

UNIVERZITA KARLOVA

Filozofická fakulta

Ústav anglického jazyka a didaktiky



Bakalářská práce

Lenka Kalvodová

Production and perception aspects of weak form words in Czech-accented English

Produkční a percepční aspekty slabých forem v české angličtině

Praha, 2021

Vedoucí práce: doc. Mgr. Radek Skarnitzl, Ph.D.

Poděkování

Děkuji panu doc. Mgr. Radku Skarnitzlovi, Ph.D. za podporu, trpělivost a konstruktivní kritiku poskytnuté během psaní této práce. Děkuji rovněž za možnost naučit se pod jeho dohledem pracovat s programy a nástroji, které mi umožnily bakalářskou práci vypracovat.

Acknowledgements

I would like to thank doc. Mgr. Radek Skarnitzl, Ph.D. for the support, patience and constructive criticism offered throughout the writing process. I appreciate greatly having had the opportunity to master the software and tools needed for this thesis under his supervision.

Prohlášení autora

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.

Souhlasím se zápůjčením bakalářské práce ke studijním účelům.

Praha, 4. prosince 2021

Jméno, příjmení

Author's Declaration

I declare that the following BA thesis is my own work for which I used only the sources and literature mentioned, and that this thesis has not been used in the course of other university studies or in order to acquire the same or a different type of diploma.

I have no objections to the BA thesis being borrowed and used for study purposes.

Prague, 4 December 2021

Name, surname

Abstrakt

Tato bakalářská práce popisuje tendence v produkci a percepci slabých forem gramatických slov u českých mluvčích angličtiny na pokročilé úrovni. Teoretická část práce se věnuje srovnání řečového rytmu angličtiny a češtiny. Dále popisuje větný přízvuk v angličtině, vč. role změn na synsémantikách. Porovnává produkci a percepci angličtiny u rodilých a nerodilých mluvčích a systematicky popisuje slabé formy synsémantik. Nakonec práce zdůrazňuje potřebu rozvoje percepčních i produkčních dovedností v oblasti slabých forem synsémantik u nerodilých mluvčích angličtiny. První sekce praktické části s pomocí analýzy nahrávek pokročilých českých mluvčích angličtiny podává přehled oblastí, ve kterých se produkce slabých forem synsémantik českými mluvčími liší od produkce rodilých mluvčích. Je potvrzena hypotéza, že čím silnější český přízvuk v angličtině mluvčí má, tím méně slabých forem synsémantik a s nimi souvisejících procesů v souvislé řeči používá. Druhá sekce praktické části je založená na dvou percepčních testech, v jednom z nich posluchači posuzují míru subjektivního porozumění („comprehensibility“) a v druhém míru cizího přízvuku („accentedness“). Výsledky potvrzují výskyt tendencí i v percepci slabých forem a dalších procesů v souvislé řeči. Míra subjektivního porozumění promluvám anglické souvislé řeči se českým posluchačům posuzuje obtížněji než míra cizího přízvuku. Délka studia angličtiny však nemá vliv na schopnost posluchače posoudit výše uvedené charakteristiky. Čeští posluchači angličtiny vnímají různou úroveň obtížnosti subjektivního porozumění a přízvukovosti souvislé řeči, tj. implicitně jsou si vědomi přítomnosti slabých forem synsémantik a dalších procesů v souvislé řeči, jelikož tyto ovlivňují výše zmíněné faktory. Výsledky výzkumu jsou nakonec srovnány s předchozími studii, především s podobným výzkumem provedeným na mluvčích s jinými mateřskými jazyky.

Klíčová slova

souvislá řeč, gramatická slova, synsémantika, slabé formy gramatických slov, produkce, percepcie, vnímání, technika spojitého masky, angličtina s českým přízvukem

Abstract

The present BA thesis is concerned with perception and production aspects of weak forms of grammatical words in advanced Czech-accented speakers of English. The theoretical part first compares speech rhythm in Czech and English. Discussing the components of rhythm in English, it focusses on sentence stress, including the role of changes on grammatical words. Production and perception are described as observed in native versus non-native speakers. A systematic overview of weak forms of grammatical words is provided. Finally, the need for perception and production training in non-native speakers is emphasised. The first half of the analytical part, focussing on production, examines recordings of proficient Czech-accented speakers. It uses the data to present an overview of areas in which their production of weak forms and other connected speech processes differs from native-like. The hypothesis claiming that the more accented the speaker, the fewer weak forms and associated connected speech processes can be found in their production is confirmed. The second half of the analytical part is based on two perception tests, one asking the listeners to assess the level of comprehensibility and one the level of accentedness. The results showcase the presence of patterns in the way Czech listeners perceive weak forms and associated connected speech processes. Comprehensibility proves to be subjectively easier for Czech listeners to assess than accentedness, no matter their academic seniority in terms of studying English. Czech listeners' being aware of accentedness and comprehensibility levels implies their perceptiveness to weak forms of grammatical words and associated connected speech processes, as these constitute both aforementioned qualities of speech. Finally, these results are compared to and integrated into previous findings on this topic.

Keywords

connected speech, connected speech processes, grammatical words, function words, weak forms, production, perception, matched-guise technique, Czech-accented English

List of Abbreviations

BA – Bachelor of Arts

CS – connected speech

CSP – connected speech processes

EFL – English as a foreign language

FSP – functional sentence perspective

IPA – International Phonetic Alphabet

L2 – second/non-native language

SSBE – Standard Southern British English

WF – weak form (of a grammatical word)

WFW – weak form word

List of Tables

Table 1. Overview of weak form words. Please note that the lists of examples are not exhaustive.	25
Table 2. Percentage of native-like realisation of CSPs.	52

List of Figures

Figure 1. Linking on grammatical words, both groups.....	38
Figure 2. Linking on grammatical words, all speakers. The vertical line in all similar figures divides less accented speakers on the left and more accented on the right side.	38
Figure 3. Linking on lexical words, both groups.	39
Figure 4. Linking on lexical words, all speakers.	40
Figure 5. Vowel reduction to schwa, both groups.	41
Figure 6. Vowel reduction to schwa, all speakers.....	41
Figure 7. Linking occurring exclusively in combination with the elision of [h], both groups.	42
Figure 8. Linking occurring exclusively in combination with the elision of [h], all speakers.	43
Figure 9. Linking occurring exclusively in combination with the elision of [h] in chosen speakers.....	43
Figure 10. Consonant elision in <i>and</i> exclusively, all speakers.	45
Figure 11. Unreleased consonants, both groups.	46
Figure 12. Unreleased consonants, all speakers.....	47
Figure 13. Coalescence, both groups.	48
Figure 14. Coalescence, all speakers.	48
Figure 15. Glottalisation, both groups.	49
Figure 16. Manipulation of the recording of speaker F08, Sentence 04 (“There was a young man there.”).....	54
Figure 17. Manipulation of the recording of speaker F14, Sentence 21 (“To be, or not to be.”).	55

Table of Contents

1	INTRODUCTION	9
2	THEORETICAL BACKGROUND	10
2.1	RHYTHM	10
2.2	STRESS	13
2.3	CONNECTED SPEECH	16
2.3.1	CONNECTED SPEECH PROCESSES.....	18
2.3.2	PRODUCTION AND PERCEPTION OF CONNECTED SPEECH.....	21
2.4	WEAK FORMS OF GRAMMATICAL WORDS	23
2.5	CZECH-ACCENTED ENGLISH	25
2.6	THE SIGNIFICANCE OF INSTRUCTION ON WEAK FORMS FOR L2 ACQUISITION	28
2.7	HYPOTHESES	32
3	PRODUCTION	34
3.1	METHOD	34
3.2	RESULTS	36
3.2.1	LINKING ON GRAMMATICAL WORDS.....	37
3.2.2	LINKING ON LEXICAL WORDS.....	39
3.2.3	VOWEL REDUCTION	40
3.2.4	ELISION OF [H] WITH SUBSEQUENT LINKING	41
3.2.5	CONSONANT ELISION.....	44
3.2.6	UNRELEASED CONSONANTS.....	45
3.2.7	ASSIMILATION	47
3.2.8	GLOTTALISATION	49
3.3	SUMMARY AND DISCUSSION	50
4	PERCEPTION	53
4.1	METHOD	53
4.2	RESULTS	56
4.2.1	ACCENTEDNESS.....	56
4.2.2	COMPREHENSIBILITY	57
4.3	SUMMARY AND DISCUSSION	58
5	GENERAL DISCUSSION AND CONCLUSIONS	59

5.1	SUMMARY AND INTERPRETATION.....	59
5.2	LIMITATIONS.....	61
5.3	FUTURE WORK.....	62
6	RESUMÉ	63
7	REFERENCES.....	67
8	SOURCES.....	68
9	APPENDICES	69
9.1	APPENDIX 1.....	69
9.2	APPENDIX 2.....	70
9.3	APPENDIX 3.....	73
9.4	APPENDIX 4.....	74
9.5	APPENDIX 5.....	75

1 INTRODUCTION

The present thesis is concerned with production and perception of weak forms of grammatical words and associated *connected speech processes* (CSPs) in proficient Czech students of English. In the first part, the theoretical background, findings up to date on rhythm and stress will be summarised, along with a comparison of Czech and English in this regard. An overview of current knowledge on connected speech in English and its constituent processes will be given, focussing on weak forms of grammatical words. This will be complemented by what has been gathered so far on the production of these in native-like English versus in accented varieties, including Czech-accented English. The last part of the theoretical background will be devoted to the importance of instruction on weak forms and associated connected speech processes in the acquisition of English as a second language.

The aim of the research part of this thesis is to establish whether there are any patterns in production and perception of weak word forms in Czech-accented proficient learners of English. An analysis of connected speech phenomena occurring in the environment of grammatical words, or the lack thereof, in the speech of these learners will be presented.

The part of this thesis focussing on production will attempt to establish whether and if so, then how do Czech-accented speakers perceive weak forms and associated CSPs.

Findings from the analytical part of the thesis will be compared with and linked to the findings explored in the theoretical background, focussing on the importance of weak forms as part of connected speech in English language acquisition.

2 THEORETICAL BACKGROUND

2.1 RHYTHM

A definition of *rhythm* and its types has often been used in linguistics that is based on, as Nolan & Jeon (2014) explain, a metaphor. This yields a categorisation of rhythm that is straightforward and easy to comprehend, and therefore attractive. As Roach states, “one of the most familiar distinctions in phonetics is that between stress-timed and syllable-timed languages” (1982, pp. 73-79). This means that it should, in theory, be possible to divide all languages into two groups, based on whether their rhythm is created mainly through peaks stressed to approximately the same degree, or through syllables of the same ‘weight’. However, further research, which will be summarised in the discussion below, has reminded us that the metaphor we have extended over speech rhythm, while a useful aid, cannot be reduced to a universal categorisation of rhythm patterns into binary categories.

The first issue arises already at the definition of rhythm, which differs widely amongst researchers, often even based on completely different criteria (Turk & Shattuck-Hufnagel, 2013, pp. 93, 95). This is not to claim that the rhythm categories typically used in linguistic education – stress-timed, syllable-timed, and, possibly, mora-timed languages – are an incorrect notion (note: this thesis works with the notion of timing as defined by Nolan & Jeon: measurable durations of different segments of speech (2014, p. 1)). However, sufficient attention should be given to the criteria based on which a language is categorised as such, and to the fact that these are ‘labels’ rather than decisive categories. Turk & Shattuck-Hufnagel (2013, pp. 93-94) and Nolan & Jeon (2014) point out that while many researchers and other professionals claim strong hypotheses on rhythmic language profiles, these are often based solely on a notion of alleged periodicity in timing. Acknowledging the serious implications of the claim that speech rhythm is the product of regular, isochronous periodicity in timing, Turk & Shattuck-Hufnagel have examined the theory of speech rhythm being controlled via a periodic structure during either its production or its perception. They point out that “normal conversational speech is not periodic on the surface, i.e., that no constituent recurs at regular temporal intervals” (Turk & Shattuck-Hufnagel, 2013, p. 94). In other words, speech that is not recited poetry, chanting, or other deliberately rhythmical form, is not periodical in nature.

Turk & Shattuck-Hufnagel explore various bases for rhythm in order to assess: (1) how timing is present in rhythm and (2) how rhythm is created. This is done with the thought in mind that speech does involve an element of timing, and is therefore perceived as – at the very least roughly – rhythmical (2013, pp. 94-95). They explore three main hypotheses of how speech involves timing. The first option is that rhythm is based on actual temporal periodicity. The second option is *abstract phonological rhythm*, based on prominence of elements with relative salience, which in turn affect timing of utterances. The third option is rhythm as a surface property, reflecting various factors (e.g. duration, intensity), where patterns are not isochronous (Turk & Shattuck-Hufnagel, 2013, pp. 95-98). The last option proposed means that the elements of speech are ordered in a hierarchy, but not necessarily involving isochronous periodicity in the sense of complete temporal regularity of element occurrence. Their conclusion lays in between the options: speech is ‘rhythmic’ in the sense that its “serially ordered elements of speech are structured into hierarchical grouping and prominence structures involving stronger and weaker elements” (Turk & Shattuck-Hufnagel, 2013, p. 99), which is due to an innate human tendency towards isochrony in speech. It is, however, far from isochronous (Turk & Shattuck-Hufnagel, 2013, pp. 99-100). What should therefore be stressed here is the *tendency towards* isochrony, not perfect isochrony itself.

In other words, Turk & Shattuck-Hufnagel conclude that while rhythm reflects many factors, and it is partially a ‘product’ of what human perception is accustomed to, it is not a matter of surface isochrony (2013, p. 104). They side with the claim that while languages do have different rhythmical profiles based on differences in relative prominence of certain segments, these profiles *cannot* be classified into two or three categories. Rather, languages can be placed at different points of a continuum, which we can imagine as spanning from what is generally established as stress-timed on one side of the spectrum to syllable-timed languages on the other (2013, p. 105). They also draw our attention to the fact that there is no proof so far as to why stress- and syllable-timed languages are perceived differently (2013, pp. 105-110). Discussing other elements of rhythm, Turk & Shattuck-Hufnagel conclude that “[the factors for rhythmic differences] are large in number and may operate differently in different languages” (2013, p. 112).

Nolan & Jeon extend “the alternative [claim] that languages exhibit contrastive rhythm subsisting merely in the alternation of stronger and weaker elements” (2014, p. 1), but

add that while this is “initially plausible, particularly for languages with a steep ‘prominence gradient’, i.e. a large disparity between stronger and weaker elements, [...complete] alternation is poorly achieved even by a ‘stress-timed’ language such as English” (2014, p. 1). Additionally, Nolan & Jeon claim that, rather than being ‘concerned’ with creating and maintaining rhythm, through whichever means that may be, languages mainly allow for and support syntagmatic contrast between units (2014, p. 1). While their claim that speech is inherently arhythmic, or even antirhythmic, may seem forward, it is in line with the previously discussed recent research. In fact, their theory could be considered less forward than the strict categorisation of languages into clear-cut groups of stress- and syllable-timed, as it offers more space to the naturally “non-binary” (2014, p. 3) identity of all languages in terms of rhythm. Similarly to Turk & Shattuck-Hufnagel (2013), Nolan & Jeon discuss two definitions of rhythm: 1. temporal rhythm based on isochrony, and 2. rhythm based on structural alternation of stronger and weaker elements. After considering similar potential sources of rhythm, they, too, conclude that there is no sufficient evidence gathered to support the isochrony-based definition (despite its being widely accepted), and support the definition of rhythm as ordered patterns of prominence (2014, p. 2). They explain that languages differ in the distribution of the factors (such as duration, intensity, pitch obtrusion) that make up their respective rhythm profiles, the weighing being different for each language. This means that Nolan & Jeon do not discard the labels ‘stress-’ and ‘syllable-timed language’ completely. They re-define the labels not as binary categories, but as opposite ends of a continuum. The ends, however, are not absolute in nature, either, as all languages have some degree of prominence variation between successive segments (2014, pp. 4-5), meaning that there is no purely stress- or syllable-timed language.

According to Nolan & Jeon, wherever on the continuum a language is, it is neither rhythmical in the isochronous sense, nor is it ‘perfectly’ contrastively rhythmical. The established classification of speech as rhythmical and its division into the usual two groups has its roots in a metaphor that has been applied to languages at the two ends of the proposed continuum, and then extended into a widely accepted categorical division. The fact that the metaphor of rhythm (temporal or prominence-based) roughly fits the way languages are organised is “accidental” and “imperfect” (2014, pp. 6-9). What this suggests is that we should re-examine our understanding of language rhythm classification, extending it beyond the core of the (admittedly helpful) metaphor.

Turk & Shattuck-Hufnagel's (2013) and Nolan & Jeon's (2014) explanation of rhythm typology is significant for this thesis in that it allows a greater freedom in how we think about speech rhythm and compare it in different languages. As Kukačka points out, languages still have “more similarities than differences” (2018, p. 17), and their rhythm profiles do not fall into binary, opposite groups. A further proof of: (1) that there is not a dichotomy between rhythm profiles of different languages, and (2) that there is no completely regular isochrony in speech rhythm in general has been found for example already in Roach's 1982 study. He measured inter-stress intervals in various languages with the conclusion of no significant difference, which means that “the distinction between stress-timed and syllable-timed languages may rest entirely on perceptual skills acquired through training” (1982, pp. 73-79).

Overall, the established theories of stress- and syllable-timed languages are indeed helpful and based on a generally correct observation, but it should be kept in mind that these are approximate profiles and that a particular speech rhythm profile is a balance of multiple elements. Roach (1982, pp. 73-79) also agrees with the authors mentioned above in that rhythm, including seeming isochrony, is partially a product of human perception of speech. The same tendency, which causes listeners to claim isochrony “even in sequences of inter-stress intervals that are manifestly far from equal” (1982, pp. 73-79) in perception, should ideally (in native-like speech) manifest itself in production, too. This is why production and perception are presented as very much intertwined in this thesis.

The issue of Czech and English having different rhythm profiles, which results in issues in both production and perception of English by Czech learners, will be further discussed below in Section 2.5. Sections 2.2 through 2.4 will present an overview of elements that play a role in creating the rhythm profile of English.

2.2 STRESS

English is, as established in Section 2.1, of a mostly stress-timed nature, meaning that its rhythm (i.e. the way the segments of speech are organised) is from a large part constituted by prominence contrasts at the level stress groups (Brown & Kondo-Brown, 2006, p. 2). In English, stressed syllables function as regular peaks of prominence – relatively,

in comparison to the rest of the utterance (cf. syllables of roughly the same prominence as all other syllables in Czech, a syllable-timed language). A marker of relative prominence, *stress* itself is a suprasegmental speech property composed from different variables. Collins & Mees list intensity (perceived as ‘loudness’), pitch variation (changes of pitch level in comparison to neighbouring syllables), vowel quality and vowel duration as the quantifiable components of stress (2013, p. 129). Besides differing slightly in the terminology used for the components that make up stress (including e.g. components that are less easily measurable), authors also distinguish between types of stress up to different levels. One of the distinctions typically proposed is between several degrees, or types, of stress. Volín & Johaníková first define stress in line with the rhythm continuum theory presented above:

“Human speech comprises chains of units that can be most of the time assessed as mutually contrastive in prominence. A unit can be perceived as stronger or weaker (i.e., more salient or less salient) than the one that precedes or follows” (2018, p. 181),

but add that a distinction between stressed and unstressed units may not be sufficient to describe certain languages. This is why they support working with four categories of stress level in a unit of speech: *primary stress*, *secondary stress*, *full unstress*, *weak unstress* (Volín, Johaníková, 2018, p. 181). The same categories are proposed e.g. in J. C. Wells’ Longman Pronunciation Dictionary, which explains that a stressed syllable always contains a full (strong) vowel (as opposed to a weak one) no matter if the stress is primary or secondary (2007, p. 844). The employment of primary and secondary stress can be seen for example in the seven-syllable word “interchangeability” [,ɪntə,tʃeɪndʒəˈbɪləti] (Wells, 2000, p. 9), where primary (stronger) stress (ˈ) lies on the fifth syllable, and secondary stress (ˌ) on the first and third syllables. Unstressed syllables can be further divided into weak unstress, containing a full (strong) vowel, or full unstress, containing a reduced (weak) one. Illustrated on the same word, the remaining syllables contain a weak unstress, represented either by vowel reduction to schwa [ə] or by [ɪ] (as opposed to the full /ɪ/). The two categories of unstress will be of particular interest in this thesis, as they are connected closely to weak forms and CSPs.

Another one of the principal divisions is that between *word stress* and *sentence stress*. Collins and Mees define word stress, or lexical stress, as “a stress in the isolated word” (2013, p. 20), meaning a stress in one syllable. Lexical stress in Czech is fixed on the first syllable, while, as Skarnitzl & Rumlová point out, in English it is free and bears a contrastive function (2019, p. 113). This is typically illustrated by examples of polysemy, such as in:

perfect (n., adj.)	/ 'pɜ:fɪkt /
(v.)	/ pə'fekt /,

where the placement of word stress distinguishes between parts of speech. In addition to this, a segment might undergo changes (typically in vowel quality, length, etc.), as illustrated above. Brown & Kondo-Brown define sentence stress as “the stress or pattern of stress groups in [an utterance]” (2006, p. 2).

In terms of sentence stress, a word in an utterance can be stressed or unstressed. In the latter case, changes occur in the word in native-like speech, which help maintain native-like speech rhythm. This is why sentence stress is of more interest to this thesis, which will discuss the role of sentence stress in maintaining speech rhythm in more detail below.

As explained by Collins & Mees, “English native speakers tend to distinguish between words which have some lexical meaning and those which only express grammatical relationships in sentences” (2013, p. 20). The first kind, called *lexical* (alternatively *autosemantic* or *content*) *words*, carry semantic information. The second kind, *grammatical* (*function*) *words*, connect the information units expressed by lexical words, creating relationships between the pieces of information. Alameen & Levis also point out that grammatical words have a large overlap with closed-class words (2015, p. 161) – this morphological, rather than semantic, point of view can be helpful in understanding the role of such words. Shockey explains that English “depend[s] heavily on stress as a bearer of meaning” (2003, pp. 20-21) – not only does lexical words’ being stressed help the listener comprehend the message, but it also subconsciously directs them to the semantic content of the utterance. This means that function words typically need not receive as much stress as lexical words, since the kind of information they carry (not semantic, but organisational information) can be inferred, as opposed to that carried by lexical words. Bearing little stress (relatively, compared to other words in their

environment within the utterance) and semantic load, grammatical words typically undergo processes which alter them to a certain degree without interfering with the intelligibility (for the definition of intelligibility, see Chapter 2.5) of the utterance. This, in turn, helps the utterance maintain the rhythm that is natural (native-like) for that language. This is typically the case in languages that are closer to the stress-timed end of the continuum, where the unstressed and/or altered forms constitute the roughly regular ‘filling’ in between the peaks of prominence. The peaks are constituted by lexical words, the semantic load therefore supported by all that makes up the prominence of the peak. Section 2.3 will provide an overview of the above-mentioned processes that occur on grammatical words, allowing them to fit in between the peaks.

2.3 CONNECTED SPEECH

Every word in English has what is called a *citation form* pronunciation. Typically, this is the pronunciation that would be found in a dictionary, when the word is isolated, stressed and the speaker is enunciating carefully. Shockey points out that a single English word can have more than one citation form even within one variety (Shockey, 2003, p. 2); this can be for example due to different division into syllables, or diachronic changes in pronunciation between generations. Certain words, however, are subject to a number of processes when in the environment of other words as part of an utterance. This results in what is called *connected speech* (CS).

Historically, connected speech has been seen by some as informal, even ‘lazy’ or ‘sloppy’, but this is far from the truth (see e.g. Collins & Mees, 2013, p. 21). Connected speech is a natural result of the utterance’s being accommodated to the rhythm profile of the language – if it were not for connected speech, rhythm would be disrupted, inhibiting perception. Alameen & Levis state that when producing connected speech, “[n]ative speakers do not seem to know that they are producing speech that differs from citation form.” (2015, p. 164), and point out that connected speech is not associated solely with certain situations, such as casual speech. Barańska & Zajac have proven that there is no significant difference in CS production in different speech situations in non-native speakers, either. The rhythm profile of a language needs to be maintained no matter the context of the utterance (2014).

Maintaining speech rhythm, the primary function of CS, is carried out through *connected speech processes* (CSPs). CSPs is a term that encompasses modifications that take place in some words when these are part of connected speech. Similarly to the relationship between word stress and sentence stress, as described above, Alameen & Levis explain that there are processes which occur within words or within utterances, when parts of an utterance interact with each other (2015, p. 160). The latter kind of modifications is of interest to this thesis, and Shockey divides them into phonetic and phonological modifications. Phonetic modifications (typically reductions) occur in a single word or expression when it “is not necessary to make the [physical articulatory] effort to achieve a maximal pronunciation after the first token” (Shockey, 2003, p. 3). This is in line with the *functional sentence perspective* (FSP, theme-rheme structure), as when the semantic item shifts to the theme position, it can undergo modifications in form, having already been perceived and understood. This is something that occurs in all languages, to a varying degree. Shockey explains that phonological reductions, on the other hand “occur in predictable environments and [...] appear to be controlled by cognitive mechanisms rather than by physical ones” (Shockey, 2003, p. 3). These occur systematically, regardless of previous mention (or the lack thereof) and are language-specific (Shockey, 2003, pp. 3, 4, 13).

Both Alameen & Levis (2015, p. 159) and Shockey (2003) state that the processes that affect words in connected speech are important mainly because of their significance when combined. CSPs in English involve a range of processes that occur, as established above, in all of native-like speech, but Shockey reminds us of two key facts concerning the nature of CSPs. First, that the motivations of CSPs are heterogenous, including grammatical and articulatory constraints (Shockey, 2003, p. 77), to which Alameen & Levis add extralinguistic factors (2015). Second, that CSPs as produced by speakers form a continuum, not a binary distinction consisting of the CSP either being present or absent (2003, p. 77), which shall be kept in mind in the research part of this thesis (Chapter 3). Alameen & Levis support the claim, also stating that “speakers do not always produce a specific CSP in the same way” (2015, p. 164). This means that there are two layers of a continuum to be taken into consideration: (1) from connected speech to citation form, and (2) the more subjective variation between speakers (idiolects).

2.3.1 CONNECTED SPEECH PROCESSES

Section 2.3 has defined connected speech processes (CSPs) in general; this section will present the classification that will be used in this thesis. This is due to the fact that, as for example Alameen & Levis (2015, p. 161) point out, there is still a lack of extensive research on CSPs and the terminology is far from united.

The following section offers an overview of CSPs as referred to in this thesis. A scheme of how these have been used in the research material for the present thesis is included in Appendix 2. I believe the following overview to be useful in illustrating CSPs directly on the material used in the empirical part of this thesis.

Schwa

Vowel reduction to schwa ([ə]) should occur in grammatical words whenever they are in a weak form context (as will be explained below in Section 2.4). This is illustrated in Sentence 4., Appendix 2: “There was a young man there.” [ðə wəz ə 'jʌŋ 'mæn ðe:]. The first instance of the word “there”, because of its function, would include vowel reduction to schwa in a native-like realisation, while the second one, in full (citation) form would have the full vowel [ε:]. Volín et al. stress the importance of schwa in CS, as it is essential in creating rhythm, while also constituting 1/3 of total vowels occurring in speech (2013, p. 32-33).

Glottalisation

While *glottalisation* occurs in many contexts, for instance replacing some plosives in certain accents, the concern of this thesis is with glottalisation where it cannot be replaced by any other speech sound in native-like pronunciation. Such environment occurs when a voiceless stop is followed immediately by a sonorant, such as in Sentence 12.: “He got me her number and email.” ['gɒʔ mi].

Assimilation

This label groups together multiple kinds of processes similar in nature. They can occur in two directions: regressive (anticipatory), wherein a sound is influenced by the following sound, or progressive (preservatory), where ‘direction’ of the change is the opposite.

Dentalisation is a type of coarticulatory process wherein the alveolars /t/, /d/, /n/, when followed by dental fricatives (/θ/, /ð/), undergo a change in place of articulation towards dental. For example in Sentence 24.: “he said that” [seɪ̯d̪ t̪æt], the /d/ in “said” is dentalised; in other words it is coarticulated with the dental that follows it. Progressive assimilation of manner (as described below in this paragraph) then takes place on the following dental. Its manner of articulation changes from fricative to plosive, due to the influence of the preceding plosive. The result is two consonants identical in manner and place of production – dental plosives. The amount of articulatory effort for the segment is therefore reduced. Additionally, the first plosive is unreleased, as per the mechanism described below in the paragraph dealing with unreleased consonants.

Assimilation of place of production typically affects these alveolars: /t/, /d/, /s/, /z/, /n/, which are modified by the following phoneme. The second phoneme can be a bilabial (/p/, /b/, /m/), one of these postalveolars: /ʃ/, /ʒ/, the palatal /j/, or one of these velars: /k/, /g/, /ŋ/. The place of articulation of the first phoneme changes to that of the second phoneme, e.g. “this shoe” / ðɪs / + / ʃu: / changing into [ðɪʃ ʃu:]. Note that this type of assimilation is not examined in the research carried out for this thesis; it was nonetheless included because of its frequent occurrence.

Assimilation of manner of production, too, typically affects dental fricatives, when preceded by alveolar plosives or alveolar nasals, changing their manner of production are produced to a plosive/nasal, respectively. This means that in Sentence 24., “that” will be pronounced as in [seɪ̯d̪ t̪æt], having changed its manner of articulation from a fricative to a plosive. In the same way, “there” in Sentence 14. will be produced as in [bi:ɹ̪ ñe:], changing its manner of articulation from a fricative to a nasal.

Coalescence (also called *coalescent* or *fusional assimilation*) is the merging of two phonemes into one. Alveolar stops /t/ and /d/ merge with /j/ into [tʃ] and [dʒ] respectively, as in Sentence 1: “told you” [təʊldʒ_u]. Alveolar fricatives /s/, /z/ merge with /j/ into [ʃ], [ʒ] (e.g. Sentence 16. “as your” [æʒ_ə]).

Unreleased consonants

Consonants are typically released based on the speech sounds in their immediate environment. If followed by a pause and/or a glottal stop, consonants are released fully. Plosives (/p/, /b/, /t/, /d/, /k/, /g/) can either be fully released, or unreleased. When one plosive is directly followed by another, it is typically not released at all, so as not to disrupt

the speech rhythm. Sentence 21., where the first consonant is unreleased, illustrates this: “not to” [nɒt̚tə].

Linking

Linking is a CSP occurring in word boundaries in some contexts where the second word starts with a vowel. Depending on the ending of the first word, these are then classified as described below. The result is that the two words that are linked are pronounced as one unit, with no ‘pause’ in between. The opposite would be a ‘pause’ with aspiration in between the words, which is called a *glottal stop* (?). Linking is desirable in native-like connected speech, as a glottal stop in the same context requires more articulatory force and disrupts the speech rhythm. There are four ways of linking that are of interest to this thesis.

(1) *Pseudo-resyllabification* occurs when a word ending in a consonant is followed by a word beginning with a vowel. These are then pronounced as one unit, without a glottal stop in between. The final consonant of the first word could therefore theoretically be perceived as the first consonant of the second word, hence the name of this process. An example would be “Can I” in Sentence 2., ideally produced as [kən_aɪ] – as opposed to [kən ʔaɪ].

(2, 3) *Transient w* and *transient j* are inserted in connected speech in between two words whose citation form they are not part of. Due to their manner of articulation (they are glides, or, approximants), they occur naturally in certain environments. Transient *w* is inserted between /u/, /u:/ and diphthongs towards [ʊ] in the first position and a vowel in the second, as in Sentence 2.: “you about” [ju_(w)əbaʊt]. Transient *j* is inserted when /ɪ/, /i:/ and diphthongs towards /ɪ/ are followed by a vowel, e.g. Sentence 10.: “be on” [bi:_(j)ɒn].

(4) *Linking [r]* should not be confused with *intrusive [r]*, which is inserted in pronunciation in between two vowels by speakers of some varieties. Intrusive [r] is, however, not part of the first word in its spelling. Linking [r], on the other hand, is /r/ that is not realised at the end of the first word in its citation form by speakers of non-rhotic accents. It is, however, realised in native-like pronunciation when followed immediately by a word starting in a vowel. Isolated, “there” in Sentence 17. [ðəɹ_ɪz] would be realised as [ðɛ:] (in the same manner as it is at the end of Sentence 4.). For speakers of rhotic accents, this would fall under pseudo-resyllabification, but has been included

as a separate phenomenon, since most Czech speakers are taught a non-rhotic accent of English at school.

Consonant dropping (elision)

There are a number of contexts in which consonants are elided in native-like connected speech, out of which two are of interest in this thesis. Firstly, when an alveolar plosive (/t/, /d/) is preceded and followed by another consonant, it is elided. This is the case within a morpheme (e.g. “postcard” [pəʊskɑ:d], “handsome” [hænsəm]), across a morpheme boundary (Sentence 5.: “must look” [məs lʊk], Sentence 3. “and discuss” [ən dɪ'skʌs]). Secondly, it is the elision of other consonants within a morpheme, for example of /r/ in Sentence 8.: “from Paris” [fəm 'pærɪs].

Elision (h-dropping) + linking

Elision of word-initial /h/ in unstressed grammatical words (as defined in Section 2.4 below) that causes them to be linked to the preceding word has been singled out as a separate CSP for the purposes of this thesis. This is because the other types of elision do not facilitate linking, and, simultaneously, the linking here could not occur were it not for the h-dropping. For example in Sentence 8., “does he” could not have been produced as [dəz_i] if the pronoun retained the /h/, as /z/ + /h/ is a context that does not allow for any linking.

It should be noted that there are other changes words undergo in spoken form, but since these are not part of CS, they are not of interest here. Other researchers might choose a slightly different division and labelling (e.g. Alameen & Levis (2015) propose six categories of CSPs), but the core stays the same.

2.3.2 PRODUCTION AND PERCEPTION OF CONNECTED SPEECH

The reason that this thesis concerns itself with both the production and the perception of *weak form words* (WFWs), which are part of CS, is that these are very much intertwined. This subsection will provide an overview of some strategies used by the speakers and listeners to produce and perceive CS, as well as of how these are examined by researchers. Section 2.5 will then focus on these topics in Czech-accented English specifically.

As for the production of CS, Shockey (2003) reminds us that all speech is the balance of the minimal effort principle and the need to communicate the message. Speakers, whether they produce native-like CS or not, are not aware “they are producing speech which differs from citation form, and, in fact, deny it when asked” (Shockey, 2003, pp. 73-74). Using acoustic display of speech via spectrograms, Shockey has proven that there is no difference in unscripted versus scripted speech in this regard. The spectrogram tool has proven useful in working with materials for both parts of this thesis research.

Perception is linked closely to production, especially in the process of L2 (second language) acquisition. As Alameen & Levis state, native speakers (NSs) produce speech in a way that L2 learners are not used to employing themselves (2015, p. 165). This inhibits L2 speakers’ perception of it, too, as they are not aware, even subconsciously, of the sounds (realisations) of CS.

Alameen & Levis are in favour of the bottom-up strategy as an explanation for CS perception when it comes to examining and teaching it (2015, p. 165). This brings more attention to the components of speech, including their forms other than the citation ones, that serve as building blocks for CS. To find out how to best describe this as a method, Alameen & Levis examined how NSs segment speech (including assessing word boundaries). They conclude that “the listener compares a representation of the actual speech stream to stored representations of words[, which is where] the presence of CSPs may create lexical ambiguity due to the mismatch between the lexical segments and their modified phonetic properties” (2015, p. 165). The main cues are lexical, complemented by cues on syllable and word boundaries (phonotactic) (Alameen & Levis, 2015, pp. 165-166). The other approach – top-down – rather focusses on the speech as a whole, with the listener trying to single out semantic information from the utterance.

This means that the main issue in CS comprehension is word recognition, and Alameen & Levis add that this is an issue for both beginner and advanced students (although beginners struggle more) (2015, p. 166). Perception skills of advanced Czech EFL learners will be examined in the research part of this thesis in order to identify possible patterns in skill level with increasing proficiency, or, rather, decreasing accentedness.

Shockey proposes *normalisation* as the main strategy in CS perception. Normalisation is a learnt process, whereby the listener singles out all variable factors of the utterance

(rate and coarticulation, speaker-dependent variables, environment-related variables) to identify the words. She offers the example of how a listener gets accustomed to a speaker they have not heard before (2003, p. 89). The same technique is used (albeit subconsciously) for getting accustomed to native-like CS in general.

Furthermore, Shockey concludes that phonology, not only acoustics and phonetics, is crucial in speech perception. Having conducted three gating experiments, she has proven that we do not store all (reduced or otherwise altered) variants in our mental lexicon, but only the citation form (2003, p. 94). This could also be because of the vast range of alterations an English word can undergo in CS (e.g. if we only consider regional varieties). Shockey's experiments assume that normally, CS is so fast that we do not register it in segments. This only happens subsequently in *gating*, whereby we initially only focus on the beginning segment, and then re-introduce the following ones gradually, in gates. This also means that shorter words are recognised later than longer words, which have 'enough' segments by themselves. An important finding is that we also correct our perception of a previous segment that has been altered by CSPs if we hear a clue to its form later in the utterance. All this suggests 'active phonology' in the perception of CS and Shockey explains this by saying that "[w]e develop language-specific algorithms for interpretation of phonetic input which are congruent with production algorithms (phonological rules or processes)" (2003, p. 94). This is significant for this thesis as proof of CS perception mechanisms being something that can be acquired.

2.4 WEAK FORMS OF GRAMMATICAL WORDS

CSPs in English occur very frequently in weak forms of grammatical words. As established above in Section 2.2, grammatical words express relationships between the content (meaning) that is expressed by lexical words. Therefore, grammatical words can be reduced to a great degree, resulting in what is called a *weak form*. A weak form (WF) is relatively less prominent – it is seemingly, as Volín & Johaníková put it, "living in the shadow of the surrounding auto-semantic words" (2018, pp. 181-182). An important note is that not all grammatical words have a weak form, e.g. the preposition "on" is always realised with a full vowel ([ɒn]) in SSBE (Standard Southern British English). Shockey discusses the interesting concept of 'icons' – words or phrases that are used so frequently within a certain discourse context (e.g. group of speakers – in a sociolect) that they are reduced up to an extreme degree, which is an example

of the other side of the scale. These are often proper names (Shockey, 2003, p. 42), but it could be said that some clusters of weak form words work in the same way. “Could you have” in Sentence 14. [kədʒu_(w) əv] undergoes a high degree of alteration, but because weak form words are limited to a relatively low number and therefore predictable, the message is fully comprehensible.

As explained in more detail above, CS affects all of speech. However, its effect on English grammatical words, whose weak forms are a constituent of CS, is more extensive due to their nature. Similarly to the other phenomena mentioned so far, WF realisations form a continuum, especially because of their occurring through and being combined with (other) CSPs. Windsor Lewis defines weak forms as a pronunciation variant that would, in most cases, be less loud, less prominent/unstressed, shorter and containing a different set of phonemes than the corresponding citation form (1967, p. 42). Shockey adds that they are frequent in occurrence, and, unlike contractions, they are entire words (2003, p. 46).

There are several factors that determine which grammatical words have a weak form (Ito, 2006, p. 18; Shockey, 2003). Shockey discusses frequency, discourse background, rate, membership in a linguistic unit, phonetic and phonological criteria, and morphology. What follows in this paragraph will summarise Shockey’s (2003, pp. 14-19, 47) notes on these factors. First, the more frequent the word, the more likely it is to be reduced, but frequency itself is not a decisive criterion. Second, as for discourse, what is important for this thesis is that the rheme (newly introduced information as explained by FSP) is never reduced, and that the discourse situation has no connection to the amount or form of WFs. A grammatical word, therefore, when used utterance-finally (typically in preposition stranding), will always occur in its full form. Third, speech rate, similarly, is often a correlate of WFs, but not a requirement. Table 1 below reflects the fact that some words are only realised as a WF when they function as a certain part of speech (e.g. existential *there* – cf. as an adverb). In addition to this, the morphological character of the word plays a role – for instance monomorphemic words tend to be reduced into a WF. As mentioned in the CSP overview above, certain phonemes are more liable to alteration, for example alveolars tend to be assimilated, and /h/ elided. This is a question of both phonetic character and phonological criteria of the sound (Shockey, 2003, pp. 14-19, 47).

There are many WF classifications deserving our attention, such as the relatively comprehensive one presented by Collins & Mees (see 2013, pp. 21-23). According to Volín et al., there are approximately 40 grammatical words that regularly occur in contexts requiring a WF pronunciation (2013, p. 32). Other recent classifications worth directing our attention to include Volín & Johaníková (2018, p. 18), or Kukačka's compilation using multiple sources (2018, pp. 24-29). Table 1 will provide an overview of WF categories, along with examples of possible weak realisations, as used in the research part of this thesis (see Appendix 2).

Table 1. Overview of weak form words. Please note that the lists of examples are not exhaustive.

pronouns	<i>he, his</i> [ɪz], <i>you, your, who</i> [u]
auxiliary verbs	<i>be</i> [wəz], <i>do</i> [dəz], <i>have</i> [əv]
modal verbs	<i>can</i> , [kən], <i>should</i> [ʃəd], <i>must</i> [məs]
prepositions	<i>to</i> [tə], <i>from</i> [fɒm], <i>for</i> [fə]
conjunctions	<i>and</i> [ən], <i>or</i> [ə], <i>as...as</i> [əz]
determiners	<i>a</i> [ə], <i>an</i> [ən], <i>the</i> [ðə]
quantifiers	<i>some</i> [səm]
<i>there</i> (existential and similar types)	[ðə]
infinitival <i>to</i> particle	[tə]

2.5 CZECH-ACCENTED ENGLISH

To begin this section, a number of key terms, which are also used in the research part of this thesis, need to be defined. Skarnitzl & Rumlová (2019) define *accentedness* as the overall strength of deviations from native-like pronunciation, but note that the scale of accentedness is wide, especially in terms of its effect on the communication process. Objective understanding of speech is described through *intelligibility*. *Comprehensibility*, on the other hand, is the “the subjective ease of processing of [...] speech”. (Skarnitzl & Rumlová, 2019, p. 109). This means that speech can be intelligible (the listener can make out the message), but its comprehensibility can be decreased

(the listener needs to employ more cognitive force than usual to make out the message), for example due to accentedness.

This section will present an overview of features in English that are, chiefly due to their absence or different realisation in Czech, challenging for Czech learners of English. Shockey also points out a fundamental difference in how L2 versus native speakers treat speech: the latter need not focus on the form, but only on the message (2003, p. 1). For L2 speakers, on the other hand, these differences can result in accented speech on the production side, as well as in difficulties in perception of native-like speech.

With the definition of speech rhythm as presented above in mind, we can say that Czech and English could be placed near opposite ends of the rhythm continuum – English bearing more traits of a stress-timed language, while Czech is closer to the syllable-timed end. This means that the two languages achieve rhythm via different means, which poses challenges to Czech speakers of English. This is because even though, as established above, no speech rhythm is completely regular, listeners in general are very sensitive to any deviations in rhythm. Our perception relies on native-like realisation of rhythm, and disrupting it impedes our understanding of the message. In addition to the difference in rhythm typology, word stress is fixed in Czech, and while in English it is not arbitrary, the rules of its organisation are quite complex.

Besides having a different speech rhythm, English also employs phonemes that pose a challenge to Czech speakers. As for vowels, Skarnitzl & Rumlová identify the main differences between the two systems as: (1) vowel length being distinctive in Czech, while in English it is not, and (2) the English vowel system being more complex (2019, p. 110). Volín & Johaníková have compared the speech of strongly accented Czech speakers to that of NSs in order to identify segmental and suprasegmental patterns. They have proven that, in general, Czech-accented English often lacks reductions – weakened or elided consonants and reduced vowels (2018, pp. 182-183). Vowel reduction in particular is a salient feature of English CS, as it is frequent in occurrence and contributes significantly to native-like rhythm. Volín et al. draw our attention to the fact that the native-like use of schwa “contributes considerably to the sound differences between native and non-native speech” (2013, p. 31). Schwa ([ə]) is one of the English phonemes that is missing from the Czech phoneme inventory.

Although it is not difficult for Czech speakers to produce, they often do not employ it where desirable (e.g. using a full vowel in WFs). Research conducted by Skarnitzl & Rumlová has found that Czech speakers tend to use Czech equivalents of English vowels. For example, “while Czech has only one vowel pair /a/–/a:/ in the entire open region, there are four vowels occupying this space in English, /æ, ʌ, ɑ:, ɒ/” (2019, pp. 110). They add that Czech speakers also often replace /æ/ with [ɛ]. As mentioned above, since vowel length is distinctive in Czech, speakers often try to use length to distinguish between English vowels. English vowels however, actually differ in quality (Skarnitzl & Rumlová, 2019, pp. 109-111).

Another group of phonemes that Skarnitzl & Rumlová see as problematic in terms of accentedness (although not necessarily comprehensibility) are the “notorious” dental fricatives, which are very frequent in CS (2019, p. 112). Dental fricatives do not occur in Czech, which is why L2 speakers tend to replace them with phonemes they are familiar with – typically /d/ for /ð/ and /s/ for /θ/. Skarnitzl & Rumlová discuss other realisations typical for Czech-accented English, as well as noting a general lack of linking and other CSPs, and a characteristically flat pitch range (2019). These features will be used as the criteria to group the speakers in the research part into more and less accented.

Volín & Johaníková (2018) explain that weakening is difficult to master for those whose mother tongue is not stress-timed. They summarise the result as “Czech speakers tak[ing] the WFWs as individual regular bricks to be put next to some other bricks rather than as inconspicuous filling that holds the other bricks together” (2018, pp. 193). They add that when grammatical words are too prominent, they may not only attract unwanted attention, but also alter the typical rhythmic pattern and be difficult to process. Lower comprehensibility can then lead to stigmatisation. Having examined a read text by six Czech university students proficient in English on the one hand, and six NSs on the other, they have compared their realisations of chosen WF words. They have found that, on the whole, WFWs produced by Czech speakers were 1/3 longer in duration, often because the latter group pre-glottalise in vowel-initial words, adding to the duration of the words (Volín & Johaníková, 2018). Selected detailed results of this study will be discussed below in Chapter 5, providing a useful comparison and context for the results of the research part. Another study by Volín et al., focussing on reduction to schwa in Czech-accented English, has compared spectral slope features to provide quantitative evidence on the realisation of schwa by Czech-accented speakers, compared

to NSs. They found that Czech-accented schwas have a more prominent sound manifesting itself as narrower formant bandwidths, showing a more gradual decrease of energy towards higher frequencies (2013, p. 37). The potential problems in reduction, therefore, are manifold, including vowel quality, length of the whole word, prominence (stress).

2.6 THE SIGNIFICANCE OF INSTRUCTION ON WEAK FORMS FOR L2 ACQUISITION

This section will summarise what has been proven up to date on WF and CSP instruction to L2 learners, focussing on issues in production and perception, potential difficulties for the learners, and proposed pedagogical methods.

As often pointed out in recent publications (e.g. Brown & Kondo-Brown, 2006; Alameen & Levis, 2015), little research has been conducted so far into CSPs and, specifically, WFs (e.g. Ito, 2006). This is despite the importance of these phenomena in L2 acquisition – as summarised in the sections above, CS is ubiquitous and WF words (while they can sometimes be inferred) are essential for understanding the relationship between the items of information presented by lexical words. Alameen & Levis point out the discrepancy between the amount of research up to date on segmental and suprasegmental features separately on the one hand and CS, “in which segmental and suprasegmental features interact strongly” on the other (2015, p. 159). On the whole, despite their proven – essential – role in both the perception and the production of speech, CSPs and WFs are marginalised in curricula and teaching materials.

The reason as to why WFs and associated CS should be taught to L2 students could be summarised into a very simple argument, presented by Brown & Kondo-Brown: that these phenomena actually do occur in English, and frequently (2006, p. 5). To elaborate, it has been established above that they occur in all speech (except for e.g. deliberate enunciation). As for WFs in particular, the argument offers itself that while it is important to know what the topic is (i.e. content words), this is little use when the relationships between these are unknown. Alameen & Levis argue that CSP training can improve skills both in the short and the long time horizon (2015, p. 170). This encompasses improvement

in production and perception skills, including eliminating accentedness, but also metalinguistic skills that can build a solid basis for proficiency.

Intelligibility, comprehensibility and accentedness, as defined in Section 2.5, are all affected by the absence of WFs and CSPs. On the one hand, Alameen & Levis state that the way L2 speakers produce CS can pose a challenge to the intelligibility of their production from the side of native speakers (2015, p. 160). It could be argued that an intelligibility issue arises even for other L2 speakers, who speak a different, albeit also accented, variety. Mayers adds that while intelligibility may not be affected, especially the lack of WF reduction in L2 speakers' production can lead to semantic misunderstandings, e.g. making a grammatical word seem of particular importance when it is actually not (1981, pp. 422-423). On the other hand, native-like production of WFs can cause problems in perception to L2 speakers of English, beginners in particular. The logical strategy therefore seems to be to first focus on intelligibility, rather than native-like realisation, since, as Alameen & Levis point out, the former is more realistic (2015, pp. 170-171), and only then – if desirable – on eliminating accentedness.

The argument that WF and CSP instruction is beneficial for both production and perception is mentioned in a number of studies (e.g. Brown & Kondo-Brown, 2006, p. 5; Alameen & Levis, 2015). These state that learners develop more profound skills overall when activities focussed on production (output) and on perception (input) are either balanced, or when output is practiced more. Izumi has set out to find a psycholinguistic foundation for what he calls 'the output hypothesis' and explain why output is essential in acquiring a more complex understanding of the language. He uses the explanation that while comprehensible input only requires "semantic processing", "syntactic processing" is necessary for comprehensible output to support this hypothesis. In other words, when only practicing input, L2 learners are not required to assess (and, ideally, therefore overcome) their own weak points in the knowledge of the target language. Output, on the other hand, 'forces' them to transcend their own limited abilities in order to make their output be understood (2003, pp. 168-191). It seems, therefore, to be mainly through output practice that EFL learners can develop the metalinguistic consciousness necessary for a better grasp on the language, including the fact that, as in the case of Czech speakers, the target language has a largely different rhythm system. As mentioned above, this does

not mean that input should be disregarded completely in favour of output – it should be given equal, or almost equal, attention. However, what seems of utmost importance is that the two types of skills practiced be intertwined and tailored to the students' background and level.

A number of issues in WF/CS acquisition that are typical for L2 learners have been pointed out. Production-wise, a lack of WFs or CSP renders processing demanding for the recipient, which can lead to negative attitude towards the speaker. On the other hand, L2 speakers can experience problems in processing native-like speech with CSPs and WFs. Shockey mentions that L2 learners experience a (longer) 'processing lag', as they rely on syntactic-semantic cues to process what they hear (2003).

In effortless communication, production and perception should be seamlessly intertwined (Volín & Johaníková, 2018, pp. 182-193; Ito, 2006, p. 19). It seems, however, that in teaching there is more focus on issues in perception, since, as Alameen & Levis state, "it is presumed that L2 learners will assume a passive role when using the target language" (2015, pp. 167-168). Izumi adds that EFL classes are often "teacher-centred" and comprehensible enough for the students to follow, meaningful output is not required from the students (2003, p. 169). This seems to be a substantial problem leading to many consequences. Kennedy & Blanchet note that having experienced difficulties with CS/WFs is discouraging for L2 learners in general (2014, p. 91), no matter the target L2 (the target L2 of interest in their study was French, which also contains a number of CSPs).

With the above in mind, a number of teaching methods have been proposed, since, as Ito points out, the currently established teaching activities are "restricted in terms of their effectiveness in that they do not involve any meaningful communication" (2006, p. 23). As for input, he adds that even "authentic" [listening] materials, which might at first seem effective, often fail to prepare learners for using the language outside the classroom (2006, p. 17). This could be because even if the students understand the message of such a recording, they are often not asked to interact with it any further, their contact with such a recording remaining passive. Another issue could be the fact that, as it is often the case in Czech schools, the speakers in the materials all have the same accent (SSBE), which is not the case in real life interactions. Ito's suggestion of using listening materials where understanding a reduced form is the prerequisite

to comprehending the message and completing the task could be another solution to the ‘passivity problem’ (2006, p. 24). Kennedy & Blanchet propose a bottom-up approach for perception (2014, pp. 92-93) in hopes of teaching the students about the nature of CS and how to identify its components. This would, once again, give the students a metalinguistic understanding of any speech they hear, rather than relying on much extra information about the context of a recording they are about to hear. This could be practiced in a classroom setting using such recordings, before moving on to the much desired practice in real-life interactions.

As per Izumi’s (2003) argument explained above, output is essential in naturally allowing the learners to stretch their language abilities to solve potential communication problems. What is so special about production, in Izumi’s opinion, is that it cannot be ‘faked’ the way perception can be (2003). This means that there is no room for guessing in production, and it is active in its very nature, forcing the learners to use the language. As for specific areas and strategies, Barańska & Zajac (2014) call for more attention to vowel quality instruction in Polish EFL learners. This could be applied to Czech speakers, too, considering the above-mentioned findings by Volín & Johaníková about English vowels as produced by Czech learners (Volín & Johaníková, 2018).

In general, the consensus is that CS/WFs should be taught in a comprehensive way, aiming at the students’ understanding of the phenomena, rather than just supplying isolated examples. While both perception and production are important, and production seems to be particularly effective in CSP/WF acquisition, the ideal is to combine the two skills in a meaningful way (e.g. Ito, 2006, pp. 24-25). In such tasks, learners have to both produce the phenomena and rely on the perception of these to be able to participate in the interaction. The word ‘interactive’ has been used plenty in language teaching, but it should not be forgotten that its true nature is in the union of perception and production in a way that serves meaningful communication. A useful sidenote is mentioned by authors including Tauroza (1993) and Mayers (1981, p. 426) – WFs should never be overenunciated, even if they are the focus of instruction, as that would be paradoxical. Instead, Mayers suggests using idioms and other commonly used phrases that include WFs (1981, p. 425).

Having explored issues in CSP/WF instruction as well as arguments for its importance, this last paragraph will focus on some possible reasons for the lack of CSP/WF instruction in EFL. One issue could be the teachers, both native and non-native ones, themselves not being familiar with WFs (Ito, 2006, pp. 21-22). Brown & Kondo-Brown explain that the consequence is often that while even non-native teachers do teach the phonemes of the target language, that is English, they teach the citation forms. In that case, communication in real-life contexts may fail, the efforts of both teachers and students notwithstanding (2006, p. 6). Another possible reason is the lack of (appropriate) teaching materials. Ito notices that, even if teachers find the time and materials to take on WFs/CS in the classroom setting, the materials are often unrealistic, especially at lower proficiency levels (2006, pp. 21-22), where other aspects of the target language knowledge are still limited. It would therefore seem that not much has changed since Windsor Lewis (1967) noted that materials for students often include pronunciation guides in IPA (International Phonetic Alphabet), which is far from useful for, let us say, elementary school students. In addition to that, the pronunciation guides often contain isolated citation forms, which are of no use in this issue, either.

2.7 HYPOTHESES

Based on the findings gathered above, two null hypotheses were established for the part of this thesis concerning production, and one for perception. These serve as the basis for a total of six working hypotheses.

The two null hypotheses for the research part on production are the following:

H0(1): The production of weak forms of grammatical words and associated CSPs in Czech-accented speakers of English as a foreign language exhibits no deviations from native-like realisation.

H0(2): The level of the speakers' accentedness has no relationship to the amount of weak forms of grammatical words and associated CSPs in their production.

The following working hypotheses concerning production have been deduced from the null hypotheses above:

H1: Czech-accented EFL (English as a foreign language) speakers show deviations from native-like production of weak forms of grammatical words and the associated CSPs.

H2: Patterns in certain words or environments can be observed in the production specified in H1.

H3: The more accented the speaker, the fewer weak forms and CSPs they employ in their speech.

The null hypothesis for perception is the following:

H0(3): There are no patterns in the perception of weak forms of grammatical words (and associated CSPs) in Czech-accented speakers of English as a foreign language.

H0(3) serves as a basis for the following working hypotheses regarding perception:

H4: The patterns in perception as specified in H0(3) correspond to the patterns in production as specified in H0(1) and H0(2).

H5: Comprehensibility is more difficult to assess for Czech-accented learners of English than accentedness.

H6: The more advanced the student, the more successful they are at assessing accentedness and comprehensibility correctly.

The research described in the analytical part of this thesis will strive to prove these hypotheses true or false. In doing so, it will explore and map any potential patterns in production and perception of WFs and associated CSPs in the speech of Czech-accented proficient speakers of English.

ANALYTICAL PART

3 PRODUCTION

3.1 METHOD

The purpose of this part of the thesis is to provide an overview of the tendencies present in the way advanced Czech learners of English produce weak forms of grammatical words. To analyse this, recordings have been taken of the research subjects reading out loud a text specifically prepared for this aim. The full text, consisting of 24 unrelated, numbered sentences, is available in Appendix 1. The main reason for using sentences artificially constructed for the purposes of this experiment was that these assured the occurrence of all types of words that are typically listed under weak-form grammatical words (see Section 2.4 above). If spontaneous speech were to be used instead, a limited number of different weak-form words might have occurred. The claim that the grammatical words in the 24 sentences are typically pronounced in their weak forms and/or with connected speech processes is based on SSBE. It should be pointed out that there may be more pronunciation possibilities both within one variety of English, as well as in different varieties. SSBE has been chosen for reference based on its widespread use in the Czech educational system. However, any realisation employing native-like use of WFs and/or CSPs, as opposed to its citation form, is accepted. For instance in Sentence 3, “and discuss” and “want to” can both be produced in two ways classified as WFs in connected speech – either (a) with an unreleased plosive, or (b) with the same plosive elided:

<i>and discuss</i>	a) [ənd [̚] dɪ'skʌs]
	b) [ən dɪ'skʌs]
<i>want to</i>	a) [wɒnt [̚] tə]
	b) [wɒn tə]

The target words include pronouns, adverbs, prepositions, conjunctions, particles, auxiliary and modal verbs, articles and determiners. Most of these only occur in the text in what would be produced as a reduced form in a native-like realisation, but specimens such as *there* are included in both a weak form and a full form context (“there” in Sentence 4, Appendix 1). The weak forms and associated CSPs predicted

to occur in CS in the 24 sentences are summarised in Appendix 2, divided into groups where applicable, and colour-coded in both the sentences and their transcription.

This part of the study was conducted on 34 volunteers beginning the first year of their studies at the time of the recording. All speakers were female; however, this was not intentional and bears no importance for the research. All were students of either the English Studies (Anglistika-amerikanistika) or the English Translation Studies (Angličtina pro mezikulturní komunikaci) BA (Bachelor of Arts) programmes, for which they had previously passed an entrance exam. This ensured that all participants would have, with variation, a proficient knowledge of the English language, and were expected to have reached a B2/C1 level according to the CEFR for languages. The most challenging aspect of constructing the text was therefore assuring that the sentences not be perceived as unnatural by the subjects of the study. Although the speakers had not yet received any instruction concerning weak-form words or connected speech as part of their studies at the Faculty of Arts at the time of the recording, they were presumed to be linguistically aware. The sentences were constructed using a variety of lexical words to complement the large concentration of grammatical words. On the other hand, the sentences used relatively simple vocabulary that the subjects would have been familiar with, as the lexical words played no role in the research and the subjects possibly getting caught on these would have affected the important parts of the recording. The participants were given a printed sheet containing the 24 sentences in the form listed under Appendix 1. They were asked to read these out loud in a recording booth after having familiarised themselves with them. The recordings were obtained in the sound-treated recording studio at the Institute of Phonetics in Prague, using an AKG C4500 B-BC condenser microphone at a sampling rate of 32 kHz with 16-bit quantisation.

The recording session yielded 34 separate files with each speaker reading out all 24 sentences. These 34 files were divided into two groups based on the level of the speakers' accentedness: 'more accented' and 'less accented'; this division having been approved with a second opinion from an experienced phonetician. Material from a number of speakers had to be omitted due to unnatural speech, which included extensive stuttering, atypical intonation, or exaggerated enunciation. Although the subjects had been familiarised with all steps of the procedure, the unfamiliarity of the recording booth environment might have affected the output in some. The speakers were assessed in terms

of the level of accentedness in their speech and divided into more and less accented, forming two preliminary groups. It should be noted that assessing the speakers on accentedness, although without analysing any specific phenomena in their speech just yet, always means sub-consciously taking into account WFs and associated CSPs, as these constitute accentedness. It could therefore be argued that the division into more- and less-accented group and the comparison of these in terms of WF production is to a certain extent circular. However, this does not interfere with the research purpose of the part of this thesis dealing with production, which is to map the tendencies in all speakers and the specifics of possible correlation with accentedness. Ultimately, 12 speakers, whose speech was natural, consistent (relatively, compared to the others) and on further ends of the ‘accentedness spectrum’, were chosen for each accentedness group.

A TextGrid file was aligned with each of the 24 separate sound files using the P2FA (the Penn Phonetics Lab Forced Aligner). Some of the alignment had to be corrected manually in Praat (Boersma & Weenink, 1992). This occurrence is in no way negative for the research – quite the opposite, as the fact that the automated aligner was unable to recognise some word boundaries hints at the connected nature of the speech in the recordings. Another occasion where the TextGrid files called for manual adjustment was whenever the speakers repeated or omitted a segment (the likes of the relativiser *that* in subordinate clauses for the latter case). Having aligned the TextGrid transcription with each recording, a new point tier was added, with a point for every phenomenon of interest. Each individual realisation of a phenomenon has been analysed using careful listening and labelled in a tier point. A system of abbreviations and IPA symbols has been established to keep the marking consistent. These entries have then been extracted into a table file and divided based on the type of the phenomenon examined. The refined results allowed for clear results to be processed in R (R Core Team, 2021) and visualised using the ggplot2 package (Wickham, 2016), providing for the graphs in Section 4.2.

3.2 RESULTS

This section provides an overview of the data gathered in the recording sessions described in Section 3.1, analysing them according to the different phenomena of CS they represent. The more general type of graphs, such as in Figure 1, is labelled based on the division

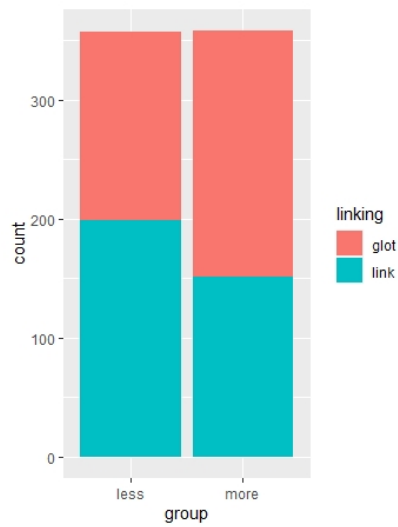
of the results between the less and the more accented speaker groups. The graphs portraying all individual results (e.g. Figure 2) are organised according to the two groups, the line in the middle dividing them into less accented speakers on the left and more accented on the right side. The individual speakers themselves are listed in the order of the codes they were assigned. In some graphs, the number of total realisations of a phenomenon differs for the more and less accented group, or between the speakers. This is due to two reasons: (1) that each speaker, in their own idiolect, has exhibited a unique combination of connected speech phenomena (or lack thereof) throughout the text, and (2) that some speakers omitted certain segments altogether by accident.

3.2.1 LINKING ON GRAMMATICAL WORDS

This subsection examines how many speakers in each group employed linking in their speech, as opposed to a glottal stop. The text included four different ways of linking that would take place in native-like speech in that environment – pseudo-resyllabification, transient *w*, transient *j*, and linking *r*. As outlined in Appendix 2, these are ways in which a grammatical word is linked to the preceding word, be it grammatical (“was as” in Sentence 16, Appendix 2) or lexical (“good as” (Sentence 16)). The reason for examining linking on grammatical and lexical words separately is the presumption that the speakers may work differently with clusters of grammatical words typically including linking (as there is a limited number of those and they are used often). Seemingly, they store such clusters as if idioms, including the way they are pronounced, i.e. with linking. Linking on lexical words, on the other hand, may occur in an unlimited number of contexts, making it improbable for the pronunciation of these to be remembered as a unit.

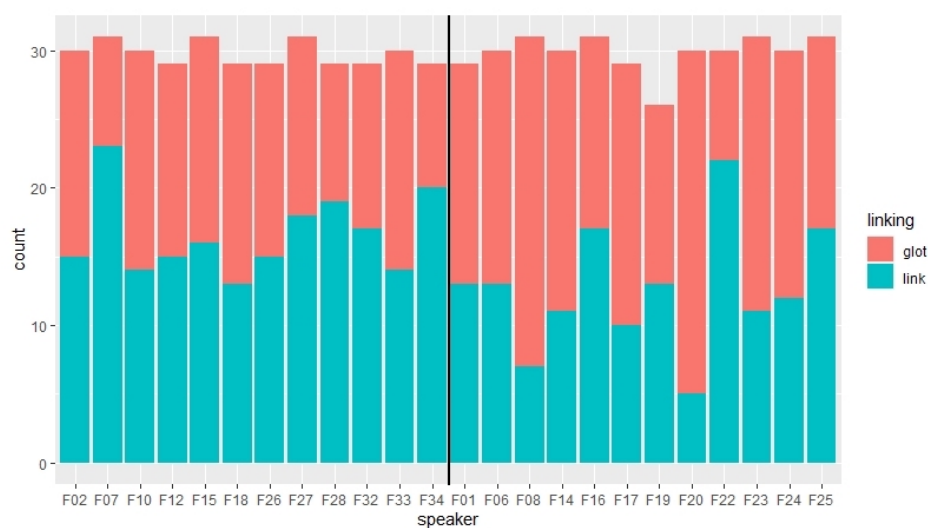
As a whole, out of the 715 slots for linking on grammatical words in all recordings, the speakers linked 350 times (in 48.9%) and glottalised 365 times (in 51.1%). As shown in Figure 1, there is a difference of approximately 15 percentage points between the less and the more accented group.

Figure 1. Linking on grammatical words, both groups.



This is further visible in Figure 2, which shows that the results are mixed in both groups. Nonetheless, the consistency of the results is lower within the more accented group, with results from speakers F08 and F20 significantly below average. While speaker F20 was average in all other processes, speaker F08 overall exhibited a tendency towards a lack of connected speech processes.

Figure 2. Linking on grammatical words, all speakers. The vertical line in all similar figures divides less accented speakers on the left and more accented on the right side.



3.2.2 LINKING ON LEXICAL WORDS

In linking to lexical words, be it from a lexical or a grammatical word that precedes it, the speakers show an overall lack of tendency towards linking. As mentioned in Subsection 3.2.1, this could be due to the fact that the speakers would not have a linked pronunciation of any of the segments in the text memorised. This means that their realisation would have relied solely on the level of connected speech processes incorporated into their speech. This can be compared in terms of information value especially with the results in Subsection 3.2.1 above. Overall, only 52 (28.4%) out of the total of 183 instances had linking where it would be expected to take place in native-like speech, with 131 (71.6%) cases of glottalisation.

The graphs in Figures 3 and 4 illustrate that, while the less accented group did realise more cases as linking, the difference is not clear-cut. Both groups have speakers who did not link in any case (F10, F33 in less accented and F08, F23, F24 in more accented). With 8 slots per speaker for possible linking on lexical words in the whole text, results such as F07 are not necessarily proof of exception. The results vary greatly in both groups, but no speaker in the more accented group realised more than 50% of the instances as linking, which signals a tendency of less linking on lexical words with increased speaker accentedness.

Figure 3. Linking on lexical words, both groups.

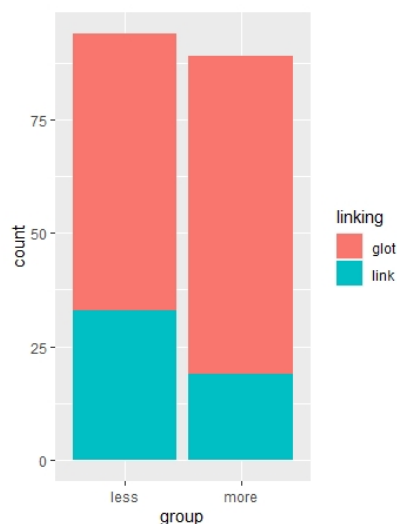
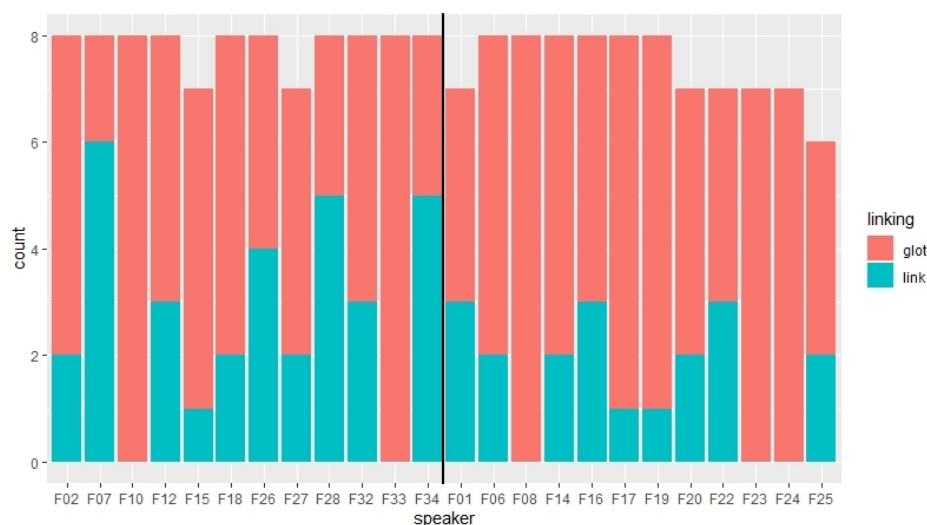


Figure 4. Linking on lexical words, all speakers.



3.2.3 VOWEL REDUCTION

Speakers from both groups exhibited a very wide range of the level of vowel reduction, hence the ‘intermediate’ category in Figures 5 and 6. For example, in the original labelling done in Praat (Boersma & Weenink, 1992), it was recorded that the vowel in *to* was realised not only as either [u:] or [ə], but also elsewhere on the spectrum from close-back to mid-central. Since these realisations were so numerous, a third category has been added.

Out of the total 1311 instances in all recordings, vowels have been reduced to schwa in 527 cases (40.2%) and realised as full vowels in 727 cases (55.4%), with 57 cases (4.4%) in the intermediate category. Figure 5 shows that, overall, the more accented group of speakers reduced vowels markedly less frequently (by some 20 percentage points). Figure 6 also portrays a notably regular pattern of the proportion of vowels reduced decreasing with increasing speaker accentedness.

Figure 5. Vowel reduction to schwa, both groups.

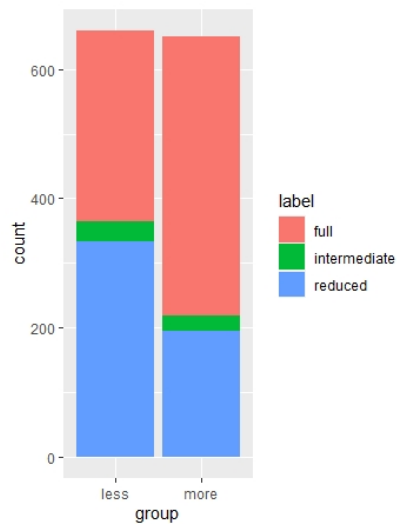


Figure 6. Vowel reduction to schwa, all speakers.



3.2.4 ELISION OF [h] WITH SUBSEQUENT LINKING

The graphs below offer a visualisation of a connected speech phenomenon that has been singled out as such because the linking in the phonetic environment cannot occur without the elision of /h/, e.g. as in Sentence 6 (“What have you been doing?”), where the first two words can either be realised as ['wɒt əv] or as ['wɒt həv]. Out of the total 327 slots, elision of [h] and linking was realised in 33 (10.1%) cases, while no elision and therefore no linking in 294 (89.9%) cases. As shown in Figure 7 below, speakers

in both groups elided the [h] in grammatical words and subsequently linked these to the preceding in a low number of instances. The only exception was speaker F26, who has exhibited average to above-average results in the other partial results, too. Figure 8 shows that there is, however, a pattern in that the more accented the speaker, the less elision and linking occurred in their speech. Many speakers of the more accented group did not elide at all (while this is only the case in two speakers from the less accented group).

Figure 7. Linking occurring exclusively in combination with the elision of [h], both groups.

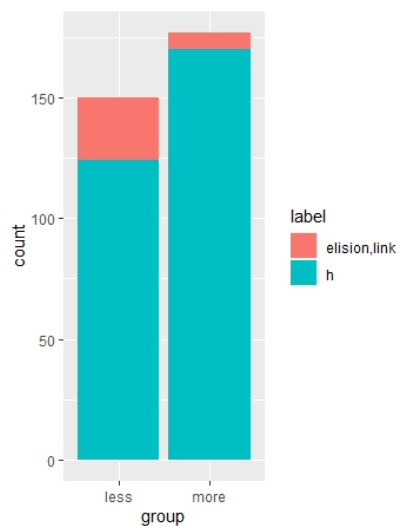
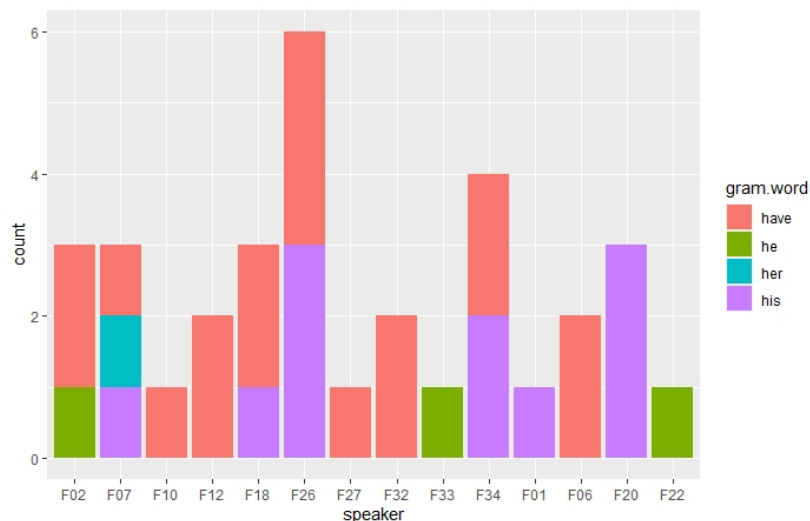


Figure 8. Linking occurring exclusively in combination with the elision of [h], all speakers.



Figure 9 illustrates the distribution of linking for the four grammatical words beginning with [h] that have been incorporated into the text. It demonstrates the wide spectrum of different realisations of [h] elision and linking in these four words. Overall, the results vary greatly, with *have* and *his* elided and linked the most. On the other hand, *her* is only elided/linked once in the 52 instances it occurs in in all of the recordings, and *he* three times out of 25 instances.

Figure 9. Linking occurring exclusively in combination with the elision of [h] in chosen speakers.



3.2.5 CONSONANT ELISION

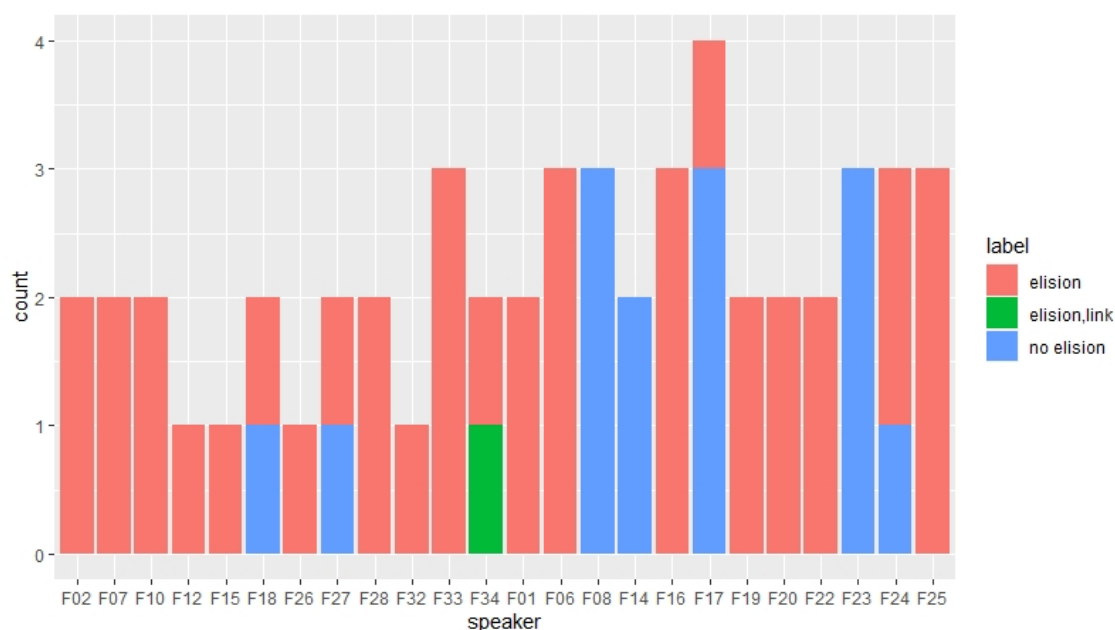
This subsection deals with the elision of consonants other than /h/. Unlike in the case of /h/ elision, the consonants elided here may facilitate linking, but it is not a condition. The consonants elided in the ideal case were /d/, /t/, /r/, /h/ in the words “and”, “must”, “from”, “who”. The text contained what is realised in native-like speech either as consonant elision by itself, or as consonant elision that makes subsequent linking possible.

Two of the consonants – /r/ in from and /h/ in who – were not elided by any speaker at all, which is why the following overview will focus on the other two cases.

In the 56 slots in the words “and” and “must”, there were a total of 27 cases (48.2%) of elision. 12 speakers from the less accented group elided, and 15 from the more accented, which is notable, as the opposite would have been expected. There was a single realisation of elision combined with linking, which will be discussed in more detail below. Finally, a total of 28 speakers (50%), with 12 speakers in the less accented group and 16 in the more accented one, omitted the elision altogether.

The option with linking would have been realised in native-like speech on *and* in Sentences 5 ([ən ʌm]) and 12 ([ən ʌ'i:meɪ]), so the ideal total number would be 48 realisations. There were, however, 10 speakers who omitted this segment altogether, by accident or due to nervousness, so the total number of slots is 38. There is a single native-like realisation (“elision,link”), from the less-accented speaker F07 in Sentence 12. This means that the other speakers either only elided the consonant but did not link (“elision” in Figure 10), or did neither (“no elision”). 27 speakers (71.1%) only elided and 10 speakers (26.3%) did not elide at all, which typically resulted in a released consonant followed by a pause.

Figure 10. Consonant elision in *and* exclusively, all speakers.



3.2.6 UNRELEASED CONSONANTS

The native-like realisation of the text would include unreleased consonants, namely the alveolar plosives /t/ and /d/, and the bilabial plosive /m/. In general, the plosives /p/, /b/, /t/, /d/, /k/, /g/, and the nasal plosive /m/ are unreleased in native-like speech when followed by a plosive (/p/, /b/, /t/, /d/, /k/, /g/) or an affricate (/tʃ/, /dʒ/). This produces two of the possibilities shown in Figure 11 (the accented “rel”, or the native-like “unrel”). When the alveolar stops are succeeded by dental fricatives (/θ/, /ð/), the latter is typically dentalised. The dentalisation would not occur otherwise, which is why these two phenomena are summarised under one pronunciation variant (“unrel, dent”). This means that there are two native-like pronunciation possibilities: (1) unreleased, or (2) unreleased and dentalised, and one accented possibility: (3) released. Overall, 108 (75.5%) of all consonants were realised as unreleased and 22 (15.4%) as unreleased with subsequent dentalisation, both of which were native-like options in the given contexts. 13 cases (9.1%) were released, therefore not native-like.

Figure 11. Unreleased consonants, both groups.

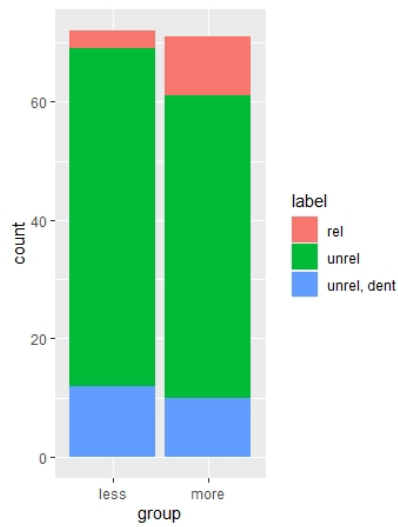


Figure 12 shows that the one instance of what would be “unrel, dent” in native-like speech – [bʌt̚ d̚ər] – in Sentence 17 has been produced so by all speakers from the less accented group, but not all from the more accented one. Both speakers F08 and F24 have consistently ranked as more accented even within their group, especially speaker F08. As for the other slots, there are twice as many released realisations in the more accented group, but upon a closer look the figures are only three versus six, so it cannot yet be said whether the pattern is conclusive.

Figure 12. Unreleased consonants, all speakers.



3.2.7 ASSIMILATION

Three types of assimilative processes have been analysed in the subjects' speech – coalescence (coalescent/fusional assimilation), assimilation of manner, and dentalisation (as a type of coarticulation).

Out of the total 167 slots for coalescence, it was realised in 77 cases (46.1%) across both groups, as opposed to 90 realisations without coalescence (53.9%). As Figures 13 and 14 show, the more accented the speaker, the less coalescence they employed in their speech. The more accented group is also considerably less consistent in comparison with the other group.

Figure 13. Coalescence, both groups.

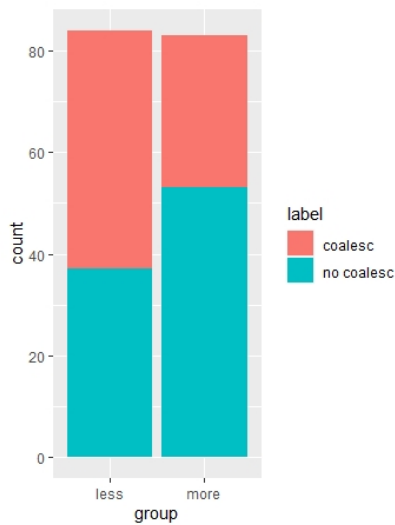


Figure 14. Coalescence, all speakers.



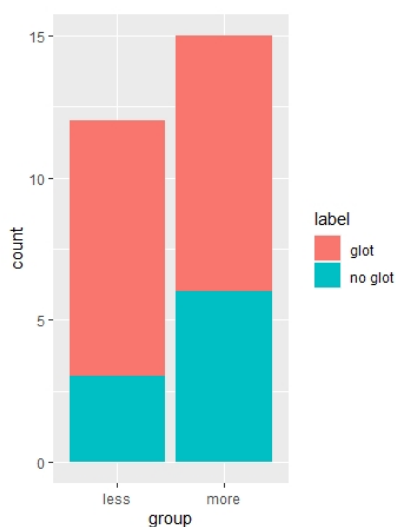
As for assimilation, the total results for both groups have shown that 20 (21%) out of the total 95 slots have been realised as assimilated. Less accented speakers employed full assimilation three times more often than the more accented. 26 slots (27.4%) had no assimilation whatsoever. 49 instances (51.6%) have been realised with dentalisation – an option halfway towards assimilation, included because of the width of the assimilation spectrum and the number of these ‘intermediate’ realisations.

Sentences 5, 10 and 14 specifically would contain a combination of assimilation and dentalisation in native-like, which has been marked as “assimilation”. Where the second consonant was not assimilated, but only dentalised, the label “dentalisation” has been used. This realisation was more common in the more accented group, with 34 (69.4%) out of the total 49 realisations of dentalisation.

3.2.8 GLOTTALISATION

The glottalisation analysed in this study is not, as is often the case, a matter of a variant, where the speaker would have a choice between e.g. [t], a flap [ɾ], or a glottal stop [ʔ]. It has been incorporated into two environments where <t> has the native-like pronunciation of [ʔ] no matter the speakers’ accents – namely preceding the bilabial nasal /m/. The usual alternative in most speakers was a released [t]. Out of the 28 predicted possibilities, glottalisation has been realised in roughly one third in all speakers. As Figure 15 shows, less accented speakers glottalised in 9 out of 12 (75%) times, while more accented in 9 out of 15 (60%) times. The number of slots differs, once again, in between the two groups solely due to omissions and alterations of the text by the individual speakers.

Figure 15. Glottalisation, both groups.



Overall, glottalisation appeared in Sentence 12 (“got me”) 23 times out of 24 (95.8%). On the other hand, there were no instances in the similar context of Sentence 23 (“what does it mean”). This is quite intriguing, since the two environments are similar. An explanation could be that “me” is a grammatical word, ideally produced as a weak unstressed form, while “mean” is a lexical, stressed word.

There were three other instances of glottalisation that had not been predicted but were native-like in each given context. In Sentence 1, speakers F08 and F20 glottalised the <t> and in “it” at the very end. In Sentence 17, F19 glottalised in “but” and subsequently inserted a pause. In Sentence 24, F28 glottalised the final <t> in “that”. With the exception of F28, all of the speakers who employed glottalisation in contexts not predicted are from the more accented group. In Sentences 17 and 24 the glottalisation could be interpreted as a sign of hesitation, where a pause had been inserted by the speakers.

3.3 SUMMARY AND DISCUSSION

The ‘production’ part of the research was successful in having gathered enough material to, one, provide an outline of weak form and CSP production pattern in Czech-accented proficient speakers of English and, two, prove true or false the hypothesis that the more accented the speaker, the fewer weak forms and connected speech phenomena occur in their speech (H3).

As a whole, the students linked on grammatical words in ca. 50% of the possible times, as opposed to inserting a glottal stop. What is important for the aforementioned hypothesis (H3) is that the ‘more accented’ group clearly had less linking as well as consistency in terms of the results than the other group.

On lexical words, only some 30% of all speakers linked, with less accented speakers linking more in general. Although there are not enough results to claim a conclusive nature of this result, it is, in and of itself, in line with the hypothesis.

The rate of vowel reduction in weak forms was some 40% on average, with additional 4% in the intermediate range. The average rate in the less accented speakers was considerably higher than in the more accented group – by 20 percentage points.

Together with the trend of gradual decrease of vowel reduction with increasing accentedness, this helps prove the hypothesis true.

Elision of [h] with subsequent linking has occurred at a very low rate (10%) in both groups. In line with the hypothesis, the phenomenon occurred at a higher rate with the falling accentedness of the speakers.

Out of the four different consonants that are elided in native-like speech, neither /r/ nor /h/ was elided in a single instance, therefore they are excluded from the total statistics. On average, all speakers elided /d/ or /t/ in some 50% of the slots. What is notable is that this is the only phenomenon going against the hypothesis H3, with 15 speakers eliding the consonant in the more accented group and only 12 in the less accented one. While this goes against the hypothesis, the claim to an exception would require the support of more large-scale research. Where elision with subsequent linking (realised so by a single speaker) is the native-like realisation, roughly two thirds elided only, which is, however, still a 'positive' result, indicating some awareness of CS.

The data on unreleased consonants have shown that around 90% of all realisations have been native-like, which is a remarkable result in comparison to the other phenomena examined. Once again, the less accented group employed a slightly higher percentage of native-like realisations.

The proportion of coalescent assimilation across both groups is some 45%, with more consistency in the less accented group. As for assimilation of place or manner, only some 20% of speakers produced the native-like variety, with three times as many instances in the less accented group. Dentalisation, an intermediate variant, was employed by roughly 50% overall, in the majority by more accented speakers. While dentalisation is closer to native-like speech than no assimilation at all, it was still the less accented speakers that employed the native-like option in more cases. All partial results prove true hypothesis H3 in that the more accented the speaker, the less coalescence.

Glottalisation in the contexts where it is the only native-like option was realised by one third of all speakers. What is interesting is that there were three other instances of native-like optional glottalisation, indicating that even though the students had

no formal knowledge of this phenomenon, they would have been familiar with it enough to use it in a native-like manner. Overall, the less accented group glottalised in an average of 75%, the more accented of 60%, proving the hypothesis.

With the exception of consonant elision, the realisation of all other phenomena confirms the hypothesis that the more accented the speaker, the fewer CSPs and weak forms occur in their speech.

The level to which Czech-accented speakers employ CSPs in a native-like manner ranges as widely as from 10 to 90%, showcasing that there are, indeed, more problematic parts of connected speech that deserve attention in EFL teaching. The percentages of native-like realisations for each phenomenon are summarised in Table 2 below; intermediate realisation percentage is included in brackets where applicable.

Table 2. Percentage of native-like realisation of CSPs.

linking on grammatical words	51.1%
linking on lexical words	28.8%
vowel reduction to schwa	40.2% (4.4%)
h-elision with subsequent linking	10.1%
consonant elision (/t, d/ only)	48.21%
unreleased consonants	90.9%
coalescence	46.1%
assimilation	21.1% (69.4%)
glottalisation (mandatory)	33%

4 PERCEPTION

4.1 METHOD

The 34 recordings originally taken for the purposes of the production analysis described in Chapter 4 were used as the basis for constructing a perception test. The purpose of the perception test was to examine whether and, if so, then how do advanced Czech speakers of English perceive accentedness and comprehensibility. The two properties were represented by the presence or absence of weak forms and CSPs in the perception test recordings. The material from the original recordings was used to construct a duplicate pair for each of the 24 sentences. The technique used is a type of the matched-guise technique, which would typically involve one speaker in multiple languages or varieties. In this case, these remain the same, with only the speech rhythm either natural or disrupted. Each pair contains one of the sentences manipulated to (1) a ‘better’ version with native-like employment of WFs, as well as CSPs, and (2) a ‘worse’ version, completely lacking both. In summary, by ‘better’ it is meant ‘less accented and therefore more comprehensible’, and vice versa. These manipulations were created using two software programmes – Adobe Audition and Praat (Boersma & Weenink, 1992).

A major part of the alterations was done through deleting or adding cycles of sound waves. Although Adobe Audition offers the function of automatically aligning the two ends of a wave after having cut out a segment, the cuts were always made either at: (1) the nodes where the wave crosses the medium line, or (2) at the crest or the trough of the sound wave. A quantity of *n times one whole period* was always used to ensure a smooth transition around the altered segment. This method was used mainly to alter vowel quality. Full vowels were ‘reduced’ to schwa by deleting cycles in vowels, achieving WFs. Citation forms, conversely, were created by adding cycles in a vowel sound. Another possibility of achieving a sound was to copy it from elsewhere in the recording of the same speaker. This, however, presented the challenges of e.g. different pitch level or intonation at different parts of the recording, requiring more editing for a smooth integration of the pasted segment. In a similar manner, ‘periods’ of silence have been inserted to achieve e.g. a lack of linking, manifesting itself as a glottal stop, where connected speech would typically have pseudo-resyllabification or other types of linking. Another type of sound inserted into the ‘worse’ versions was plosion on fortis consonants, where the ‘better’ versions employed linking (specifically pseudo-resyllabification, as the environment is $\{C\}_ \{V\}$). In this case, the ‘copy-paste’

method has proven useful, as it would have been much more difficult to achieve the plosion through any other method.

Pitch has been altered, especially in and around the sections that had been cut or added. The basic work for pitch alteration was laid in Adobe Audition, and subsequently refined in Praat's 'Manipulation object' function (Boersma & Weenink, 1992). This has allowed for any abrupt transitions to be smoothed out gradually by adjusting the curve and adding more point boundaries to it where necessary. This combination of methods prevented what listeners would describe as 'cracks' in the recordings.

Finally, Adobe Audition was used to treat any minor inconsistencies left. Amplitude was adjusted where necessary to: (1) smooth out transitions between the original recording and the pasted parts, and (2) make segments end or begin more gradually, especially in the environment of linking. The manual Spot Healing Brush tool has been used in Adobe Audition's spectrogram to modify the signal directly in the spectral domain. This tool has allowed for the energy of the signal to be reduced in a gradual manner, thereby smoothing out any high pitch white noise and 'cracks'.

A pair of a 'worse' and a 'better' version has been created for each of the 24 sentence recordings. The 24 recordings used were taken from a total of 10 speakers. The speakers were only chosen based on how close their output had already been to either the ideal of connected speech, or to the opposite, in order to facilitate the manipulation and to keep as much of the original sound as possible. The pairs created for the 24 sentences were then grouped into individual sound files of a pair each. Figures 16 and 17 below each show one sentence in a 'better' and a 'worse' manipulation, in that order, viewed in Audacity (Audacity Team, 2021). The waveform for each of the two manipulations is visibly different, showcasing the alterations made.

Figure 16. Manipulation of the recording of speaker F08, Sentence 04 ("There was a young man there.").

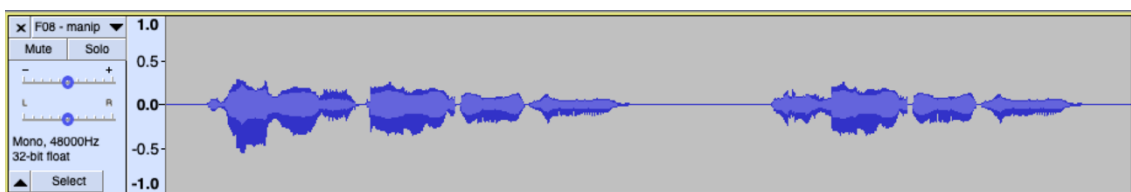
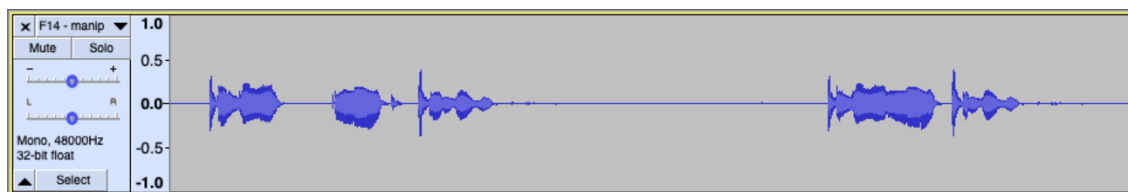


Figure 17. Manipulation of the recording of speaker F14, Sentence 21 (“To be, or not to be.”).



Due to social distancing measures associated with the COVID-19 epidemic, it was not possible to conduct the perception test in person at the Department of Phonetics. Instead, two Praat (Boersma & Weenink, 1992) perception tests were created using the scripts in Appendices 3 and 4 and administered individually via email. There were two groups of listeners to allow for potential comparison, the first group (hereinafter referred to as ‘Group 1’) being 15 first-year students of the English Studies BA programme at the Faculty of Arts, Charles University. Group 1 listeners were in their second semester of a phonetics and phonology course at the time of taking the perception test. The second group (hereinafter referred to as ‘Group 2’) was made up of 8 third-year students, who had already completed the same year-long course in English phonetics, including an overview of the aforementioned topics. What is important, however, is that none of the participants were familiar with the topic of this thesis or the specifics of the research conducted in the ‘perception’ part. The only information they were given is that the research is for a BA thesis in phonetics.

One of the tests gathered data on whether and, if so, then how Czech students of English at an advanced level perceive accentedness in English (therefore referred to as “accentedness test” further on in this thesis). The second test, administered to the same subjects, gathered data on their perception of comprehensibility in English (hereinafter referred to as “comprehensibility test”). The test set was accompanied by a set of detailed instructions (see Appendix 5), guiding the listeners through the whole process. The tests themselves included introductory and concluding texts as well as the main question repeated on the screen with each recording pair. After having run the script as per the instruction sheet and read the introductory commentary, the listeners heard the sentence pairs in a randomised order. The recordings were the same for both tests and in each test, all sentence pairs were played, as well as four control pairs.

The control pairs were duplicates of Sentences 1, 14, 17, 23, added under a fictive code (there was no speaker “F35”) and included in the randomised order, serving to assess the consistency of the answers. The listeners’ task was straightforward – to click on one of the two buttons displayed (“first” or “second”), according to which realisation of the same sentence they perceived as: (1) more accented in the accentedness test, and (2) more comprehensible in the comprehensibility test. Each sentence pair was only played once in the test. Another sentence pair was only played when the listeners clicked on a “next” button, allowing them to pace the test themselves. They were also prompted to take a short break throughout the test if necessary. All of the above was the same in both tests, and the order in which the tests were filled out was of no importance, which was stressed in the instructions. After finishing both perception tests, the subjects extracted their results in Praat (Boersma & Weenink, 1992) and submitted these via email. Based on the results submitted by the listeners, where Praat includes the response time, picking the response for one sentence pair took anywhere between 6 to 12 seconds. This is ultimately important to establish that filling out the two tests would have taken approximately 10 minutes, and the whole procedure along with downloading the test package and submitting the results via email would have taken around 15 minutes. This had been predicted and should mean that the experiment was short enough for the listeners to remain attentive and pick their responses with intention. The results received from the listeners have been gathered in two separate table files corresponding to the groups.

4.2 RESULTS

4.2.1 ACCENTEDNESS

There are only 14 sets of results for the accentedness test from the Group 1, as one result set had to be disqualified in this test. Although the control sentence pairs corresponded in 4 cases out of 4, the result itself was 0 out of 24, which was noticeably different from any other result sets. Otherwise, the ‘test’ sentence pair answers that the 14 participants in this group gave to the four test sentence pairs correspond with that in the pairs proper in 3.4 out of 4 cases, providing for a high level of reliability. There is only one result set with 2 out of 4 pairs corresponding; the rest of the result sets correspond in 3 or 4 test pairs out of 4.

On average, Group 1 listeners have recognised the more accented version from a sentence pair in the accentedness test in 19.4 out of 24 sentence pairs, in other words in 80.8% of cases. The averages for separate sentences ranged from 0.6 to 1 out of 1 per sentence. Eight separate sentences have been recognised correctly as more accented by all listeners, with Sentences 1 and 23 recognised twice (as the original and the duplicate).

Group 2 provides 8 result sets for the accentedness test, and on average the listeners recognised 20.4 out of 24 (85%) sentences correctly as more accented. The test pair responses match in average in 3.4 out of 4 cases, once again with only one set in eight having 2 out of 4 corresponding pairs only. The sentence averages range from 0.6 to 1 out of 1 recognised correctly as more accented. There are 9 sentence pairs that have always been assessed correctly. This includes, once again, Sentences 1 and 23, as well as Sentence 2.

The results in the accentedness test are similar for Groups 1 and 2, with 80-85% of sentence pair manipulations recognised correctly as more accented, and some 3.4 test pairs out of 4 corresponding to the original pairs in terms of the response.

4.2.2 COMPREHENSIBILITY

Group 1 provides 15 sets of results for this perception test. With only two result sets that have 2 out of 4 control pairs corresponding, the overall test pair correspondence rate is, similarly to the accentedness test, 3.4 out of 4. It is the results themselves, however, that are interesting in comparison to those from Subsection 5.2.1. The overall average for this group in picking the more comprehensible sentence realisation is 14.7 out of 24 sentences (61.3%). This is considerably less than the 80.8% in the accentedness test. There were also only two sentence pairs (Sentences 2 and 22) that had an average score of 1, with the other sentence pairs ranging from 0.3 to 0.8 out of 1.

The control pair score for the eight participants from Group 2 is, yet again, very similar – 3.3 out of 4 responses corresponded with those for the original sentence pair. Two respondents had 2 out of 4 test pairs correspond, while the rest had 3-4 out of 4. The results, though, are even more notable than in Group 1, as the total average was 13.8 out of 24 (57.5%) sentences assessed correctly. Sentences 1, 2 and 22 were each

assessed correctly in 100% of cases once, and Sentence 23 twice. What is an exception with this perception test is that there are two sentence pairs (4 and 6), where no listener was able to pick the more comprehensible realisation correctly. Like in Group 1, this is a much lower average than the 85% in the accentedness test. In addition to this, the average comprehensibility score in the third-year students is, albeit only slightly, lower than that in their first-year colleagues.

Throughout the four tests, there were sentence pairs that were consistently assessed correctly in all or most cases. These included Sentence pairs 1 (always in both its original and duplicate form), 23 (in 5 out of 8 times), as well as 2 and 22.

4.3 SUMMARY AND DISCUSSION

In the ‘accentedness test’ the results of Groups 1 and 2 were very close to each other in various aspects. On average, the speakers recognised 80-85% realisations as the more accented out of a manipulated pair. The individual scores (out of 24) in both groups varied from 16 to 23 in Group 1 and 17 to 23 in Group 2. This result is, once again, very similar. It was also the same sentences that got classified correctly by all speakers in both groups, namely Sentences 1 and 23.

The results of the ‘comprehensibility test’ were notably different. The speakers only recognised ca. 60% of the sentences as more comprehensible in the pair. The average also spanned wider for individual sentences, some in the ‘comprehensibility test’ only assessed correctly in 33% (as opposed to ‘accentedness’ with a minimum of 57%). Sentences 1 and 23 were, once again, among those assessed correctly in all instances. This test, however, also had two pairs that had a 0% recognition score, indicating a lower consistency of the results.

Although Group 1 provided 14 to 15 result sets and Group 2 only provided 8, their results in both perception tests were so close that it can be said that their academic seniority played no part in their determining accentedness and comprehensibility.

5 GENERAL DISCUSSION AND CONCLUSIONS

5.1 SUMMARY AND INTERPRETATION

The aim of this section is to interpret the results of both segments of the research part and point out any significant findings. The results will be compared to the working hypotheses as well as to the results of other research on the topic.

The two null hypotheses for production (H0(1, 2)) have been disproven, as there are tendencies in the production of weak forms of grammatical words (and associated CSPs) in Czech-accented speakers of English as a foreign language (H0(1)), and as these are connected to the level of the speakers' accentedness (H0(2)).

As predicted in hypothesis H1, Czech-accented EFL speakers show deviations from the native-like production of WFs and associated CSPs. No single CSP was present in any of the speakers' production in all the contexts native-like speech would have them. The percentage of native-like CSP use ranged as widely as from 10 to 90%, which is characteristic of non-native speech. Elision with subsequent linking was only realised in a native-like way in 10% of all slots, while consonants were unreleased (or unreleased with subsequent dentalisation) in a native-like manner in an average of 90% of occurrences. What is also typical for L2 speakers is a number of intermediate realisations. This was represented in various 'degrees' of vowel reduction, consonant elision lacking subsequent linking, or dentalisation instead of assimilation. This proves the argument from Alameen & Levis, presented in Section 2.3 – that CSPs and WF realisations form a continuum and are realised differently by different speakers (2015).

A similar study has been done by Barańska & Zajac, focussing on WFs in advanced English learners from Poland (whose language is closely related to Czech). They found that the most problematic and salient phenomenon is vowel reduction, stating that most speakers did not alter vowel quality where desirable (2014, pp. 281-282). This, in combination with the fact that an intermediate category needed to be aged to summarise the results in this thesis, could mean that vowel reduction in particular is problematic for speakers of Slavic languages in general.

The hypothesis that the speakers' production includes patterns in certain words or environments (H2) has been proven in some CSPs.

Linking occurred on grammatical words by some 30 percentage point more often than on lexical words. This could be due to the fact that there is a limited number of grammatical words (as opposed to lexical), and therefore the speakers will have heard all or most of them used with linking before. On the other hand, they might not have heard the lexical words used in that environment before, or had no reason to remember the particular combinations. Nonetheless, the speakers would have certainly been familiar with these lexical words at their level.

The percentage of h-elision realised varied significantly on the four h-initial grammatical words. “Have” and “his” underwent this change considerably more often. The tendency for “have” could be explained by the fact that it is a very frequent auxiliary verb (e.g. Roach, 2009, pp. 89-96; Collins & Mees, 2013, p. 23; Volín & Johaníková, 2018). Therefore, the speakers might have a subconscious passive awareness of the reduced way “have” is pronounced in native-like speech. There seems to be no explanation for the fact that the /h/ in “his” was elided often, while in “her” only in 1 out of 52 instances. Similarly, in “he” was only elided in 3 out of 25 cases. To assess whether the difference between “his” on the one hand and “he”, “her” on the other is arbitrary and specific to this study would require more data.

Consonant elision only occurred on dental plosives /t/, /d/ in “must”, “and”, but never on /r/ in “from” and /h/ in “who”. This is in line with Kukačka’s thesis, which, having focussed on several select WFWs only, including “from”, found that the word is, indeed, often pronounced in its citation form by Czech speakers (2018, p. 51). A question is whether the result is influenced by the nature of the elided sounds, or, which is more probable, by their position. Elision of word-final dental plosives is relatively more frequent than the other two types, which may have influenced the speakers’ production. While glottalisation did occur more frequently in a grammatical word followed by another word of the same kind (as opposed to a lexical word), the number of total slots (28, with one environment type each only) is too low to be conclusive.

Hypothesis H3, saying that the more accented the speaker, the less WFs and CSPs they employ, has been proven true, with one sole process going against it. Consonant elision was realised in a native-like manner by more speakers from the more accented group. The actual numbers (15 as opposed to 12), however, are not necessarily conclusive and more evidence would be needed to make a strong claim of an exception. Otherwise, the less accented group typically produced by 15-20 percentage points more CSPs

(this was the case in linking on grammatical words, vowel reduction to schwa, and glottalisation). The most significant difference was in assimilation of place or manner, which was employed by some 50 percentage points more often by less accented speakers.

Another similarity to the study of Barańska & Zajac (2014), is that, while for most phenomena the less accented speakers did produce more native-like realisations, the overall difference between more- and less-accented speakers was not substantial.

The null hypothesis for perception (H0(3)) has also been disproven, since patterns in the perception of WFs (and associated CSPs) in Czech-accented speakers of English have been found.

H4, saying that the patterns in perception of WFs and associated CSPs in Czech-accented speakers of English correspond to those in production (i.e. deviations from native-like execution and correlation with level of accentedness in listeners), could be considered disproven. The reasoning for this would be that, while native-like production ranged from 10 to 90%, 'correct' perception only from 60 to 85%. This suggests that perception of WFWs and associated CSPs is more developed in Czech-accented speakers of English than production is.

The hypothesis claiming that comprehensibility is more difficult for Czech speakers of English to assess than accentedness (H5) has been proven correct. The listeners assessed accentedness correctly in 80-85% of cases, with comprehensibility only at 60%.

Kukačka has found in his thesis that having taken phonetics and phonology classes had no effect on the results in WFW production (2018, p. 51). The same conclusion was reached in the perception part of this thesis, where H6 (the more advanced the student, the more successful they are at assessing accentedness and comprehensibility correctly) was disproved.

5.2 LIMITATIONS

This section will consider limitations to the research presented in Chapters 4 and 5. Overall, both parts of the research task were successful in having gathered enough quality

material to determine general tendencies. Nonetheless, there were both limitations that are typical for this kind of research task, as well as some linked to the COVID-19 pandemic, which affected mainly the perception part of the research task.

In production (Section 3.1), the numbers of realisations for the phenomena differ in various speakers. This is because of a number of reasons. First, some speakers employed contractions, although the text included full forms only. A similar problem was encountered in the study by Volín & Johaničková (2018) and could be considered a natural occurrence regardless of the speakers' nationality. Second, some of the speech was unnatural in that the speakers stuttered extensively or employed excessive enunciation. Finally, some speakers left segments out by accident, which caused a different number of entries for different phenomena.

Due to the COVID-19 epidemiological situation, the perception test (Section 4.2) had to be administered online, as opposed to in the phonetics laboratory. However, although this required more effort from both sides, all participants managed to run the test and extract the results correctly.

5.3 FUTURE WORK

Although the overall amount of material for the production part was satisfactory to determine general tendencies, the amount of data on two of the phenomena was not necessarily enough to draw conclusions. Specifically the production of unreleased consonants followed by obligatory dentalisation, and glottalisation would require more research in order to establish significant tendencies.

While the perception test did have relatively clear results that correspond with other similar studies (Barańska & Zajac, 2014), it would be useful to confirm this on more participants, preferably in a standard phonetics laboratory setting.

The research has proven what had been stated in the theoretical part of the thesis, which is the fact that none of the phenomena discussed form binary categories. Rhythm profiles as well as realisations of WFs/CSPs should rather be assessed as placed at a certain point of a continuum.

6 RESUMÉ

Tato bakalářská práce se věnuje tématu slabých forem gramatických slov v české angličtině; konkrétně tomu, jak tyto formy produkují a vnímají mluvčí na pokročilé úrovni. Cílem práce bylo zjistit, zda se v produkci či vnímání slabých forem gramatických slov českými mluvčími vyskytují nějaké tendence, a případně je zmapovat. Práce tímto navazuje na podobný výzkum zaměřený na slabé formy u mluvčích, jejichž mateřským jazykem je např. polština či japonština.

Teoretická část práce představuje teorii rytmu řeči jako kontinua a poukazuje na zjednodušení spojené s binárními kategoriemi rytmu (mimo které bývají zmíněny ještě jazyky s rytmem tvořeným morami, tzv. „mora-timed languages“). Rozdělení jazyků podle toho, zda je v nich rytmus utvářen primárně pomocí slabik stejné váhy (sylabické jazyky) či pomocí pravidelně rozmístěných segmentů s přízvukem (prozodické jazyky) je zavedené i v odborných publikacích. Tato práce se však staví za klasifikaci jazyků na kontinuu sahajícím od jazyků sylabických po jazyky prozodické. Jazyk je tedy blíže k jednomu konci kontinua, tj. má více znaků dané skupiny; rozdělení však není binární. Po ustanovení definice rytmu angličtiny jako většinou prozodické a češtiny naopak většinou sylabické se práce věnuje přízvuku jakožto suprasegmentální (prozodické) vlastnosti řeči. Popisuje, čím je v angličtině tvořen, a dělí jej na kategorie, především pak na přízvuk slovní a větný. Tato práce se zabývá především přízvukem větným, který leží na některých slovech v promluvě. Vysvětluje rozdíl mezi lexikálními slovy, jakožto nositeli významu, a gramatickými slovy (syntaktiky), která významové jednotky spojují do logických vztahů. Dále je popsána souvislá řeč v angličtině, kde jsou právě syntaktika podrobena změnám za účelem dodržení přirozeného řečového rytmu (který je v angličtině převážně prozodické povahy). Výsledkem těchto změn jsou slabé formy gramatických slov, které se liší od forem citačních (slovníkových). S pomocí příkladů použitých přímo ve výzkumné části jsou popsány tyto změny probíhající na syntaktikách a další procesy, které se k nim váží v souvislé řeči. Jelikož práce zkoumá produkci a vnímání slabých forem syntaktik a s nimi spojených procesů v souvislé řeči, následuje úvod do produkce a percepce angličtiny obecně, dále souvislé řeči v angličtině a nakonec konkrétně slabých forem syntaktik. Tento úsek představuje různé strategie produkce a percepce, jakožto i metody jejich zkoumání. Následuje konkrétní přehled anglických gramatických slov, která mají slabou formu. Od obecného srovnání rodilých a nerodilých mluvčích, co se produkce a percepce souvislé řeči a jejích komponentů týče, se práce přesouvá k podrobnému přehledu poznatků o českých

mluvčích v tomto ohledu. Na základě dostupné literatury podává přehled procesů, které jsou pro české mluvčí problematické, a vyvozuje důsledky. Závěrem teoretické části práce je zdůrazněna potřeba rozvoje percepčních i produkčních dovedností nerodilých mluvčích angličtiny právě v oblasti slabých forem synsémantik.

Výzkum popsáný v praktické části práce je rozdělený na dvě sekce podle toho, ve které části komunikačního procesu slabé formy synsémantik zkoumá: produkce a percepce. Výzkum zjišťuje, že se produkce slabých forem synsémantik českými mluvčími liší od produkce rodilých mluvčích, a mapuje tyto rozdíly. Vytváří tak poměrně obsáhlý přehled produkce slabých forem gramatických slov v angličtině českými mluvčími. Dále výzkum potvrzuje hypotézu, že čím silnější český přízvuk v angličtině mluvčí má, tím méně slabých forem synsémantik a s nimi souvisejících procesů v souvislé řeči používá. Co se týče percepce slabých forem a dalších procesů v souvislé řeči, výzkum potvrzuje, že se v ní u českých mluvčích vyskytují určité tendence. Míra subjektivního porozumění („comprehensibility“) promluvám anglické souvislé řeči se českým posluchačům posuzuje obtížněji, než míra cizího přízvuku (tj. odchylky od produkce rodilých mluvčích, „accentedness“). Hypotéza předpokládající, že čím dál pokročil posluchač ve studiu anglické fonetiky a fonologie, tím lepší budou jeho výsledky v posuzování míry subjektivního porozumění i míry cizího přízvuku, není potvrzena. Výsledky skupiny pokročilejších studentů se nijak markantně neliší od výsledků méně zkušených studentů. I tak však výzkum týkající se percepce potvrzuje, že čeští mluvčí angličtiny vnímají obtížnost subjektivního porozumění a přízvukovosti souvislé řeči. Konkrétně to znamená, že vnímají vliv přítomnosti slabých forem synsémantik a dalších procesů v souvislé řeči. Tyto závěry byly vyvozeny z výzkumu rozděleného na dvě části.

První část, zaměřená na produkci slabých forem a dalších procesů souvislé řeči, analyzuje nahrávky 24 mluvčích z řad studentů prvního ročníku filologických oborů. Obsahem nahrávek je 24 čtených vět s vysokou koncentrací zkoumaných fenoménů. Mluvčí byli rozděleni na dvě skupiny podle síly cizineckého přízvuku, což umožnilo srovnání produkce v závislosti na míře přízvuku. Hlavní část podkapitoly o produkci je tvořena podrobnou analýzou výsledků následujících procesů: vázání na slovech lexikálních i gramatických, redukce samohlásek na schwa, elize /h/ s následným vázáním, elize (vynechávání) souhlásek, nevypuštěných (bezdetenzních) souhlásek, různé typy souhláskové asimilace, použití rázu. Všechny procesy kromě elize souhlásek potvrzují hypotézu, že čím silnější cizinecký přízvuk mluvčí má, tím méně slabých forem synsémantik a dalších procesů souvislé řeči jeho projev obsahuje. Míra výskytu

se výrazně liší mezi jednotlivými procesy – od 10 % u elize /h/ s následným vázáním po 90 % u nevypuštěných souhlásek. To naznačuje, že určité procesy jsou pro české mluvčí náročnější, a tudíž by se nabízelo zaměřit se na ně ve výuce. Závěry výzkumné části zaměřené na produkci potvrzují všechny tři pracovní hypotézy: (1) projev českých mluvčích se liší od projevu rodilých mluvčích, co se týče slabých forem synsémantik a dalších procesů souvislé řeči; (2) tyto změny tvoří vzorce, které jsou v práci popsány; (3) čím silnější cizinecký přízvuk mluvčí má, tím méně slabých forem synsémantik a dalších procesů náležitých souvislé řeči používá.

Druhá část výzkumu, zaměřená na percepci týchž fenoménů, použila nahrávky pořízené pro první část výzkumu. Pomocí kombinace několika programů z nich vytvořila pár dvou nahrávek pro každou ze 24 vět, simulující vždy jednu realizaci rodilým mluvčím a jednu mluvčím se silným cizineckým přízvukem. Tyto úpravy lze shrnout jako přítomnost (či naopak nepřítomnost) výše zmíněných procesů. Z těchto 24 dvojic byly sestaveny percepční testy, jejichž úkolem bylo zjistit, zda čeští posluchači vnímají přítomnost slabých forem synsémantik a dalších procesů v souvislé řeči. Toto je zjištěno prostřednictvím analýzy jejich vnímání míry subjektivního porozumění („comprehensibility“) a míry cizího přízvuku („accentedness“). Kvůli epidemiologickým nařízením ve spojitosti s virem COVID-19 nebylo na jaře roku 2021 možné percepční test provádět fyzicky v laboratoři Fonetického ústavu. Byly proto sestaveny dva „distanční“ percepční testy („accentedness test“ a „comprehensibility test“) ve formě skriptu v Praatu (Boersma & Weenink, 1992). Každý z testů posluchačům přehrál v náhodném pořadí každou ze 24 dvojic právě jednou. Testy byly vždy doplněny ještě o 4 kontrolní dvojice, náhodně začleněné do pořadí. V testu „comprehensibility“ vybírali posluchači provedení věty, které je snadnější na porozumění („more comprehensible“), v testu „accentedness“ provedení, které se vyznačuje nižší mírou cizineckého přízvuku („less accented“).

Percepční testy měly dvě skupiny respondentů – 15 studentů prvního ročníku studijního oboru Anglistika-amerikanistika („Skupina 1“) a 8 studentů třetího ročníku téhož oboru („Skupina 2“). Porovnání percepčních dovedností studentů dle pokročilosti ve studiu vyvrací hypotézu, že čím pokročilejší student, tím úspěšnější je v posouzení míry cizího přízvuku a míry porozumění promluvě. Skupina 1 rozeznala verzi věty se silnějším přízvukem průměrně v 80 % případů, Skupina 2 průměrně v 85 % případů. Posluchači obou skupin byli také konzistentní, co se týče individuálních výsledků pro celý test (16-23 pro posluchače Skupiny 1, 17-23 pro Skupinu 2). Stejně byly i některé

věty, které správně posoudili všichni posluchači obou skupin. Průměrné výsledky pro míru porozumění souvislé řeči všech posluchačů jako celku se znatelně lišily. Celkové průměrné skóre obou skupin bylo sice téměř identické (cca. 60 %), avšak tento výsledek je znatelně nižší než průměr prvního testu. Také individuální výsledky posluchačů obou skupin byly na širším spektru a některé věty nebyly správně posouzeny ani jedním mluvčím. Toto potvrzuje hypotézu, že míra obtížnosti porozumění promluvě („comprehensibility“) se českým studentům posuzuje obtížněji než míra cizího přízvuku v promluvě („accentedness“). Celkově tedy lze říci, že se vzorce vyskytují u českých studentů i v percepci slabých forem a souvislé řeči. Na rozdíl od míry cizineckého přízvuku v produkci však nemá na percepci vliv úroveň vzdělání studentů v oboru.

Práce v závěru shrnuje a interpretuje výsledky obou částí výzkumu, včetně rekapitulace potvrzených či vyvrácených hypotéz. Výsledky jsou dále srovnány s předchozími studii pojednávajícími o příbuzných tématech. Dále se práce věnuje případným omezením provedeného výzkumu, a to především distanční formě a velikosti vzorku druhé části. Nakonec práce zmiňuje možnosti navazujícího výzkumu, např. s větším vzorkem pro část o percepci, a začlenění výsledků do výzkumu již provedeného u mluvčích s českým a jiným cizím přízvukem.

7 REFERENCES

- Alameen, G. & Levis, J. (2015). Connected speech. In: Reed, M. & Levis, J. (Eds.), *Handbook of English Pronunciation*, 159–174. Oxford: John Wiley & Sons.
- Barańska, A. & Zajac, M. (2014). The influence of attention to language form on the production of weak forms by Polish learners of English. *Research in Language*, 12(3), 279–290.
- Collins, B., & Mees, I. M. (2013). *Practical Phonetics and Phonology: A Resource Book for Students* (3rd ed.). Routledge. <https://doi.org/10.4324/9780203080023>
- Brown, J. D. & Kondo-Brown, K. (2006). Introducing Connected Speech. In: Brown, J. D. & Kondo-Brown, K. (Eds.) *Perspectives on Teaching Connected Speech to Second Language Speakers*, 1-15. Manoa, Honolulu, HI: National Foreign Language Resource Center, University of Hawaii at Manoa.
- Ito, Y. (2006). The significance of reduced forms in L2 pedagogy. In: Brown, J. D. & Kondo-Brown, K. (Eds.) *Perspectives on Teaching Connected Speech to Second Language Speakers*, 17–26. Manoa, Honolulu, HI: National Foreign Language Resource Center, University of Hawaii at Manoa.
- Izumi, S. (2003). Comprehension and production processes in second language learning: in search of the psycholinguistic rationale of the output hypothesis. *Applied linguistics* 24(2), 168-196. Oxford University Press.
- Kennedy, S & Blanchet, J. (2014). Language awareness and perception of connected speech in a second language. *Language Awareness*, 23(1-2), 92-106. <https://doi.org/10.1080/09658416.2013.863904>
- Kukačka, A. (2018). *The Production of Weak Forms of Function Words by Non-native English Speakers with a Special Focus on Prepositions* (dissertation). Brno: Masaryk University. <https://is.muni.cz/th/oq2jj/>
- Mayers, R. P. (1981). A new approach to the teaching of weak forms. *ELT Journal*, 35(4), 421–426. <https://doi.org/10.1093/elt/XXXV.4.421>
- Nolan, F. & Jeon, H.-S. (2014). Speech rhythm: a metaphor? *Philosophical Transactions of the Royal Society*, 369:1658. <https://doi.org/10.1098/rstb.2013.0396>
- Roach, P. (1982). On the distinction between 'stress-timed' and 'syllable-timed' languages. In: Crystal, D. (Ed.), *Linguistic Controversies*, 73-79. Accessed 22 June 2021 at <http://www.personal.reading.ac.uk/~llsroach/phon2/frp.pdf>
- Shockey, L. (2003). *Sound Patterns of Spoken English*. Oxford: Blackwell Publishing.
- Skarnitzl, R. & Rumlová, J. (2019). Phonetic aspects of strongly-accented Czech speakers of English. *Acta Universitatis Carolinae Philologica 2: Phonetica Pragensia*, 109–128. <https://doi.org/10.14712/24646830.2019.21>

Tauroza, S. (1993). Recognizing words in continuous speech: how important are word-final consonants?. *ELT Journal*, 47(13), 211-218. Oxford University Press.

Turk, A. & Shattuck-Hufnagel, S. (2013). What is speech rhythm? A commentary on Arvaniti and Rodriguez, Krivokapić, and Goswami and Leong. *Laboratory Phonology*, 4(1), 93–118. <https://doi.org/10.1515/lp-2013-0005>

Volín, J. & Johaníková, T. (2018). Weak structural words in British and Czech English. In: Volín, J. & Skarnitzl, R. (Eds.), *The Pronunciation of English by Speakers of Other Languages*, 181–195. Newcastle upon Tyne: Cambridge Scholars Publishing. <https://www.cambridgescholars.com/resources/pdfs/978-1-5275-0390-8-sample.pdf>

Volín, J., Weingartová, L. & Skarnitzl, R. (2013). Spectral characteristics of schwa in Czech accented English. *Research in Language*, 11(1), 31–39. <https://doi.org/10.2478/v10015-012-0008-6>

Wells, J. C. (2000). *Longman Pronunciation Dictionary*. Harlow: Longman.

Windsor Lewis, J. (1967). Phonetic transcription and weak-forms. *ELT Journal*, 22(1), October 1967, 41–44. <https://doi.org/10.1093/elt/XXII.1.41>

8 SOURCES

Audacity Team (2021). Audacity: Free Audio Editor and Recorder. Version 3.0.0.

Boersma, P. & Weenink, D. (1992). Praat: doing phonetics by computer (version 6.1.38). Amsterdam: The University of Amsterdam.

R Core Team (2021). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

9 APPENDICES

9.1 APPENDIX 1

The following 24 sentences were read out by the participants of the ‘production’ part of the research:

1. I would have told you about it.
2. Can I talk to you about something?
3. When did you want to meet and discuss it?
4. There was a young man there.
5. We must look in the locker and in the drawer.
6. What have you been doing?
7. The people who arrive at five PM are too late.
8. When does he arrive from Paris?
9. She should have asked for his permission.
10. You shall be on the list.
11. He can't have gone behind your back!
12. He got me her number and email.
13. There is a lack of answers.
14. Could you have been there?
15. There were some men who knew them.
16. His university was as good as your college.
17. No, but there is an umbrella.
18. Do you like her more than Jane?
19. It's a gift from us for him and his wife.
20. The parents were nice to them.
21. To be, or not to be.
22. Out of everyone here, I am the best.
23. What does it mean to us?
24. He said that his brother was an artist.

9.2 APPENDIX 2

The following scheme contains the predicted distribution of weak word forms and connected speech processes in the text from Appendix 1. This is based on SSBE and is not the sole native-like pronunciation for neither SSBE, nor any other variety.

Schwa

Glottalisation

Assimilation

- Dentalisation (as a type of coarticulation)
- Assimilation of place or manner
- Coalescence (= coalescent/fusional assimilation)

Unreleased consonants

Linking

- Linking: transient *w*
- Linking: pseudo-resyllabification
- Linking: *r*
- Linking: transient *j*

Consonant dropping/elision

Elision (h-dropping) + linking (where the linking can only occur together with elision)

1. I would have told you about it.

[aɪ wəd əv 'təʊldʒ u(w)əbaʊt ɪt]

2. Can I talk to you about something?

[kən aɪ 'tɔ:k tə ju(w)əbaʊt sʌmθɪŋ]

3. When did you want to meet and discuss it?

['wen dɪdʒ ə wɒnt tə 'mi:t ən dɪ'skʌs ɪt]

4. There was a young man there.

[ðə wəz ə 'jʌŋ 'mæn ðe:]

5. We must look in the locker and in the drawer.

[wi mʌs 'lɒk ɪn ðə 'lɒkə ən ɪn ðə 'drɔ:ə]

6. What have you been doing?

['wɒt əv jə bi:n 'du:ɪŋ]

7. The people who arrive at five PM are too late.
[ðə 'pi:pl̩ hu(w)ə raɪv ət faɪv pi:(j)'em a: tu: 'leɪt]

8. When does he arrive from Paris?
['wen dəz i(j)ə raɪv frəm 'pærɪs]

9. She should have asked for his permission.
[ʃi:ʃəd əv 'ɑ:skt fər ɪz pə'mɪʃn]

10. You shall be on the list.
[ju:ʃəl 'bi:(j)ɒn ðə 'lɪst]

11. He can't have gone behind your back!
[hi 'kɑ:nt əv 'gɒn | bə'hændzə 'bæk]

12. He got me her number and email.
[hi 'gɒt mi(j)ə 'nʌmbər ən 'i:meɪl]

13. There is a lack of answers.
[ðər ɪz ə 'læk əv 'ɑ:nsəz]

14. Could you have been there?
[kədʒu(w) əv 'bi:n ðe:]

15. There were some men who knew them.
[ðə wə səm 'men u 'nju: ðəm]

16. His university was as good as your college.
[hɪz u:nɪvɜ:səti | wəz əz 'gʊd əz jə 'kɒlɪdʒ]

17. No, but there is an umbrella.
['nəʊ | bʌt ðər ɪz ən 'ʌm'brelə]

18. Do you like her more than Jane?
[dʒə 'laɪk ɜ: mɔ: ðən 'dʒeɪn]

19. It's a gift from us for him and his wife.
[ɪts ə 'gɪft frəm 'ʌs | fɔ: 'hɪm ən ɪz 'waɪf]

20. The parents were nice to them.
[ðə 'peərənts wə 'naɪs tə ðəm]

21. To be, or not to be?
[tə 'bi:(j)ə 'nɒt tə bi]

22. Out of everyone here, I am the best.

[aʊt əv 'evriwʌn 'hiə 'aɪ(j)æm ðə 'best]

23. What does it mean to us?

['wɒt dəz ɪt 'mi:n tʊ(w)əs]

24. He said that his brother was an artist.

[hi 'sed ðæt hɪz 'brʌðə wəz ən 'ɑ:tɪst]

9.3 APPENDIX 3

The script for the ‘accentedness’ part of the perception test.

```
"ooTextFile"
"ExperimentMFC 6"
blankWhilePlaying? <no>
stimuliAreSounds? <yes>
stimulusFileNameHead = "recordings/"
stimulusFileNameTail = ".wav"
stimulusCarrierBefore = ""
stimulusCarrierAfter = ""
stimulusInitialSilenceDuration = 1 seconds
stimulusMedialSilenceDuration = 0
stimulusFinalSilenceDuration = 0.2 seconds
numberOfDifferentStimuli = 28
  "F01-07" " |first|second"
  "F01-13" " |first|second"
  "F01-15" " |first|second"
  "F01-24" " |first|second"
  "F02-11" " |first|second"
  "F06-08" " |first|second"
  "F06-09" " |first|second"
  "F06-17" " |first|second"
  "F06-19" " |first|second"
  "F06-20" " |first|second"
  "F07-22" " |first|second"
  "F08-01" " |first|second"
  "F08-03" " |first|second"
  "F08-04" " |first|second"
  "F08-06" " |first|second"
  "F14-05" " |first|second"
  "F14-21" " |first|second"
  "F14-23" " |first|second"
  "F19-12" " |first|second"
  "F19-14" " |first|second"
  "F25-18" " |first|second"
  "F26-02" " |first|second"
  "F26-10" " |first|second"
  "F26-16" " |first|second"
  "F35-01" " |first|second"
  "F35-14" " |first|second"
  "F35-17" " |first|second"
  "F35-23" " |first|second"
numberOfReplicationsPerStimulus = 1
breakAfterEvery = 5
randomize = <PermuteAll>
startText = "In the following experiment you will hear short sentences read out loud.
Each sentence will be realised two times.

Your task will be to choose which one of the two realisations
(of the same sentence) seems to be pronounced with a stronger foreign accent.

This decision is purely subjective; neither the semantic contents of the sentences
nor the identity, gender etc. of the speakers matter in any way.
There are multiple speakers in the whole recording.

You will only hear each sentence pair (i.e. the same sentence pronounced twice) once;
please pick according to your first impression. Click on one of these buttons
- FIRST or SECOND - based on which of the two realisations seems to have a stronger
foreign accent. Do NOT focus on any other aspects than the strength of foreign accent.

Click anywhere on the screen to start the test."

runText = "Choose which realisation has a stronger foreign accent."
pauseText = "Please take a short break now.
The next batch of sentence pairs will start playing after you left-click the mouse."
endText = "Thank you for taking the time to complete this experiment!"
maximumNumberOfReplays = 0
replayButton = 0.2 0.45 0.05 0.15 "Zopakovat" ""
okButton = 0.7 0.85 0.05 0.15 "OK" ""
oopsButton = 0 0 0 0 "" ""
responsesAreSounds? <no> "" "" "" "" "" 0 0 0
numberOfDifferentResponses = 2
  0.3 0.8 0.7 0.8 "angličtina" 30 "" "1"
  0.3 0.8 0.5 0.6 "ruština" 30 "" "2"
numberOfGoodnessCategories = 0
  0.25 0.35 0.10 0.20 "1 (poor)"
  0.35 0.45 0.10 0.20 "2"
  0.45 0.55 0.10 0.20 "3"
  0.55 0.65 0.10 0.20 "4"
  0.65 0.75 0.10 0.20 "5 (good)"
```

9.4 APPENDIX 4

The script for the ‘comprehensibility’ part of the perception test.

```
!ooTextFile"
"ExperimentMFC 6"
blankWhilePlaying? <no>
stimuliAreSounds? <yes>
stimulusFileNameHead = "recordings/"
stimulusFileNameTail = ".wav"
stimulusCarrierBefore = ""
stimulusCarrierAfter = ""
stimulusInitialSilenceDuration = 1 seconds
stimulusMedialSilenceDuration = 0
stimulusFinalSilenceDuration = 0.2 seconds
numberOfDifferentStimuli = 28
  "F01-07" "|first|second"
  "F01-13" "|first|second"
  "F01-15" "|first|second"
  "F01-24" "|first|second"
  "F02-11" "|first|second"
  "F06-08" "|first|second"
  "F06-09" "|first|second"
  "F06-17" "|first|second"
  "F06-19" "|first|second"
  "F06-20" "|first|second"
  "F07-22" "|first|second"
  "F08-01" "|first|second"
  "F08-03" "|first|second"
  "F08-04" "|first|second"
  "F08-06" "|first|second"
  "F14-05" "|first|second"
  "F14-21" "|first|second"
  "F14-23" "|first|second"
  "F19-12" "|first|second"
  "F19-14" "|first|second"
  "F25-18" "|first|second"
  "F26-02" "|first|second"
  "F26-10" "|first|second"
  "F26-16" "|first|second"
  "F35-01" "|first|second"
  "F35-14" "|first|second"
  "F35-17" "|first|second"
  "F35-23" "|first|second"
numberOfReplicationsPerStimulus = 1
breakAfterEvery = 5
randomize = <PermuteAll>
startText = "In the following experiment you will hear short sentences read out loud.
Each sentence will be realised two times.

Your task will be to choose which one of the two realisations (of the same sentence)
seems more comprehensible to you, i.e. is easier to understand, to listen to.

This decision is purely subjective; neither the contents of the sentences
nor the identity of the speakers matter in any way.
There are multiple speakers in the whole recording.

You will only hear each sentence pair (i.e. the same sentence pronounced twice) once;
please pick according to your first impression.
Click on one of the buttons - FIRST or SECOND - based on which of the two realisations
seems more comprehensible to you. Do NOT focus on any other aspects.

Click anywhere on the screen to start the test."

runText = "Choose which realisation is more comprehensible,
i.e. is easier to understand, to listen to."
pauseText = "Please take a short break now.
The next batch of sentence pairs will start playing after you left-click the mouse."
endText = "Thank you for taking the time to complete this experiment!"
maximumNumberOfReplays = 0
replayButton = 0.2 0.45 0.05 0.15 "Zopakovat" ""
okButton = 0.7 0.85 0.05 0.15 "OK" ""
oopsButton = 0 0 0 "" ""
responsesAreSounds? <no> "" "" "" "" 0 0 0
numberOfDifferentResponses = 2
  0.3 0.8 0.7 0.8 "angličtina" 30 "" "1"
  0.3 0.8 0.5 0.6 "ruština" 30 "" "2"
numberOfGoodnessCategories = 0
  0.25 0.35 0.10 0.20 "1 (poor)"
  0.35 0.45 0.10 0.20 "2"
  0.45 0.55 0.10 0.20 "3"
  0.55 0.65 0.10 0.20 "4"
  0.65 0.75 0.10 0.20 "5 (good)"
```

9.5 APPENDIX 5

The following instructions were attached along with the perceptions tests themselves to the email sent to the participants of the perception test.

“Please read through these simple instructions once before you start, then refer to them throughout the perception test.

1. Download the zipped file called “Perception test”. Move it e.g. to your desktop and unpack it through the usual right-click option. Having unpacked the ZIP, you now have a folder called “Perception test”. This folder contains (1) two files, which are called “script A.praat” and “script C.praat”, and then (2) a subfolder called “recordings”, which contains 28 .wav sound files. Keep all the parts where they are, no need to interact with anything else than what is described in the points below.

2. Check carefully that the sound on your computer is on at a level where you can hear speech clearly.

3. You can take the two perception tests in any order you wish. They do not have any connection to each other whatsoever. For the purposes of this manual, let us say you take ‘test A’ first. Open Praat Objects, then choose “Open” -> “Read from file” in the menu on top and choose the file called “script A.praat” from the folder “Perception test” you’ve downloaded. The chosen file will appear in your “Objects” list as “ExperimentMFC script_A”. Select it by clicking on it once, then click on the “Run” button to the right of it.

4. Follow the simple instructions of the test, all you need to do is click. When you have finished the test (“Thank you for taking the time to complete this experiment!” has appeared on your screen), simply close that window and return to the Praat Objects menu. Do not worry, it will look exactly the same as before the test.

5. Now you have returned to the Praat Objects menu, which still contains “ExperimentMFC script_A”. Select it by clicking, then click on “Extract results” on the right. An object called “ResultsMFC script_A” will appear.

6. Select the new object (“ResultsMFC script_A”) by clicking on it. Then, in the top menu, click “Save” -> “Save as text file...”. Choose a location on your computer to save this file, which you will then send back to me.

Repeat steps 2. to 6. with the second perception test, everything is the same, except for the name of the test (which contains the letter C instead of A, or vice versa) and the questions in the test itself.

7. You should now have two files saved in a chosen location on your computer. Their names will be “ResultsMFC script_A” and “ResultsMFC script_C”. The last step is simply sending the results as an email attachment to lenka.kalvoda@gmail.com.”