Filozofická fakulta
Univerzita Karlova

Bakalářská práce

Praha, 2020   Roksolana Fedorenko

Filozofická fakulta

Univerzita Karlova

Ústav anglického jazyka a didaktiky

Bakalářská práce

Roksolana Fedorenko

Phonetic Aspects of Fluency in Read and Spontaneous Speech

Fonetické aspekty plynulosti čtené a spontánní řeči

Praha, 2020   doc. Mgr. Radek Skarnitzl, Ph.D

Acknowledgements

I would like to thank my supervisor, doc. Mgr. Radek Skarnitzl, Ph.D, for guiding me through the process of writing this BA thesis, his support, useful instruction and advice, and readiness to help whenever there was a need.

Prohlašuji, že jsem bakalářskou práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

I declare that the following BA thesis is my own work for which I used only the secondary literature that is listed in the resources. This thesis was not used as a part of any other university study, nor was it used to gain a different university degree.

Abstract

The aim of this BA thesis was to compare linking and prosodic phrasing parameters in the recordings of students of English and American studies. We examined the speech of B2 Czech speakers of English with strongly accented pronunciation to see whether the completion of English Phonetics and Phonology course had an impact on students' performance of these aspects of fluency. We analyzed two sets of recordings by 8 speakers, one was recorded when participants were in their first year of studies, another a couple of years later. We analyzed recordings from Prague Phonetic and LINDSEI corpora, where we counted the number of linked vs glottalized words, analyzed the structuring of prosodic phrases, and measured their temporal characteristics such as the length of prosodic phrases and speech rate. The use of linking was found to be significantly higher and prevailing over glottalization in later recordings. Participants were successful in structuring logic prosodic phrases in both sets of recordings. No dynamic changes in temporal characteristics of prosodic phrasing were observed. These results show that training the pronunciation of linking leads to a long-term improvement of skills. However, the same cannot be stated about the prosodic phrasing performance within the framework of this study.

Key words: English, fluency, connected speech, linking, prosodic phrasing, speech rate, Czech speakers

Abstrakt

Hlavním cílem této bakalářské práce bylo porovnat vázání a některé prozodické parametry v nahrávkách studentů anglistiky-amerikanistiky. Zkoumali jsme řeč českých mluvčích angličtiny – na úrovni B2 – se sílným přízvukem, za účelem zjištění, zda absolvování kurzu fonetiky a fonologie angličtiny mělo vliv na zlepšení těchto aspektů v plynulosti řeči. Analyzovali jsme dva soubory nahrávek 8 mluvčích, první soubor byl nahrán, když účastníky byli v prvním ročníku studia, a druhý o několik let později. Zpracovávali jsme nahrávky Pražského Fonetického a LINDSEI korpusu, ve kterých jsme spočítali korelaci mezi vázanými a glotalizovanými slovy, analyzovali jsme strukturování prozodických frází, změřili jsme jejich temporální charakteristiky, jako třeba délku prozodických frází či mluvní tempo. Bylo zjištěno, že použití vázání se výrazně zlepšilo a začalo převládat nad glotalizací. Žádné dynamické změný v temporálních charakteristikách prozodického frazování nebyly nalezeny. Výsledky tohoto výzkumu ukazují, že cvičení výslovnosti vázání vede k dlouhodobému zlepšení návyků. Nicméně se, v rámci tohoto výzkumu, nedá použit stejné tvrzení ohledně prozodického frázování.

Klíčová slova: angličtina, plynulost, souvislá řeč, vázání, prozodické frazování, mluvní tempo, české mluvčí

Table of Contents

List of Figures

1. Introduction

Acquiring fluency may be challenging for language learners as it is a complex notion that works on both segmental and suprasegmental levels of speech. Thus, when teaching the pronunciation of English, the same attention should be paid to both levels. The importance of teaching the pronunciation of suprasegmentals seems to be underestimated. We decided to conduct this study to show that specific training of suprasegmentals may be effective for improving fluency. This BA thesis focuses on some fluency aspects of read and spontaneous speech of Czech speakers of English, particularly on linking and prosodic phrasing.

The aim of this study was to describe the use of linking and prosodic phrasing phenomena and to find whether some progress in the use of these fluency aspects occurred. In other words, we wanted to investigate whether the speech of students becomes more fluent after they complete English Phonetics and Phonology course in their first year of studies.

The material used in this research consists of 16 recordings by 8 Czech speakers of English with strongly accented speech, all of them studied English philology. It was especially interesting to compare linking scores of participants as Czech speakers of English have a natural tendency to glottalize initial vowels in connected speech, the way they pronounce it in their L1. We measured the number of glottalized vs. linked words and the distribution of linking in different phonetic contexts. We also examined some features of prosodic phrasing, such as the structuring of prosodic phrases (whether they are built in accordance to the lexical and syntactic structure of a sentence), their length (in syllables) and speech rate (in syllables per second).

The following theoretical part provides useful information for our study. It introduces some relevant basic terms and concepts, and previous research concerning fluency, foreign accent, prosody, and connected speech. Material section provides a detailed description of the material used. The analysis of the obtained data and results of the study are presented in sections Linking and Phrasing, which are then further discussed and concluded in General Discussion and Conclusions sections.

2. Theoretical Background

2.1 Fluency in L2

The nature of fluency is not easy to identify, as it is an abstract notion, which can be applied in multiple fields of studies. Fluency is a complex metacognitive phenomenon, it is "the subjective experience of ease or difficulty associated with completing a mental task" (Oppenheimer 2008, p.237). Our focus should be put on the definition of fluency in the linguistic area. Fluency in second language acquisition has proved itself to be a problematic field of study. Questions considering how fluency can be measured, which aspects of speech production have an impact on it and how it is related to first language performance are some of the central issues that linguists are interested in. Second language performance is a complex psycholinguistic process consisting of two main stages: speech planning and speech production. The former has a huge impact on the latter.

It is useful to examine the overall process of L2 speech production in order to better understand the role fluency plays in it. Three central phenomena lie in the basis of second language performance: complexity, accuracy and fluency. There is a certain tension between form (complexity and accuracy) and fluency as all these areas of L2 performance require attention and working memory involvement (Skehan, 2009, p.511). As speakers pay more attention to one of these areas, they end up with having lower performance in others. Fluency can be improved by planning and structuring the speech, which means modifying the first stage of language performance. Planning helps the non-native speaker to retrieve less frequent words and to make different lexical selections. Structuring one's speech can also be beneficial for second language performance as it helps the speaker to operate within helpfully limiting parameters and as a result more attention is available for the ongoing performance (Skehan, 2009, p.519).

Fluency is primarily a temporal and intonational phenomenon. Temporal variables of fluency can be divided into two groups. The first one includes speech rate, the mean length of runs (an average number of syllables produced in utterances between pauses) and pauses, speech rate and phonation-time ratio (the percentage of time spent speaking as a percentage proportion of the time taken to

produce the speech sample). The second one contains frequency of filled and unfilled pauses and other disfluencies (Kormos & Dénes, 2004). Such performance phenomena as lexical and non-lexical fillers, false starts, self-corrections and repeats are considered to be filled pauses. They are often used in order to relieve pressure of online planning, because their production helps the speaker to earn additional time in order to arrange mental planning with the physical aspects of speech production or for choosing the appropriate form for the message that is being communicated (Gráf 2017, p.65). In general, such occurrences are called hesitation phenomena. They have a direct impact on the evaluation of fluency of nonnative speakers of English. Speakers who use hesitation pauses less frequently are considered more fluent than those who produce more of them. On the one hand, hesitation phenomena may help the speaker to better formulate his speech, on the other hand, this may cause some negative impact on the receiver and his comprehension of such utterance. Gráf (2017) in his study of disfluencies focused on the phenomenon of repeats in Czech speakers of English (repeats are segments of speech which are involuntarily repeated in close proximity without adding any propositional content to the message), which is one of the most frequently occurring types of disfluency. Most typically repeated units are pronouns, articles, prepositions and contracted forms. Repeats may be viewed from two perspectives: as a type of disfluency (when they are not used as a speech planning strategy) and as a fluency-enhancing strategy, providing speakers with extra time that they may use to plan their speech (Ibid, p.76).

Silence in speech production, mainly pausing, is important and must be taken into consideration when examining fluency. Native and non-native pausing differs. Native speakers naturally pause mostly at the end of a clause. The non-native speakers quite often struggle with advanced vocabulary and complex syntactical structures. As a result, they tend to pause in places that seem to be semantically or syntactically unjustified. When comparing fluent and non-fluent L2 learners, fluent students tend to pause at grammatical junctures while disfluencies tend to occur in clusters in the speech of non-fluent L2 learners (Kormos et al., p.149). Fluent speech production is directly related to automaticity. High level of automaticity of speech production mechanisms helps one to speak fluently. Non-fluent speakers cannot rely on a sufficient number of automatic sequences

and apply conscious rule-based mechanisms, and when they struggle to be highly accurate, their speech becomes very slow (Ibid, p.160). Fluency cannot be explained only in terms of a learner's linguistic knowledge. It reflects a learner's ability to produce language rapidly and smoothly under the cognitive pressures of real-time processing. Ewa Guz (2015) proposes that some aspects of L2 fluency may be related to a person's individual speech characteristics which pervade his oral performance in any language and are not specific to the learner's performance in a particular language. Most people have come across speakers who are more and less fluent in their native language. Thus, certain aspects of language production can be typical of the speaker's individual style and, as a result, they can have an influence on both L1 and L2. Therefore, the higher the speech rate is in one's L1, the higher it would be in their L2. Speech rate seems to be the only area of speech production that remains the same across different languages a person speaks. According to Guz (2015) hesitation pauses that disrupt the syntactic congruence of clauses are typical of L2, but they almost never occur in L1. This phenomenon reflects the degree of cognitive effort related to speech production (Ibid, p.240). As a result, L2 performance has a lower speech and articulation rate, a greater number of filled and silent pauses, a greater duration of filled and silent pauses and shorter fluent runs between pauses.

In their study considering fluency, Pinget, Bosker, Quené & de Jong (2014) investigated the problem of fluency from a slightly different point of view. According to this perspective we can distinguish three aspects of fluency: cognitive fluency, utterance fluency and perceived fluency. Cognitive fluency reflects the speaker's ability to efficiently plan and execute speech, utterance fluency can be measured in a speech sample, and perceived fluency refers to the judgement listeners make about the fluency of a speaker (Ibid, p.350). The authors suggest that a speaker's individual L1 fluency characteristics can have an impact on his L2 fluency. The first language of a speaker is seen as a basis for his L2 production. Simply, the most fluent speech one can produce is a speech made in his first language. But it does not necessarily mean that there should be a strong connection between L1 fluency and L2 fluency of the speaker. The experiment made by Pinget et al. showed that "taking L1 fluency into account does not result in better prediction of fluency ratings" (Ibid, p.361). It had a

quantitative character rather than qualitative and, as a result, there was a supposition that the perception of foreign accent is also based on the qualitative analysis of a range of other factors such as sentence intonation and correct stress placement. This leads us to the next problematic area, the connection between the perception of fluency and foreign accent.

2.2 Foreign Accent

Different language varieties can be put into two categories: standard and nonstandard varieties. Standard varieties are those that are codified, they provide us with norms of correct usage of vocabulary, grammar as well as pronunciation, for example Standard American English in the US, Received Pronunciation in the U.K., or Cultivated Australian English in Australia. Those standards are often used in broadcasting, administrative matters, literature and, most importantly, in teaching. Nonstandard varieties are more likely to be associated with specific linguistic communities as their language has some distinctive features. They differ in terms of pronunciation and language use. Thus, foreign accents belong to the latter group as they are not categorized as standard. Why is our speech accented when using a foreign language? Difficulties with learning a foreign language occur because one's L1 tends to have a great influence on production in the L2. Of talking about the phonetical aspects specifically, articulatory organs are used to the production of sounds that are typical for L1 and they have to be properly trained in order to articulate L2 sounds in an appropriate way. The perceptual training can improve not only the perception of L2 but also its production (Lengeris 2012, p.34). Foreign accent is caused by articulatory and prosodic interferences, and those can have an impact on the way native speakers perceive accented speech (Reitbrecht & Hirschfeld, 2015, p. 0166. 2). Not only foreign accent illustrates differences between the way vowels, consonants, lexical stresses or intonation-related phenomena are realized, but it is also closely related to the question of fluency. For this reason, foreign accent has a huge impact on the perception of one's speech. It influences both the process of communication between the individuals and the way they evaluate each other. Foreign accent can be one of the signals that helps listeners to infer to which social group speakers belong and, as a consequence, they can attribute some stereotypic traits that are associated with those group memberships (Dragojevic, Giles, Beck & Tatum 2017, p. 385). The stronger the

accent of a speaker is, the less positive evaluation he is likely to get from his listeners. Most foreign accents tend to be negatively stereotyped and speakers with a strong foreign accent are more likely to be attributed negative stereotypic traits, because their speech is more difficult to process (Ibid, p.386). As a result, the evaluation of speakers with foreign accent is prejudiced, because they are immediately categorized according to their accent. Language attitudes refer to evaluative reactions to different language varieties (e.g., accents, dialects) and are organized along two primary evaluative dimensions: status (e.g., intelligent, competent) and solidarity (e.g., friendly, warm) (Ibid, p.386). When talking about foreign accent, we should take into consideration Oppenheimer's saying that "objects are fluent when they have been seen frequently, have been seen recently and/or have been seen for a long period of time" (Oppenheimer 2008, p.237). This assertion helps us to identify the connection between the perception of fluency in speech and foreign accents. As foreign accents for the native speaker may be non-frequently heard phenomena, they can disrupt the feeling of speech being fluent. How can we measure a foreign accent? In the study held by Flege (1988) the connection between the perceived degree of fluency and foreign accent judgements was investigated. The scholar suggested that removing pauses from sentences spoken by nonnative speakers would not improve the foreign accent scores accorded to their sentences. The suggestion was based on the fact that nonnative speakers tend to produce sentences containing more and longer pauses than the natives. The results of the study have shown that removing pauses had a small effect on the scores.

2.3 Prosody

The production of speech consists of two main aspects: the segmental (pronunciation of sounds on a level of a single segment, for example vowels and consonants) and suprasegmental or prosodic one (the one that extends the level of a segment; such as stress, rhythm, intonation and phrasing). Prosody plays one of the key roles in communication and can have an impact on speech's intelligibility. It is central to understanding spoken language and it contributes to creating a basic skeleton that helps us to hold an auditory linguistic sequence in memory. (Frazier, Carlson, Clifton 2006, p. 248). Languages differ in their prosodic structures and realizations. When producing a sentence, we focus not only on the grammatical structure and the semantic meaning of it, but we also

pay attention to the prosodic phrasing. Through the use of prosody we can convey information that cannot be shared in orthographic dimension. That is why it is important to pay attention to the way sentences are divided into prosodic units. Prosody can be highly connected to the lexical and syntactic structure of a sentence, but it is the speaker who chooses how to phrase a sentence. Though prosody is not governed only by syntactic structures, some traces of interrelation between these two phenomena may be identified. "Prosodic phrasing converges with syntactic structures to transmit the syntagmatic organization of elements of speech to a listener. This applies to the level of words, of phrases and of sentences." (Kohler 2010, p.110). Though it is connected to the syntactic structure of a sentence, "prosodic constituency reflects syntactic constituency, when possible. Limits on a transparent one-to-one mapping of syntax and prosody arise because prosodic structure, unlike syntactic structure, is not recursive and thus the two types of structure often cannot coincide completely." (Clifton et al. 2002, p. 105). Therefore, there are other than syntactic factors that influence prosodic structure of a sentence. Some of these factors can be speech rate, semantic focus, discourse structure and rhythm (Shattuck-Hufnagel 2000, p.203). Speech prosodic structure can work as a planning framework for phonological and phonetic speech planning. Thus, the actual prosody may deviate from theoretical predictions. The actual realization of speech is governed to some extent by the prosodic structure. It may differ from speaker to speaker and can also vary from utterance to utterance produced by the same speaker. (Ibid, p.209)

Prosodic parameters (or suprasegmentals) as well as segmental parameters contribute to the perception of foreign accent in L2 speech and they are likely to be affected by L1 (Piske 2012, p.47). One of important suprasegmental elements in the English language is the primary stress at the level of phrases, also called prominent, or somewhat inaccurately, sentence stress. The primary phrase-level stress indicates new and contrastive information, while old or given information is presented as unstressed (Hahn 2004, p.202). For example, in a dialogue between two speakers (A and B), new information receives primary stress (marked with ´), while given or old information remains unstressed, prosodically suppressed (marked with an underline):

A: Are you alone?

B: I am néver <u>alone.</u>

A study conducted by Hahn (2004) examined the interconnectedness of primary stress and intelligibility within a nonnative discourse. Results of the study have shown that primary stress significantly contributes to the intelligibility of speech (Ibid, p.218). As primary stress belongs to the group of suprasegmentals, it can be considered that their role in the process of speech comprehension is as important as that of segmentals. Such findings are fundamental for this paper as it also focuses on one of the suprasegmentals which is prosodic phrasing.

2.3.1 Prosodic Phrasing

Prosodic phrasing can affect our understanding of sentences and it can also affect syntactic processing. It refers to the division of words into units which are typically interconnected with both semantic and syntactic structure of an utterance. Such units are defined as prosodic phrases; other terms include intonation phrases, prosodic groups, sense groups, breath groups, tone units, and thought groups. Boundaries between individual prosodic phrases are marked using melodic and temporal cues: a nuclear tone (typically rising within a longer utterance and falling at its end) and final deceleration or lengthening. What is the function of prosodic phrasing? It is closely related to semantics as it helps the speaker to organize the meaning of an utterance in order to convey it in a proper way. Prosodic phrasing shapes the final semantic outcome of speech. The two examples of different prosodic phrasing can illustrate this function (Kohler 2010, p.112):

1. I read a review of nasality | in German ||

2. I read a review | of nasality in German ||

The two sentences presented above are identical in writing, they have the same syntactic structure and consist of the same words. However, their semantic meaning can change with different prosodic phrasing. Most interestingly, it can be realized in both spoken and read domain. Such an example explicitly shows us the important role of prosodic phrasing in speech. There are some aspects of prosody that are obligatory, others are not as obligatory, they are just a matter of style and preference. Some syntactic structures require prosodic boundaries in particular locations (the end of initial

subordinate clause, flanking an appositive structure). Thus, they guide the analysis of a syntactic structure. Non-obligatory prosodic aspects, on the other hand, are not dictated by grammar and it is up to the speaker whether to put a break or not, where the break is possible but not required. (Frazier et al, p. 245) Prosodic boundaries help to identify syntactic boundaries when analyzing speech. If we want to keep the track of prosodic boundary positions, we should assume that what the listener parses is a phonologically and prosodically structured input (Clifton, Carlson & Frazier 2002, p. 107).

As was mentioned previously, prosodic phrasing does not only occur in spoken speech as it can be also related to silent reading. For example, commas give rise to the same component as intonational boundaries do in speech or the locations of obligatory prosodic boundaries correspond to locations of long reading times in eye-movements. (Frazier, Carlson & Clifton 2006, p.247) This type of prosody is also called implicit prosody and it is different from the prosody generated by the prosody of spontaneous speech, because in spontaneous speech speakers try to generate prosody from the information structure they tend to deliver. (Jun 2010, p. 1220) When reading a written text unprepared and when the speakers know that they are being recorded, their focus would be put on reading as fluently as possible. The best way to achieve a "good performance" would be to produce each content word prominently and to follow phonological and structural limitations by putting a prosodic break more frequently. (Ibid, p.1221) Another aspect of comparing reading aloud and silent reading is that silent reading is faster. Thus, the speakers organize prosodic structures at faster speech rates. "Prosody generated during silent reading may reflect the semantic and pragmatic meaning of the sentence but prosody generated during oral reading might not." (Ibid, p. 1223) Generally, prosody helps us to better comprehend and structure an utterance. It works as a framework, which helps to process language and to memorize it. That is why "highly constrained or even stereotyped prosodic or rhythmic patterns seem especially characteristic of language that must be held in memory, for example, nursery rhymes for young listeners, and epic poems for all listeners." (Frazier et al., p.248).

2.3.2 Prosody of fluent speech

The previous paragraph focused on the way prosody works on the syntactic level. Standard linguistic approach constructs the text from sentences and investigates their intonation patterns. As a

result of this process, fluent speech is seen as a sequence of independent sentences. There is a higher prosodic organization which governs the natural flow of connected speech. Tseng, Pin, Lee, Wang & Chen (2005) suggested in their study that fluent connected speech consists of larger units of multi-phrase speech paragraphs or independent unrelated phrases and sentences. In their study they worked only with read speech not the spontaneous one as it is typically influenced by processes connected with speech planning. They came up with the idea that fluent speech prosody differs from sentence prosody and developed a multi-phrase model for it. There is a framework of multiple-phrase grouping (a higher governing junction above individual sentences by which existing linguistic definitions still apply), which is based on perceived unit located inside different levels of boundary breaks across the flow of spontaneous speech. It accounts for how speech patterns are perceived in fluent speech via phrase grouping, and how it provides prosodic specifications to individual phrases under grouping in addition to phrasal intonation. (Ibid, p.286) Their study also discovered that multi-phrase prosody units are independent from structural specifications, which means that the placement of boundaries does not always correspond to syntactic boundaries. "Boundary breaks are used by the speaker for planning in speech production, and as forecasting apparatus for processing by the listener as well." (Ibid, p.298). The comparative study carried out by Cavallieri, Wigglesworth & Maxwell (2019) examined the relationship between prosodic durational signals and speech fluidity. Participants were Spanish speakers of English, so the prosodic features of Spanish and English were investigated of both read and spontaneous speech. One of the aspects the scholars focused on were pause phenomena (the study differentiated between fluent and dysfluent pauses). The results of the study have shown that the use of pauses is related to the task being performed (more pausing in spontaneous speech) and that the speech style is idiosyncratic and has an impact on the use of pauses. (Ibid, p.2627).

2.4 Connected speech

Once words are integrated into continuous speech their pronunciation may differ from that when they are being spoken in isolation. The reason for this phenomenon is that speakers tend to save as much energy as possible while speaking, thus simplifying the process of articulation. The needs of the speaker and the listener are different in speaking. Speakers want to communicate in the easiest

possible way, with the least effort. As a result, their articulatory system tries to minimize the effort and is likely to take shortcuts. However, speakers tend to move from one sound to the next in the shortest and easiest way. They can blend sounds together when it is possible, sometimes even change or omit them. Speakers cannot be too lazy in pronunciation, otherwise they risk not being understood by the listener. On the contrary, the best option for the listener is to hear clearly pronounced words with the least changes. That is why there should be some balance between the needs of the speaker and the listener, otherwise the process of communication is likely to fail. For the speaker, the process of speech production is future-oriented, it is driven by certain purposes and is prospectively organized Lindblom 1990, p.404). For listeners, the process of speech perception is both signal-derived and influenced by the dynamic state of the processing system, or signal-complementary processes (Ibid, p.407) as listeners not only need to focus directly on the utterance they hear, but they also need to apply their own knowledge to process the speech. Listeners need the input (speech) to resonate with their internal knowledge to be more successful in decoding the signal. Ideally, the aim of the speaker should be to produce the utterance that will be contrastive enough, so that the listener would be able to easily identify individual lexical items (Lindblom 1990, p.404-405).

The comprehension of connected speech may be challenging for language learners if they are not familiar with its typical features. That is why it is highly important to provide them with knowledge about the processes that are typical of connected speech. Native speakers of English may change the pronunciation of words when they blend them together in connected speech and there is a mechanism for doing this. So, what happens when two words standing next to each other come into contact? Some interesting phenomena may occur. "The pronunciation of words in connected speech may leave vowel and consonant sounds relatively intact, as in some types of linking, or connected speech may result in modifications to pronunciation that are quite dramatic, including deletions, additions, or changes of sounds into other sounds, or combinations of all three in a given word in context." (Alameen & Levis, 2015, p.159) Such phenomena frequently occur in the speech of native speakers, with them doing it subconsciously. Connected speech processes (CSPs) are natural and essential for speech production. There is a direct connection between the frequency of CSPs

occurrence and the speech style and social distance between the speakers. Speakers tend to pay less attention to articulation processes in casual conversation, but the situation changes once they are using speech for some formal occasions. Similarly, social distance has an impact on the frequency of CSPs as when speakers belong to the same social group, the comprehension of speech will become easier, thus they will be able to pay less attention to distinctive articulation. (Ibid, p.164). The comprehension of connected speech may be challenging for L2 learners because they develop their listening skills on rather artificial language. As a result, they are often unfamiliar with native-like pronunciation of spoken language. One of the main problems with comprehension is that L2 listeners sometimes are not able to segment the speech in a proper way because they tend to compare the actual connected speech to stored words. Thus, "the presence of CSPs may create lexical ambiguity due to the mismatch between the lexical segments and their modified phonetic properties." (Ibid, p.166) It is necessary to provide L2 learners with the information about CSPs phenomena and to train their abilities to both produce and comprehend connected speech by supplying them with as many varieties of the foreign language as possible. Without training in listening to authentic native speech, learners of English are likely to go through a frustrating experience of not understanding native speaker conversations. Quite often, teaching English as a foreign language puts a heavy focus on understanding grammar, learning new vocabulary, practicing learned material through dialogues and artificial conversations. The problem here is that in most cases learners does not encounter native-like natural mode of speech, because they are often provided with listening materials that are spoken in a clearly pronounced and articulated speech which is far from casual conversations in terms of production. Native connected speech differs from that taught in the classroom as due to the law of economy, speakers tend to draw sounds together with the purpose of saving time and energy. Their speech organs are 'lazy' for taking a new position for each sound. Connected speech is not typical for some concrete geographic area or exclusive variety of English. It is a process that happens across different standards and regional dialects and in all languages, which makes it a prominent feature of natural speech. Generally, dynamic speech is the product of series of alternations. Temporal pressures that a speaker feels when producing a speech lead to alternations in pronunciation. As a result, certain

segments, syllables, or even entire constituents are affected and they tend to be absorbed in fluent speech. Hieke (1984) proposes a threefold classification of absorption: linking (as in syllabic restructuring); levelling (as in vowel reduction, flapping, assimilation); loss (such as of vowels or syllables). All the phenomena listed above seem to be a result of one common process – absorption.

2.4.1 Linking

As our work is going to examine linking processes that occur in speech, this paragraph will focus on it. Linking is one of the most prominent connected speech processes and one of the most frequently occurring. It happens when the last sound of a word smoothly connects to the initial sound of the following one. Its function in connected speech is to make two words sound like one without changes in segmental identity, and it can result in resyllabification (consonants are attached to syllables other than those from which they originally came) of the segments without changing them (Alameen et al. 2015, p.162). Linking helps us to better differentiate between native and nonnative English speech in proficiency testing and is one of the parameters for evaluating fluency in English. (Hieke 1984, p.348) It is important to point out that linking happens within one thought group (a group of words that form a unit of meaning), and much less frequently across the boundaries of thought groups (units of meaning). For example, the sentence *tell me the story about what happened yesterday* consists of two thought groups, with the first one being *tell me the story* and the second one being *about what happened yesterday* accordingly. In this case, linking is less likely to occur between the words *story* and *about* ([stɔːri₍ⱼ₎əbaʊt]) as they belong to different thought groups. The postmodification is likely to be emphasized in order to make it clear what particular story the speaker wants to hear. However, some native speakers, with fast speech rate, may link even between different thought groups, which may make their utterance more difficult to comprehend.

Five linking phenomena can be identified:

- pseudo-resyllabification

- linking [r]

- intrusive [r]

- transient [j]

- transient [w]

Pseudo-resyllabification occurs when a word-initial vowel is preceded by a word-final consonant. The listener may have an impression that syllabic boundaries do not coincide with word boundaries. For example, the phrase *got into his car* [ˈgɒt.ɪn.tʊ.ɪs.ˈkɑː] becomes [ˈgɒ.tɪn.tʊ.ɪs.ˈkɑː] as a result of the above mentioned process (Volín 2006, p.64).

As RP is a non-rhotic accent, the /r/ sounds are not pronounced in a word-final position. However, they may be pronounced in connected speech to link words to each other. Such [r]s are used "to avoid intervocalic glottal stops and to prevent two vowels from a direct contact" (Ibid, p.65) What is the difference between linking and intrusive /r/? Linking /r/ appears in contexts were the historically attested final /r/ sound is introduced when followed by a vowel-initial morpheme. "This *r* is generally presumed to be retained or inserted either to serve as a 'hiatus-breaking' element, or to provide a

The intrusive /r/ looks almost identical to the linking /r/ but there is a difference between these two. While in the case of linking one, the sound /r/ was historically present in a word, in the case of intrusive /r/ there was no such a sound. "Intrusive /r/ arises essentially from the natural tendency to give identical treatment to words with identical endings" (Wells 1982, p. 223; cited in Gick 1999). Gick (1999, p. 32) provides several examples that demonstrate the contrast between linking and intrusive /r/ (the first one is an example of the linking /r/, the second of the /intrusive /r/):

a. 'mar' [mɑː] -> 'mar is' [mɑːriz]

    'ma' [mɑː] -> 'ma is' [mɑːriz]

b. 'lore' [lɔː] -> 'lore is' [lɔːriz]

    'law' [lɔː] -> 'law is' [lɔːriz]

c. 'coder' [kəʊdə] -> 'coder is' [kəʊdəːriz]

    'coda' [kəʊdə] -> 'coda is' [kəʊdəːriz]

The transient [j] is used in contexts were a word-initial vowel is preceded by /iː/, /ɪ/, /eɪ/, /aɪ/, or /ɔɪ/. The j-like sound that is put between the two vowels has only a transitory function that helps

to make a soft transition between the two sounds. In the transcription, the transient [j] is transcribed as (j). Examples:

- *the end* [ðiː(j)end]

- *they are* [ðeɪ(j)ɑː]

- *my own* [maɪ(j)əʊn]

The transient [w] has similar qualities to those of the transient [j], as it is used to connect two vowels. "It is a glide from close back area in the vocalic space to another vowel. We can hear it if word-final /ʊ/, /uː/, /aʊ/ or /əʊ/ are linked to a vowel of the following word" (Volín, p.66). It is transcribed with the (w) symbol. Examples:

- *two of* [tuː(w)əv]

- *know it* [nəʊ(w)ɪt]

- *now escape* [naʊ(w)ɪˈskeɪp]

What happens when a word-initial vowel is not linked to the previous sound? The phenomenon called the glottal stop comes into play. Glottalization is the complete or partial closure of the glottis during the articulation of another sound. Thus, glottalization and linking are "simply opposite strategies for repairing onsetless syllables, with linking resulting in weakening of word boundaries and glottalization having a boundary-strengthening effect." (Šimáčková, Podlipský & Kolářová 2014, p. 679). Hieke (1984) when describing the process of consonant attraction, mentions the predictable conditions under which glottal stop occur before the syllable-initial vowel (p.346):

1. In deliberate speech for reasons of extra clarity or aesthetics.

2. Where special stress assignment overrides absorption phenomena.

3. Where plus juncture is phonemic and thus obligatory.

4. After silence (with no prior syllable to draw on).

It is interesting that word-initial glottalization may have an impact on intelligibility in specific contexts and, of course, it affects the way we evaluate nonnative speech. Czech speakers of English are highly influenced by their native language. In spoken Czech, a glottal stop is a frequently occurring phenomenon, "a reliable boundary marker". (Šimáčková et al, p.681) As a result, Czech

speakers of English tend to glottalize word initial vowels in their L2. Such a tendency marks their speech as accented and has an impact on its rhythm and perception. As to the rhythm, the insertion of a glottal stop before a weak vowel conflicts with what we expect to hear from native speakers. The main problem lies on the level of overall rhythm in the two languages. English is a stress-timed language (stressed syllables occur at approximately regular intervals), the Czech language is syllable-timed (each syllable takes approximately the same amount of time to be pronounced). Here, the conflict between L1 and L2 may be observed. Czech speakers of English employ the same rhythmic rules they are used to in their L1 in their L2. This may end in giving prominence to wrong words and may even cause some pragmatic problems in particular contexts (Ibid, p. 679).

2.4.2 Linking and English Teaching

Are there some methods that can help to improve the way L2 speakers link? An experiment conducted by Sardegna (2011) investigated the impact the teaching intervention has on improving the way speakers link. The course created by Sardegna instructed students how to use pronunciation learning strategies to improve their oral skills in language production outside the class. The students were provided with an extensive repertoire of pronunciation strategies and received written and oral feedback both during the classes and during the individual meetings with their instructors. The training focused on improving students' pronunciation features such as segmental, suprasegmental, and connected speech features, with linking included. After participating in the course, which lasted for four months, students' score increased significantly after being instructed about linking and practicing pronunciation strategies for improving it. They showed a long-term improvement with linking while reading aloud. "Differences in improvement during instruction could not be predicted on the basis of students' gender, language background, or length of stay in the target language community before instruction, but it could be predicted on the basis of students' entering proficiencies with linking. Students with a higher level at the beginning of the course improved less than those with a lower level." (Ibid, p.115) Sardegna came up with five factors that may improve the way students link (p.116):

1. Entering proficiency level with linking.

2. Degree of improvement with linking during the course.

3. Quantity, quality, and especially frequency of practice with linking in covert rehearsal.

4. Strong intrinsic and extrinsic motivations to improve.

5. Prioritization of linking over other targets for focused practice.

The improvement of linking is a process that requires a huge amount of time and effort invested by both the student and the teacher. The more students work on their own, the better the improvement will be. But this also requires control and support on the teacher's side. The collaboration in this case can be highly effective. Sardegna's study is highly valuable for our research as it is closely connected to our hypothesis.

3. Material

For this study, 16 recordings of speech of 8 female speakers were chosen from two corpora: the Prague Phonetic Corpus and the Czech subcorpus of the Louvain International Database of Spoken English Interlanguage (LINDSEI). They were native speakers of Czech and their English pronunciation was evaluated as strongly accented. All the speakers were students of English and American Studies at the Faculty of Arts, Charles University, and their age varied from late teens to early twenties. Recordings taken from the Prague Phonetic Corpus were recorded as a required part of English Phonetics and Phonology course that the students had in their first year of studies. The speakers were asked to read a standard BBC news report. An average reading duration was around 4 minutes. The speakers were given sufficient time for preparation. The recording took place in a recording studio at the Institute of Phonetics of the Faculty of Arts, Charles University, using the AKG C4500 B-BC condenser microphone at a sampling frequency of 32 kHz and 16-bit resolution.

Another part of the recordings, those from the LINDSEI subcorpus, was recorded a couple of years later, when the students were in their third or fourth year of studies, using the same recording studio and equipment. These were dialogues that consisted of two parts. The first part was an improvised speech on a specific topic chosen by the interviewee. In the second part interviewees were asked to describe the picture they were shown. The level of English of the speakers was defined as B2 by the interviewer. As the duration of recordings was over 10 minutes, the recordings of dialogues had to be cut into shorter passages. An average duration of each passage was about 2 minutes. Two passages from the middle part of each recording were chosen for future analysis.

Connected speech phenomena such as linking and prosodic phrasing, which were discussed in the previous chapter, were analyzed by means of listening using the Praat software. Both phenomena were labelled manually.

4. Linking

4.1 Method

Linking was marked in the TextGrid using specific labels for each of the phenomena:

- the occurrence of the glottal stop in places were linking was expected was marked with '?'

- resyllabification was marked as 'res'

- transient w, j were marked with 'w' and 'j' accordingly

- linking and intrusive r were marked as 'r'

Glottal stops were not marked if they occurred after a pause. The data were extracted using a script in Praat and then exported to tables using the Microsoft Excel and sorted according to the speech style (read/spontaneous).

We were mostly concerned in the opposition of linking vs glottalization, that is why the numbers of linked words were given in comparison to the number of glottalized words. As we wanted to see whether there is any contrast in the distribution of different types of linking, linking phenomena were put into six categories according to the phonetic environment in which they occurred:

1. obstruent (obstruent + vowel)

2. sonorant (sonorant + vowel). The /r/ sound was put into a separate category as it causes different linking phenomenon

3. r (/r/ sound + vowel)

4. i (/i:/, /ɪ/, /eɪ/, /aɪ/, /ɔɪ/ + vowel)

5. u (/u:/, /ʊ/, /aʊ/, /ɔʊ/, /əʊ/ + vowel)

6. schwa (/ə/ + vowel)

The annotated data from Excel tables were afterwards used for further processing in the statistic software R, where they were visualized using the ggplot2 package (Wickham, 2009).

4.2 Results
4.2.1 General Results

First, we will have a look at the overall distribution of linking and glottalization in both read and spontaneous speech. As can be seen in **Figure 1**, the tendency to link words increased among all the speakers. In read speech, the number of linking is considerably low, while glottalization is a prevailing phenomenon. The data is significantly different for spontaneous speech, where the number of linked words is prevailing for most of the speakers.
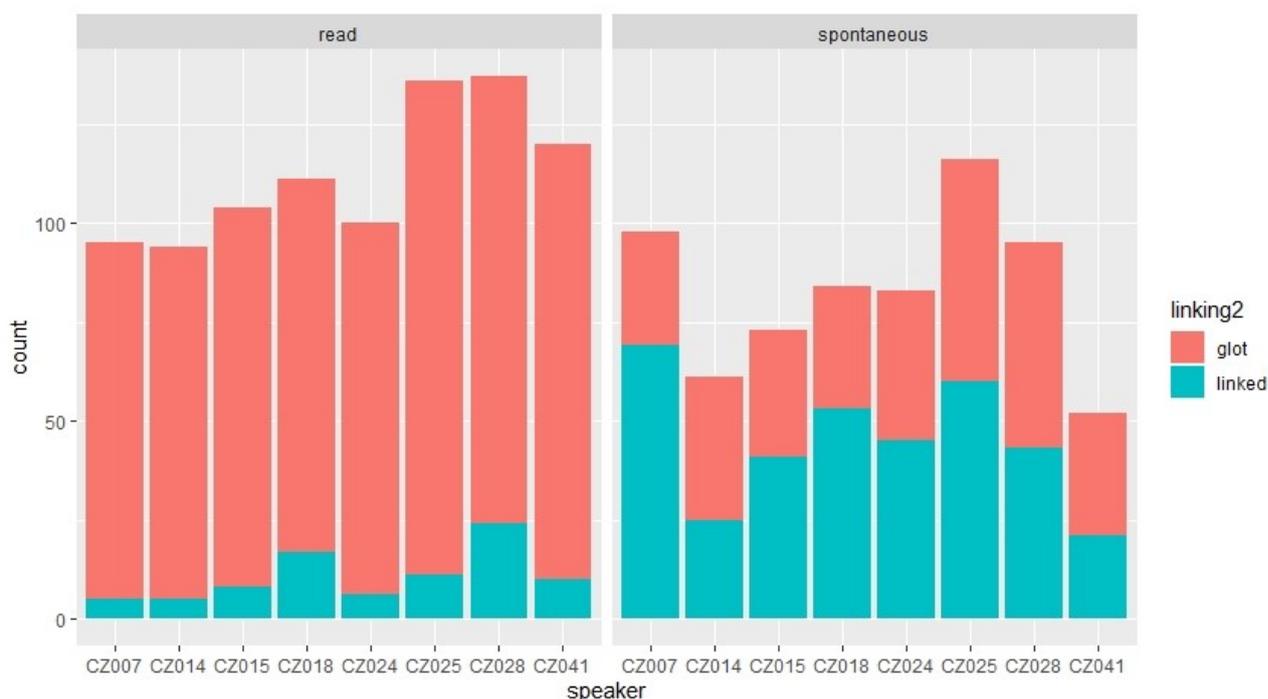

Figure 1. Bar chart showing the overall number of linking vs glottalization for all speakers in reading and spontaneous speech.

For better comparison, **Figure 2** demonstrates the same data in percentage. All the speakers had a relatively low linking rate in read speech, between 5% and 15%. Such a tendency had dramatically changed in a couple of years where we can see that the miminum linking rate is 40% and the maximum is around 70%.
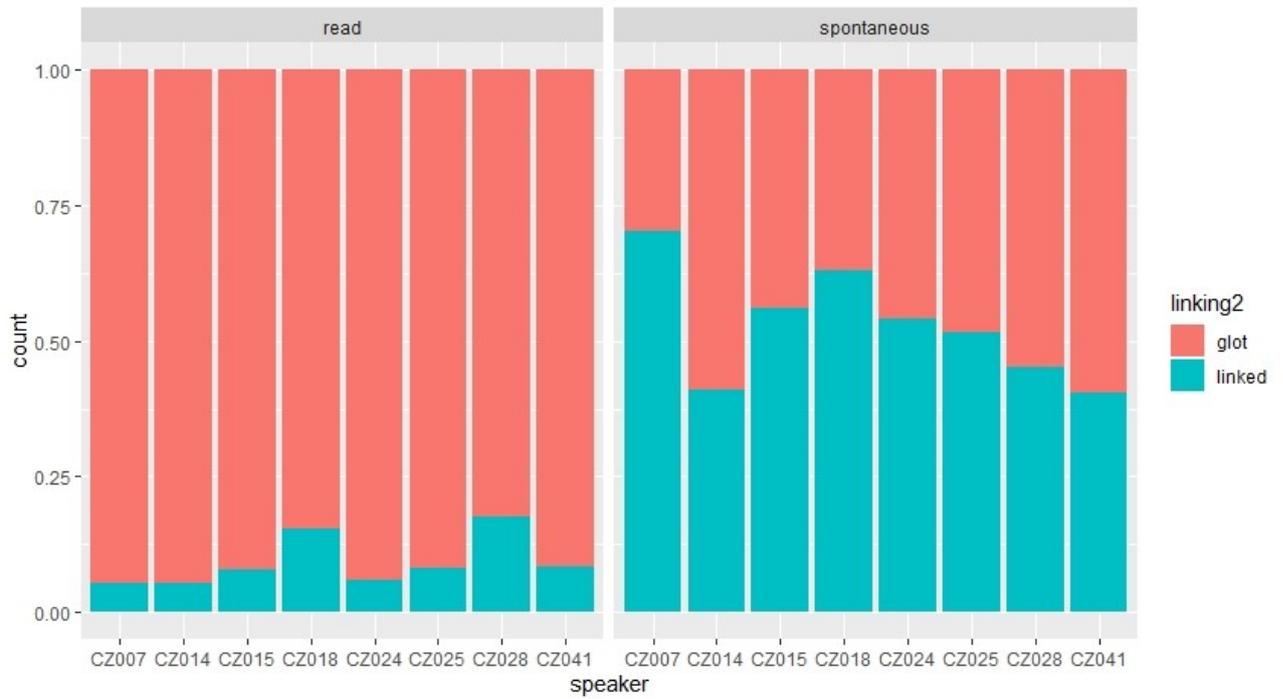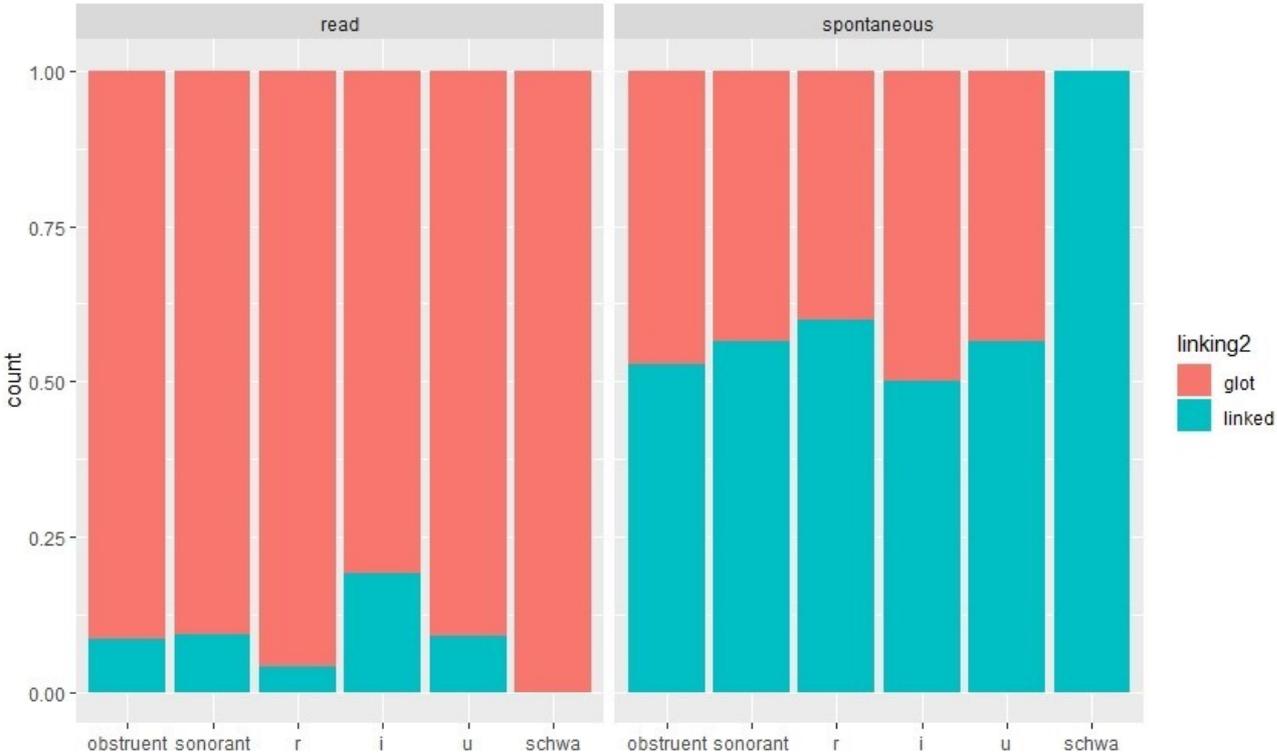
Figure 2. Percentage bar chart demonstrating the distribution of linking and glottalization for all speakers in reading and spontaneous speech.
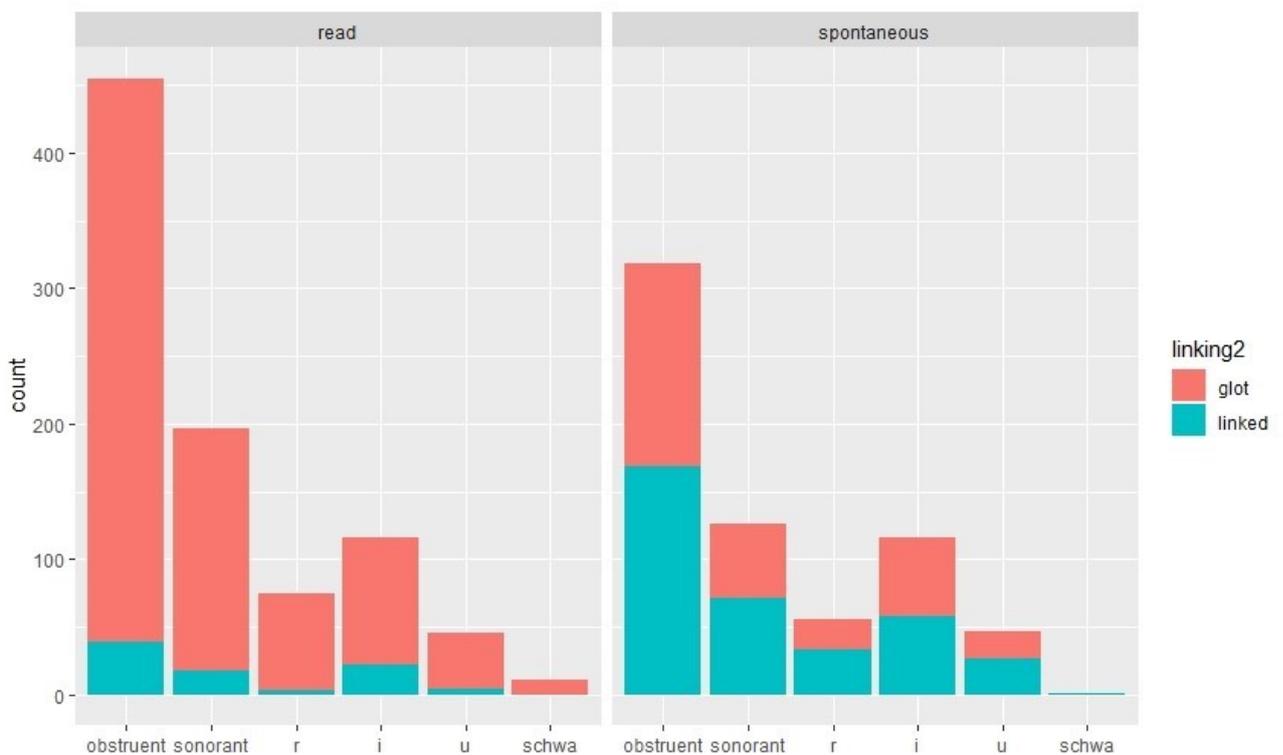
Figure 3. Bar plot showing the number of linking vs glottalization according to the phonetic environment for all speakers.

The comparison of the distribution of linking in different phonetic environment illustrated in

Figure 3, shows that the speakers almost did not use linking r and transient w in read speech. The

results were slightly better for resyllabification, which occurs with obstruents and sonorants, and there

seems to be no significant difference in resyllabification rates for these two groups of sounds. The data demonstrating the occurrence of linking in the phonetic environment / ə/ + vowel will not be commented upon as there was only one occurrence of it in spontaneous speech.
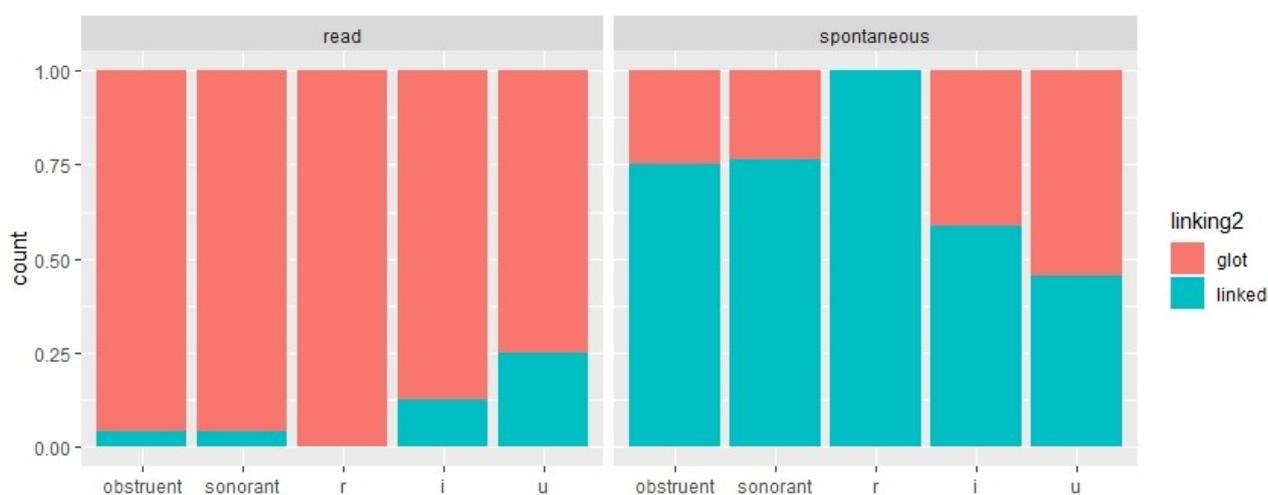

Figure 4. Percentage bar chart demonstrating the distribution of linking vs glottalization for all speakers according to the phonetic environment.

For better demonstration, **Figure 4** illustrates the same data in percentage. As can be seen, the progress in the use of resyllabification and the transient w is almost the same. The percentage bar chart shows that the speakers in their first year of studies most frequently used the transient j type of linking, 15% of linked words to 85% glottalized. The linking r type had the lowest rate, 5% of linked words to 95% glottalized. The opposite tendency may be observed in the recordings taken after a couple of years. There is a considerable progress in the use of linking r and transient w, from 5 % to 60% of linked words and 10% to 55% accordingly. The improvement in the use of transient j is a bit less significant, from 15% to 50% of linked words. The same tendency may be observed for resyllabification, the improvement is 10-11% to 52-55% (obstruents and sonorants).

4.2.2 Individual Results

After we have observed the general tendencies in the use of linking and glottalization in both speech styles, it would be interesting to see the way in which individual speakers used these connected speech phenomena. Speaker CZ007 demonstrated a significant progress as shown in **Figure 5.** The biggest improvement is seen in resyllabification, from 2.5% of linked words to 75-77% (with obstruents and sonorants). The improvement in the use of transient j and w is 12.5% to 60% and 25% to 45% respectively. 100% success in using the linking r is not demonstrative as quantitatively there were only 10 occurrences of this phenomenon in read and 5 occurrences in spontaneous speech.
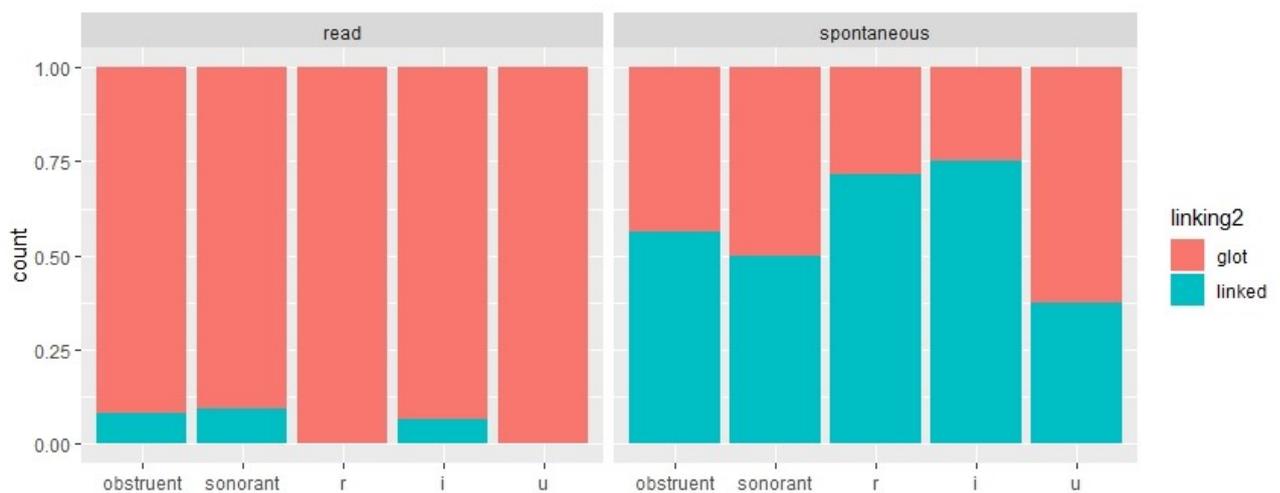

Figure 5. Percentage bar chart demonstrating the distribution of linking vs glottalization for speaker CZ007 according to the phonetic environment.

The progress of speaker CZ014 shown in **Figure 6** is not as significant. The only type of linking used by the speaker in read speech is that of resyllabification, which occurs with obstruents and sonorants. All linking indices for spontaneous speech are under 50%, which means that glottalization remained prevailing over linking phenomena. The biggest improvement can be observed in the transient j type. It was not used by the speaker in read speech, but the picture has changed in spontaneous speech with 45% of linked words. The resyllabification rate has also increased, from 2.5% to 42% of linked words with obstruents and 12.5% to 40% with sonorants. The increased use of the linking r type is not as significant, with only 25% of linked words.
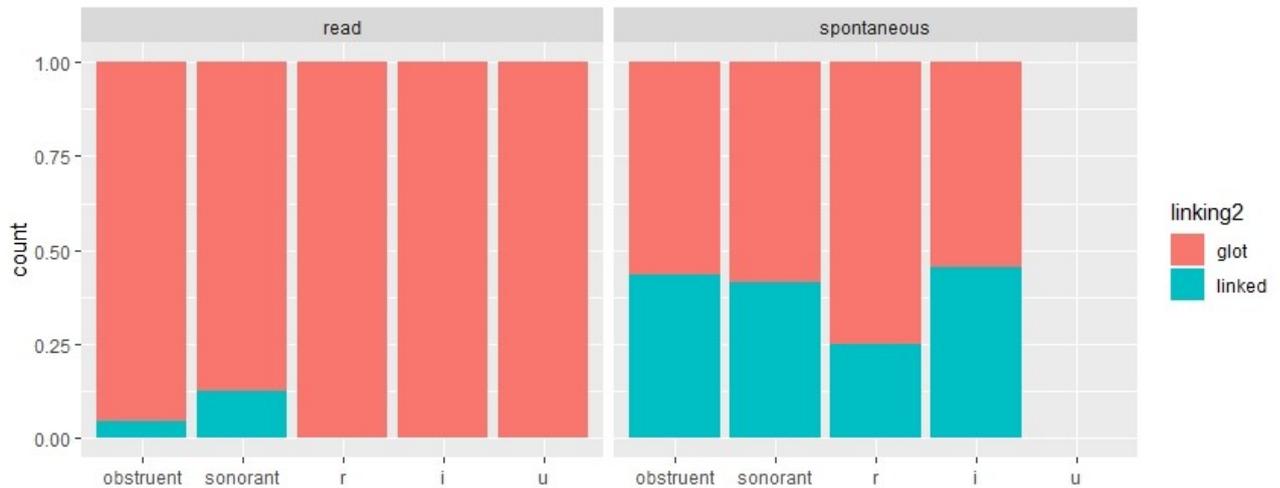
Figure 6. Percentage bar chart demonstrating the distribution of linking vs glottalization for speaker CZ014 according to the phonetic environment.
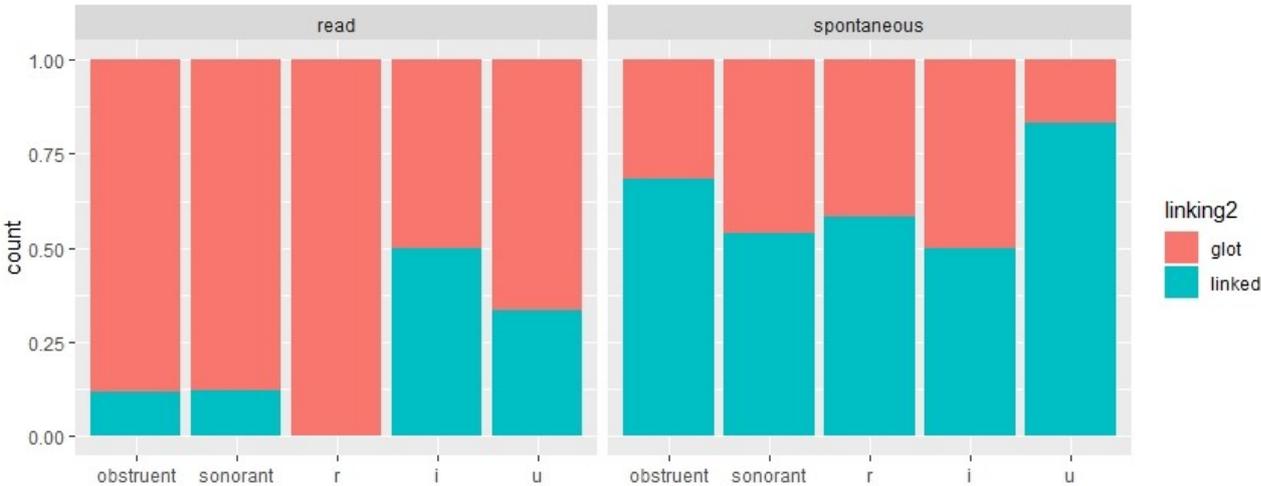
The distribution of linking for speaker CZ015 is illustrated in **Figure 7**. The results for the transient w type cannot be taken into consideration because quantitatively there was only 1 instance of this phenomenon in the speech. A considerably large difference may be seen in the use of linking r and transient j, 0% to 70% and 6.25% to 75% accordingly. The improvement in the use of resyllabification is slightly less significant, with 8.5% to 55% with obstruents and 10% to 50% with sonorants.

Figure 7. Percentage bar chart demonstrating the distribution of linking vs glottalization for speaker CZ015 according to the phonetic environment.

The results for speaker CZ018 demonstrated in **Figure 8** show the same tendency in the use of linking r. The speaker did not use this type of linking in the first year of studies, but the rate increased to almost 60% in spontaneous speech. There is no change in the use of transient j. There is quite a significant difference in the use of transient w, with 35% to 85%. Resyllabification rates have

also increased in a considerably large number, with 12% to 65% with obstruents and 12.5% to 52.5% with sonorants.

Figure 8. Percentage bar chart demonstrating the distribution of linking vs glottalization in percentage for speaker CZ018 according to the phonetic environment.

Speaker CZ024 has the lowest resyllabification rate in read speech (obstruent + vowel) with only 1.25% of linked words, the numbers are slightly better with sonorants, 11.5% as can be seen in **Figure 9**. The significant change is observed in the distribution of linking r, the rates have changed from 13.5% to almost 75% of linked words. While the transient w type was not used by the speaker in read speech, the distribution of this phenomenon is 38.75% in spontaneous speech.
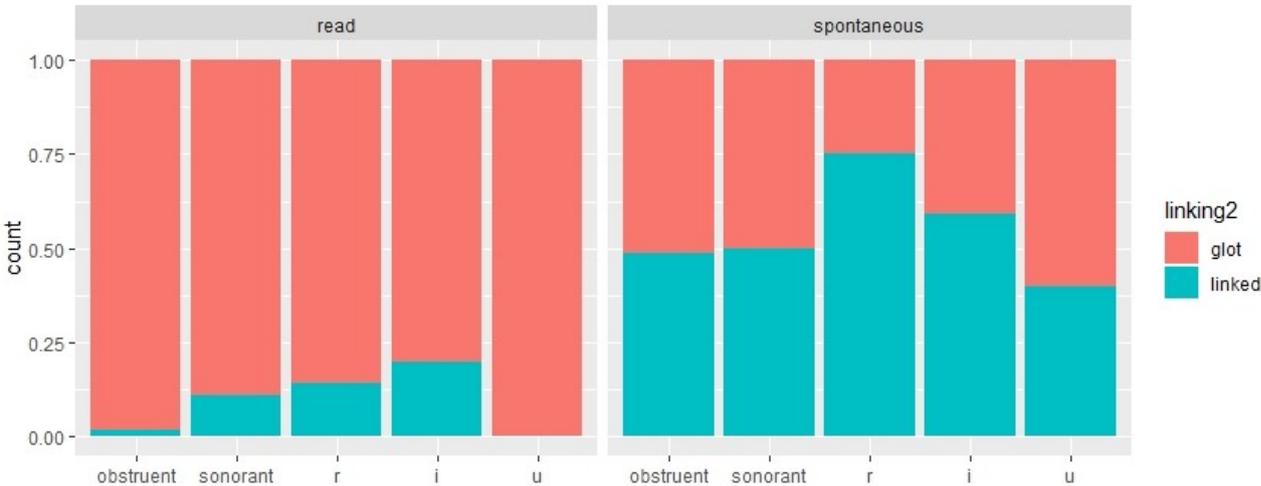
Figure 9. Percentage bar chart demonstrating the distribution of linking vs glottalization for speaker CZ024 according to the phonetic environment.

The transient j type of linking was prevailing in read speech of speaker CZ025 with 15% of linked words, as shown in **Figure 10**. Resyllabification and linking r were used in less than 10% of

cases, the transient w was not used by the speaker. A considerable improvement is observed with the use of linking r and transient w, from 7.5% to 85% and 0% to 75% accordingly. There is almost no improvement in the use of transient j. Resyllabification rates have increased from 10% to 52-60% (obstruents and sonorants).
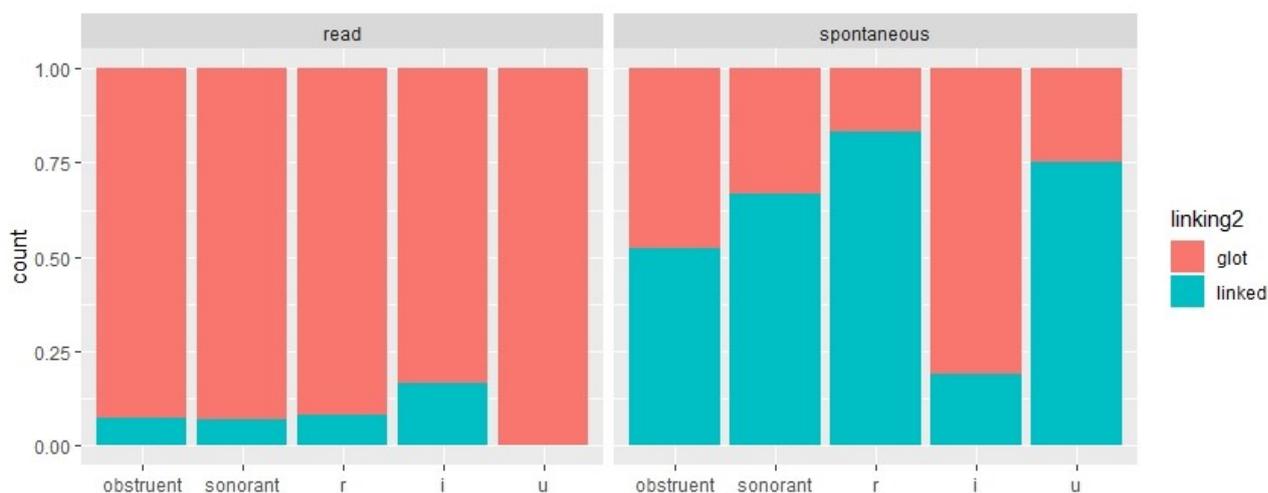
Figure 10. Percentage bar chart demonstrating the distribution of linking vs glottalization for speaker CZ025 according to the phonetic environment.

Speaker CZ028 has a relatively high rate of the transient j distribution in read speech, around 45%, as demonstrated in **Figure 11**, but almost no improvement may be observed in the spontaneous speech index, which increased only to 50%. The most considerable change occurred in the distribution of the transient w type of linking, which increased from 0% to 65%. Resyllabification occurred more frequently with sonorants, around 60%, in comparison to 32.5% of linked words with obstruents (in spontaneous speech). The progress in the distribution of linking r is relatively high, from 7.5% to 60% of linked words.
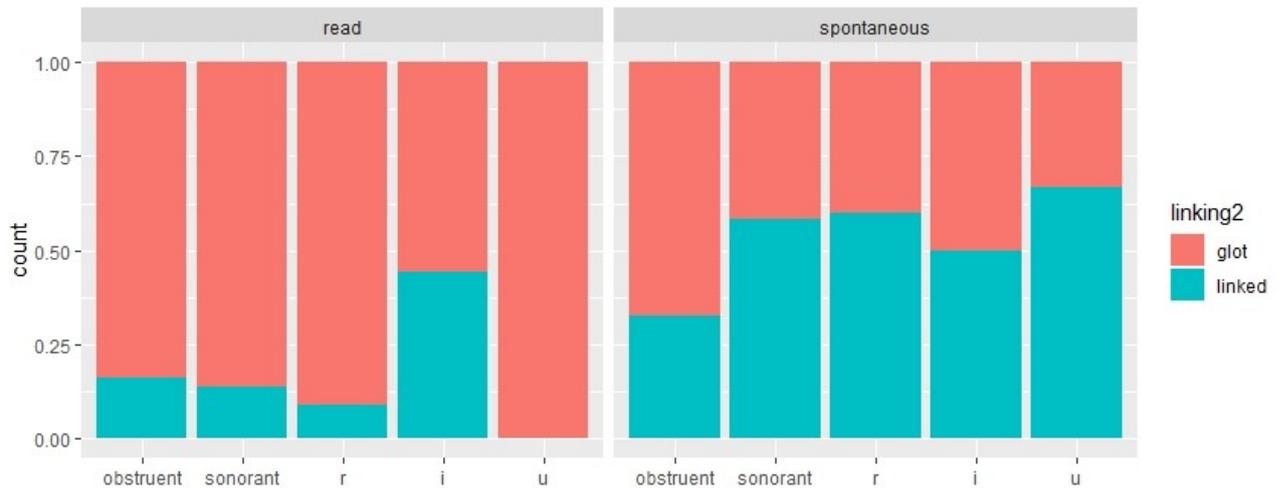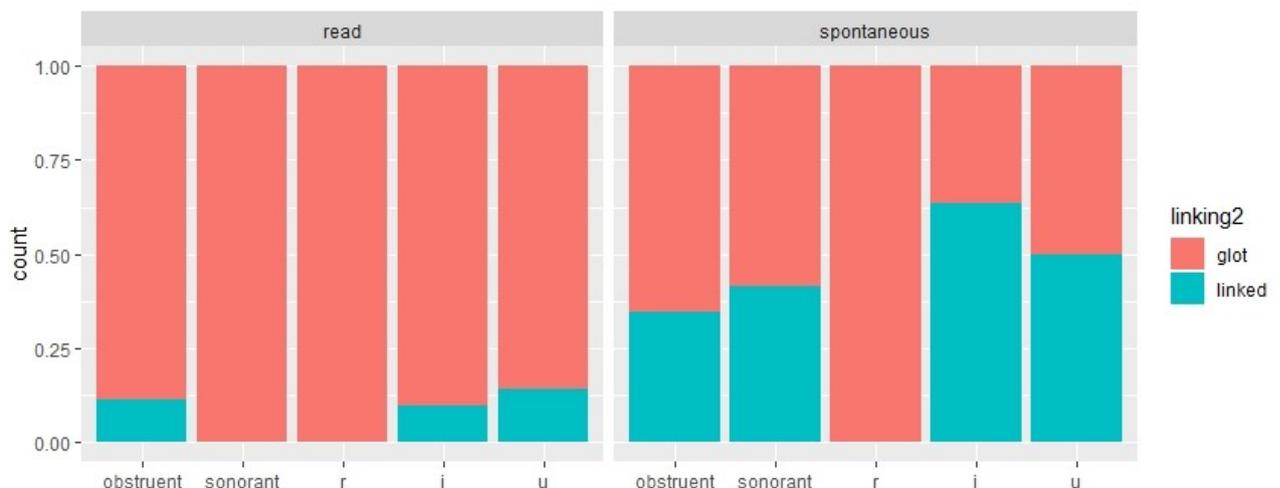
Figure 11. Percentage bar chart demonstrating the distribution of linking vs glottalization for speaker CZ028 according to the phonetic environment.

Linking rates for speaker CZ041 are demonstrated in **Figure 12**. The noticeable difference from other speakers is the 0% distribution of resyllabification in the sonorant + vowel environment in read speech. There is also no occurrence of linking r in spontaneous speech which was not observed in rates of other speakers. The maximum progress is in the use of transient j, the rate increased from 10% to 52% of linked words. The improvement of the transient w use should not be taken into consideration as quantitatively there were only two occurrences of this phenomenon in spontaneous speech. The overall progress of linking is relatively low.

Figure 12. Percentage bar chart demonstrating the distribution of linking vs glottalization for speaker CZ041 according to the phonetic environment.

4.3 Discussion

After a closer examination of the obtained data, several trends may be observed. In their first year of studies, all the students had a low level of linking performance. 5 out of 8 speakers did not use the linking r, as well as 5 out of 8 speakers did not use the transient w. None of the speakers used the intrusive r in read speech. However, resyllabification phenomena were used by each speaker and only 1 out of 8 speakers did not use the transient j type of linking. As linking rates for read speech were extremely low, we can suppose that participants did not have any preliminary training in linking. Thus, resyllabification and transient j are linking phenomena that were naturally used by the speakers. Recordings that were taken a couple of years later helped us to trace whether there were any changes in the way participants link words. The results have proved our expectations. Linking rates of all participants became considerably higher. While in read speech glottalization prevailed over linking, the opposite tendency is observed in spontaneous speech rates. The progress is seen in all types of linking that were demonstrated in the results section. Sardegna (2011) examined the importance of pronunciation learning strategies for improving linking in SLA, and the results of our study are similar, as linking scores of all participants have increased significantly. We also wanted to see if there is any contrast in the improvement of the distribution of linking according to the phonetic context. There are two phonetic contexts in which the most significant progress occurred (when comparing individual rates of participants): /r/ sound + vowel (linking r) and /uː/, /ʊ/, /aʊ/, /ɔʊ/, /əʊ/ + vowel (transient w). These types of linking were the least frequently used by participants in read speech but had the highest rates in spontaneous speech. The frequency with which such phonetic contexts occur in English language is comparatively low (as opposed to the resyllabification and the transient j). Thus, there is a lesser probability for language learners to hear and imitate these types of linking. After they become familiar with the existence of such phenomena, they start to use them actively.

5. Phrasing
5.1 Method

Boundaries of prosodic phrases were identified on the basis of both temporal and melodic features such as final lengthening and melodic movement, or a pause in some cases. We did not include pauses and hesitations into prosodic phrase boundaries. Prosodic phrases built with accordance to the syntactic or semantic structure of a sentence were marked as 'log' (logical), those that did not meet these requirements were marked as 'illog' (illogical). The number of syllables per prosodic phrase was extracted using the script in Praat. Syllables were counted as the number of vowels. Speech rate was measured by dividing the number of syllables in a prosodic phrase by the duration of the prosodic phrase in seconds.

The variables measured were:

- number of logical/illogical phrases

- number of syllables per prosodic phrase

- speech rate in syllables per second

The data were extracted using a script in Praat and then exported to tables using the Microsoft Excel and sorted according to the speech style (read/spontaneous). The annotated data from Excel tables were afterwards used for further processing in the statistic software R, where they were visualized using the ggplot2 package (Wickham, 2009).

5.2 Results

As can be seen in **Figure 13**, the overall tendency observed among all speakers is that they have a high rate of logically built (with accordance to the syntactic and semantic structure of a sentence) prosodic phrases.
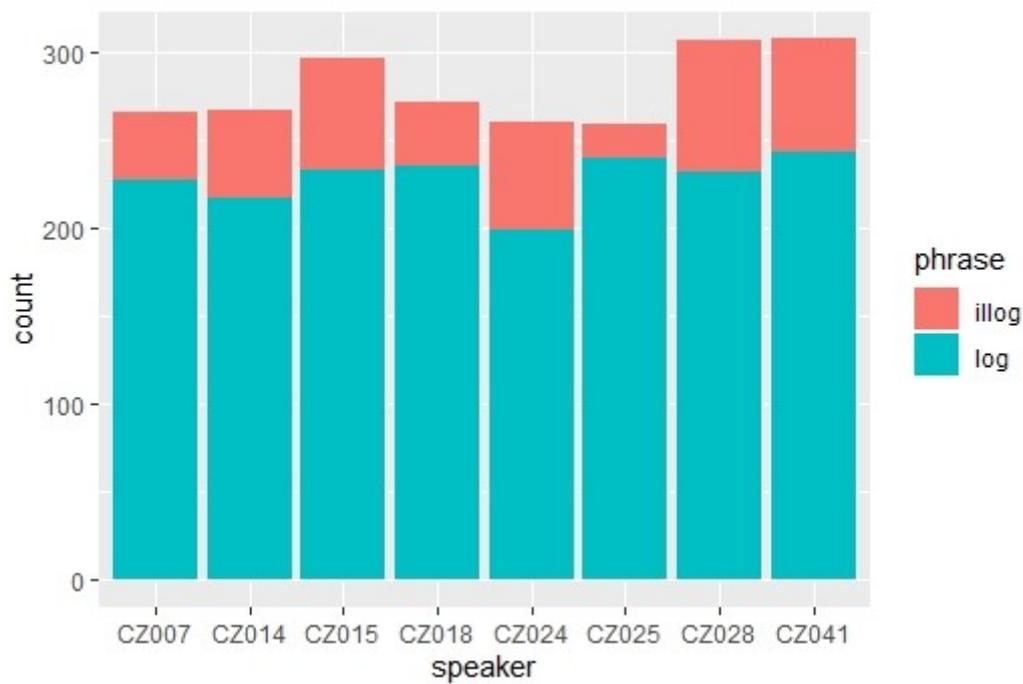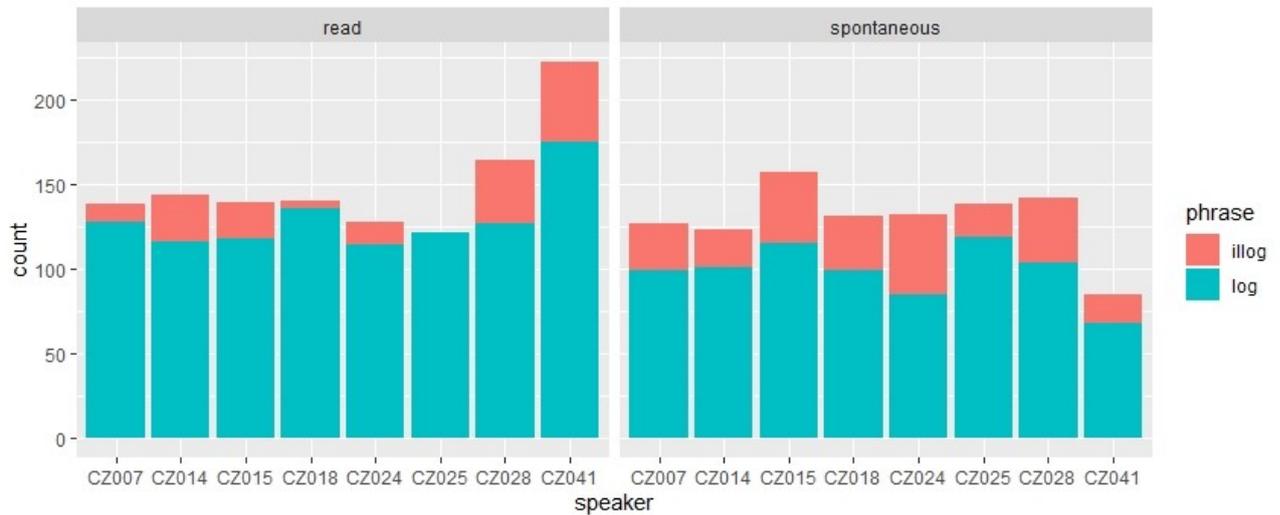

Figure 13. Bar plot showing the overall distribution of logical ("log") and illogical ("illog") phrases in both speech styles.

To see whether there is a significant difference in the distribution of logical/illogical phrases between read and spontaneous speech, we will look at **Figure 14.** We can see that the speakers are slightly less successful in building logical prosodic phrases in spontaneous speech, though the difference is not significant. This may happen due to the difference in speech production, as when producing spontaneous speech, speakers have less time for speech planning, thus they tend to have a higher number of illogical phrases.

Figure 14. Bar plot showing the distribution of logical and illogical phrases for read and spontaneous speech.

Figure 15 demonstrates temporal comparison of prosodic phrases. The most frequent phrase length for read speech is 5 syllables and 3 syllables for spontaneous speech. The overall trend for both types of speech is that speakers tend to build relatively short prosodic phrases. Speakers CZ028 and CZ041 have an extremely high number of phrases in read speech when comparing with other speakers, which shows that their speech was very fragmented.
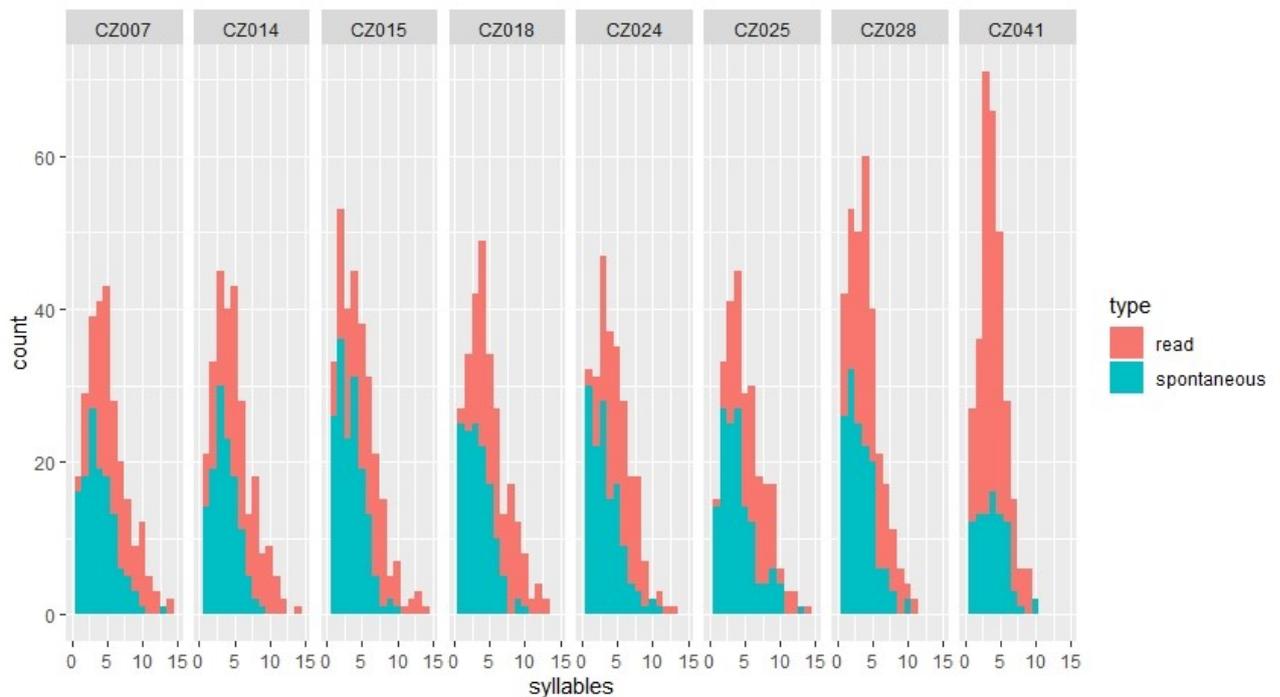
Figure 15. Histogram of syllable counts per prosodic phrase for read and spontaneous speech.

The following graphs will illustrate the comparison of speech rate distribution in both speech styles. There is no significant difference in speech rate scores between read and spontaneous speech

as can be read from the boxplot demonstrated in **Figure 16**. Only 3 (CZ 007, CZ024 and CZ028) out of 8 speakers had a considerably higher speech rate in read speech, which is only 37.5% of the whole data. Thus, such result is insignificant to draw any general conclusions. The median for all speakers in both speech styles is between 3 and 5 syllables/second.
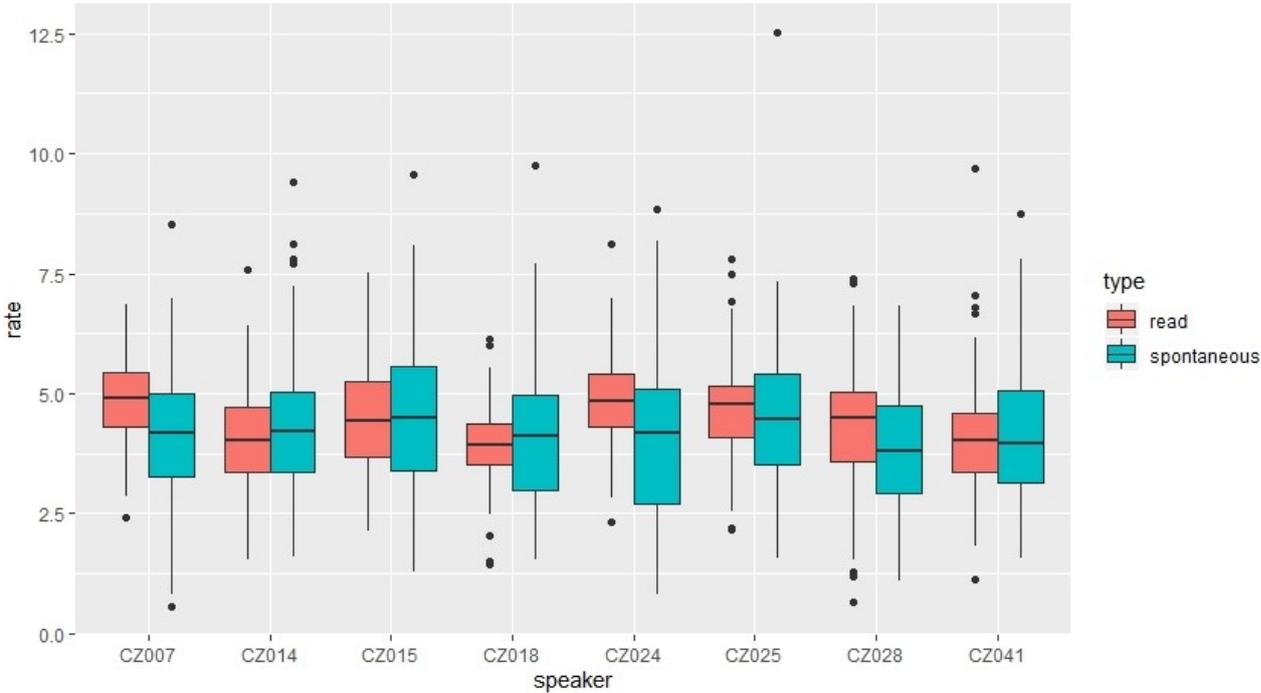
Figure 16. Boxplot demonstrating the speech rate for read and spontaneous speech. 95 % of the data are situated within the range of vertical line, the other 5% of extreme data are indicated with marks. The box demonstrates the middle portion of the data. The horizontal line in the center of the box is the median (the middle of a data set).
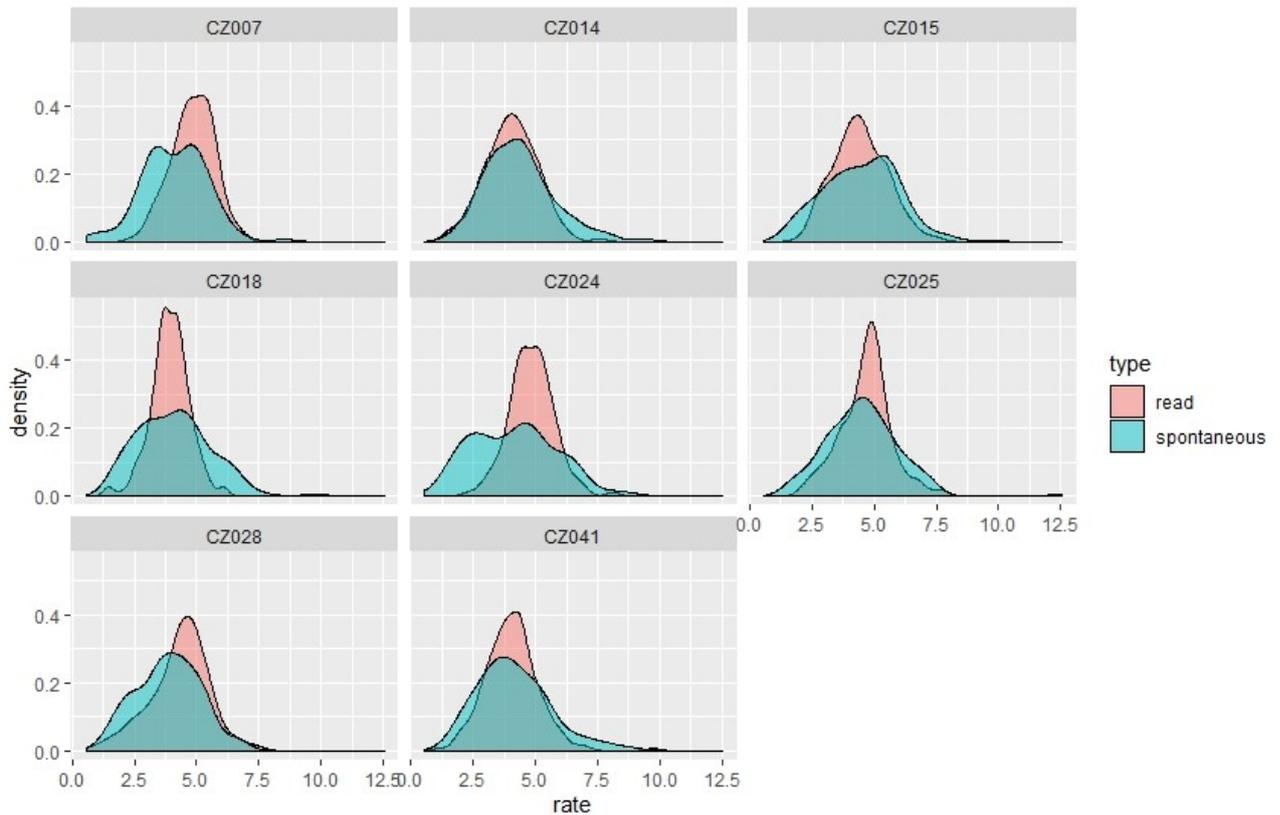
Figure 17. Density plot comparing the speech rate distribution in both speech styles.

Figure 17 provides a better visual comparison. Read speech plots look comparatively similar as they all have a steep curve and only one peak, which illustrates that the speech rate is relatively stable, the average ranging from 3 to 6 syllables/sec. The same cannot be said about spontaneous speech plots, as the curves are smoother and wider in range, which means that the speech rate is not as stable as in the previous case. Overall, there is a higher amount of data with less than 3 syllable/sec for spontaneous speech. The data for the speakers CZ007 and CZ024 are significantly different as there are two maxima of data set, 3 and 5 syllables/sec for the speaker CZ007; 2.5 and 4.5 for the speaker CZ024.

5.3 Discussion

In the results section, we compared the way participants build prosodic phrases, the length of prosodic phrases, and the speech rate in both sets of recordings. While participants had low linking rates in read speech, they were successful in building logic prosodic phrases, with accordance to the lexical and syntactic structure of a sentence. The rates were slightly lower for spontaneous speech, but the difference was not that significant to draw any conclusions. As to the comparison of the length of prosodic phrases, the general trend that was observed is that participants tended to build longer prosodic phrases in read speech and shorter ones in spontaneous speech (the most frequent length 5 syllables/phrase in read to 3 syllables/phrase in spontaneous speech). There may be two possible explanations of this phenomenon:

- participants had sufficient time for preparation before they read the news report, while spontaneous speech was improvised, thus it involved more complex cognitive processes as a result, participants built shorter phrases
- participants were less skillful in speaking

When measuring speech rate, we excluded pauses and hesitations and divided the number of syllables in a prosodic phrase by its duration in seconds. This means that our data is closer to articulation rate, as speech rate includes pauses. The median of both data sets varied from 3 to 5 syllables/second, with no significant difference between the two types of speech. When comparing the absolute data, the spontaneous speech was more varied, with higher density of data with speech rate lower than 3 syllables/second. Not only participants built shorter phrases in spontaneous speech, but they also spoke with a less stable speech tempo. However, the change observed is not that considerable to claim that their speech became less fluent. As the production of spontaneous speech requires more time for speech planning, it has a negative impact on the speakers' prosodic performance.

6. General Discussion

The obtained data show several trends that will be commented upon. Let us first discuss linking. The pronunciation of linking in the first year of studies was highly influenced by participants' L1, they tended to glottalize initial vowels as they do it in Czech. A different situation was observed in recordings taken a couple of years later. The pronunciation training of linking has proved to be effective, as the results for later recordings show that all the speakers have significantly improved their linking skills. However, we may suppose that the results might have been different if we were comparing recordings of the same speech style. This remains an open question. We may suggest that if participants had enough time to plan their speech, they would have performed better. In spontaneous speech, they were focused on different aspects of speech planning, as the dialogues were improvised, thus this might have influenced their performance of linking.

When we compare linking with prosodic performance, we may see that while participants had low linking rates in their first year of studies, they were successful in building logical prosodic phrases. There are several possible explanations for this phenomenon:

- during their previous studies students were trained in prosodic phrasing

- prosodic phrasing is easier to identify and imitate than linking

- there is little difference between Czech and English prosodic phrasing

However, no significant differences in the distribution of logical phrases (structured with accordance to the syntactic and semantic structure of a sentence) between read and spontaneous speech were observed, as well as no dynamics was seen in temporal characteristics. The rates were slightly, but not considerably, lower for spontaneous speech. The speech style has a strong influence on prosodic performance. Thus, we cannot say whether there was any worsening of prosodic phrasing skills.

Our hypothesis – that English Phonetics and Phonology course that students have in the first year of studies improve their fluency performance – was partially confirmed by this study. There was a great progress of linking skills; the results for prosodic phrasing were rather neutral, but participants have shown a high level of proficiency in structuring logical prosodic phrases in both speech styles.

While there seems to be no connection between these two aspects of fluency, some may still be found. Both linking and structuring of prosodic phrases seem to depend on the speech style to a lesser degree than temporal characteristics of prosodic phrasing do. That is why it is advisable to conduct a study that would compare the data of the same speech style to examine whether specific pronunciation training has an impact on temporal characteristics of prosodic phrasing.

7. Conclusions

In this study we wanted to investigate the effectivity of pronunciation training of some of the suprasegmentals such as linking and prosodic phrasing. The nature of this study was descriptive. We compared the recordings of the same speakers that were taken at different stages of their studies. This helped us to see whether there were some dynamics in the distribution of these fluency aspects. We expected to observe a progress in both cases.

The theoretical background of this thesis provided useful information about the concept of fluency and its multiple aspects, prosodic phrasing and connected speech phenomena, and previous research in these areas. The methodological section described the process of material compilation, its preparation for future analysis, the methods used when working with the recordings and analyzing the obtained data. The results of this study have shown that specific pronunciation training of suprasegmentals has a positive impact on the fluency performance. The significant increase has been observed in the use of linking phenomena, while prosodic phrasing characteristics remained on almost the same level. However, the results with prosodic phrasing may be problematic as in this study we compared recordings of different speech styles, read and spontaneous, and while this has little effect on the performance of linking, it may influence temporal characteristics of prosodic phrasing. A further research is needed to prove or disprove this hypothesis.

8. References

Alameen G. & Levis, J. M. (2015). Connected Speech. In: Reed M. & Levis M. J. (eds.), *The Handbook of English Pronunciation*, 159-174. Oxford: Blackwell.

Cavallieri, I. R., Wigglesworth, G. & Maxwell, O. (2019). Second Language Fluency: Re-thinking Utterance Fluency from a Phonetics-Phonology Interface. *Proceedings of the ICPhS Conference*, 2019, Melbourne, 2625-2629.

Clifton C. Jr., Carlson, K., Frazier, L. (2002). Informative Prosodic Boundaries. *Language and Speech*, 45 (2), 87-114.

Dragojevic, M., Giles, H., Beck A.-C. & Tatum, N. T. (2017). The Fluency Principle: Why Foreign Accent Strength Negatively Biases Language Attitudes. *Communication Monographs*, 84 (3), 385-405.

Flege J. E. (1988). Factors Affecting Degree of Perceived Foreign Accent in English Sentences. *Journal of the Acoustical Society of America,* 84 (1), 70-79.

Frazier, L., Carlson, K., & Clifton, C. Jr (2006). Prosodic Phrasing is Central to Language Comprehension. *Trends in Cognitive Sciences*, 10(6), 244–249.

Gick, B. (1999). A Gesture-Based Account of Intrusive Consonants in English. *Phonology* 16, 29-54.

Gráf, T. (2017). Repeats in Advanced Spoken English of Learners with Czech as L1. *Acta Unuversitatis Carolinae: Philologica*, 65-78.

Guz, E. (2015) Establishing the Fluency Gap Between Native and Non-Native Speech. *Research in Language*, 13 (3), 230-247.

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

Hieke, A. E. (1984). Linking as a Marker of Fluent Speech. *Language and Speech*, 27 (4), 343-354.

Jun, S-A. (2010). The implicit prosody hypothesis and overt prosody in English. *Language and Cognitive Processes*, 25 (7-9), 1201-1233.

Kohler, K.J. (2010). The Transmission of Meaning by Prosodic Phrasing: A Comparison of French with English and German Using No Ls and Hs. *Phonetica*, 67, 100-124.

Kormos, J., & Dénes M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32, 145–164.

Lengeris, A. (2012). Prosody and Second Language Teaching: Lessons from L2 Speech Perception and Production Research. In: J. Romero-Trillo (ed.), *Pragmatics and Prosody in English Language Teaching*, 25-40. Dordrecht: Spiegel.

Lindblom, B. (1990). Explaining Phonetic Variation: A Sketch of the H&H Theory. In: Hardcastle, J. & Marchal, A. (eds.), *Speech Production and Speech Modelling,* 403-439. Dordrecht: Kluwer Academic Publishers.

Oppenheimer, D. M. (2008). The Secret Life of Fluency. *Trends in Cognitive Sciences,* 12 (6), 237-241.

Pinget, A-F., Bosker, H.R., Quené H. & de Jong, N. H. (2014). Native Speakers' Perceptions of Fluency and Accent in L2 Speech. *Language Testing,* 31 (3), 349-365.

Piske, T. (2012). Factors Affecting the Perception and Production of L2 Prosody: Research Results and Their Implications for the Teaching of Foreign Languages. In: J. Romero-Trillo (ed.), *Pragmatics and Prosody in English Language Teaching*, 41-59. Dordrecht: Spiegel.

R Core Team (2017). R: *A language and environment for statistical computing (version 3.3.2)*. Vienna: R Foundation for Statistical Computing. Retrieved from http://www.Rproject.org.

Reitbrecht, S. & Hirschfeld, U. (2015). The Impact of Fluency and Hesitation Phenomena on the Perception of Non-Native Speakers by Native Listeners of German. *Proceedings of the ICPhS Conference*, 2015, Glasgow, 0166.1-4.

Sardegna, V. G. (2011). Pronunciation Learning Strategies that Improve ESL learners' Linking. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference*, September 2010, Ames, IA: Iowa State University, 105-121.

Shattuck-Hufnagel, S. (2000). Phrase-Level Phonology in Speech Production Planning: Evidence for the Role of Prosodic Structure. In: Horne, M. (ed.), *Prosody: Theory and Experiment,* 201-229. Berlin: Kluwer Academic Publishers.

Skehan, P. (2009). Modelling Second Language Performance: Integrating Complexity, Accuracy, Fluency, and Lexis. *Applied Linguistics*, 30 (4), 510–532.

Šimáčková, Š., Podlipský, V.J., & Kolářová, K. (2014). Linking Versus Glottalization: (Dis)connectedness of Czech-Accented English. *Concordia Working Papers in Applied Linguistics,* 5, 678-692.

Tseng, C., Pin, S., Lee, Y., Wang, H. & Chen, Y. (2005). Fluent Speech Prosody: Framework and Modeling. *Speech Communication*, 46, 284-309.

Volín, J. (2006). *IPA-Based Transcription for Czech Students of English*. Praha: Karolinum.

Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. New York: Springer

9. Resumé

Hlavním cílem této bakalářské práce bylo porovnat nahrávky studentů anglistiky-amerikanistiky a zjistit, zdá se plynulost jejich řeči zlepšila po tom, co absolvovali kurz fonetiky a fonologie angličtiny v prvním ročníku studia. Konkrétně jsme se zaměřili na takové aspekty plynulosti řeči, jako je například vázání a prozodické frázování. U vázání jsme spočítali korelaci výskytů mezi vázanými a glotalizovanými slovy. U prozodického frázování jsme analyzovali logiku jeho strukturování a také temporální charakteristiky: délku a mluvní tempo. Nejvíc nás zajímaly rozdíly mezi úspěšností použití těchto jevů v různých fázích studia..

Teoretický podklad této bakalářské práce je představen v kapitole 2. Jsou zde popsané základní koncepty a také výsledky předchozích studií, které jsou relevantní pro náš výzkum. V uvedené kapitole začínáme vysvětlením obecného pojmu plynulosti řeči. Plynulost je velmi komplexní fenomén, který se dá aplikovat v rámci různých oblastí vědy. Jelikož se zabýváme plynulostí řeči, zajímá nás, jak je tento koncept vnímán ve fonetice. Plynulost řeči je primárně temporálním a melodickým jevem, který je součástí mluveného projevu. Abychom lépe pochopili její podstatu v rámci osvojování druhého jazyka, musíme také vzít v úvahu, že plynulost souvisí s dalšími aspekty produkce řeči, jako složitost jazyka a přesnost artikulace. Čím více pozornosti mluvčí věnuje jedné z těchto tří kategorií, tím menší úspěšnost budeme pozorovat ve dvou dalších. Proto je plynulá řeč přímo spojená s automatičností. Vysoká úroveň automatičnosti řeči vede ke zlepšení plynulosti. Vnímání plynulosti řeči může být ovlivněno cizineckým přízvukem, podkapitola 2.2 popisuje podstatu vzniku cizineckého přízvuku a také vysvětluje, proč může mít negativní vliv na evaluaci plynulosti jazyka.

Následující dvě podkapitoly, 2.3 a 2.4, jsou věnovány detailnímu popisu prozódie a souvislé řeči, jelikož v této bakalářské práci byly zkoumány jevy, které prozódie a souvislá řeč zahrnuje. Prozódie je suprasegmentálním aspektem řeči, který zastává důležitou roli v procesu komunikace, a může mít velký vliv na srozumitelnost řeči. Pomocí prozódie má mluvčí možnost vyjadřovat informaci, kterou nelze sdělit v ortografické dimenzi. Proto je velmi důležité věnovat pozornost

principům strukturování prozodických frází. Prozodické frázování je členění slov do jednotek, které jsou spojené se sémantickou a syntaktickou strukturou promluvy. Pomocí prozodických frází mluvčí systematizuje význam mluveného projevu tak, aby ho mohl předat správným způsobem. Hranice prozodických frází se da určit podle melodických (klesající intonace na konci fráze) a temporálních charakteristik (finální zpomalení). Prozodické frázování používáme nejenom při mluveném projevu, ale také při čtení. Prozodické frázování funguje celkově jako rámec, který přispívá k lepšímu zpracovávání řeči a k její memorizaci.

Neméně důležitým aspektem anglické plynulé řeči je vázání. Vázání je jedním z nejpodstatnějších a nejčastějších procesů, které vznikají ve souvislé řeči. Výslovnost izolovaných slov se může lišit od výslovnosti použité v souvislé řeči. Hlavním důvodem, proč dochází k takové změně, je to, že se mluvčí snaží vyslovovat slova maximálně šetrným způsobem, tzn. si usnadňuje proces artikulace. Kvůli tomu, že mohou nastávat takové změny, se může proces porozumění řeči stát obtížným pro nerodilého mluvčího angličtiny. Aby nedocházelo k problémům s porozuměním souvislé řeči, je zásadní, aby se při výuce jazyka studenti seznámili s jevy, jež jsou typické pro souvislou řeč. Je velmi důležité věnovat pozornost trénování výslovnosti vázání, protože právě vázání je jedním z takových parametrů, podle kterého lze ohodnotit úroveň plynulosti řeči mluvčích. Při vázání dochází k resylabizaci (souhlásky se připojují k slabikám, které patří k jinému slovu) segmentů bez jakékoliv dálší změny. Slovo, které začíná samohláskou, se spojuje s předchozím. V závislosti na fonetickém kontextu můžeme rozlišovat několik druhů vazání. Jelikož výslovnost češtiny se v tomto smyslu hodně liší od angličtiny, je velmi zajímavé sledovat, jakým způsobem čeští mluvčí angličtiny používají právě tyto zkoumané jevy. V češtině jsou počáteční samohlásky glotalizované. Jelikož je výslovnost v cizím jazyce ovlivněna rodným jazykem, je pro anglickou řeč českých mluvčích se silným přízvukem typická glotalizace počátečních samohlásek. Je to jeden z hlavních důvodů, proč nás zajímalo, zda se výslovnost vázání zlepšila u studentů po absolvování fonetického kurzu.

Náplní kapitoly 3 je popis materiálu, který byl použit v rámci výzkumu. Vybrali jsme 16 nahrávek 8 českých mluvčích angličtiny se sílným přízvukem, s jazykovou úrovní B2, všichni byly studentky anglistiky. Jedna část nahrávek byla vzatá z Pražského fonetického korpusu. Nahrávání

bylo požadovanou součástí kurzu fonetiky a fonologie angličtiny, který absolvovali studenti prvního ročníku. Účastníci během nahrávání četli BBC zprávy. Druhá část nahrávek se skládá z LINDSEI dialogů, které byly nahrané o několik let později. Z dialogů bylo vybíráno více úseků, a to vždy z prostřední části nahrávky. Vázání a prozodické členění bylo analyzováno poslechem pomocí softwaru Praat, kde každý z výše zmíněných fenoménů byl ručně označen.

Zaměřili jsme se na korelaci počtu výskytů mezi glotalizací a vázáním slov. U prozodického členění jsme zkoumali strukturování prozodických frázi a jejich členění v souladu se syntaktickým a sémantickým významem věty. Fráze, které byly postavené v souladu s těmito podmínkami byly označené jako logické, ty, které nesplňovaly tyto požadávky, jsme oznacovaly jako nelogické. Hranice frází byly určené podle temporálních a melodických charakteristik. Také jsme změřili jejich délku a mluvní tempo. Délka byla spočítaná jako počet slabik ve frázi. Mluvní tempo bylo vypočítáno ve slabikách/sekundu. Veškeré ůdaje byly extrahováne pomocí skriptu a přenesené do Microsoft Excel tabule, aby byly dále zpracované ve statistickém programu R.

Výsledky výzkumu ukázaly, že mezi dvěma soubory nahrávek je významný rozdíl v distribuci vazání. Ve čtené řeči, která představuje návyky studentů v prvním ročníku studia, převládajícím jevem byla glotalizace, a nejvyšší počet vyskytu vazání byl jenom 15%. V nahrávkách spontánní řeči se situace výrazně slepšila, vazání stalo převažujícím fenoménem, a to s výskyty od 50% do 60%. Výsledky prozodického členění jsou velmi odlišné. Žádné dynamické rozdíly mezi soubory nahrávek nebyly nalezené. V obou případech byly v postavení logických prozodických frází všechny studentky vysoce úspěšné. Výrazné změný temporálních charakteristik se také neobjevily. Výsledky temporálních parametrů můžou být ale problematické, jelikož souvisí s typem řeči. V rámci této studie jsme porovnávali čtenou a spontanní řeč, což může být jedním z důvodů, proč se výsledky temporálních charakteristik dá považovat za neprůkazné.

Šestá kapitola prezentuje diskusi. Částečně se potvrdil předpoklad, že kurz fonetiky a fonologie angličtiny, který studenti absolvují v prvním ročníku studia, zlepšuje vybrané aspekty plynulosti řeči. Změny, které byly nalezeny v distribuci vázání, jsou dynamické a ukazují tendence k výraznému zlepšení. U prozodického členění se ale totéž neprokázalo. Otázkou ale je, jestli získané

výsledky ohledně temporálních parametrů prozodických frází nebyly ovlivněné typem řeči, jelikož improvizovaná spontánní řeč zahrnuje více kognitivních procesů. Proto by bylo nutné v dalším výzkumu zjistit, zdá se po specifickém trénování objevují temporální změny v prozódii, které by byly zkoumané na stejném typu řeči.