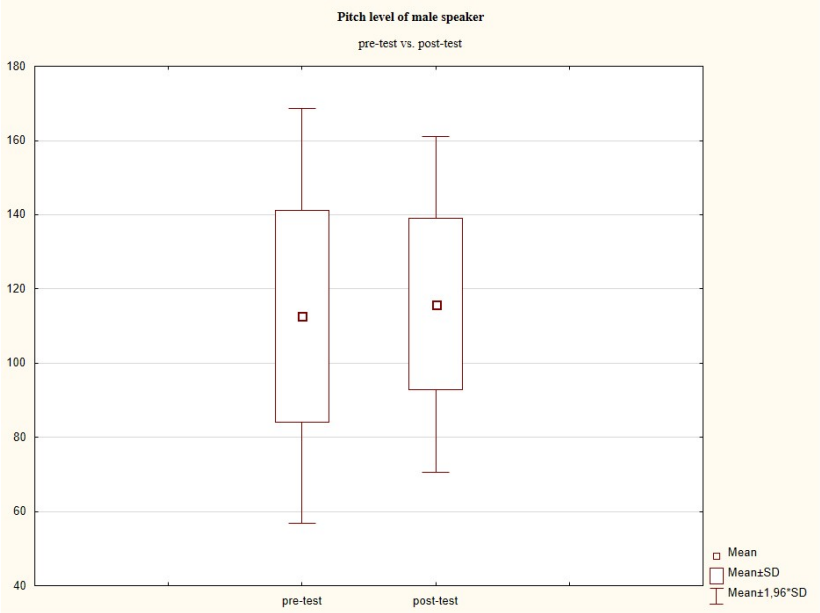
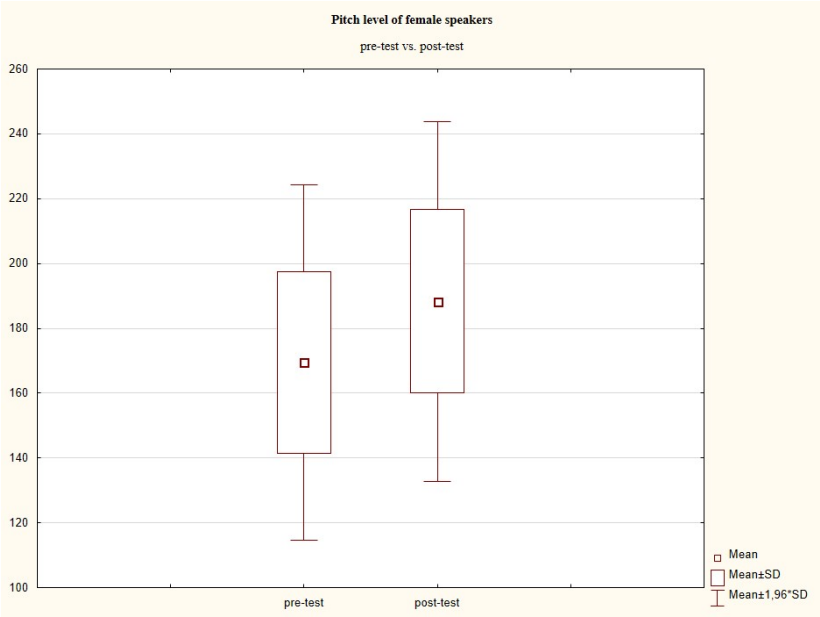
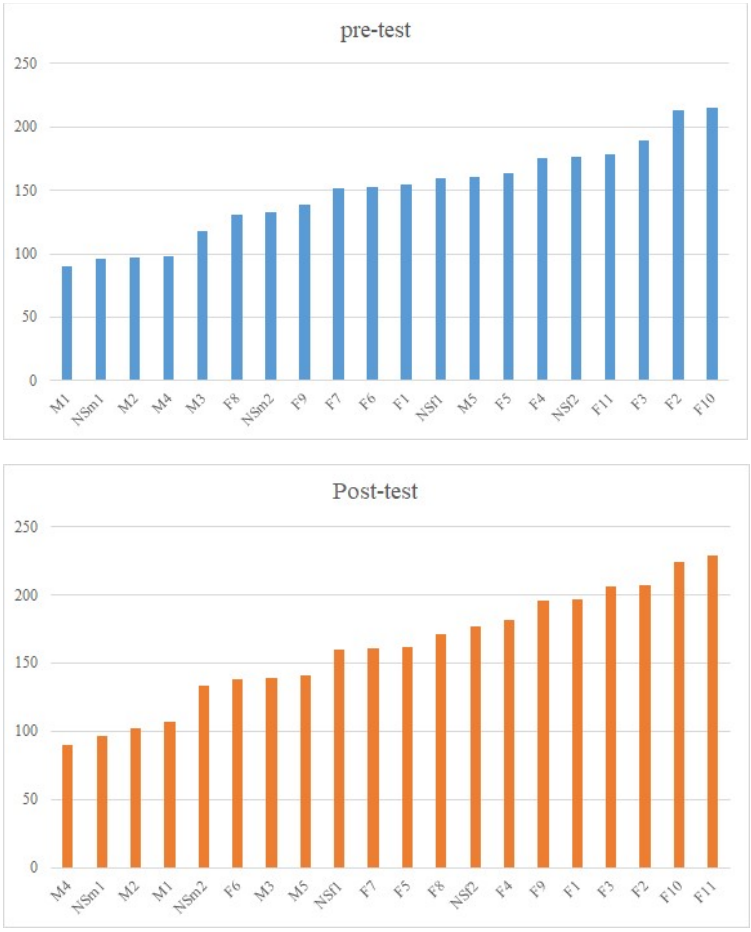


Figure 5.4: Box and whisker-plots graph for the pre-test vs. post-test pitch level values of 11 female NNSs.



For speaker **F1**, the dependent-samples t-test was calculated from only 6 sentences because in the post-test recording there was no sentence-final low point due to the last word being voiceless. The results indicate that the difference between pre-test (M = 153.80, SD = 14.10) and post-test (M = 196.62, SD = 34.66) pitch level values is statistically significant,

Figure 5.5: Mean pre-test and post-test pitch level values for individual speakers (native and non-native alike).



$t(5) = -3.75$, $p = .01$.

There was a statistically nonsignificant difference between the pre-test ($M = 213.31$, $SD = 10.06$) and post-test ($M = 206.60$, $SD = 25.24$) pitch level values of speaker **F2**, $t(6) = 0.91$, $p = .40$. Similarly, the results of the dependent-samples t-test for speaker **F3** showed nonsignificant difference between the pre-test ($M = 189.63$, $SD = 9.96$) and post-test ($M = 206.24$, $SD = 34.80$) values, $t(6) = -1.34$, $p = .23$. For speaker **F4**, the difference between pre-test ($M = 175.16$, $SD = 11.14$) and post-test ($M = 182.01$, $SD = 9.78$) pitch level values was statistically near significant and the change was towards the higher frequency, $t(6) = -2.06$, $p = .08$. There was a statistically nonsignificant change between the pre-test ($M = 163.24$, $SD = 18.16$) and post-test ($M = 162.06$, $SD = 12.32$) values of speaker **F5**, $t(6) = 0.26$, $p = .80$. For speaker **F6**, there was a statistically nonsignificant difference between the pre-test ($M = 152.93$, $SD = 30.51$) and post-test ($M = 137.71$, $SD = 74.10$) pitch level data,

$t(6) = 0.47$, $p = .66$. For speaker **F7**, there was a statistically nonsignificant difference between the pre-test ($M = 151.99$, $SD = 6.05$) and post-test ($M = 160.96$, $SD = 18.11$) pitch level data and the change was towards the higher frequency, $t(6) = -1.22$, $p = .27$. For speaker **F8**, the results show that there was a nonsignificant difference between the pre-test ($M = 130.74$, $SD = 65.40$) and post-test ($M = 170.97$, $SD = 32.75$) data because of the large standard deviation in the first recording, $t(6) = -1.70$, $p = .14$.

There was a similar problem for speaker **F9** as there was for speaker F1, but in this case the lacking sentence-final low point was in the pre-test recording. The dependent-samples t-test was again calculated only from three sentences (this is one of the two speakers that produced only 4 of the analysed sentences) and the results of the pre-test ($M = 138.40$, $SD = 52.00$) and post-test ($M = 169.20$, $SD = 27.87$) data indicate that the difference was statistically nonsignificant, $t(2) = -1.02$, $p = .41$.

For speaker **F10**, the difference between the pre-test ($M = 214.94$, $SD = 71.06$) and post-test ($M = 224.13$, $SD = 47.07$) pitch level data is statistically nonsignificant, $t(6) = -0.42$, $p = .69$. For speaker **F11**, there is also a nonsignificant difference between the pre-test ($M = 178.91$, $SD = 24.71$) and post-test ($M = 228.47$, $SD = 97.79$) pitch level data, $t(6) = -1.50$, $p = .18$.

There was a statistically nonsignificant difference between the pre-test ($M = 90.09$, $SD = 5.32$) and post-test ($M = 106.40$, $SD = 27.59$) data of speaker **M1**, $t(6) = -1.44$, $p = .20$. For speaker **M2**, the pitch level was calculated from only 4 sentences and the results indicate that the difference between the pre-test ($M = 97.25$, $SD = 11.20$) and post-test ($M = 102.38$, $SD = 17.17$) data is statistically nonsignificant, $t(3) = -3.12$, $p = .05$. There is a statistically significant difference between the pre-test ($M = 117.59$, $SD = 5.59$) and post-test ($M = 139.34$, $SD = 19.25$) data of speaker **M3**, $t(6) = -2.87$, $p = .03$. For speaker **M4**, the difference between the pre-test ($M = 98.06$, $SD = 7.48$) and post-test ($M = 90.03$, $SD = 29.43$) pitch level data is statistically nonsignificant, $t(6) = 0.66$, $p = .53$. There was a statistically nonsignificant difference between the pre-test ($M = 160.30$, $SD = 41.38$) and post-test ($M = 141.34$, $SD = 24.09$) data of speaker **M5**, $t(6) = 0.95$, $p = .38$.

5.3. Utterance Duration Results

For the third part of the analysis, the analysis of utterance duration, I ran a dependent-samples t-tests to compare the mean duration values (expressed in Milliseconds) of the 16 NNSs before and after the pronunciation instruction and single-sample t-tests to compare the individual speakers to the NS reference, which was (1710.00 ms). In some cases,

however, the mean NS reference value was calculated from the values of two faster NS production and was 1570.00 ms. The group results indicate that there was a statistically significant difference between the pre-test (M = 1371.88, SD = 132.98) and post-test (M = 1517.50, SD = 200.02) duration values, $t(15) = -3.60$, $p = .003$. The illustration of the result can be seen in figure 5.6 below and the graphical illustration of the mean values of the 16 NNSs and 4 NSs in figure 5.7. The precise mean values of all the speakers can be seen in figure 20 in the Appendix.

For speaker **F1**, there was a nonsignificant difference between pre-test (M = 1488.57, SD = 310.45) and post-test (M = 1425.71, SD = 296.19) utterance duration values towards the faster pronunciation, $t(6) = 1.26$, $p = .25$. While in the pre-test sample, the difference between this speaker and the NS reference value (1710.00) was nonsignificant, $t(6) = -1.89$, $p = .11$, the post-test sample was significantly different, $t(6) = -2.54$, $p = .04$. However, when compared to the reference value calculated from the two fastest native speakers

Figure 5.6: Box and whisker-plots graph for the pre-test vs. post-test utterance duration (expressed in milliseconds).

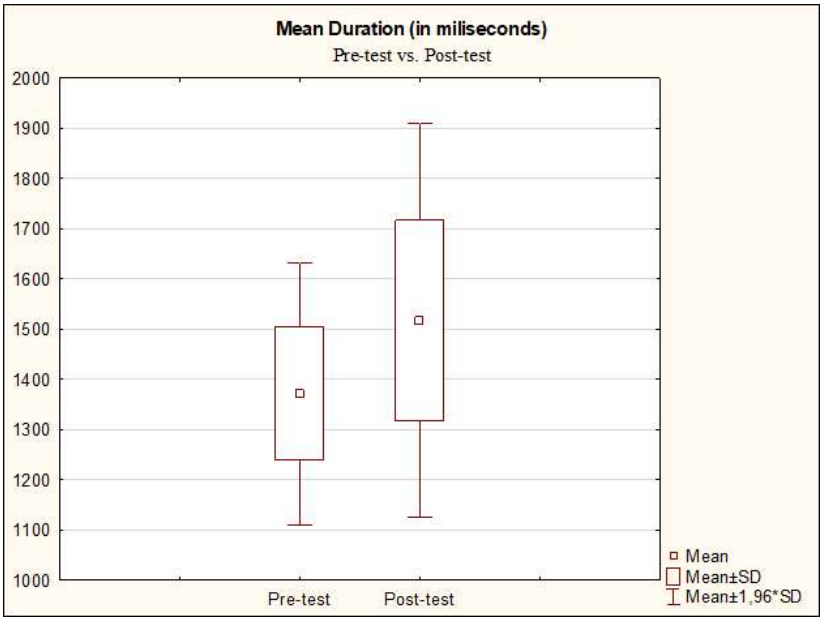


Figure 5.7: Mean pre-test and post-test utterance duration (expressed in milliseconds) of 16 non-native speakers and four native speakers.

(1570.00), the results are nonsignificant in both pre-test and post-test sample, $t(6) = -0.60$, $p = .51$ and $t(6) = -1.29$, $p = .25$ respectively.

For speaker **F2**, there was a statistically significant difference between the pre-test ($M = 1458.57$, $SD = 221.09$) and post-test ($M = 1565.71$, $SD = 189.90$) duration data towards the slower pronunciation, $t(6) = -4.78$, $p = .003$. There was a statistically significant difference between the mean utterance duration of this speaker and the NS reference (1710.00) in the pre-test sample, $t(6) = -3.01$, $p = .02$. And there was a statistically near significant difference in the post-test sample, $t(6) = -2.01$, $p = .09$. But again, in the comparison to the mean value of the two fastest native speakers (1570.00), the results indicate that the difference is statistically nonsignificant in both cases, $t(6) = -1.33$, $p = .23$ and $t(6) = 0.06$, $p = .95$ respectively.

There was a significant difference between the pre-test ($M = 1317.14$, $SD = 207.74$) and post-test ($M = 1852.86$, $SD = 463.81$) duration data for speaker **F3**, who produced slower utterance in the second recording, $t(6) = -3.61$, $p = .01$. In the pre-test sample, there was a statistically significant difference between the production of this speaker and the reference (1710.00) value, $t(6) = -5.00$, $p = .002$. What is, however, more interesting is that this speaker's pre-test production was significantly different (faster) even from the faster NS reference (1570.00), $t(6) = -3.22$, $p = .02$. The post-test sample, on the other hand, showed statistically nonsignificant difference from the NS reference (1710.00), $t(6) = 0.81$, $p = .45$.

For speaker **F4**, the difference between the pre-test (M = 1380.00, SD = 164.72) and post-test (M = 1362.86, SD = 130.48) duration data is statistically nonsignificant, $t(6)$ = 0.56, p = .60. However, in both samples, the difference between the mean duration of this speaker in both pre-test and post-test sample and the NS reference value (1710.00) is statistically significant, $t(6)$ = -5.30, p = .002 and $t(6)$ = -7.04, p < .001 respectively. Furthermore, the results indicate that the production of this speaker in both recordings is significantly faster than the mean value of the two faster NSs (1570.00), $t(6)$ = -3.05, p = .02 and $t(6)$ = -4.20, p = .01 respectively.

There is a statistically significant difference between the pre-test (M = 1318.57, SD = 168.07) and post-test (M = 1521.43, SD = 204.24) duration data for speaker **F5**, $t(6)$ = -3.57, p = .01. The results for the pre-test sample indicate that the production of this speaker is significantly faster than both the mean NS reference (1710.00) and the mean faster NS reference (1570.00), $t(6)$ = -6.16, p < .001 and $t(6)$ = -3.96, p = .01 respectively. The result for the post-test sample, on the other hand, is statistically near significant when compared to the NS reference (1710.00), $t(6)$ = 2.44, p = .05, but nonsignificant when compared to the faster NS reference (1570.00), $t(6)$ = -0.63, p = .55.

For speaker **F6**, there is a statistically significant difference between the pre-test (M = 1345.71, SD = 246.63) and post-test (M = 1650.00, SD = 336.80) duration data, $t(6)$ = -3.50, p = .01. The results for the pre-test sample indicate that the production of this speaker is significantly faster than the NS reference (1710.00), $t(6)$ = -3.91, p = .01. The result for the comparison to the faster NS reference (1570.00) is near significant due to rounding down the p-value, $t(6)$ = -2.41, p = .05. The result for the post-test sample, on the other hand, indicates that the difference between the production of this speaker and the NS reference (1710.00) is nonsignificant, $t(6)$ = -0.47, p = .65.

For speaker **F7**, the difference between the pre-test (M = 1432.86, SD = 389.35) and post-test (M = 1445.71, SD = 160.09) duration data is statistically nonsignificant, $t(6)$ = -0.10, p = .92. The results for the pre-test sample indicate that there is a statistically nonsignificant difference from both the NS reference value (1710.00) and the faster NS reference value (1570.00) which is given by the larger SD caused by the disfluencies in two sentences, $t(6)$ = -1.88, p = .11 and $t(6)$ = -0.93, p = .39. The difference between the post-test sample and the NS reference (1710.00) is statistically significant, $t(6)$ = -4.37, p = .005. Moreover, statistics show that the difference between the post-test sample and the faster NS reference (1570.00) is near significant, $t(6)$ = -2.05, p = .09.

There is a statistically significant difference between the pre-test (M = 1461.43, SD = 271.93) and post-test (M = 1755.71, SD = 306.04) duration data for speaker **F8** towards the slower production, $t(6) = -5.51$, $p = .002$. The results for the pre-test sample indicate that the difference from the NS reference (1710.00) is near significant (due to rounding down), $t(6) = -2.42$, $p = .05$. However, when compared to the faster NS reference (1570.00), the result is nonsignificant, $t(6) = -1.06$, $p = .33$. On the other hand, there is a statistically nonsignificant difference between the post-test sample and the NS reference (1710.00), $t(6) = 0.40$, $p = .71$.

The difference between the pre-test (M = 1345.00, SD = 263.00) and post-test (M = 1332.50, SD = 314.15) duration data for speaker **F9** is statistically nonsignificant, $t(3) = 0.28$, $p = .80$. The results for the pre-test data indicate that the difference from the NS reference (1710.00) is statistically near significant, $t(3) = -2.78$, $p = .07$. This, however, does not hold when compared to the faster NS reference (1570.00), $t(6) = -1,71$, $p = .19$. For the post-test data, there is a statistically nonsignificant difference from the NS reference value (1710.00), $t(3) = -2.40$, $p = .10$. It is important to keep in mind that there were only four sentence analysed for this speaker and the SD in those sentences was rather large. The result, therefore, may not be conclusive.

For speaker **F10**, there was a statistically significant difference between the pre-test (M = 1630.00, SD = 265.20) and post-test (M = 1891.43, SD = 358.91) duration data, $t(6) = -2.71$, $p = .04$. However, there was a nonsignificant difference from the NS reference (1710.00) for both the pre-test and post-test samples, $t(6) = -0.80$, $p = .46$ and $t(6) = 1.34$, $p = .23$ respectively.

For speaker **F11**, there was also a statistically significant difference between the pre-test (M = 1268.57, SD = 241.48) and post-test (M = 1488.57, SD = 266.92) duration data, $t(6) = -4.25$, $p = .01$. But the difference between the production of this speaker and the NS reference (1710.00) was significant in the pre-test sample and near significant in the post-test sample, $t(6) = -4.84$, $p = .003$ and $t(6) = -2.19$, $p = .07$. The pre-test sample was also significantly different from the faster NS reference (1570.00), $t(6) = -3.30$, $p = .02$. In the post-test sample, the result was statistically nonsignificant, $t(6) = -0.81$, $p = .45$.

There was a statistically nonsignificant difference between the pre-test (M = 1694.29, SD = 332.61) and post-test (M = 1707.14, SD = 291.25) duration data for speaker **M1**, $t(6) = -0.10$, $p = .93$. In both the pre-test and post-test sample, the difference from the NS reference (1710.00) was nonsignificant, $t(6) = -0.13$, $p = .90$ and $t(6) = -0.03$, $p = .98$ respectively.

For speaker **M2**, there was a statistically nonsignificant difference between the pre-test (M = 1157.50, SD = 266.13) and post-test (M = 1325.00, SD = 119.02) data, $t(3) = -1.81$, $p = .17$. In both cases (pre-test and post-test), however, the production was significantly different from the NS reference (1710.00), $t(3) = -4.15$, $p = .03$ and $t(3) = -6.47$, $p = .01$ respectively. Furthermore, there was near significant result for the comparison of this speaker's pre-test production to the faster NS reference (1570.00) and a significant result for the post-test sample, $t(3) = -3.10$, $p = .05$ and $t(3) = -4.12$, $p = .03$ respectively.

The difference between the pre-test (M = 1574.29, SD = 699.26) and post-test (M = 1378.57, SD = 218.44) duration data for speaker **M3** was statistically nonsignificant which was affected by a severe disfluency in the pre-test sample, $t(6) = 0.77$, $p = .47$. Due to the disfluency in the pre-test sample that cause the large SD, the difference from the NS reference (1710.00) was statistically nonsignificant, $t(6) = -0.51$, $p = .63$. The difference between the post-test sample and the NS reference was significant, $t(6) = -4.01$, $p = .01$. There was also a near significant difference between the post-test production of this speaker and the faster NS reference (1570.00), $t(6) = -2.32$, $p = .06$.

For speaker **M4**, the difference between the pre-test (M = 1210.00, SD = 296.93) and post-test (M = 1358.57, SD = 331.03) is s statistically near significant, $t(6) = -2.27$, $p = .06$. The result for the comparison to the NS reference (1710.00) is significant in both the pre-test and post-test sample, $t(6) = -4.46$, $p = .004$ and $t(6) = -2.81$, $p = .03$ respectively. However, when compared to the faster NS reference (1570.00), the difference is statistically significant in the pre-test sample and nonsignificant in the post-test sample, $t(6) = -3.21$, $p = .02$ and $t(6) = -1.69$, $p = .14$ respectively.

For speaker **M5**, the difference between the pre-test (M = 1302.86, SD = 242.88) and post-test (M = 1190.00, SD = 253.71) duration data was near significant and the change was towards a faster production in the second recording, $t(6) = 1.98$, $p = .09$. When compared to the NS reference (1710.00), the difference in both pre-test and post-test sample was statistically significant, $t(6) = -4.44$, $p = .004$ and $t(6) = -5.42$, $p = .002$ respectively. The same result was in the comparison to the faster NS reference (1570.00) for both the pre-test and post-test same, $t(6) = -2.91$, $p = .03$ and $t(6) = -3.96$, $p = .01$ respectively.

6. Discussion

The purpose of this thesis was to analyse the effect of pronunciation instruction on the fluency of advanced non-native learners of English. The analysis focussed on pitch range, which was analysed on two independent dimension, *pitch span* and *pitch level*, and the second part of the analysis focussed on the duration of utterances.

The results of the pitch span analysis of the pre-test and post-test data for all 16 speakers taken together as a group indicate that the pronunciation instruction did not significantly affect their reading fluency and that in both recordings, the pitch span was significantly narrower than that of the four NSs. From this we can conclude that the first hypothesis, in which it is assumed that the pitch range in the pre-test data will be narrower than that of the NSs, has been accepted. With the second hypothesis, which assumes that the pronunciation instruction will affect the NNSs production and move it towards the NS model, it is not as straightforward. When we consider only the mean value of all 16 non-native speakers and compare their pre-test and post-test data, the second hypothesis needs to be rejected. However, if we consider the results for individual speakers, which indicate that the pronunciation instruction did affect the production of certain speakers and moved them towards the NS model, the second hypothesis can be at least partially accepted.

Considering the change in pitch span of individual speakers, we can conclude that the pronunciation instruction had a *significant effect* only on five speakers (F1, F2, F4, F7, and M1) and, except for speaker F1, all of them changed towards the wider pitch span in their second recording. The effect of the pronunciation instruction on the remaining 11 speakers was nonsignificant.

In terms of the change towards either narrower or wider pitch span, the effect of the pronunciation instruction was inconclusive because 8 speakers changed their production towards the *wider* pitch span (F2, F3, F4, F5, F6, F7, F8, M1) and 8 speakers towards the *narrower* pitch span (F1, F9, F10, F11, M2, M3, M4, M5). It is, however, important to keep in mind that for majority of the speakers the change was very small, as can be seen in figure 5.2 in the pitch span results section.

When the pre-test and post-test data were compared to the NS reference, three kinds of results emerged. In the first kind, there were nine of the sixteen speakers (F1, F2, F4, F6, F10, F11, M2, M4, M5) who were statistically *different* from the NS reference in both the pre-test and post-test sample. In figure 5.2, it is visible how speaker F2 changed significantly from rather narrow pitch span to a considerably wider one, it is also visible that the post-

test span is significantly wider than that of the NS reference and can be understood as an example of unnatural production. This was given mostly by the position of the measurement points for the post-accentual valley and sentence-final low which were positioned together in six out of seven sentences because the accent peak was located on the last word. The position of the individual measurement points in all seven sentences can be seen in figure 17 in the appendix. And as can be seen from this figure, the accented words in some of the sentences were produced with unnaturally higher frequency than the rest of the sentence.

To get back to the different kinds of pitch span result, the second kind concerned only three speakers, F8, F9, and M3. The results indicate that these speakers were statistically *comparable* to the NS reference in both the pre-test and post-test recording. For none of these speakers, the change between the two recordings was statistically significant. However, only speaker F8 moved towards the wider pitch span in the post-test recording. It is important to keep in mind that the change was very small for all three speakers (only around 0.3 St).

The last kind of the result, which was also the most interesting one, concerned four speakers (F3, F5, F7, and M1). These speakers were significantly *different* from the NS reference in the pre-test sample but in the post-test sample, they were statistically *comparable* to the NS reference. The effect of the pronunciation instruction was statistically nonsignificant in the case of speaker F3 and F5, but significant for speaker F7 and M1. It is in the context of these four speakers that the second hypothesis can be partially accepted.

The second dimension of pitch range, pitch level, was analysed separately for male and female speakers because of the use of linear scale expressed in Hertz. As was expected from the use of this scale, the pitch level values were different for male and female speakers. The mean value of male speakers in the pre-test and post-test was 112.68 Hz and 115.88 Hz, while the female speakers reached the mean value of 169.47 Hz and 188.36 Hz.

The result for the change in pitch level of the male speakers was statistically nonsignificant. This is mainly given by the fact that there were only five male speakers whose speech production has been analysed and, therefore, the results are not conclusive. Despite that, three of the five speakers (M1, M2, M3) increased their pitch level in the second recording, but only speaker M3 had a statistically significant result. The remaining two speakers (M4 and M5) had a lower value for the pitch level in the post-instruction recording.

The result for the change in pitch level of the female speakers was statistically significant. However, only two female speakers (F1, F4) had statistically significant or near

significant difference between the pre-test and post-test pitch level values, for the remaining nine speakers the result was statistically nonsignificant. As well as for the male speaker, the general tendency among the female speakers was to move towards the higher pitch level. Only three female speakers (F2, F5, and F6) moved towards the lower value in the second recording.

In the case of the pitch level analysis, we should keep in mind that this dimension is very speaker specific and, especially when it focuses on the sentence-final low not the overall f0 of an utterance, cannot be properly compared to a mean value of more than one native speakers. It is more useful, in this case, to look at the change within individual speakers and compare each speaker to the rest of the NNSs and the four NSs to see whether there were serious movements along the scale from low to high voice quality. This can be seen nicely in figure 5.5 in the pitch level results. There are two separate graphs, one for the pre-test data and one for the data collected after the instruction. In this illustration we can see how the individual speakers changed their position on the scale in relation to the other speakers.

For speaker F5, for example, the values for pitch level were relatively the same, there was only a minor difference of 1.1 Hz, but the position on the scale changed quite a lot in relation to the change of the other speakers. In comparison, the post-test value for the pitch level of speaker M2 was higher but he remained in exactly the same position on the scale. There were three other speakers (M4, M5, F6) whose position shifted considerably towards the lower positions on the scale, which correlated with the change in the values of the pitch level towards the lower values. On the other hand, for speakers F1, F8, F9 and F11, the change was towards a considerably higher position on the scale, which was also correlated by the change towards the higher frequency. An interesting case is speaker M5, whose pre-test position on the scale was above five female non-native speakers and one female native speaker. In the post-test scale, this speaker moved down considerably but still remained above one female speaker.

Furthermore, what is interesting to note here in the discussion of pitch level is that the standard deviation from the mean pitch level value varied greatly between the individual speakers. While for some speakers (F4, F5, F7) the standard deviation from the mean value calculated from the seven sentences was below 20 Hz in both recordings, for some others (F10, F11) the standard deviation value rose to approximately 80 Hz.

The conclusion that can be drawn from the analysis of the pitch level is that even though the general tendency of the 16 NNSs was to increase their pitch level in the post-

instruction recording, the results for individual speakers were mostly statistically nonsignificant. Only three speakers (F1, F4, and M3) had statistically significant or near significant result for the change between the pre-test and post-test pitch level values. Therefore, the hypothesis that the pitch level will not be affected by the pronunciation instruction has been accepted.

For the duration of the utterances, the hypothesis was that the pronunciation instruction will affect the duration and the move would be towards the slower rate resembling the NS model. Based on the group results, and in most cases also the individual speaker results, the conclusion is that the hypothesis can be accepted.

The result for overall tendency of the 16 non-native speakers to change their speech rate towards the slower one was statistically significant. There were only four speakers (F1, F4, F9, and M5), whose utterance duration was faster in the post-instruction recording. The results were nonsignificant for three of these speakers, only speaker M5 had a statistically near significant result for the change between pre-test and post-test utterance duration. Seven of the remaining twelve speakers had a significant result for the change between the utterance duration in the pre-test and post-test recording and in all of these cases, the move was towards the slower speech rate.

In terms of the comparison to the NS reference value, it is important to mention that there were considerable difference in utterance duration between the four native speakers. Because of this, the comparison to the NS reference was dual. In the first part, I compared the NNS duration value to the mean NS reference calculated from the four speakers. This reference was 1710.00 milliseconds. When, however, there was a significant difference between the non-native speaker's value and the NS reference, I ran another statistical test comparing it to the mean NS reference calculated from the two faster durations of speakers NSf2 and NSm2. The mean value for this comparison was 1570.00 milliseconds.

As for the pitch span results, the results for the utterance duration can be classified into four categories. In the first category, the results were *comparable* in the pre-test sample and *different* in the post-test sample as compared to the slower NS reference (1710.00 ms). There were three speakers in this category. For speaker F1, this result was given by the move towards the faster speech rate in the second recording. For speakers F7 and M3, however, the comparable result in the first recording was given by the disfluencies in two of the seven sentences, which caused the considerable increase in the overall utterance duration. The second recording of these two speakers was without the disfluencies and, therefore, the result was significantly different from the NS reference. When compared to

the faster NS reference (1570.00 ms), speaker F1 was *comparable* in both recordings. The results of this comparison to the faster NS reference were interesting for the other two speakers (F7 and M3), who remained *different* in the second recording.

The second category contains only two speakers (F10 and M1), whose utterance duration was *comparable* in both recordings. For speaker F10, the utterance duration was comparable to the NS reference (the slower one) despite the significant change between the pre-test and post-test sample. For speaker M1, the comparable result in the pre-test might have been affected by the disfluency in one of the seven sentences.

In the third category, in which the pre-test duration was *different* from the NS reference (the slower one) and *comparable* in the second recording, there are four speakers (F3, F6, F8 and F9). For speakers F8 and F9, the further comparison to the faster NS reference showed that both were *comparable* in the pre-test recording, which means that they were essentially comparable to the native speakers in both recordings. For speaker F3 and F6, on the other hand, the further comparison to the faster NS reference showed that they were significantly *different* in the pre-test recording than the native speakers. For these two speakers, the pronunciation instruction had a significant effect on the utterance duration and its move towards the NS reference.

The last category included speakers who were statistically *different* from the NS reference (the slower one) in both the pre-test and post-test recording. In this category, there are seven speakers (F2, F4, F5, F11, M2, M4, and M5). For speaker F2, the further comparison to the faster NS reference showed that the utterance duration of this speaker in both the pre-test and post-test recording was *comparable* to native speakers. For speaker F5, F11 and M4, the further comparison to the faster NS reference indicated that they were significantly *different* from the native speakers only in the pre-test recording but were *comparable* in the post-test recording. In the context of these speakers we can also talk about a significant effect of the pronunciation instruction on the utterance duration and its overall move towards a native-like pattern. The remaining three speakers (F4, M2 and M5) remained *different* from the NS reference even in the further comparison to the faster NS value.

To summarise the utterance duration results, the pronunciation instruction did not have any effect on the speech rate of three speakers (F4, M2, M5), whose results were significantly different from the NS reference in both recordings and the change between their pre-test and post-test sample was nonsignificant (near significant for speaker M5). The results for two speakers (F7 and M3) are inconclusive because they indicate that the

pronunciation instruction had the opposite effect on the speech rate (they moved from comparable to different duration). This was, however, given by the disfluencies in the pre-test recording in which they were comparable, thus, the results are negatively influenced by it and cannot be taken as conclusive. The instruction had a positive effect on four speakers (F2, F8, F10, and M1), whose utterance duration was comparable to the NS reference in both recordings. Despite that, all four of them moved towards a slower speech rate in the second recording. Interesting result emerged for speakers F1 and F9, whose speech rate was comparable to the NS reference in both recordings despite their move towards a faster speech rate (although the result was nonsignificant). The most positive effect of the pronunciation instruction towards a slower, native-like speech rate was observed for five speakers (F3, F5, F6, F11, and M4), whose utterance duration was different from the NS in the first recording but comparable in the second one. So overall, the pronunciation instruction did have a positive effect on the speech rate of the non-native speakers, whose utterance duration moved towards a native-like pattern, thus the fourth hypothesis can be accepted.

7. Conclusion

As the review of literature indicates, fluency is a broad and widely discussed topic. The position of this thesis within the existing research is in the area of reading fluency that is, unfortunately, primarily occupied by research focussed on the reading fluency of native speakers, specifically on native children and their improvement of reading fluency. There is an insufficient amount of cross-linguistic studies focussing on the reading fluency of L2 adult speakers and how it can be improved with pronunciation instruction. This lack of research in this specific study area is the reason why I chose to analyse the effects of explicit classroom-based pronunciation instruction on the oral reading fluency of advanced, adult Czech learners of English. The results of the analysis bring new ideas to the existing research within the area of reading fluency.

In the review of literature, I focussed on the various definitions available in the previous research for both the concept of fluency and reading fluency. I also addressed in detail the concept of prosody and its role in reading fluency. The individual features of prosodic reading has been described as well as the essential differences between the spontaneous speech prosody and reading prosody. Furthermore, the role of punctuation in relation to the reading prosody has been discussed and the differentiation between the syntactic and emphatic prosody has been made. In the closing subsection of the first chapter of the literature review, I focussed on the difference between English and Czech intonation pattern which was based on the results of Volín et al. (2015). The second chapter of the literature review was dedicated to the pronunciation instruction and here I analysed the relationship between pronunciation instruction and the various approaches to L2 teaching. Moreover, I addressed the evidence for the effectiveness of pronunciation instruction in the context of L2 learning and described what the effective instruction should look like. The role of the teacher and the corrective feedback has been addressed and the various methods of reading fluency assessment has been described.

In the method of the thesis, I described the participants, the process of data collection, as well as the instruction process. The detailed analysis of the pre-instruction and post-instruction recordings based on the model in Patterson (2000) was followed by an extensive statistical analysis of the results and a subsequent discussion of the results.

There were four hypotheses at the beginning of this research focussing on two of the six acoustic features of fluent reading, summarised by Dowhower (1991) in the introduction to this thesis. The two acoustic features addressed in the method of this thesis were pitch

range, which was further differentiated into two independent dimension, namely pitch span and pitch level, and the second feature was utterance duration. Under the first hypothesis, it was assumed that the pitch span of the native speakers in the pre-instruction recording will be narrower than that of the four reference native speakers. The results of this thesis indicate that this assumption was correct, thus, the hypothesis has been accepted. The second hypothesis asserted that the pronunciation instruction will have a positive effect on the pitch span of the non-native speakers and that in the post-instruction recording the pitch span would be wider than in the pre-instruction recording, while also resembling the native-like pitch span. The result of the analysis, however, showed that the effect of the pronunciation instruction was nonsignificant and that the majority of the non-native speakers remained significantly different than the reference native speakers. As a result of that, the second hypothesis has been refuted by the analysis. Under the third hypothesis, it was assumed that the pitch level is a speaker specific quality that is usually stable in the context of adult speakers and, therefore, will not be affected by the pronunciation instruction. The results indeed indicate that this is the case and the hypothesis has been accepted. For the fourth hypothesis, it was assumed that the speakers will successfully employ slower speech rate in the oral reading to children, resembling the tendencies observed in the native speakers, despite the fact that the pronunciation instruction focused on the production of faster speech. And as the results of the analysis indicate, the non-native speakers produced significantly slower speech in the post-instruction recording, which in most of the cases resembled the speech produced by the reference native speakers. Thus, the fourth hypothesis has also been accepted.

One of the most serious drawbacks of this analysis is the disproportion of female and male subjects. For further research, I would recommend choosing a more proportional subject sample so that differentiation between male and female results could also be made. The second drawback of this thesis, although not as serious as the previous one, is the choice of the material for analysis. As the choice of the specific sentences for the analysis was done after the data collection, the position of the measurement points in the chosen sentences was not ideal. My recommendation for further research, therefore, would be to prepare the material in such a way as to prevent the position of the non-initial accent being located on the last word of the utterance. This would prevent the marking of the points for the post-accentual valley and sentence-final low on the same position. Last, but not least, I recommend analysing the pitch level not as a mean value of all sentence-final lows, but as

a mean f0 of the whole utterance. This would better indicate the differences among pitch ranges of individual speakers and allow the comparison to the native speakers.

Overall, this thesis is a valuable contribution to the existing research within the area of reading fluency in that it addresses the reading fluency of adult non-native speakers, analyses the effect of explicit pronunciation instruction on the enhancement of reading fluency, and focusses on a language combination that is not frequently addressed. The results bring new insights to the existing research and suggest ideas for further research.

8. References

Abadiano, Helen R., and Jesse Turner. 2005. “Reading Fluency: The Road to Developing Efficient and Effective Readers.” *The New England Reading Association Journal* 41 (1): 50-56.doi:https://webcapp.ccsu.edu/u/faculty/TurnerJ/Reading%20Fluency%20The%20Road%20to%20Developing.pdf.

Atli, Isil, and Ayfer Su Bergil. 2012. “The effect of pronunciation instruction on students’ overall speaking skills.” *Procedia – Social and Behavioral Sciences* 46: 3665-3671. doi: 10.1016/j.sbspro.2012.06.124.

Bäckström, Tom. 2020. “Fundamental Frequency (F0).” Last modified August 14. https://wiki.aalto.fi/pages/viewpage.action?pageId=149890776.

Bailly, G. and Gouvernayre, C. 2012. “Pauses and respiratory markers of the structure of book reading.” In *Thirteenth Annual Conference of the International Speech Communication Association, Portland, September 9-13, 2012*. https://hal.archives-ouvertes.fr/hal-00741667.

Beinum, F.J.K.v. 1991. “Spectro-temporal reduction and expansion in spontaneous speech and read text: Focus words versus non-focus words.” *Phonetics and Phonology of Speaking Styles*.

Binder, Katherine S., Elizabeth Tighe, Yue Jiang, Katharine Kaftanski, Cynthia Qi, and Scott P. Ardoin. 2013. “Reading Expressively and Understanding Thoroughly: An Examination of Prosody in Adults with Low Literacy Skills.” *Reading and Writing* 26 (5): 665-680. doi: 10.1007/s11145-012-9382-7.

Boersma, P. & Weenink, D. 2016. *Praat: Doing Phonetics by Computer* (version 6.0.20). Retrieved from www.praat.org.

Chafe, Wallace. 1988. “Punctuation and the Prosody of Written Language.” *Written Communication* 5 (4): 395-426. doi:10.1177/0741088388005004001.

Conlen, Madeline M. 2016. "A Linguistic Comparison: Stress-timed and syllable-timed languages and their impact on second language acquisition." *Honors College Theses.* Wayne State University.

Cowie, R., E. Douglas-Cowie, and A. Wichmann. 2002. “Prosodic Characteristics of Skilled Reading: Fluency and Expressiveness in 8-10-year-old Readers.” *Language and Speech* 45 (1): 47-82. doi: 10.1177/00238309020450010301.

Derwing, Tracey M., and Marian J. Rossiter. 2003. “The Effects of Pronunciation Instruction on the Accuracy, Fluency, and Complexity of L2 Accented Speech.” *Applied Language Learning* 13 (1): 1-17.https://www.researchgate.net/profile/Tracey_Derwing/publication/234570703_The_Effects_of_Pronunciation_Instruction_on_the_Accuracy_Fluency_and_Complexity_of_L2_

Accented_Speech/links/0f31753124e4f3a907000000/The-Effects-of-Pronunciation-Instruction-on-the-Accuracy-Fluency-and-Complexity-of-L2-Accented-Speech.pdf?origin=publication_detail.

Derwing, Tracey M., and Murray J. Munro. 2005. "Second Language Accent and Pronunciation Teaching: A Research-Based Approach." *TESOL Quarterly* 39 (3): 379-397. doi: https://www.jstor.org/stable/3588486?seq=1#metadata_info_tab_contents.

Derwing, Tracey M., Murray J. Munro, and Grace Wiebe. 1997. "Pronunciation instruction for fossilized learners: Can it help?." *Applied Language Learning* 8(2): 217–35.

Derwing, Tracey M., Murray J. Munro, and Grace Wiebe. 1998. "Evidence in Favor of Broad Framework for Pronunciation Instruction." *Language Learning* 48 (3): 393-410. 10.1111/0023-8333.00047.

Dowhower, S. L. 1991. "Speaking of prosody: Fluency's unattended bedfellow." *Theory in Practice* 30: 158–164.

Erekson, James A. 2010. "Prosody and Interpretation." *Reading Horizons* 50 (2): 80-98. https://www.researchgate.net/publication/254733314_Prosody_and_Interpretation.

Fillmore, C. 1979. "On fluency." In *Individual differences in language ability and language behaviour,* edited by C. Fillmore, D. Kempler, and W. S.-Y, Wang, 85-101. New York: Academic Press.

Flege, James E., Murray J. Munro, and Ian R. A. MacKay. 1995. "Factors affecting strength of perceived foreign accent in a second language." *Journal of the Acoustical Society of America* 97 (5): 3125–3134. https://doi.org/10.1121/1.413041.

Galante, Angelica, and Ron I. Thomson. 2016. "The Effectiveness of Drama as an Instructional Approach for the Development of Second Language Oral Fluency, Comprehensibility, and Accentedness." *TESOL Quarterly* 51 (1): 115-142. doi: 10.1002/tesq.290.

Ghorbani, Mohammad Reza, Malihe Neissari, and Hamid Reza Kargozari. 2016. "The Effect of Explicit Pronunciation Instruction on Undergraduate EFL Learners' Vowel Perception." *Language and Literacy* 18 (1): 57-70. doi: 10.20360/G2XW2K.

Godde, Erika, Marie-Line Bosse, and Gérard Bailly. 2019. "A review of reading prosody acquisition and development." *Reading and Writing* 33 (2): 1-19. doi: 10.1007/s11145-019-09968-1.

Goldman, J.-P., A. Auchlin, and A.C. Simon. 2009. "Discrimination de styles de parole par analyse prosodique semi-automatique." *Actes d'IDP,* 207-221.

Goldman, S.R., P.M. Meyerson, and N. Coté. 2006. "Poetry as a mnemonic prompt in children's stories." *Reading Psychology,* 27(4): 345-376. doi: 10.1080/02702710600846894.

Gordon, Joshua, and Isabelle Darcy. 2016. “The development of comprehensible speech in L2 learners.” *Journal of Second Language Pronunciation* 2 (1): 56-92. doi: 10.1007/s11145-019-09968-1.

Gósy, Mária, and Valéria Krepsz. 2018. “Phrase-final Lengthening of Phonemically Short and Long Vowels in Hungarian Spontaneous Speech Across Ages.” In *Challenges in Analysis and Processing of Spontaneous Speech,* 99-126. Budapest: MTA Nyelvtudományi Intézet. doi: 10.18135/CAPSS.99.

Gürbüz, Nurdan. 2017. “Understanding Fluency and Disfluency in Non-native Speakers’ Conversational English.” *Educational Sciences: Theory & Practice* 17 (6): 1853-1874. doi: 10.12738/estp.2017.6.0071.

Hirschberg, Julia. 2000. “A corpus-based approach to the study of speaking style.” *Prosody: Theory and experiment.* 335-350. Dordrecht: Springer.

Hudson, R.F., H.B. Lane, and P.C. Pullen. 2005. “Reading fluency assessment and instruction: What, why, and how?” *The Reading Teacher*, 58: 702-714.

Jam, Bashir, and Maryam Adibpour. 2014. “Intuitive-imitative Approach versus Analytic-linguistic Approach toward Teaching /V/, /F/, and /w/ to Iranian Students.” *Procedia – Social and Behavioral Sciences* 98: 757-763. doi: 10.1016/j.sbspro.2014.03.478.

Karimi, Lotfollah, and Nahid Esfandiari. 2016. “The Effect of Recast vs. Explicit Corrective Feedback on Iranian EFL Learners’ Stress Patterns Learning.” *Theory and Practice in Language Studies* 6 (6): 1166-1174. doi: http://dx.doi.org/10.17507/tpls.0606.05.

Kerr, Judith. 1968. *The Tiger Who Came to Tea*. London: HarperCollins.

Kissling, Elizabeth M. 2014. “What Predicts the Effectiveness of Foreign Language Pronunciation Instruction?:Investigating the Role of Perception and Other Individual Differences.” *Canadian Modern Language Review* 70(4): 535-558. doi: 10.3138/cmlr.2161.

Kuhn, Melanie R., and Steven A. Stahl. 2003. “Fluency: A Review of Developmental and Remedial Practices.” *Journal of Educational Psychology* 99 (1): 3-21. doi: 10.1037/0022-0663.95.1.3.

Kuhn, Melanie R., Paula J. Schwanenflugel, and Elizabeth B. Meisinger. 2010. “Aligning Theory and Assessment of Reading Fluency: Automaticity, Prosody, and Definitions of Fluency.” *Reading Research Quarterly* 45 (2): 232-253. dx.doi.org/10.1598/RRQ.45.2.4.

Ladd, Robert D. 1996. *Intonational Phonology.* Cambridge: Cambridge University Press.

Lems, Kristin. 2003. “Adult ESL Oral Reading Fluency and Silent Reading Comprehension.” PhD diss., National Louis University.

Mennen, Ineke. 2007. “Phonological and phonetic influences in non-native intonation.” In *Non-native Prosody,* edited by Jürgen Trouvain and Ulrike Gut, 53-76. Berlin: Walter de Gruyter.

National Reading Panel. 2000. “Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports from the subgroups.” Washington, DC: National Institute of Child Health and Human Development.

Patterson, David. 2000. “A linguistic approach to pitch range modelling.” Unpublished PhD diss., University of Edinburgh.

Pinnell, G. S., J.J. Pikulski, K.K. Wixson, J.R. Campbell, P.B. Gough, and A.S. Beatty. 1995. *Listening to children read aloud: Data from NAEP´s integrated reading performance record (IRPR) at grade 4.* Washington, US: U.S. Department of Education.

Piske, Thorsten, Ian R. A. MacKay, and James E. Flege. 2001. “Factors affecting degree of foreign accent in an L2: a review.” *Journal of Phonetics* 29: 191-215. doi:10.006/jpho.2001.0134.

Pival, J. 1968. “Stress, pitch and juncture: tools in the diagnosis and treatment of reading ills.” *Elementary English* 45: 458-467.

Pourhossein Gilakjani, Abbas. 2017. “English Pronunciation Instruction: Views and Recommendations.” *Journal of Language Teaching and Research* 8 (6): 1249-1255. doi: http://dx.doi.org/10.17507/jltr.0806.30.

Rasinski, T., A. Rikli, and S. Johnston. 2009. “Reading fluency: More than automaticity? more than a concern for the primary grades?” *Literacy Research and Instruction,* 48(4): 350-361.

Rasinski, T.V. 2004. “Assessing Reading Fluency.” *Pacific Resources for Education and Learning (PREL).*

Rasinski, Timothy. 2005. “The Role of the Teacher in Effective Fluency Instruction.” *New England Reading Association Journal* 41 (1): 9-12. https://search.proquest.com/openview/18825b058fa8ed8a629cdf4d8402a68f/1?pq-origsite=gscholar&cbl=34991.

Rasinski, Timothy. 2006. “Reading fluency instruction: Moving beyond accuracy, automaticity, and prosody.” *International Reading Association* 59 (7): 704-706. doi: 10.1598/RT.59.7.10.

Rossiter, Marian J., Tracey M. Derwing, Linda G. Manimtim, and Ron I. Thomson. 2010. “Oral Fluency: The Neglected Component in the Communicative Language Classroom.” *The Canadian Modern Language Review/ La revue canadienne des langues vivantes* 66 (4): 583-606. doi:10.3138/cmlr.66.4.583.

Schmidt, R. 1992. “Psychological mechanisms underlying second language fluency.” *Studies in Second Language Acquisition,* 14: 357-385. doi: 10.1017/S0272263100011189.

Schwanenflugel, Paula J, and Rebekah George Benjamin. 2017. “Lexical Prosody as an Aspect of Oral Reading Fluency.” *Reading and Writing: An Interdisciplinary Journal* 30 (1): 1-45. doi: 10.1007/s11145-016-9667-3.

Schwanenflugel, Paula J., Anne Marie Hamilton, Joseph M. Wisenbaker, Melanie R. Kuhn, and Steven A. Stahl. 2004. “Becoming a Fluent Reader: Reading Skill and Prosodic Features in the Oral Reading of Young Readers.” *Journal of Educational Psychology* 96 (1): 119-129. doi:10.1037/0022-0663.96.1.119.

Segalowitz, Norman. 2010. *Cognitive Bases of Second Language Fluency.* New York: Routledge.

Shen, Hwei-Jiun. 2003. “The Role of Explicit Instruction in ESL/EFL Reading.” *Foreign Language Annals* 36 (3): 424-433. https://doi.org/10.1111/j.1944-9720.2003.tb02124.x.

Šimáčková, Š. 2020. ′Seminar 10.′ *AFO4/AF11: Topics in Phonetics*. Palacký University Olomouc. Unpublished.

Šimáčková, Š. 2020. ′Seminar 11.′ *AFO4/AF11: Topics in Phonetics*. Palacký University Olomouc. Unpublished.

Šimáčková, Š. 2020. ′Seminar 2.′ *AFO4/AF11: Topics in Phonetics*. Palacký University Olomouc. Unpublished.

Šimáčková, Š. 2020. ′Seminar 3.′ *AFO4/AF11: Topics in Phonetics*. Palacký University Olomouc. Unpublished.

Šimáčková, Š. 2020. ′Seminar 4.′ *AFO4/AF11: Topics in Phonetics*. Palacký University Olomouc. Unpublished.

Šimáčková, Š. 2020. ′Seminar 5.′ *AFO4/AF11: Topics in Phonetics*. Palacký University Olomouc. Unpublished.

Šimáčková, Š. 2020. ′Seminar 6.′ *AFO4/AF11: Topics in Phonetics*. Palacký University Olomouc. Unpublished.

Šimáčková, Š. 2020. ′Seminar 7.′ *AFO4/AF11: Topics in Phonetics*. Palacký University Olomouc. Unpublished.

Šimáčková, Š. 2020. ′Seminar 8.′ *AFO4/AF11: Topics in Phonetics*. Palacký University Olomouc. Unpublished.

Skehan, P. and P. Foster. 1997. “Task type and task processing conditions as influences on foreign language performance.” *Language Teaching Research,* 1: 185-211.

Skehan, P. and P. Foster. 1999. “The influence of task structure and processing conditions on narrative retellings.” *Language Learning,* 49: 93-120. doi: 10.1111/1467-9922.00071.

Swain, Kristine D., Elizabeth M. Leader-Janssen, and Perry Conley. 2017. “Effects Of Repeated Reading And Listening Passage Preview.” *Reading Improvement* 50 (1): 12-18. doi: https://doi.org/10.1177/0022219415605194.

The Partnership for Reading. 2001. “Put Reading First, The Research Building Blocks for Teaching Children to Read (kindergarten through grade 3).” National Institute for Literacy. US Department of Education.

Thomson, Ron I. 2015. “Fluency.” In *The Handbook of Pronunciation,* edited by M. Reed and J. Levis, 209-226. Hoboken, New Jersey: Wiley.

Thomson, Ron I., and Tracey M. Derwing. 2015. “The Effectiveness of L2 Pronunciation Instruction: A Narrative Review.” *Applied Linguistics* 36 (3): 326-344. doi: 10.1093/applin/amu076.

TIBCO Software Inc. 2018. *Statistica.* V. 13.4.0.14. TIBCO Software Inc. PC.

Volín, Ján, Kristýna Poesová, and Lenka Weingartová. 2015. “Speech melody properties in English, Czech and Czech English: Reference and interference.” *Research in Language* 13(1): 107–123.

Young, Chase, and Timothy Rasinski. 2009. “Implementing Readers Theatre as an Approach to Classroom Fluency Instruction.” *The Reading Teacher* 63 (1): 4-13. doi: 10.1598/RT.63.1.1.

Zutell, J. and T.V. Rasinski. 1991. “Training teachers to attend to their students’ oral reading fluency.” *Theory into Practice,* 30(3): 211-217.

9. Appendix

9.1. Instruction Process

Lesson 1: Introduction to the course

During the first lesson, the students were introduced to the course and they were informed about the procedure and credit requirements. Each of the students was asked to introduce herself/himself and state the reason why they chose the course. In the meantime, each of the students was asked to silently read the short children's story *The Tiger Who Came to Tea* and the Aesop's fable *The North Wind and the Sun* for a few times and then each of them read the text out loud while being recorded in the Recording Studio at the University. They were informed that their oral reading of the children's story will be used for the analysis of reading fluency in this thesis.

Lesson 2: Pronunciation and stress

The second lesson started with a warm-up exercise in which the students listened to a sound file in which the syllables *ba-da-ga-da* were repeated and they were asked to imitate the pronunciation they have heard, with the appropriate stress placement. The same process was repeated with the syllables *ga-da-ba-da*. After the warm-up exercise, the students analysed the recordings of their friends' reading of *The North Wind and the Sun* from the previous week, focusing on the consonantal mispronunciations, intonation pattern, and stress placement.

After an exercise focused on pronunciation of individual words (such as *lemonade, sandbag, waterproof, prize, price* etc.) and pronunciation as well as connected speech in short phrases (*A long robe, As big as me, The first phase, Just a snob,* etc.), the students recorded themselves while reading short sentences containing the homographs in figure 1 below. The purpose of this exercise was to show the students that in homographs, the stress may be placed on a different syllable depending on the meaning of a word or what part of speech it represents.

Figure 1: Sentences containing homographs (in bold) taken from the tutor's handout for lesson 2 (Šimáčková 2020).

> Those who **advocate** for doctor-assisted suicide say the terminally ill should not suffer. (b) The Devil's **Advocate** is a story about some characteristically American values. (c) She kept both the **duplicate** and the original. (d) We don't want our students to **duplicate** each other's work. (e) We just need an

estimate. (f) He is taking **graduate** classes at the university. (g) A good manager knows when to **delegate**…

The remainder of the lesson focused on the rhythm of stressed syllables in nursery rhymes, such as *Hickory, Dickory, Dock* and *Wee Willie Winkie,* and in sentences with a strong beat (see figure 2). The purpose of the exercises was to show the students that stressed syllables in English sentences follow a beat that can be imitated by the sound of a metronome. In other words, the stressed syllables occur at approximately the same time in accordance to the beat of the metronome.

Figure 2: Sentences with a strong beat. The underlined syllables carry the stress, marked by "X" and the remaining syllables are unstressed, marked by "•" taken from the tutor's handout for lesson 2 (Šimáčková 2020).

Rings on her fingers and bells on her toes,
Lift it up gently and take it outside.
X • • X • • X • • X

To see a fine lady upon a white horse.
The cab will be waiting in front of the house.
• X • • X • • X • • X

Lesson 3: Rhythm and voicing

The third lesson started with a warm-up exercise, similar to the one from the previous lesson, in which the students were supposed to listen to sounds (*pa-ta-ka-ta* and *ka-ta-pa-ta*) and repeat what they have heard with the stress placed on the appropriate syllable. Following the warm-up exercise, the instructor described the difference between Czech and English lingual articulatory settings.

The rhythm in strong beat sentences was again practiced, this time the students were asked to read the rhyme (*Jack be nimble, Jack be quick, Jack jump over the candlestick. Jill be nimble, jump it too, if Jack can do it, so can you.*) (Šimáčková 2020) while listening to an audio tape of the rhyme and they were encouraged to repeat it as many times as needed for them to be able to keep up the tempo of the audio. After practicing the strong beat sentences, the instruction moved to the rhythm in regular English sentences, such as *I saw him about two years ago. The orange album is lost. She left her food untouched. The*

warriors were armed with bows and arrows (Šimáčková 2020), in which the students were to find a regular meter.

In the following exercise, the students were given a verse from Edward Lear's poem *The owl and the pussy-cat* (see figure 3a). They were asked to find the appropriate rhythm of the poem and then compare their version of it with an audio recording. Again, they practiced reading the poem along with the recording. To demonstrate their understanding of the poem's rhythm, they were given a number of sentences (in figure 3b) taken from a spoken text, which had the same rhythm as the poem. The lines, however, were not in the correct order, so the task was to match the lines from figure 3b to the appropriate line of the poem.

Figure 3a: Edward Lear's poem *The owl and the pussy-cat* taken from the tutor's handout for lesson 3 (Šimáčková 2020).

The Owl and the Pussy-cat went to sea
In a beautiful pea green boat,
They took some honey, and plenty of money,
Wrapped up in a five pound note.

Figure 3b: Lines from a spoken text with the same rhythm as the poem taken from the tutor's handout for lesson 3 (Šimáčková 2020).

a. Now interest has jumped so high.
b. The new taxes are hardly fair.
d. The news in the paper is not so good.
c. I doubt that many will ever recover

In the second part of the lesson, the topic of voicing was introduced. In the first exercise, the students opened a recording of minimal pairs (*face-phase, leaf-leave, mouth N-mouth V, race-raise, safe-save*) in Praat (Boersma and Weenink 2016) and listened to the difference in pronunciation and tried to imitate what they have heard. After they understood the difference between vowel duration before voiced and voiceless consonants, they were given a number of phrases (see figure 4) with similar minimal pairs and they were asked to change the vowel duration in them, depending on the following consonant.

Figure 4: Phrases with varying vowel duration depending on the following consonant taken from the tutor's handout for lesson 3 (Šimáčková 2020).

A dizzying **height**. A long-distance **race**. A sudden **surge**. An extensive **search**.
At a rising **rate**. I asked my boss for a **raise**. In a loose **robe**. Insurance **fraud**.

Just a minor **bruise**. Lose the **rope**. My brother **Bruce**. The atmosphere is rather **fraught**. The police made a **raid**. You'd better **hide**.

The very last exercise of the lesson was focused on the glottal stops in the words *got, about, but, can't, client*, etc. Again, the recording was opened in Praat (Boersma and Weenink 2016), where the students saw the glottal stops in the waveform and spectrogram, and they tried to imitate the sounds.

Lesson 4: Vowel duration, silent letters, and rhythm in poems and prose

The fourth lesson started with a warm-up exercise, which consisted of repeating *va-la-wa-la-va* and *wa-la-va-la-wa* sounds, as well as the rhyme *Jack be nimble…* in unisons with the recordings. The following task was focused on the minimal pairs with differing vowel duration before voiced/voiceless consonants, which was covered in the previous lesson.

The new topic of this lesson was pronunciation of words with silent letters in phrases such as *Knights do not carry knives they carry swords, Swarming bees, Whistle a tune, One calf or two calves? Oh, crumbs!, a walkie-talkie owner,* and *a subtle hint*. After this exercise, the instruction focused again on the distribution of stresses and rhythm of poems. The poem of this lesson was *The naming of cats* by T.S. Elliot. The task was to read the poem as a group in unison with the recording, and after that, the students were divided into small groups of three and each group read one verse of it, paying attention to the appropriate rhythm and stress.

The following exercise focused on the rhythm in prose (see figure 5). The students were asked to read the text, divide it into smaller units and decide where the main breaks would be positioned. Then, they marked the stressed syllables or words, listened to a recording of the proper reading of the text and tried to imitate it.

Figure 5: Rhythm in prose taken from the tutor´s handout for lesson 4 (Šimáčková 2020).

Marianne had looked forward to her tenths birthday as something special. First, she was at last going into double figures, and second her father and mother had promised she could have riding lessons, which was what she wanted more than anything else in the world.

The last exercise of the lesson focused on the correct pronunciation of [ŋ] and [n]. The students were asked to record their pronunciation of short sentences containing this consonant sound (*Stop singing! - It's not a sin to sing. Did he run home? – No, he just rang home. Don't linger and put that wedding ring on her finger.*).

Lesson 5: Connected speech and stress

The fifth lesson was the first to be transferred to distance education form. The students had access to lesson materials (texts with instructions and audio files) and were asked to go through the exercises multiple times and then record themselves doing the exercises. The recordings were then sent to the instructor's email address, so she could listen to them and provide each student with feedback.

There were six exercises in total to be recorded. The first one was a warm-up exercise consisting of several tongue twisters: *Round and round the rugged rock the ragged rascal ran. * Nothing is worth thousands of deaths. * Twelve twins twirled twelve twigs. * World wide web. * Three short sword sheaths* (Šimáčková 2020). In the second exercise, the students listened to a poem called *The sound collector* written by Roger McGough and then, they were supposed to read it several times, paying attention to the rhythm. Their recording of the poem was supposed to sound as natural as possible while keeping up the rhythm of it.

The next exercise consisted of a number of short phrases, in which the students concentrated on emphasizing the stressed syllables and reducing the unstressed ones and read the phrases in a flowing manner as a connected speech. In the fourth exercise, the students listened to a sound file, in which short, everyday phrases were firstly read word-by-word and then as a connected speech. After listening to the file, they practiced saying the phrases.

The last exercise which the students recorded consisted of phrases and sentences, which can be seen in figure 6 below. The task was to read them loud and clear, always emphasising the last word, and to decide which other words could be stressed as well in each sentence.

Figure 6: Phrases and sentences with the emphasis on the last word taken from the tutor's handout for lesson 5 (Šimáčková 2020).

Cham**pagne**.
French cham**pagne**.
A bottle of French cham**pagne**.
Two bottles of French cham**pagne**.
Two fine old bottles of French cham**pagne**.
They took two fine old bottles of French cham**pagne**.
They had taken two fine old bottles of French cham**pagne**.
It seems they had taken two fine old bottles of French cham**pagne**.

Glasses.
A pair of **glass**es.
A pair of **read**ing glasses.
A pair of horn-rimmed **read**ing glasses.
He wore a pair of horn-rimmed **read**ing glasses.
He was wearing a pair of horn-rimmed **read**ing glasses.
When I saw him, he was wearing a pair of horn-rimmed **read**ing glasses.

The very last exercise of the lesson, although not recorded, was a dialogue from a play. The students were to listen to a recording of the dialogue and answer questions about it.

Lesson 6: Stress, reduction, and linking

As all the previous lessons, the sixth one also started with a warm-up exercise, in which the students repeated tricky sentences, such as *Elizabeth's birthday is on the third Thursday of this month., Wayne went to Wales to watch walruses*. The following exercise focused on the rhythm and stress in prose. The task was to decide which syllables should be stressed and then listen to a recording and simultaneously recite it, putting the emphasis on the same words. These two tasks were recorded, along with the following ones.

The third exercise of this lesson focused on the rhythm in a song "Bubbly toes" (see figure 7). The students were asked to first circle the stressed syllables and then read the sentences out loud with a specific rhythm (• • | X • • | X • • | X • • | X). After that, they recorded themselves reading it simultaneously with the recording.

Figure 7: Bubbly toes song taken from the tutor's handout for lesson 6 (Šimáčková 2020)

It's as common as something that nobody knows
that her beauty will follow wherever she goes
up the hill in the back of her house in the wood
she'll love me forever, …I know she could.

The next part of the lesson focused on stress, linking, and reduction. In one of the exercises the students were asked to indicate the linking in a text by writing the symbol ‿ . In the following exercise, they read some sentences in a smooth manner – linked the words together while stressing appropriate lexical words and reducing the functional words. The last exercise of the lesson was a dialogue from a play introduced in the previous week. This time, however, the students were supposed to read it with the same intonation as in the recordings, either alone (reading both parts of the dialogue) or with help of a friend/sibling.

Lesson 7: Linking and Intonation

As a warm-up exercise for this lesson, the students were asked to record themselves saying swear words of their choice and after that, they recorded themselves saying 10 jokes from the provided list.

As the main topic of the lesson was linking in connected speech, the students worked with audio files containing short phrases focused on assimilation (*can buy, can go, good boy, good girl, speed boat* etc.), unreleased stops (*bad luck, good news, bit tired, lot to do*), elisions (*can´t swim, diamond ring, just one, kept going* etc.), and phrases with same sounds at word boundary (*a more ice, a more rice, cancer research, everyone knows, human nature* etc.).

The second topic of the lesson was intonation. The students were asked to listen to a sample of speech and try to imitate the intonation they have heard. In the following exercise, they practiced intonation in dialogues (see figure 8). They were asked to read the dialogue with an intonation they deemed fitted for it.

Figure 8: Intonation in dialogues taken from the tutor's handout for lesson 7 (Šimáčková 2020).

A: I'm tired.
B: Tired?
A: Yes. Tired.
B: Why?
A: Why?
B: Yes. Why?
A: Because I am. That's why.
B: That's no reason.
A: No reason?
B: No.
A: Why do you get tired, then?
B: Me?
A: Yes, you.
B: Because I do.
A: There you are, then.

The lesson ended with an exercise focusing on the intonation in a dialogue taken from the play Greg and Ginny (introduced in the previous two weeks). The students were asked to listen to the recording of the dialogue and imitate (naturally) as closely as possible the intonation they have heard.

Lesson 8: Prosody to deliver extra meaning

The eighth lesson started with a warm-up exercise focused on fast speech. The students were given a transcript of the "Big Bang Theory" theme song and they were asked to try reciting it without the recording and then along with it. The goal was to recite the whole song in 42 seconds.

In the first exercise focused on irony, the students were given a number of sentences (see figure 9) which could be understood both with and without irony and they were instructed to vary their intonation to indicate the two meanings. In a follow-up exercise, they were given three words (*really, plagiarizing, Mary)* and they were asked to say them with varying intonation in different ways (as a sentence, with surprise, shock, irony…).

Figure 9: Sentences with two meanings taken from the tutor's handout for lesson 8 (Šimáčková 2020).

> **You won a beauty contest?** Interpretation (a) That is really interesting.
> Interpretation (b) Well, that clearly isn't possible.
>
> **Her father is an architect**. Interpretation (a) New information, I didn't know.
> Interpretation (b) She's so self-important.

In the next exercises, they listened to dialogues and they were asked to spot and underline in the transcript the words with additional meaning (ironic, sarcastic, or other), read the dialogues themselves and record themselves while doing so. The last exercise of the lesson was a summary of everything covered up to this point, which means that the students were asked to read the short sentences with appropriate stresses, reduce the unstressed syllables and words, link the words together to form a connected speech, and read at a fast rate.

Lesson 9: Rhythm and intonation

After a warm-up exercise focused on fast speech using the rhyme *Mary had a little lamb*, the exercises of this lesson focused on rhythm and intonation. Similarly to the previous lessons concentrating on stresses, linking, rhythm and intonation, the students were again given a text and a recording of the text and they were asked to first identify the strong syllables and then read the text along with the recording, paying attention to linking and the correct rhythm and intonation. The same task was repeated again later, focusing on a dialogue between two people. The students first listened to a recording of the dialogue and then they were asked to imitate the rhythm and intonation of it as closely as possible.

After the students completed the reading exercise, they were asked to record themselves talking about their fears for at least 30 seconds. The impromptu monologue exercise was designed to elicit spontaneous speech samples of the students to see whether

their acquired knowledge about connected speech, intonation and rhythm in reading has been transferred to their spontaneous oral performance as well.

Lesson 10: Connected speech

The warm-up exercise for this lesson was a reading of tongue twisters in a manner of connected speech with focus on reading them as fast as and as smoothly as possible. In the following exercise the students listened to a recording in which the same sentence has been pronounced once in a manner of connected speech and then more or less word by word. They were asked to identify the difference between the two pronunciations and imitate them.

The next exercise focused on reading transcriptions of limericks (see one of them in figure 10) and the task was to pay attention to stress and rhythm, reductions, linking, and voiced-voiceless coda. The students were asked to read the limericks as they were transcribed, but this time they had no audio model to accompany the written material. In the follow-up exercise, they were given another limerick, this time without the transcription, and they were asked to read it with proper linking.

Figure 10: Limerick taken from the tutor's handout for lesson 10 (Šimáčková 2020)

[ðəzə ˈklɛvərəʊldˈmaɪzəˈhuʷəlweɪz ˈtraɪz
There is a clever old miser who always tries

dɪfrəntˈmɛθədzən ˈweɪztuʷɪˈkɒnəmaɪz
Different methods and ways to economize.

səʊi ˈsɛz wɪðə ˈwɪŋk
So he says with a wink

wʌn seɪvz ˈgælənzəˈvɪŋk
One saves gallons of ink

baɪˈsɪmpli nɒt ˈdɒtɪŋ wʌnˈzaɪz]
By simply not dotting one's I's.

In the next exercise, the students listened to three recordings containing one sentence each and they were asked to write the sentences down, indicate the linking in the sentences and then record themselves imitating the pronunciation from the recordings. The very last exercise of the lesson was a monologue (one specifically for female students and one for male students), which they were asked to perform in such a way that the listeners would understand who the speaker of the monologue is, what is the context and topic of discussion (none of these were made clear in the monologue).

Lesson 11: Prosody

As in all the previous lessons, this one also started with a quick warm-up exercise, which was followed by a recording of a dialogue and then a close imitation of it. In the next exercise, the students focused on prosodic aspects (specifically phrasing, pauses, rhythm, intonation, and linking) in an audiobook. For this exercise, which included four different tasks, they were given three different written samples and three recordings of the audiobook. In the first task, the students focused on hearing the mentioned prosodic aspects in the audiobook without looking at a written sample of it. In the second task, they were instructed to listen to the second recording while reading the written sample (see figure 11) with clearly indicated intonation breaks (vertical lines) and potential location of linking marked by the underline. The students were asked to listen to the recording, indicate where linking actually took place and then read along with the recording.

Figure 11: Prosodic reading of an audiobook with markings for intonation breaks and linking taken from the tutor's handout for lesson 11 (Šimáčková 2020).

> Mr Dursley ‖ was the director of a firm ‖ called Grunnings, ‖ which made drills.‖ He was a big, ‖ beefy ‖ man ‖ with hardly any neck ‖ although he did have a very large | moustache.‖ Mrs Dursley ‖ was thin ‖ and blonde ‖ and had nearly twice the usual amount of neck, ‖ which came in very useful ‖ as she spent so much of her time ‖ craning | over garden fences ‖ spying on the neighbours.‖ The Dursleys ‖ had a small son ‖ called Dudley ‖ and in their opinion‖ there was no finer boy ‖ anywhere.‖

In the third task, which again included both a written sample (without any prosodic marks) and a recording of it, the students were instructed to mark the intonation breaks in the text as it was done in the previous task and read it according to the model recording. In the very last task, the students had only a written sample without any prosodic marks and without a model recording, and they were asked to read the text in a consistent manner with the previous samples.

Lesson 12: Fluent oral reading

In the very last lesson of the course, the students were asked to record themselves reading the Aesop's fable *North Wind and the Sun* and the children's short story *The Tiger Who Came to Tea*. The goal was to read it in an engaging, lively way and apply the knowledge

acquired during the instruction process to their oral reading of the texts.

9.2. Analysis of the recordings:

Figure 12: Pitch contour of the third sentence for speakers F1(01), F3(01), and M1(02) with the measurement points indicated by M, V, L%.

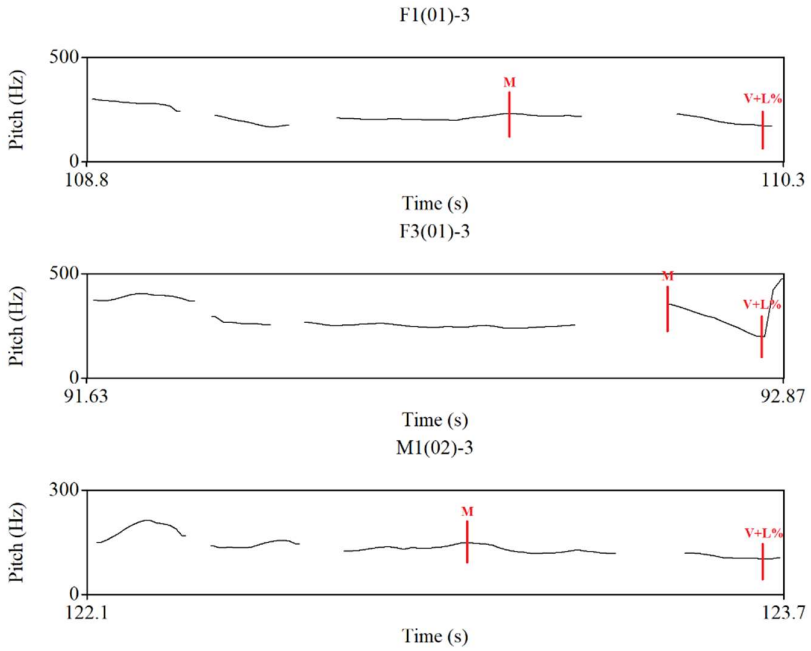


Figure 13: Pitch contour of the fourth sentence for speakers M1(01), M5(01), and F3(01) with the measurement points indicated by M, V, L%.

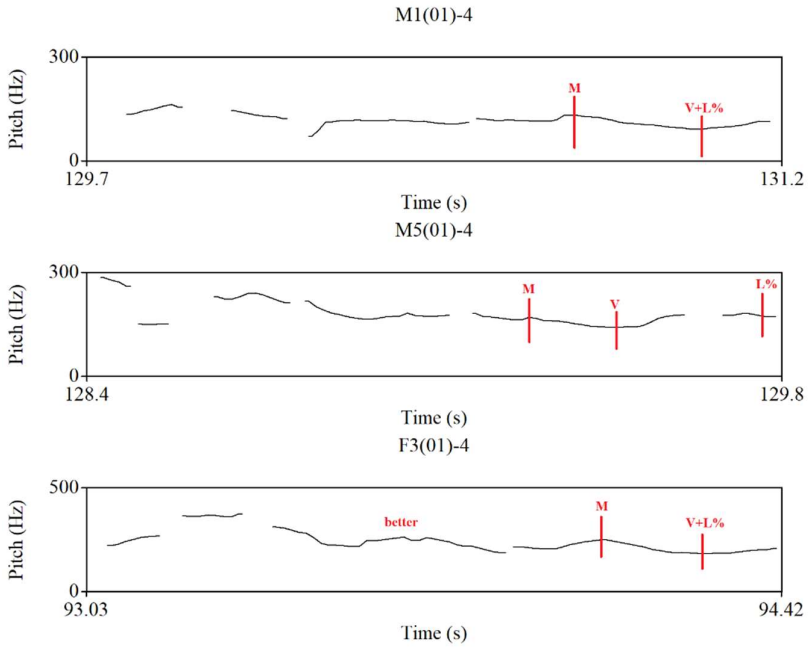


Figure 14: Pitch contour of the fifth sentence for speakers F3(02) and F4(01) with the measurement points indicated by M, V, L%.

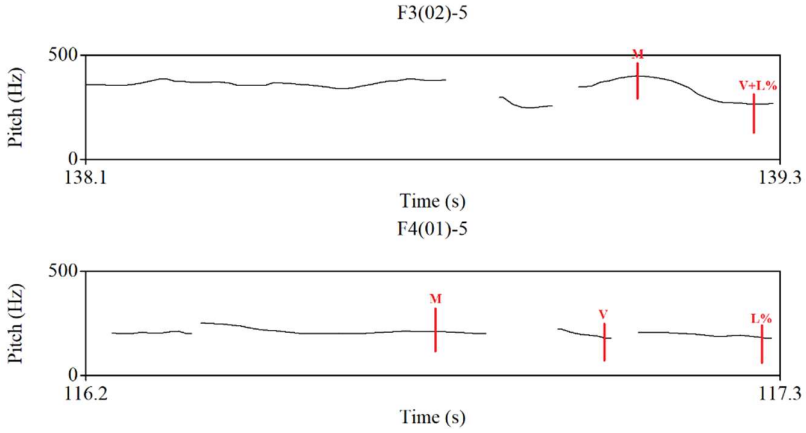


Figure 15: Pitch contour of the sixth sentence for speakers F3(01), F6(01), and F11(02) with the measurement points indicated by M, V, L%.

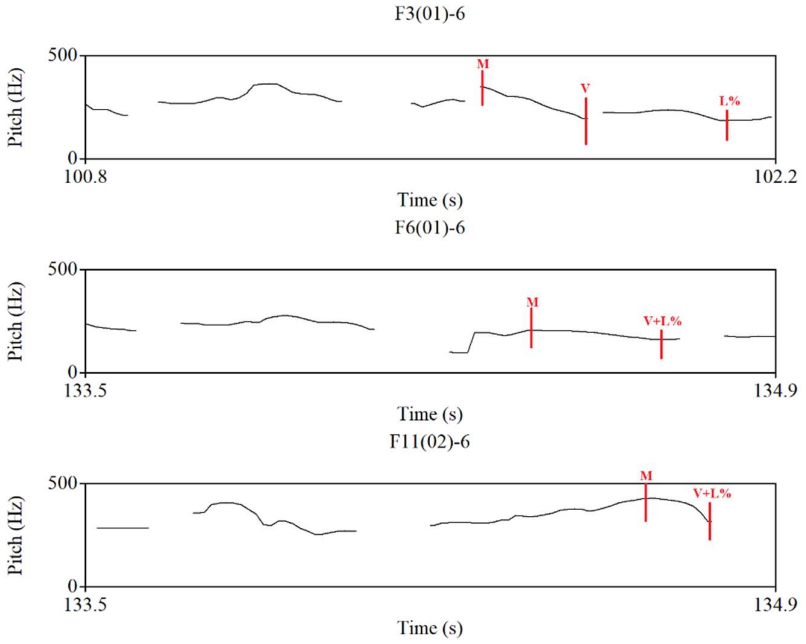


Figure 16: Pitch contour of the seventh sentence for speakers M1(02), M3(01), M5(02), and F2(02) with the measurement points indicated by M, V, L%.

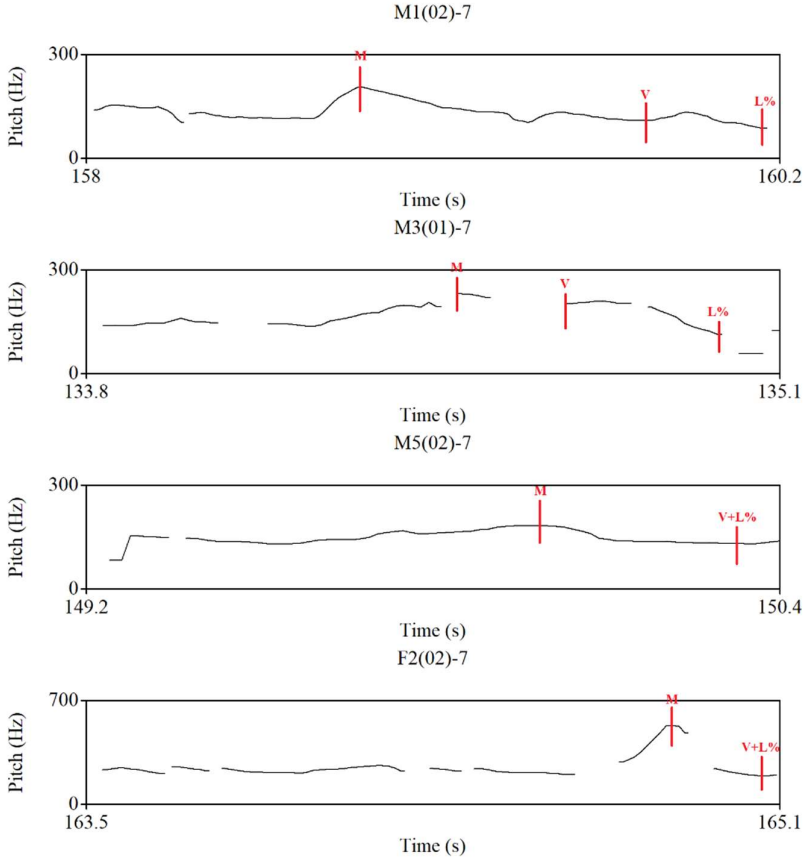
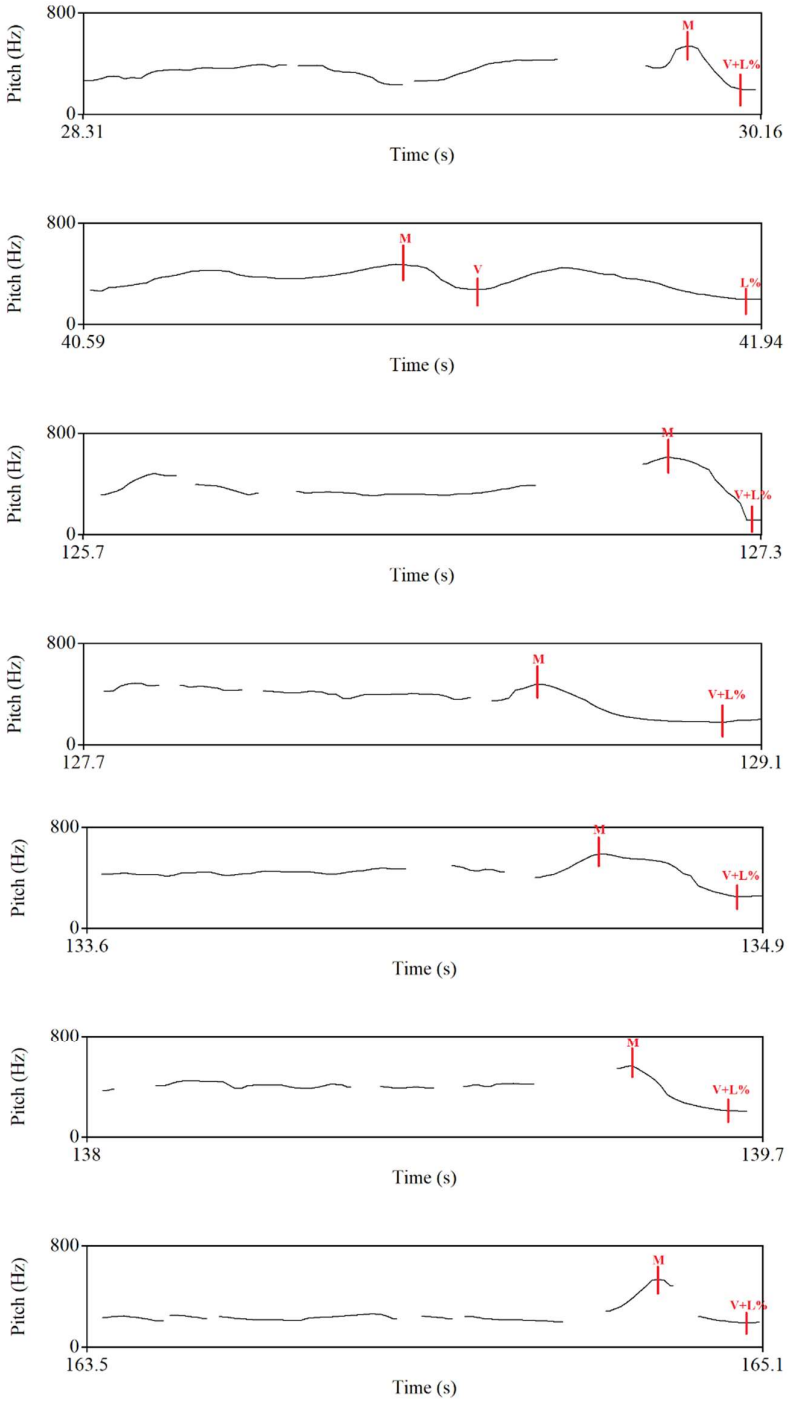


Figure 17: Position of the measurement points M, V, L% in all seven sentences for speaker F2 whose pitch span was significantly wider than that of any of the native speakers.



9.3. Reading material – the children's story

"The Tiger Who Came to Tea" by Judith Kerr (1968)

Once there was a little girl called Sophie and she was having tea with her mummy in the kitchen. Suddenly, there was a ring at the door. Sophie's mummy said: "I wonder who that can be. It can't be the milkman because he came this morning and it can't be the boy from the grocer because this isn't the day he comes. And it can't be daddy, because he's got his keys. We'd better open the door and see."

Sophie opened the door and there was a big, furry, stripy tiger. The tiger said: "Excuse me, but I'm very hungry. Do you think I could have tea with you?" Sophie's mummy said: "Of course! Come in!" So the tiger came into the kitchen and sat down at the table.

Sophie's mummy said: "Would you like a sandwich?" But the tiger didn't take just one sandwich. He took all the sandwiches on the plate and swallowed them in one big mouthful. Owp! And he still looked hungry, so Sophie passed him the buns. But again the tiger didn't eat just one bun. He ate all the buns on the dish. And then he ate all the biscuits and all the cake, until there was nothing left to eat on the table.

So Sophie's mummy said, "Would you like a drink?" And the tiger drank all the milk in the milk jug and all the tea in the teapot. And then he looked around the kitchen to see what else he could find. He ate all the supper that was cooking in the saucepans and all the food in the fridge, and all the packets and tins in the cupboard. And he drank all the milk and all the orange juice and all daddy's beer and all the water in the tap.

Then he said, "Thank you for my nice tea. I think I'd better go now." And he went.

Sophie's mum said, "I don't know what to do. I've got nothing for daddy's supper; the tiger has eaten it all." Sophie found that she couldn't have her bath because the tiger had drunk all the water in the tap.

Just then Sophie's daddy came home. So Sophie and her mummy told him what had happened, and how the tiger had eaten all the food and drunk all the drink. And Sophie's daddy said, "I know what we'll do. I've got a very good idea. We'll put on our coats and go to a café."

So they went out in the dark, and all the street lamps were lit, and all the cars had their lights on, and they walked down the road to a café. They had a lovely supper with sausages, and chips and ice cream. In the morning Sophie and her mummy went shopping and they bought lots more things to eat. And they also bought a very big tin of tiger food, in case the tiger should come to tea again.

But he never did.

9.4. Data elicited from the analysis:

Figure 18: Mean pre-test and post-test pitch span data of the NNSs with the difference between the two values calculated and the mean values of four NSs. The values are expressed in Semitones.

Pitch span of NNSs (in St)			
subj.	**pre-test**	**post-test**	**difference**
F01	7,05	5,79	-1,26
F02	4,72	15,81	-1,54
F03	7,89	10,90	-1,03
F04	4,10	6,81	11,09
F05	6,82	8,49	3,01
F06	4,11	5,70	2,72
F07	4,79	8,00	1,67
F08	8,04	8,24	1,59
F09	8,78	8,53	3,21
F10	6,69	5,16	0,20
F11	5,53	4,50	-0,25
M1	5,02	13,56	8,53
M2	3,78	2,85	-0,93
M3	7,89	7,44	-0,45
M4	6,69	6,01	-0,68
M5	6,74	5,01	-1,73

Pitch span of NSs (in St)	
NSf1	11,64
NSf2	10,24
NSm1	10,34
NSm2	9,79

Figure 19: Mean pre-test and post-test pitch level data for all NNSs and the mean values of four NSs. All data is expressed in Hertz.

Mean LEVEL (in Hz)		
subj.	**pre-test**	**post-test**
F1	155,1	196,6
F2	213,3	206,6
F3	189,6	206,2
F4	175,2	182,0
F5	163,2	162,1
F6	152,9	137,7
F7	152,0	161,0
F8	130,7	171,0
F9	138,4	195,9
F10	214,9	224,1
F11	178,9	228,5
M1	90,1	106,4
M2	97,3	102,4
M3	117,6	139,3
M4	98,1	90,0
M5	160,3	141,3

Mean LEVEL NSs in Hz	
NSf1	160,0
NSf2	176,7
NSm1	96,2
NSm2	133,1

Figure 20: Mean pre-test and post-test utterance duration data for all NNSs and the mean values of four NSs. All data is expressed in Milliseconds.

Mean Duration of NNSs		
subj.	**pre-test**	**post-test**
M1	1630	1670
M2	1160	1330
M3	1280	1420
M4	1210	1360
M5	1300	1190
F01	1490	1430
F02	1460	1570
F03	1320	1850
F04	1380	1360
F05	1320	1520
F06	1350	1650
F07	1340	1460
F08	1460	1760
F09	1350	1330
F10	1630	1890
F11	1270	1490

Mean Duration of NSs	
NS	1940
NS	1550
NS	1770
NS	1590