FILOZOFICKÁ FAKULTA
UNIVERZITY KARLOVY
V PRAZE

FONETICKÝ ÚSTAV

BAKALÁŘSKÁ PRÁCE

Veronika Pojarová

Consonantal and vocalic differences in Czech English with reduced and enhanced foreignness

Souhláskové a samohláskové rozdíly v české angličtině s potlačenou a zvýrazněnou cizostí

Vedoucí práce:
doc. PhDr. Jan Volín, Ph.D.

Praha, 2014

I would like to thank my thesis supervisor doc. Jan Volín for his most valuable advice, impeccable proofreading, and for his kind words of encouragement.

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 another type of diploma.

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.

V Praze, 12. 5.2014 ……

Veronika Pojarová

Abstract

The objective of this thesis is to identify those features of the Czech accent in English that are the most salient in the perception of the Czech listener and that may disturb the communication process. The purpose of the introductory chapter is to familiarize the reader with the subject of the foreign accent, to provide a brief summary of the current state of research and to introduce a series of empirical studies. The research part of the thesis analyzes the individual realizations of the selected speech sounds /θ, ð, ŋ, r, w, æ, ɜ:/ and ventures to draw meaningful conclusions from the results. The material analyzed consists of a total of 3568 speech sound tokens, recorded by 9 male and 19 female speakers. Each respondent produced two recordings, one in the British standard mode and another where the speaker imitated the Czech foreign accent. The 3568 tokens were individually rated and the two modes were then compared for each speaker. The results showed /r/ to be favoured by the largest number of speakers as an indicator of the Czech accent, while /θ, ð, ŋ/ often had the same rating in both modes. However, additional factors such as speaker proficiency and number of tokens from individual speech sounds must be taken into consideration before any final conclusions can be drawn from the raw data.

Keywords

consonant, vowel, foreign accent, Czech English

Abstrakt

Cílem této práce je identifikovat konkrétní rysy českého přízvuku v angličtině, které jsou nejprominentnější v percepci českého posluchače a které mohou působit rušivě v komunikační situaci. Úvodní kapitola si klade za cíl přiblížit čtenáři problematiku cizineckého přízvuku, stručně shrnuje současný stav výzkumu v této oblasti a představuje řadu empirických studií. Ve výzkumné části práce jsou analyzovány jednotlivé realizace vybraných hlásek /θ, ð, ŋ, r, w, æ, ɜː/ a následná interpretace výsledků přináší odpovídající závěry. Analyzováno bylo celkem 3568 hláskových realizací od 9 mužských a 19 ženských mluvčích. Od každého z respondentů byly pořízeny dvě nahrávky čteného textu, přičemž první byla v britském modu a ve druhé se mluvčí snažil o napodobení českého cizineckého přízvuku. Každá z 3568 realizací byla samostatně ohodnocena a poté byly srovnány oba mody pro každého mluvčího. Z výsledků vyplývá, že nejčastějším ukazatelem českého přízvuku bylo v této studii /r/, zatímco /θ, ð, ŋ/ byly často hodnoceny stejně v obou modech. Před ustanovením konečných závěrů je však potřeba vzít v úvahu další faktory, kterými jsou například jazykové dovednosti mluvčích anebo počty realizací jednotlivých hlásek.

Klíčová slova

souhláska, samohláska, cizinecký přízvuk, česká angličtina

Table of Contents

1 Introduction ..... - 8 -

1.1 Second Language Acquisition ..... - 8 -

1.2 The Foreign Accent ..... - 11 -

1.2.1 Factors Influencing the Degree of Foreign Accent ..... - 11 -

1.3 The Stigma of the Foreign Accent ..... - 13 -

1.4 The Foreign Accent As Perceived by Its Speakers ..... - 15 -

1.5 Experimental Research ..... - 18 -

1.5.1 Imitation ..... - 18 -

1.5.2 Perception ..... - 21 -

1.6 Preliminary Hypotheses ..... - 23 -

2 Method ..... - 25 -

2.1 Material ..... - 25 -

2.2 Procedure ..... - 25 -

2.3 Assessment ..... - 26 -

3 Results ..... - 29 -

3.1 Dental fricatives /θ, ð/ ..... - 30 -

3.1.1 Fortis dental fricative /θ/ ..... - 33 -

3.1.2 Lenis dental fricative /ð/ ..... - 34 -

3.1.3 Success rates for /θ/ and /ð/ ..... - 36 -

3.2 Velar nasal /ŋ/ ..... - 37 -

3.3 Alveolar approximant /r/ ..... - 38 -

3.4 Labio-velar approximant /w/ ..... - 40 -

3.5 Front open-mid unrounded vowel /æ/ ..... - 42 -

3.6 Mid-central unrounded vowel /ɜː/ ..... - 44 -

3.7 Overall success rate ..... - 46 -

4 Discussion ..... - 49 -

5 Conclusion ..... - 53 -

References ..... - 55 -

Appendix I ..... - 58 -

Appendix II ..... - 59 -

Appendix III ..... - 67 -

1 Introduction

1.1 Second Language Acquisition

One of the results of economic globalization and increasing mobility is the ever-growing need for communication among people of different cultures and different languages, which has led to the increasing interest in second language acquisition (henceforth SLA), both in theory and in practice. Apart from necessary business interactions and migration due to war, famine or poverty, the leading motivation for learning a second language (most often English) is the fact that education for native speakers of minor languages is generally provided in English, or another one of the world languages. Possibly the largest group of people with motivation to learn a second language are immigrants, for whom it is essential to learn the language of their new country in order to be able to find employment, make basic everyday transactions and eventually become a fully integrated part of the community. Doughty and Long (2008, p. 5) similarly identify the motivation and circumstances for SLA, establishing that, "[states] and citizens, scholars and laypersons alike recognize that learning a society's language is a key part of both acculturation and socialization."

It is generally believed that, as far as SLA is concerned, the lower the age of the student, the more successful the outcome of the learning process will be. This is especially true concerning pronunciation, as has been proved by many independent studies in which it was shown that phoneme discrimination abilities of children are age-dependent. It was demonstrated that the children's sensitivity to the smallest nuances in segment qualities was inversely proportional to their age, due to the fact that their native phonological system had not yet been fully acquired. Infants are particularly sensitive to changes in sound qualities, which was confirmed in a number of experiments. Such studies include that of Werker and Tees (1983), which tested the responses of English subjects of different ages to two Hindi speech contrasts. The experiment was conducted on 4, 8 and 12-year-olds, with 12 subjects in each group. The testing paradigm was a variation of that used in the infant Head Turn experiment (Werker et al., 1981), where a sequence of sounds was played to infants, who were conditioned to turn their heads towards the source of the sounds whenever they heard a change in them. Correct head turns were reinforced with the activation of a toy and praise from the assistant. The 1983 study was altered slightly in order to fit the age of the subjects, requiring subjects to press a button when the sound stimulus changed, whereupon a toy was

shown to them as a reward. The resulting data was then compared with the figures from Werker et al., 1981, and the numbers clearly show that the ability of infants to discriminate phonemic contrasts in the experiment was quite remarkable, not at all like that of older children, whose aptitude at perceiving finer distinctions between sounds was poor regardless of their age.

Working with a similar premise in mind, Flege, Munro and MacKay (1995a) conducted a very extensive and carefully controlled study on the relationship between the age at which native speakers of Italian began learning English and their production of English consonants. The 240 native Italian speakers arrived in Canada between the ages of 2 and 23, and at the time of this study had been living there for at least 15 years, with an average of 32 years, speaking more English than Italian on a daily basis. Flege et al. work with two concepts here; the age of learning (henceforth AOL) and the age of arrival (henceforth AOA). The AOA is the age at which the subjects arrived in Canada, whereas the AOL is the age at which they began learning the English language. An evaluation of their accent was based on their pronunciation of consonants in the words *pick, peak, tack, tag, tacking, tagging, cap, cab, read, raid, they, then, thought, thief.* (Flege et al., 1995a, p. 5) These word tokens were subjected to perceptual evaluation by native speakers of English, and also acoustical analysis. The age at which the learner acquires the second language proved to be crucial in all of the experiments.

Experiment 1, dealing with word-initial /ɹ/, /ð/, /θ/, showed a significant break around the 8th year as the age of learning, after which the number of cases in which the /ɹ/, /ð/, /θ/ tokens were pronounced correctly decreased rapidly. The native English speakers judged the alveolar approximant /ɹ/ as "distorted", rather than as a different sound, such as /l/ or /w/. The interdental fricatives /ð/, /θ/ were most often heard as /d/ and /t/ respectively. Experiment 2 showed that the native Italian subjects whose AOL was 15 to 21 years (but not 3 to 13 years) were less consistent in differentiating between word final /t/ and /d/, and instead pronounced the voiceless variant. Experiment 3 expanded on the previous, this time testing the subjects' word-final /p/, /b/, /k/, /g/. The results showed that a significant deviation from the standard pronunciation can be observed among native Italian subjects whose AOL was higher than 17 years for /k/, /g/ and higher than 19 years for /p/ and /b/.

Together, the results revealed that some native Italian late learners of English had difficulty producing a contrast between fortis and lenis in word-final stops, and also had

problems correctly pronouncing word-initial /ɹ/, /ð/, /θ/. With these results, the study concludes that the age of L2 (second language) learning plays a significant role in the speaker's ability to correctly pronounce the L2 sounds, in this case L2 being English. Flege et al. offer a general summary of possible causes of errors in L2 pronunciation, which are as follows:

1. Inadequate phonetic input, e.g. learners most often hear the L2 spoken with a foreign accent.

2. Psychosocial factors, e.g. "ethnic identification", as described with reference to Segalowitz and Gatbonton (1977). The level of motivation to sound like a native English speaker varies and can play an important role in determining the strength of the resulting accent. The premise is that the more nationalistic the speaker's views are, the less likely he is to acquire the standard pronunciation of English speech sounds.

3. Motivation can be smaller, especially if pronunciation errors are non-impeding.

4. Habit formation, i.e. especially habits kept from the early days of L2 learning.

5. Motoric difficulty, connected with the loss of the motoric ability to produce new sounds, especially those which are articulatorily difficult, once the L1 phonological system has been fully established.

6. Incorrect perception. The ability to perceive differences between English sounds and those L2 sounds most similar to them is weaker the higher the age of learning is.

7. Phonetic system effects. The interaction of two different phonological systems can lead to the deviation of sounds in both systems (L1 and L2)

1.2 The Foreign Accent

In the study of the L2, L1 intrusion affects the speaker's output on the level of grammar, lexis and pronunciation, the latter being arguably the most easily noticeable. What is perceived as a foreign accent is essentially a set of characteristic and consistent deviations from the standard pronunciation. Such deviations are often manifested in a differing manner or place of articulation, difference in voicing, vowel openness or segment duration, and it is exactly these segmental aspects that will be studied in this thesis.

1.2.1 Factors Influencing the Degree of Foreign Accent

The salience of a foreign accent is subjective and cannot be objectively measured, yet most native speakers perceive it in a similar way and share comparable judgments. There are many factors which can and do affect the perceived strength of the foreign accent, and Flege, Munro and MacKay (1995b) attempted to identify them. In their earlier study, Flege et al. (1995a) established that the AOL plays an important role in the strength of the speaker's accent. This only supported the already existing concept of a critical period in foreign language learning, which Lenneberg (1967) first proposed as a period, lasting from approximately two years of age to puberty, after which language acquisition becomes more and more difficult due to biological factors, namely the lateralization of brain functions.

In their 1995b study, Flege et al. utilized the same group of subjects as they had in their 1995a experiment, that is the 240 native Italian speakers, with an additional 24 native English speakers who functioned as a control group. Prior to the actual experiment, the native Italian speakers were asked to fill in a language background questionnaire consisting of 34 items, exploring for example:

- the amount of formal instruction in the L2
- musical ability and the ability to imitate
- strength of concern for pronunciation
- motivation
- overall L2 input / use

The age of 12 proved to be a kind of breaking point; 90% of native Italian speakers whose AOA, and therefore also AOL, was lower than this age considered English to be their

better language, whereas only 26% of the subjects who's AOA and AOL were higher than 12 years gave this answer. When answering the question of which of the two languages, English or Italian, they would be less willing to lose, the majority chose English (86% of speakers with their AOL lower than 12 years and 79% whose AOL was higher), showing that all speakers needed English for everyday communication purposes, which was confirmed by their frequency of use; even late arrivals spoke more English than Italian, although the difference decreased with the growing AOL.

Based on the speakers' realizations of five sentences *(I can read this for you; The red book was good; Paul ate carrots and peas; The good shoe fit Sue; He turned to the right),* the 10 native English listeners determined the strength of the accent according to their own perception by positioning a lever on a response box. The results show that the first group judged by the listeners as "foreign-accented" had an average AOL of 7,4 years. "Authentic" pronunciation was achieved, according to the native English listeners, by 78% of the native Italian subjects whose AOL was lower than 4 years, 61% with an AOL of 4-8 years, 29% with and AOL between 8-12 years, only 6% with an AOL higher than 12 years, but no one whose AOL was higher than 16 years.

According to the results obtained from this experiment, the AOL has the most significant impact on the strength of the foreign accent. Of the other factors, the relative frequency of use of the two languages also proved to be significant, the length of residence in Canada was somewhat less significant (accounting for only 1,9% of the variance for males, 1,6% for females).

1.3 The Stigma of the Foreign Accent

Although today's Western society is one ruled by political correctness, where great emphasis is placed on tolerance and intercultural communication, a world where awareness campaigns are being led against racism and discrimination, it is still one whose people harbor negative attitudes to foreign accents. Tolerant though we believe ourselves to be, certain prejudices and stereotypes are so deeply rooted in our subconscious that we sometimes make judgments without realizing it. These negative attitudes account for the existence of a stigma related to speaking with a foreign accent. Prejudices regarding foreign accentedness may inadvertently have a negative effect on both the speaker and the listener. The non-native speaker whose pronunciation deviates from the local standard will find that his manner of speech may present him with a number of difficulties, possibly in the form of social barriers, psychological problems or hindrances in the actual communication process. The listener, on the other hand, is confronted with the discomfort of having to make a greater effort in order to comprehend the utterance, facing the possibility of embarrassment due to misunderstanding.

The stigma of the foreign accent has been studied predominantly from the point of view of the listener, Derwing and Munro (2009) mention a number of studies aimed at listener sensitivity, for example Flege (1984), who found that phonetically untrained listeners could consistently discriminate between native English and French-accented speakers based on 30ms of speech. They further mention their own series of studies (Munro, Derwing & Burgess, 2003), where they discovered that listeners are able to reliably identify native versus non-native accents in speech which is played backwards. The predominance of listener-oriented studies is, according to Derwing and Munro, justified by the reliability and consistence of listener's judgments, which are "the only meaningful window into accentedness and comprehensibility." (2009, p. 478). Derwing and Munro identify three basic factors which are to be taken into account when analyzing the listener's reaction to hearing foreign accented speech:

1. accentedness: the extent to which the speaker's accent differs from the local standard
2. comprehensibility: a measure of the effort needed in order to comprehend what the speaker is saying. This factor is significant, as it is often the case that the listener is less willing to communicate with the L2 speaker for fear of having communication problems (Derwing & Munro, 2009, p. 487).

3. intelligibility: the final result, showing how much of the utterance the listener has actually understood. For this factor it is difficult to determine one reliable method of assessment; one possible way mentioned by Derwing and Munro is based on the percentage of words correctly transcribed by listeners.

Derwing and Munro proceed to examine the social aspects of accent, namely the problem of accent-based stereotypes and prejudices. Although accent-based prejudices are generally considered to be negative, there are certain accents which are often associated with sophistication and can be intentionally adopted by speakers who wish to be seen in a certain light. One of the examples presented here is that of actor Maurice Chevalier, who is said to have exaggerated his French accent, because it was considered "charming" by many people. The desire to keep one's identity through one's foreign accent is often argued to be a key motivation factor for speaking with a foreign accent, previously identified as "ethnic identification" by Flege, Munro and MacKay (1995b). Derwing and Munro do not share this view, maintaining that losing a foreign accent and identity preservation are fully compatible. To back this opinion, they introduce the research of Timmis (2002), who interviewed 400 students studying English as a foreign language and found that 67% of them aspired to pronunciation without a foreign accent. In a similar study of 100 adult students of English as a foreign language, Derwing (2003) found that 95% wanted to sound like a native speaker.

Although the point of view of the listener is important, it is also essential to take into account the L2 speaker's experience, which should not be neglected, especially from the psychological point of view. The speaker of an L2 is inevitably subject to judgment from the listener, and hence may be exposed to traumatizing experiences. This aspect was further explored by Gluszek and Dovidio (2010b), who focused on the experiences of speakers with non-native accents. They hypothesized that speakers are subject to negative experiences in two areas: a) perception of their stigmatization, and b) problems in communication. Linked to both of these aspects is then the speaker's sense of belonging to the community, which they argue is one of the strongest human needs.

In Study 1, an online survey was conducted for participants with both native (126 participants) and non-native accents (77 participants), and contained three sets of questions. The first, titled Perceived Stigmatization measure, was only for non-native speakers of English, in this section they rated the strength of their accent and then indicated on a Lickert scale of 1 to 7 the extent to which they agreed with statements such as *"I think that in certain*

situations I am being discriminated against because I have an accent" (Gluszek and Dovidio, 2010b). The second section, the Conversational Problems scale, assessed general problems in communication, both for native and non-native speakers. The third part, Difficulties in Communication, explored the experiences of non-native speakers in communication situations, where speakers with non-native accents answered for themselves, whereas speakers with native accents of English answered as they supposed a non-native speaker would have done. The results confirmed all four initial hypotheses:

1. for non-natively accented speakers, Perceived Stigmatization correlated with Conversational Problems and Difficulties in Communication
2. those who rated their accents as stronger scored higher on the Perceived Stigmatization scale
3. speakers of European accents reported less discrimination than Asians or Latinos
4. non-natively accented people experience greater difficulties in communication

Study 2 examined the difference between having a non-native accent and a regional accent in the mind of the speaker. Although both types of accent are generally rated negatively, the Gluszek and Dovidio hypothesize that speakers with regional accents will not be prone to having feelings of not belonging in the United States.

1.4 The Foreign Accent As Perceived by Its Speakers

Speakers of an L2 are aware of their own accent, caused by the intrusion of their L1. It is often the case that this leads to the speakers becoming self-conscious, lowering their confidence and willingness to communicate in the L2. Every L2 accent has its specific features, based on the phonological structure of the speaker's native language, some of the deviations from the standard pronunciation being more salient than others. It is no difficulty even for a layperson to identify the nationality of the speaker based on the intrusion of the L1, provided that they are, at least on a basic level, familiar with the given L1. The prominence of these features has a direct effect on the perceived strength of the foreign accent. In 1959, Soudek published a paper commenting on some of the most frequent deviations from the standard pronunciation of English by Czech native speakers. The standard for reference was *Received Pronunciation* and the study was conducted with 37 Czech students of English at the

Faculty of Arts' English Department, at Komenský University in Bratislava. Unfortunately, there is no information regarding the age of the students, nevertheless we are informed that 29 of them (78,3 %) had been learning English for longer than three years, while the remaining 8 students (21,6 %) had been learning English for less than three years. During the years 1956-1959, Soudek observed the 37 students, each of whom took part in the so-called conversation exercises over a period of two months. During these sessions, their pronunciation of English was carefully analysed by Soudek, and the most frequent deviations from the RP standard were then presented in his paper. The results showed that the most common deviations are as follows:

1. Fortis articulation of final lenis consonants was very common, so that for example [bed] became [bet]. This was observed for [b], [d], [ð], [g], and [z], as well as the written endings *–s* and *–ed*, even following a voiced sound. A staggering 97,2% of students pronounced final lenis consonants in this manner.

2. The speech sound [ð] was replaced with [dz] by 64,8% of students, 18,9% used simply [d] and 10,8% pronounced [dh]. Its voiceless counterpart, [θ], was pronounced as [s] by 91,8% of the students.

3. The standard English glottal fricative [h] was pronounced as its voiced counterpart, [ɦ], by 94,5% of the students.

4. 94,7% of students pronounced the "clear" [l] in all contexts, that is even when an RP speaker would pronounce the "dark" [ɬ]. The remaining 5,4%[1] of students always pronounced the dark [ɬ], even if front of a vowel.

5. 72,9% of students pronounced either [ŋk] or [ŋg] word-finally, as opposed to the standard [ŋ] pronunciation.

6. Aspiration of the voiceless plosives [p], [t], [k] was completely neglected by 86,4% of students.

7. The standard English [æ] was pronounced as [e] by 89,1% of students.

8. The mid-central [ə] was pronounced as [e] mainly in word-initial positions and in the definite and indefinite articles by 78,3% of students. In word-final positions of words ending in <-or> and <-er>, [ə] was replaced with either the vibrant [r] or the retroflex approximant [ɻ] by 62,1% of students.

[1] The fact that 94,7% and 5,4% do not add up to precisely 100%, is known to us, however these are the figures presented in Soudek's original paper.

Unfortunately, the students were not tape recorded. This was in order to create as stress-free an environment as possible, so that the speakers would not concentrate on their articulation. However, this also means that the results are without rigorous basis and are entirely subjective, because Soudek "had to rely entirely on his own acoustic impressions." (Soudek, 1959, p. 75). In fact, objective results or absolute precision were never the aim of this study, as it was only meant to be a series of general observations.

This thesis will employ acoustic measurements in order to objectively identify the basic characteristics of Czech pronunciation of English; that is, those which distinguish it from other accents of English, and even more specifically those which are shared in our speakers' perception of their accent. When a native speaker of Czech is asked to imitate the Czech accent in English, it is very likely that he will work with these most frequent deviations from the standard English pronunciation in order to achieve the most authentic Czech accent of English. Studies concerned with the perception and imitation of foreign language accents have yielded results which have relevance to the present study and which will be introduced in the following section.

1.5 Experimental Research

In this section, five empirical studies will be introduced and subsequently compared, taking into account their relevance to this thesis. The studies chosen for this purpose are: "Imitation of Suprasegmental Patterns by Non-Native Speakers of English" (Haak & Darling, 1993); "Dialectal Feature Imitation in Norwegian" (Dommelen, 2011); "The same but different – three impersonators imitate the same target voices" (Zetterholm, 2003); "Perception of French vowels by American English adults with and without French language experience" (Levy & Strange, 2008); "Perception of foreign accent in spontaneous L2 English speech" (Yuan et al., 2010).

The first three are concerned with imitation as a means of comparison, and the remaining two study the perception of foreign sounds and accents.

1.5.1 Imitation

In their paper, Haak and Darling (1993) employed The Tennessee Test of Rhythm and Intonation Patterns (TRIP) in an attempt to objectively distinguish the prosodic patterns of native and non-native English speakers. The TRIP consists of 25 items, each of which has a specific rhythm and intonation, using the syllables "ma". In the first experiment, the test was administered to 12 native speakers of Asiatic languages and 12 native speakers of English, who imitated the items. The experimenters then evaluated the responses according to the guidelines in the test manual and arrived at a final percent-correct score for each test subject. For non-native speakers this ranged from 80% to 100%, for native speakers from 84% to 100%. It has been agreed upon by the experimenters themselves that this difference is not at all significant. In truth, this study has very little to do with language, and only shows the proficiency level of the speakers to imitate melody and rhythm. It has been included here to show how imitation can be used to acquire objective data by comparing the model item and the speaker's replication of it.

An interesting study was conducted by Zetterholm in 2003, who worked with the premise that voice impersonators pick out the most perceptually significant features in a voice and imitate them. The main point of interest was whether three different male impersonators (the first two professional, the third an amateur) of the same target speaker would choose the same features. The material used were recordings of 9 target speakers, 22 imitations by the three impersonators, and the impersonators' own voices. Unfortunately, only one

impersonator imitated all of the target voices, so the analysed data are not complete. First, the data was analysed acoustically by several phoneticians at Lund University, where the listeners made judgments of pitch, voice quality, speech tempo, dialectal markers and individual phonetic habits. Their conclusion was that all three impersonators had predominantly chosen the same characteristic features of the target speakers and that all three had manipulated their pitch level so that it would correspond to that of the target speaker, although the impersonator III was not as successful in achieving a sufficient level of variation. Voice quality was another important feature which the two professional impersonators managed to mimic successfully, according to the listeners. All three impersonators seemed to take into consideration the tempo, rhythm and pausing of the target speakers, to the point where some imitations were slightly exaggerated. The subsequent acoustic analysis confirmed that the imitators altered their F0, the two professionals doing so with more variation.

One target speaker was examined more closely, and it was shown that impersonators I and II attempted to alter their natural voice quality in order to imitate the speaker's sonorous and nasal voice. They also copied the speaker's intonation pattern, which impersonator III did not, and they successfully copied the perceived slow speech tempo by means of a slower articulation rate, whereas impersonator III was unable to adhere to the initial articulation rate and sped up. However, all three impersonators managed to capture the clear articulation and the trilled [r]. This fact could suggest that segmental aspects of speech are easier to imitate than the suprasegmental, which is noteworthy as this thesis will focus on imitation on the segmental level.

"Dialectal Feature Imitation in Norwegian" (van Dommelen et al., 2011) examines the ability of native speakers of Norwegian, specifically speakers living in Stavanger and Trondheim, to imitate each other's dialects. The speakers were played a recording of the dialect they were to imitate, and were then asked to read 60 sentences in that dialect. Finally, the speakers read the same sentences, but this time in their own dialect. The quality of the imitation was determined by the speaker's realization of preaspiration in vowel – voiceless stop contexts, which is considered typical feature of the Stavanger but not the Trondheim dialect. It was discovered that both dialects have preaspiration, its average length 65ms in the Stavanger and 50ms in the Trondheim dialect. When imitating, the Stavanger speakers reduced their preaspiration duration by 13ms, and the Trondheim speakers increased theirs by 10ms. The articulation rate was also taken into account, and the result was that it was similar for both groups of speakers, and that both decreased their rate of articulation when imitating,

quite possibly due to increased speaking effort. For our purposes, the most important finding is that all speakers were on some level aware of the dialects' differing degree of preaspiration and capable of altering their articulation in order to imitate it.

These three papers were chosen to illustrate the use of imitation when studying a foreign language accent, a speaker's perception of a different dialect or of another speaker's phonetic habits. What all three have in common is that the test subjects were in all three cases asked to imitate sounds or speech that they had heard shortly before; Haak and Darling had their speakers repeat items from the TRIP after hearing them, Zetterholm's imitators mimicked their target speakers, and van Dommelen et al. played recordings to their speakers based on which they could then reproduce the given dialect. In this thesis, however, the speakers were asked to imitate the Czech foreign accent in English based solely on their experience and their own idea of what the Czech accent sounds like. In each of the presented studies, the acquired recordings were then compared with other recording which were chosen carefully to provide a point of reference; Haak and Darling compared the subjects' responses with the original items on the TRIP, judging deviations from the original as incorrect. Zetterholm's results were based on the comparison of the imitators' natural voices and the actual imitations, which were examined together with the target speakers' voices. Van Dommelen and his colleagues compared the imitations with the recordings done in the speaker's own dialects, thanks to which it was possible to determine whether the recorded preaspiration was part of the speaker's phonetic habits or an imitation of the given dialect. In this thesis, the recordings of the Czech student's imitations of the Czech foreign accent will be compared with recordings in which they spoke without attempting to mimic the Czech accent.

The conclusions of the studies can help with the forming of a hypothesis for this thesis. Zetterholm concludes that the imitators focus on the same characteristics of the target voice, and van Dommelen et al. show that all speakers made use of the typical differences in preaspiration when imitating the other dialect. Based on this we may hypothesize that our Czech speakers of English will choose to focus on the same features of the Czech foreign accent and that these features will be the most dominant. In Zetterholm, some features were exaggerated to some extent, which is a very natural thing to do when imitating, and therefore it can be expected that the Czech students may also exaggerate certain features. Our expectations are therefore as follows:

1. our subjects will choose to imitate a selection of features which most noticeably distinguish the Czech accent from the British standard

2. the features chosen will be the same, or very similar, for all the participating students

1.5.2 Perception

The remaining two empirical studies to be presented in this section, Levy and Strange (2008) and Yuan et al. (2010), explore the factors affecting the listener's perception of the foreign accent.

In their 2008 study, Levy and Strange aimed to determine how great a role is played by the listener's linguistic experience with a foreign language in the process of discriminating between two of the given language's speech sounds, the said sounds being absent from the listener's L1 phonemic inventory. Previous studies had suggested as much, and had also shown that as far as vowel perception is concerned, consonantal context is of utmost importance. The Levy and Strange study investigated American English adults' perception of the Parisian French vowel pairs /u-y/ , /œ-u/, /i-y/, /i-u/ and /i- œ/, based on the extent of their experience with the French language. The material used was recorded by three female native speakers of Parisian French and consisted of nonsense words containing the 9 French vowels[2]. The nonsense words /rabVp/ and /radVt/ supplied the vowels with differing consonantal contexts and were then set in a carrier sentence. The subjects, 20 native speakers of American English, were divided into two groups of 10- those without any previous experience with French and those who had studied French with a mean of 7 years of language instruction. A control group of native speakers of French was also included in the experiment. During the testing, the subjects rated 6 blocks of 24 trials, each consisting of three phrases where they were to determine whether the middle phrase contained the same vowel as the preceding or the following phrase. The results revealed the following:

1. participants with previous L2 experience were generally more successful at discriminating the vowel contrasts, their accuracy being very close to that of the natives.

[2] Although the target speech sounds for the experiment were only the 4 vowels /u, y, œ, i/, all 9 French vowels were recorded in order to minimize the risk of distortion caused by excessive lip rounding.

2. vowel pairs were not equally difficult to distinguish, the most problematic pair being /u-y/

3. consonantal context proved to be an extremely significant factor, especially for the inexperienced group. The /u-y/ contrast was especially problematic in the /dVt/ alveolar context, whereas the opposite was true of the vowel pair /i-y/.

For the purpose of the present study, a number of Levy & Strange's results are of relevance for the following reasons. The importance of previous L2 experience is demonstrated by our choice of subjects, who were chosen for their high level of English on the premise that the more familiar they are with the given English phoneme, the more capable they will be of producing it in its British standard form as well as in a Czech accented variation. Another important factor to be taken into account is consonantal context or, in our case, context in the broader sense. It might also be supposed that context plays an important role not only in perception but also production. As will be shortly discussed in Chapter 3, the correct pronunciation of a phone may be, to some extent, dependent on the word it appears in. For example, the pronunciation of [ð] in a grammatical word such as *the* which a) has a high occurrence rate and b) is practically empty of meaning, is more likely to be neglected than the pronunciation of [ð] in the word *father*.

Yuan et al. (2010) studied the effect of the listener's L1 on his perception of the strength of various foreign accents. The material used consisted of spontaneous English speech by subjects who were asked to speak about themselves for approximately 20 seconds. The subjects were native speakers of one of the eight selected languages, which encompassed three tone languages (Cantonese, Mandarin, Vietnamese), four stress languages (German, French, Spanish, Russian) and one pitch accent language (Japanese). The experiment was conducted in the form of an online perception test, in which eight university students of English, all native speakers of Mandarin Chinese, judged the strength of the accents presented to them on a four-point scale from "1" as an either a negligible or non-existent accent, to "4" for a very strong accent with limited intelligibility. Yuan et al. referred to this evaluating group as the "Mandarin judges." The same test was taken by three native speakers of American English, the "native judges." The average assessment scores of both groups of judges were compared, yielding the following results:

1. on average, compared to the native judges, the Mandarin judges rated the accents as milder on every one of the four levels. These results suggest that sensitivity to the

strength of a foreign accent could be influenced by whether or not the speaker's L2 is the native language of the listener.

2. the average strength of the accents as perceived by the Mandarin judges was dependent on the L1 language of the speakers. It was shown that the Mandarin judges were least sensitive to the Mandarin and Cantonese accents, while the highest perceived degree of foreignness was assigned to French, Spanish and Russian accented utterances. This could possibly be attributed to the fact that Mandarin is the L1 of the listeners and Cantonese is a language structurally similar to it, suggesting that structural similarities of the speaker and listener's native languages may have an effect on the perceived degree of foreign accent.

In the next part of the study, Yuan et al. conducted an acoustic analysis of the material used for the perception test, with the aim of identifying the main cues which could have played an important role in the assessment of the foreign accent. The acoustic features assessed were all prosodic, including F0 variation, the number and total duration of pauses and the average duration difference between stressed and unstressed vowels (Yuan et al., 2010, p. 3). Regression trees were built for each group of judges, and the conclusions drawn from comparing these trees were that English and Mandarin judges relied on different cues, but that speaking rate and pauses were important for both groups in determining the strength of the foreign accent. Possibly the most interesting conclusion is that F0 variation and deviation from the standard L1 pronunciation were important cues for the English judges, but not so for the Mandarin judges, suggesting that they either cannot perceive the cues or are unable to utilize them when assessing the strength of foreign accent in English.

1.6 Preliminary Hypotheses

Based on the current state of research on the subject of the foreign accent and the findings of the studies presented here, we can formulate several predictions as to the outcome of the present study.

1. When asked to imitate the Czech accent in English, native Czech speakers will focus on what they consider to be the key features accounting for the most perceptible differences between the Czech accent from the British standard.

2. The features will be shared by all or most of the speakers in their conception of the Czech foreign accent.

3. There will be a discernable difference in the strength of their accent in the standard mode and in the Czech mode.

2 Method

2.1 Material

The material analysed in the present study consists of recorded speech by 28 native speakers of Czech. The total duration of the speech material is 34.46 minutes, with an average duration of 73,84 seconds per speaker. The speakers, 9 male (DVRK, HSKR, KACT, KOSK, MRES, PVLK, SLAB, VIDR, VLES) and 19 female (BCKA, BNDA, BSTA, CMFA, FJTA, HBNA, HBTA, HMNA, HSKA, KVTA, MCHA, MLCA, MLKA, RMSA, SKDA, SNKA, TCHA, VLKA, ZLKA), are students of English at the Institute of Translation Studies, Faculty of Arts at Charles University in Prague, their ages ranging from 19 to 22 years. The students were asked to read a BBC news bulletin twice (see Appendix I for full text), and before each reading they were given several minutes to familiarize themselves with the text and prepare for the recording session. The speakers were individually recorded with an electret microphone IMG ECM 2000, soundcard SB Audigy 2 ZS, 32-kHz sampling frequency and 16-bit resolution, in an acoustically treated room. During the first reading, their task was to suppress their Czech accent and aim for the British standard of Received Pronunciation. For the second recording, they were asked to imitate the Czech foreign accent as authentically and accurately as possible. The result is a total of 56 recordings, 28 in the British standard mode and 28 in the Czech mode.

2.2 Procedure

The speech material was processed in Praat (http://www.fon.hum.uva.nl/praat/) in the following manner. Each recording was cut into breath groups, dividing the text into meaningful wholes. If the newly acquired sound was shorter than 1.2 seconds, it was attached to either the left or right adjacent sound, depending on which resulting whole sounded more coherent. The utterances were then phonetically segmented on the tiers "phone" and "word" using the Penn Phonetics Lab Forced Aligner (https://www.ling.upenn.edu/phonetics/p2fa/). The output of the aligner was transcription of the utterances based on American English pronunciation, which had to be for our purposes manually changed to reflect the British standard pronunciation. The automatic segmentation places segment boundaries only approximately, and so their manual adjustment was needed. This was done in accordance with Machač and Skarnitzl's guide to phonetic segmentation (2008).

The segmentation procedure is quite straightforward, save for the transcription of words with an underlying /r/ which is not pronounced in non-rhotic varieties of English; such words were transcribed with the /r/ regardless of whether its phonetic realization was present or not. As a result, words such as *minister*, which is typically pronounced as /mɪnɪstə/ in the British standard, was transcribed as /mɪnɪstər/. The presence of the underlying /r/ becomes apparent in the phrase *The prime minister of Kenya* the linking /r/ is generally pronounced in the British standard. This method of transcription was employed in order to take into account all possible realizations of the analysed speech sounds, which in the case of /r/ includes complete elision.

2.3 Assessment

For the purpose of this study, the speech sounds analysed were 5 consonants, /θ, ð, ŋ, r, w/, and 2 vowels, /æ, ɜ:/. These speech sounds were chosen by the author of the present study based on the results from Soudek's earlier research (1959) and also on her own experiences with the Czech accent. With the exception of /r/, none of the sounds are present in the Czech phonemic inventory, and although /r/ is a Czech phoneme, it differs from the British standard in manner of pronunciation; a trill in Czech and an approximant in British English. Therefore the assumption may be made that the selected speech sounds will problematic for the Czech native speakers.

For every analysed phone, an auxiliary tier containing intervals with the target speech sound was generated using a script. The recordings were methodically examined for each of the relevant speech sounds, and the author's rating of the pronunciation was recorded directly in the auxiliary tier's intervals corresponding to the analysed phoneme. The determining factors for the author's judgements were primarily the acoustic impression, and additionally the visual manifestation of the speech sound on the spectrogram. For 6 of the 7 speech sounds, a three point scale was employed. On the scale, "0" represented zero deviation from the British standard, "2" indicated an apparent deviation from the standard, and the rating "1" was reserved for ambiguous cases and was used sparingly. For the phone /r/, and additional fourth rating was introduced; an "x" for cases where there was a legal elision. A brief description of specific problems appearing during the assessment process and of the final solutions can be found below.

The phones /θ, ð/ are often mispronounced as /s, z, t, d, f/ by Czech speakers of English, where the deviation from the standard is perceptually very distinct. However, there were cases when the native English speaker would in all probability replace the canonical dental fricative /ð/ with and alveolar plosive /d/ or even an alveolar flap /ɾ/. This affects grammatical words in an utterance with an overall high tempo, and the resulting speech sound is not perceived as foreign in the context of the utterance. This was the case in the phrase *but the prime minister said*. The primary auditory impression was positive and the /ð/ speech sound realization would have been scored with a "0", but upon inspecting the spectrogram a full explosion was discovered. For the sake of consistence it was necessary to rate such tokens with a "2".

When analyzing the pronunciation of /θ/ and /ð/, it is necessary to bear in mind that they have not always had the phonemic relationship they do today, and should therefore be analyzed separately. Historically, the fortis dental fricative /θ/ had no lenis counterpart in Old English, and /ð/ appeared only as an allophone between two voiced speech sounds. The lenis dental fricative became a phoneme in Middle English, as part of the development of phonemic voiced fricatives /v, z, ð/. The Old English speech sounds /f, s/ became voiced in the Middle English period due to the influx of French loanwords, while for the development of /ð/ the most important factor was the voicing that occurred in "very lightly stressed words, especially function words like *is, was, of, his, the, then, that,* and *they*." Millward (2012, p. 150). Very much like in Middle English, today the fortis /θ/ is predominantly found in lexical words, whereas its lenis counterpart /ð/ is most often reserved for function words. There is only a handful of minimal pairs /θ/ and /ð/, for example *thigh* and *thy*, *either* and *ether*, and some near-minimal pairs, such as *breath* and *breathe* (Smith, 2007). However, the distinction between the two phonemes does not carry a large functional load. The claim that /θ/ and /ð/ are not in a true phonemic relationship may be further supported by the fact that Czech speakers of English often substitute /s/ for / θ/ but /d/ or /dz/ for / ð /. For a more detailed description, refer to Soudek (1959) in Chapter 1. For these reasons, / θ, ð / were analyzed both together and separately.

The rating of /ŋ/ realizations was limited to the word final position in the suffix *-ing*, as this is the context in which Czech speakers frequently add a velar plosive after the velar nasal, so that the resultant pronunciation of *fleeing* is often [fliːɪŋk], in extreme cases [fliːɪŋg] or even [fliːɪnk]. In the Czech phonemic system, /ŋ/ is an allophone of /n/ in velar contexts, so the pronunciation of /ŋ/ in words such as *finger* is not generally expected to cause trouble.

One of the most striking differences between the British standard and the Czech pronunciation of English concerns the realization of /r/ segments. In the British standard, /r/ is canonically pronounced as a post-alveolar approximant, whereas Czech speakers usually pronounce an alveolar trill. However, it was necessary to introduce a fourth marker, "x", owing to the fact that the distribution of /r/ in the British standard pronunciation is such that "the phoneme occurs only before vowels." (Roach, 2009, p. 50) In all other cases, for example words such as *minister*, the rhotic speech sound is not realized but remains as an underlying component, manifesting as a so-called linking "r" when followed by a vowel, e.g. *minister is* pronounced [mɪnɪstər ɪz].

Due to the presence of vowel reductions in English unstressed syllables, it was necessary to eliminate grammatical words from the selection of words containing the vowel /æ/, e.g. *that, can*, as it was highly likely that they would predominantly appear in their reduced forms.

Once all the 7 speech sounds had been individually rated, the scores were extracted from the auxiliary tiers using a script, the output of which was a list of speakers, the mode (British standard or Czech), the words containing the target speech sounds and the speech sound's rating in each word. The results derived from the acquired data will be presented in the following chapter.

3 Results

The raw data were processed for each analysed speech sound individually and the output for each of these was a table containing the following data. For every speaker, the sum of "0, 1, 2" ratings was recorded; first for the Standard mode, then for the Czech mode, henceforth STD and CZE respectively. The total number of STD and CZE tokens was calculated. Furthermore, the percentage of "0" ratings (for a full description, refer to chapter 2) was calculated for each speaker, both in the STD and CZE modes. For the speech sound /r/, the additional rating "x" was included in the calculations. Due to the fact that the elision of /r/ in specific phonemic contexts is a feature of the British standard pronunciation, the rating "x" was unified with "0" when calculating the percentage of "0" ratings. This made it possible to obtain the success rate of each individual speaker; the size of the difference between the STD and CZE percentage values directly correlates with the success rate of the speaker, i.e. his ability to differentiate between the British standard pronunciation and the Czech accent. The motivation for utilizing the difference in percentages as opposed to simply calculating the difference between "0" ratings for STD and CZE was the fact that not every speaker had produced the same total number of tokens. This discrepancy is a result of omissions or word changes. Table 1 below is provided as a demonstration. Refer to Appendix II for a complete set of tables.

	STD			CZE			Total tokens		Number of "0" Ratings		
Speaker	***0***	***1***	***2***	***0***	***1***	***2***	STD	CZE	% STD	% CZE	Success rate
BCKA	5	0	12	0	1	16	17	17	29,4	0,0	29,4
BNDA	13	1	3	0	1	16	17	17	76,5	0,0	76,5
BSTA	2	2	13	0	0	18	17	18	11,8	0,0	11,8
CMFA	9	3	5	2	0	15	17	17	52,9	11,8	41,2
DVRK	9	3	5	3	2	12	17	17	52,9	17,6	35,3
FJTA	6	2	10	4	0	13	18	17	33,3	23,5	9,8
HBNA	0	3	14	0	0	17	17	17	0,0	0,0	0,0
HBTA	2	0	15	0	0	17	17	17	11,8	0,0	11,8
HMNA	2	0	15	1	1	15	17	17	11,8	5,9	5,9
HSKA	4	1	12	1	4	12	17	17	23,5	5,9	17,6
HSKR	4	5	8	1	3	13	17	17	23,5	5,9	17,6
KACT	3	0	14	4	0	13	17	17	17,6	23,5	-5,9
KOSK	13	2	2	2	0	15	17	17	76,5	11,8	64,7
KVTA	1	0	16	0	0	17	17	17	5,9	0,0	5,9

Table 1: This evaluation of "0, 1, 2" tokens for the speech sound / θ, ð / includes sums of individual ratings for STD and CZE modes, a total number of tokens for each mode, the percentage of "0" ratings for STD and CZE, and finally the difference between the two percentual values (i.e. the success rate).

The resulting tables then served as a basis for further analysis. The speakers were filtered based on their success rate, making it possible to determine the three best speakers (i.e. the most successful at making a distinction between the STD and CZE modes) and the three worst speakers. Finally, the speakers were ranked according to their success rates and the data were subsequently visualized, first for each analysed speech sound and then collectively. A detailed description of the results for individual speech sounds can be found in the following sections.

3.1 Dental fricatives /θ, ð/

The speech sounds /θ, ð/ (labelled /th, dh/ in the graphs) were first analysed together, yielding a total of 954 word tokens. Figure 1 shows how the "0, 1, 2" ratings are distributed throughout the STD and CZE modes. It can be observed that the distribution is extremely uneven. The Chi Square test showed these results to be well above chance level [χ^2 (2, n=954) = 912, p< 0,001].

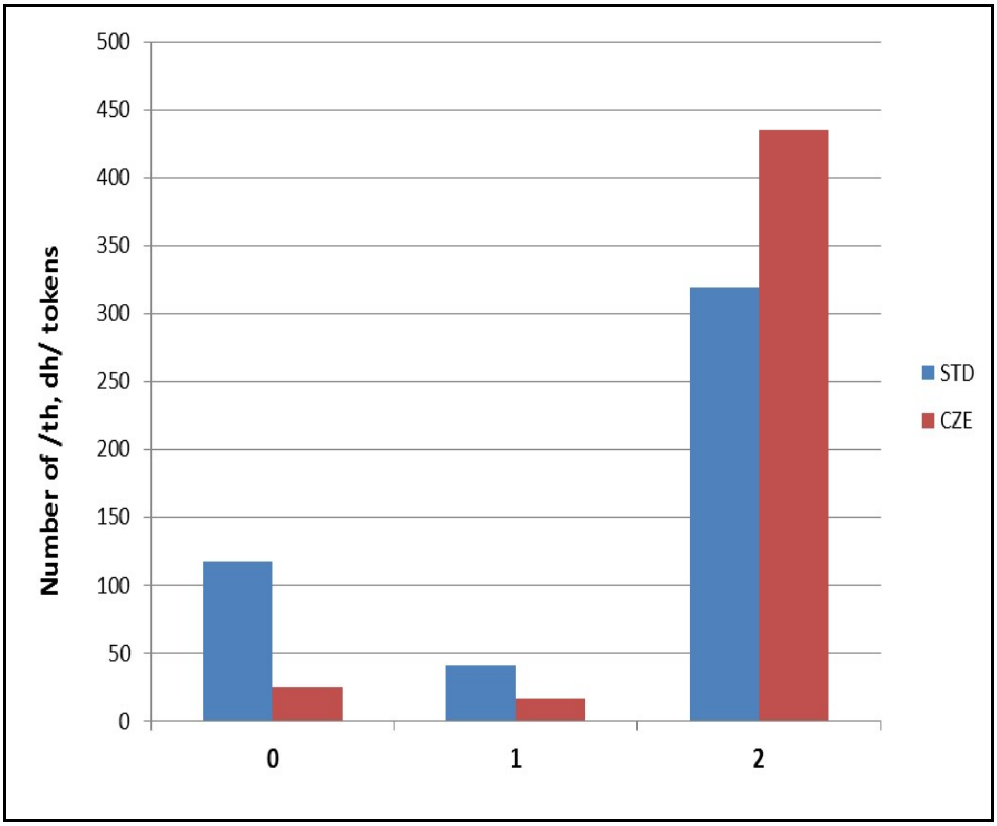

Figure 1: The distribution of "0, 1, 2" ratings for all / θ, ð / word tokens for the STD and CZE modes.

The three speakers with the highest success rates, BNDA, KOSK and VIDR (henceforth best speakers) showed great aptitude for imitating the Czech accent and very clearly differentiating between the two modes. The visualization of the data in Figure 2 shows that "0", denoting zero deviation from the standard, is a dominant value in the STD mode, whereas in the CZE mode it is present only marginally and the prevailing value is "2", where the speaker produced a completely different speech sound.

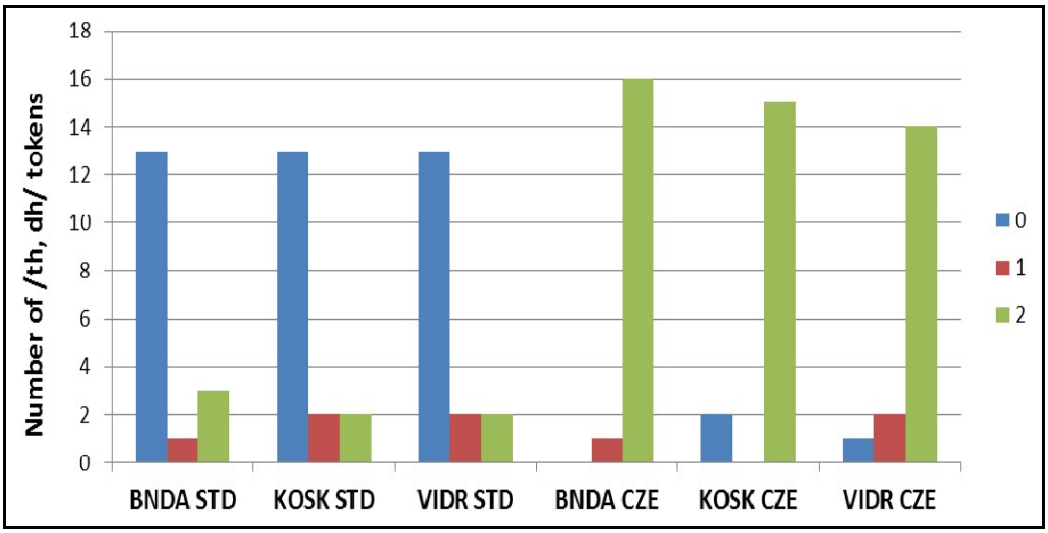

Figure 2: Comparison of / θ, ð / ratings in the STD and CZE modes for the three best speakers.

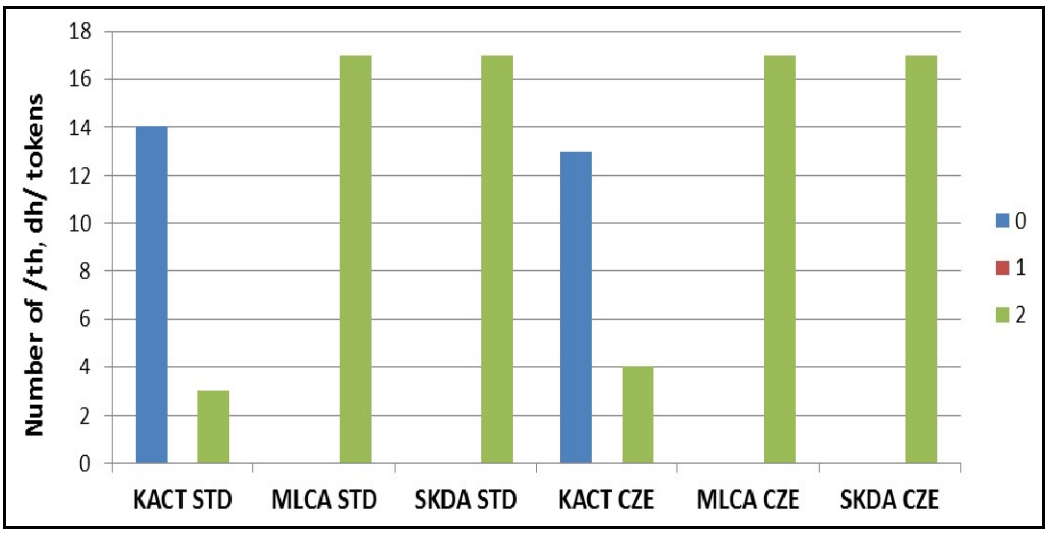

Figure 3: Comparison of / θ, ð / ratings in the STD and CZE modes for the three worst speakers.

On the other hand the three least successful speakers, KACT, MLCA and SKDA, whose scores are depicted above in Figure 3. MLCA and SKDA produced "2" in 100% of the

tokens, both in the STD and CZE modes. The reason for this is most likely their inability to produce the canonical version of the / θ, ð / speech sound at all, and therefore are unable to differentiate between the standard and Czech accented mode.

During the rating of the individual / θ, ð / tokens it was interesting to observe that the canonical pronunciation was prevalent in lexical words, whereas grammatical words often scored "2" or "1", i.e. even in cases where lexical words scored "0". This possibly indicates that whether a word is pronounced canonically or not is based on the word itself. The list of all words appearing in the text was broken up into groups as depicted in Figure 4 (below). Taking a look at the graph, we can see that the word *within* was pronounced correctly 34% of the time, *ethnic* scored "0" 25% of the time, whereas the words *the, that* and *with* (contained in the category *others*) were pronounced correctly in only 7% of instances. The Chi Square test showed these results to be well above chance level [χ^2 (8, n=954) = 98,9, p< 0,001]. Upon taking a closer look, it is perhaps surprising to see that the word with the highest percentage of correctly pronounced tokens, *within*, is a preposition, with only structural meaning. This will be examined further in Chapter 4.

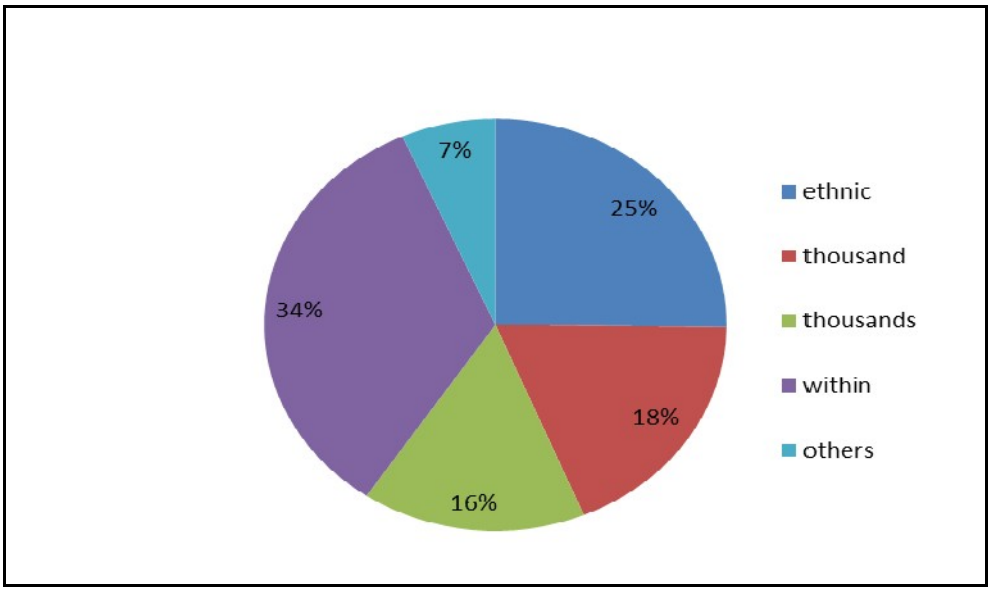

Figure 4: Percentage of canonical realizations of / θ, ð / for all 954 tokens with dependence on the words they occur in. The category "others" contains the words *the, that, with*.

As discussed previously in Chapter 2, there is room for the argument that /θ, ð/ are not, historically and functionally speaking, in a true phonemic relationship. It is therefore advisable to examine the results for each of the two speech sounds separately.

3.1.1 Fortis dental fricative /θ/

The distribution of "0, 1, 2" ratings for the speech sound /θ/ is depicted in Figure 5. In the STD mode, the count of "0" ratings is higher than in CZE mode, whereas the number of "2" ratings is higher in CZE mode. The Chi Square test showed these results to be well above chance level [χ^2 (2, n=189) = 47,1, p< 0,001].

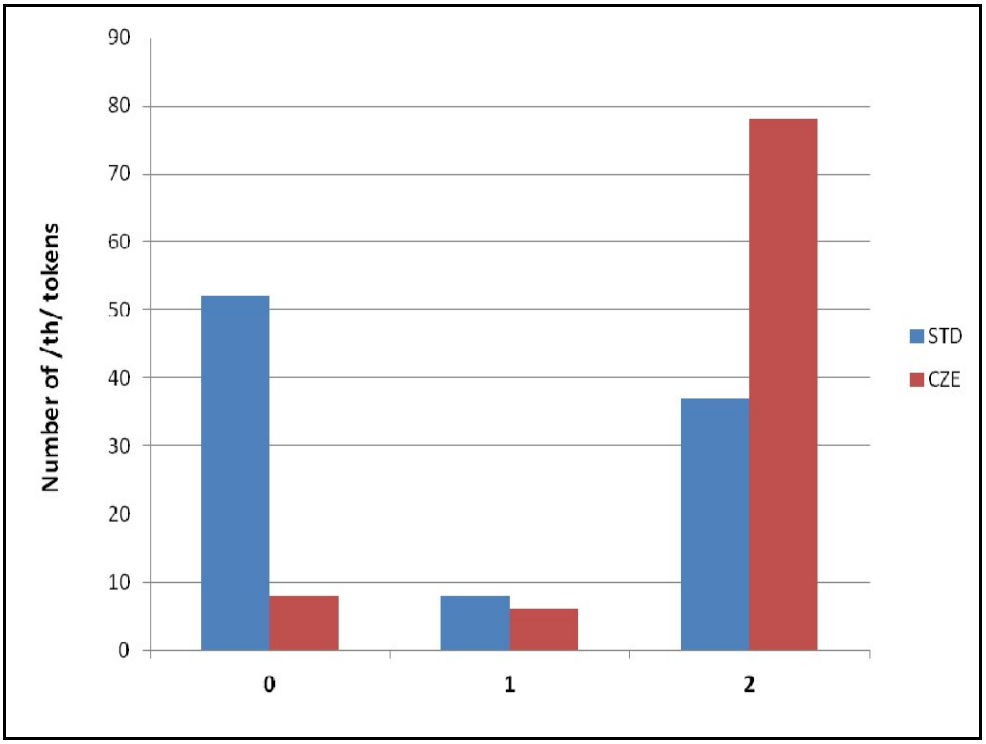

Figure 5: The distribution of "0, 1, 2" ratings for all / θ / word tokens for the STD and CZE modes.

The data for the three most successful speakers can be seen below in Figure 6. In the STD mode, the speakers produced the words *ethnic, thousand* and *thousands* with the canonical pronunciation of /θ/ and received three "0" ratings each. In the CZE mode, the realizations of the three participants were all rated "2". Figure 7 shows the results for the three least successful speakers. The speakers did not produce a single "0" rating, not even in the CZE mode, indicating that they may not be able to produce the canonical realization of /θ/ at all. SKDA and SNKA each produced three non-canonical realizations with the "2" rating in both of the modes, and ZLKA differed only in one count of the "1" rating (indicating an ambiguous realization) in the STD mode.

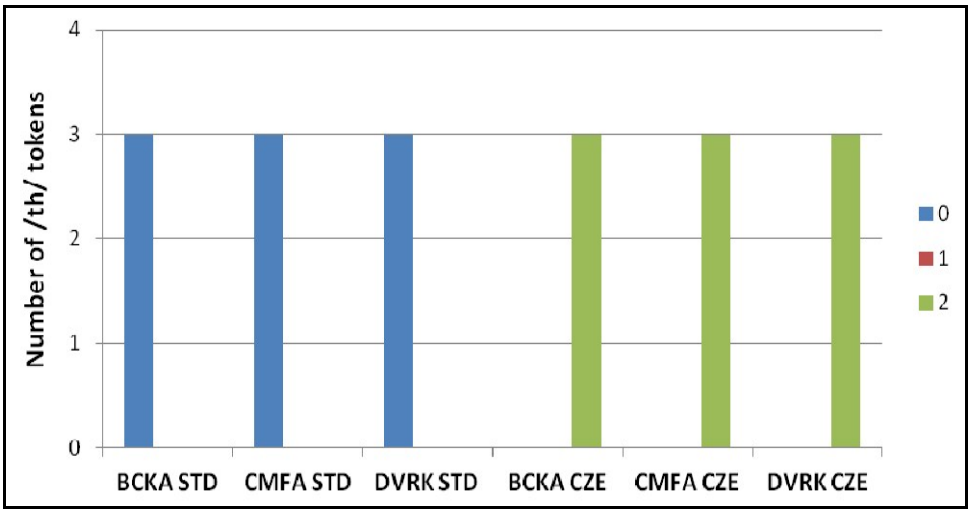

Figure 6: Comparison of / θ / ratings in the STD and CZE modes for the three best speakers

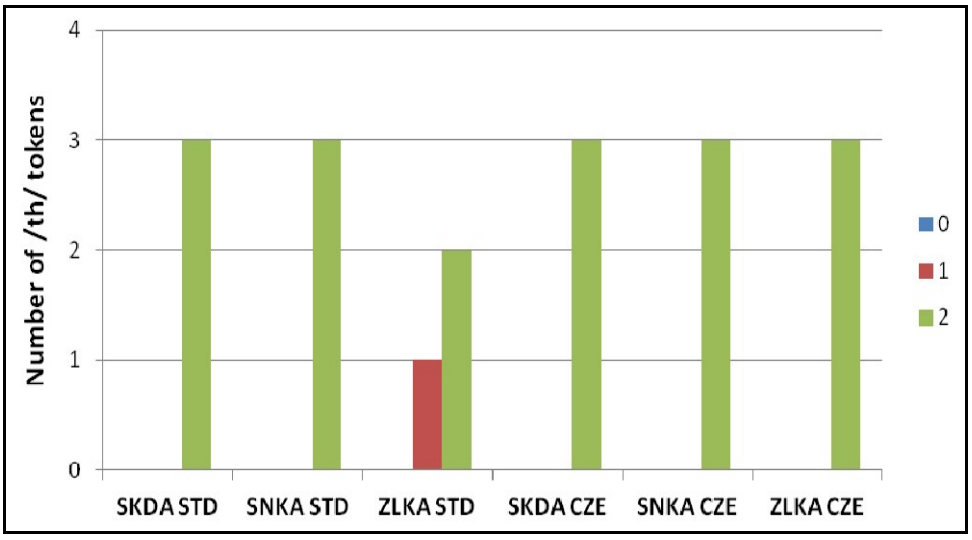

Figure 7: Comparison of / θ / ratings in the STD and CZE modes for the three worst speakers

3.1.2 Lenis dental fricative /ð/

The numbers of "0, 1, 2" ratings for the lenis dental fricative /ð/ are illustrated for both STD and CZE modes below in Figure 8. There are almost five times as many "2" ratings in the STD mode than there are "0" ratings, suggesting that the speech sound /ð/ may be essentially difficult to pronounce. The Chi Square test showed the results for /ð/ to be well above chance level [χ^2 (2, n=765) = 49,4, p< 0,001].

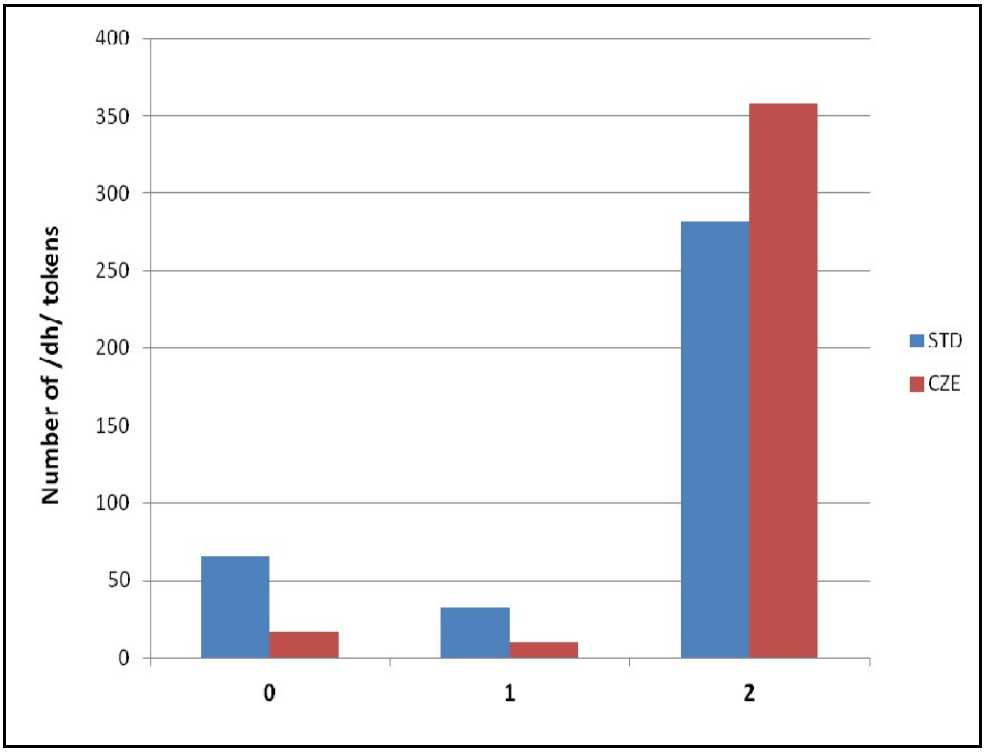

Figure 8: The distribution of "0, 1, 2" ratings for all / ð / word tokens for the STD and CZE modes.

Figure 9 (below) shows the results of the three most successful speakers in the speech sound category /ð/. Of the total 14 tokens per speaker per mode, BNDA scored 11 "0" ratings, VIDR and KOSK scored 10 each. It can be seen that even the most successful speakers had at least two counts of "2" ratings in the STD mode, which could further support the premise that /ð/ causes problems to Czech speakers. In the CZE mode, BNDA received 13 "2" ratings, VIDR and KOSK had 12 non-canonical realizations (rated "2") each.

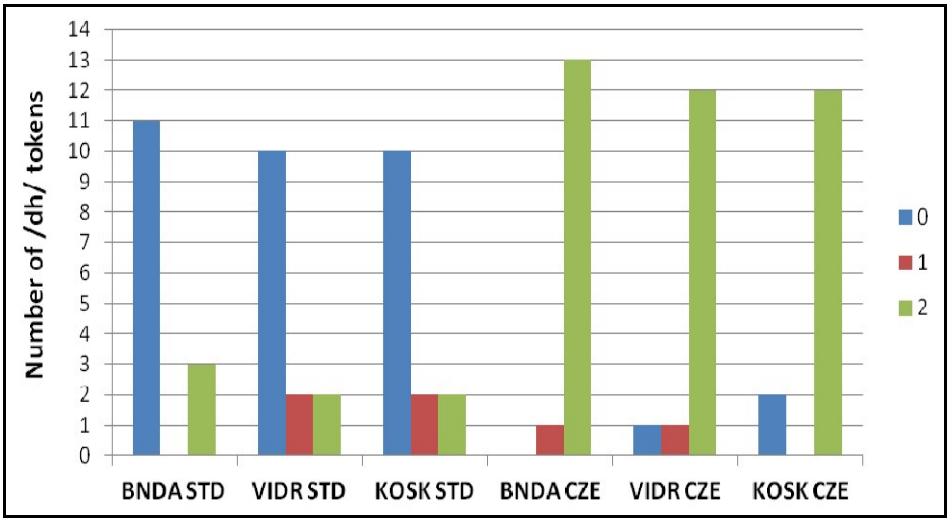

Figure 9: Comparison of / ð / ratings in the STD and CZE modes for the three best speakers

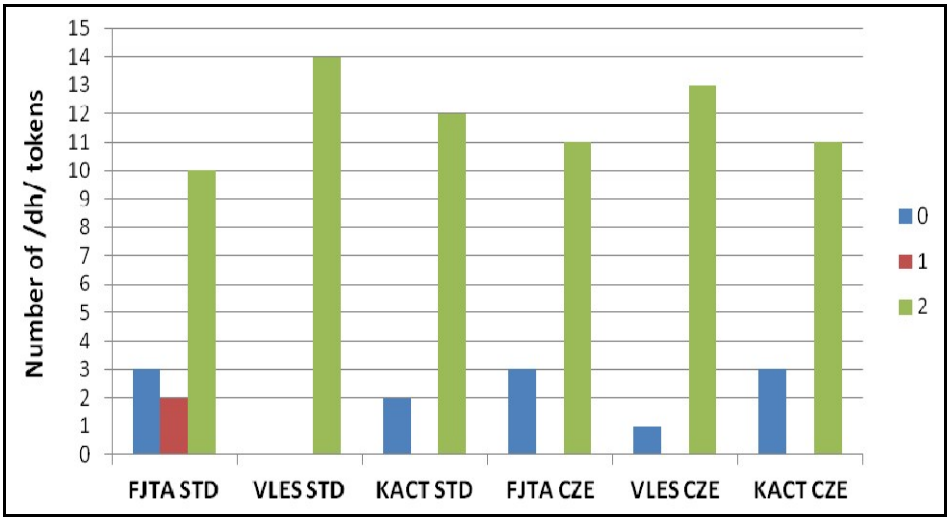

Figure 10: Comparison of / ð / ratings in the STD and CZE modes for the three worst speakers

The three worst speakers are shown in Figure 10. The count of "2" ratings per speaker per mode was at least 10, indicating that the non-canonical realization of /ð/ prevailed in both the STD and the CZE mode. Interestingly, VLES and KACT had a higher count of "0" in the CZE mode than in the STD.

3.1.3 Success rates for /θ/ and /ð/

Figure 11 summarizes the success rates of speakers for both /θ/ and /ð/. For a detailed explanation of success rates, refer to Chapter 3.7. As can be seen in the graph, over 80% of speakers had a success rate under 24 for /ð/, while for /θ/ it was less than 30%. The most successful category contains over 20% of speakers for /θ/, but less than 10% for /ð/. The Chi Square test showed these results above chance level, both for /θ/ [χ^2 (4, n=100) = 31,3, p< 0,001] and for /ð/ [χ^2 (4, n=100) = 166,7, p< 0,001].

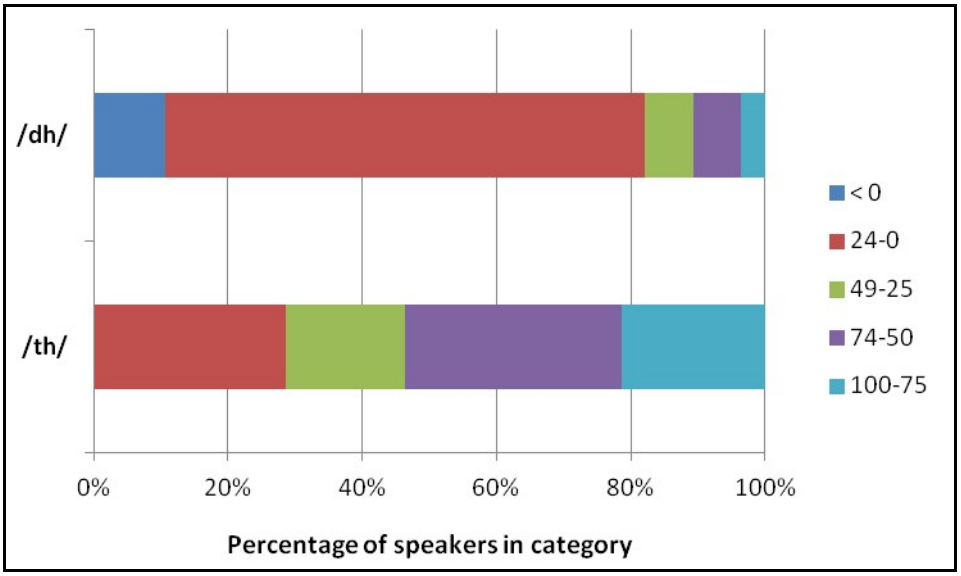

Figure 11: The percentage of speakers in given success category for speech sounds /θ/ and /ð/.

3.2 Velar nasal /ŋ/

For the velar nasal /ŋ/, labelled /ng/ in the graphs, a total of 278 realizations were rated and analysed. The results in Figure 12 show that although the "2" rating was dominant, there was a greater percentage of "0" realizations than in the previous / θ, ð /, which indicates that the velar nasal is less prone to causing problems for Czech speakers of English. According to the Chi Square test, the results are well above chance level [χ^2 (2, n=278) = 27,4, p< 0,001].

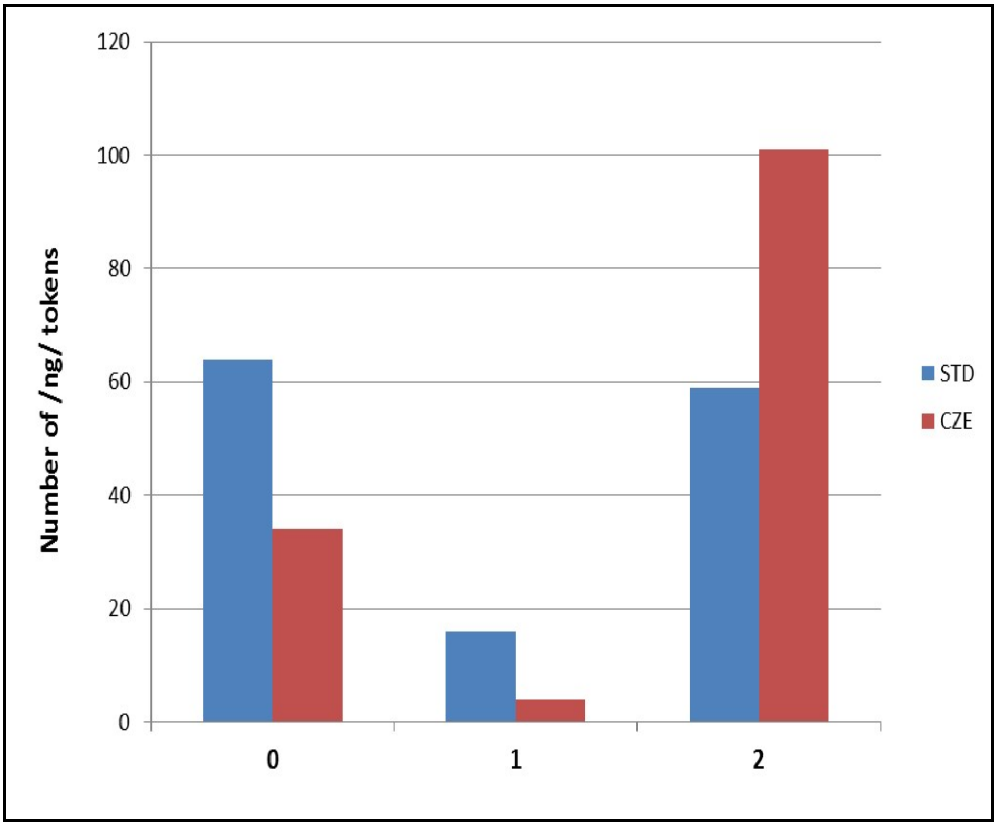

Figure 12: The distribution of "0, 1, 2" ratings for all / ŋ / tokens in the STD and CZE modes.

The most successful speakers were very consistent in the production of the canonical as well as the Czech accented version of / ŋ /, which can be observed in Figure 13 below. On the other hand, Figure 14 shows that the least successful speakers scored "0" in the CZE mode more times than they did in the STD, suggesting that they may not have understood the task assigned to them.

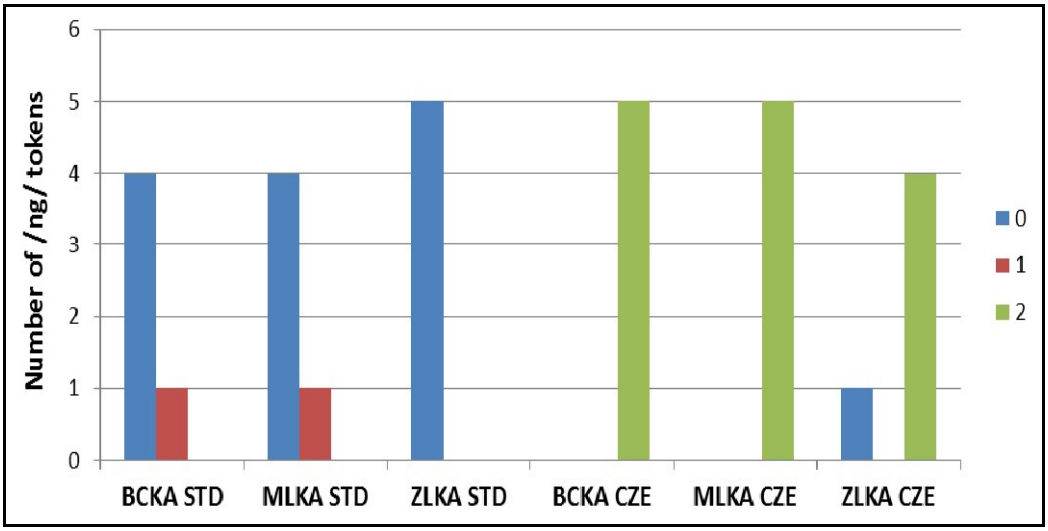

Figure 13: The distribution of "0, 1, 2" ratings of / ŋ / tokens in the STD and CZE modes for the best speakers.

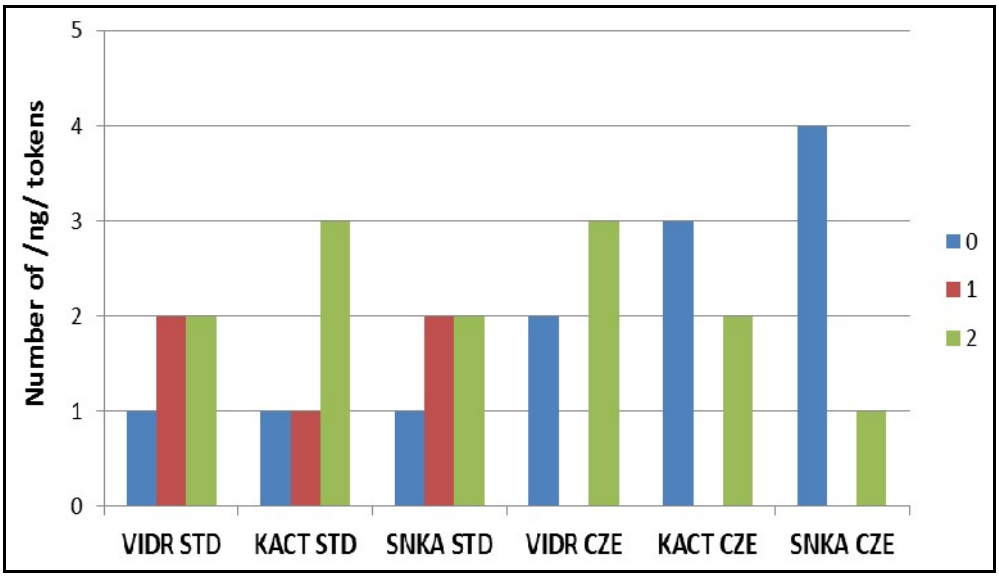

Figure 14: The distribution of "0, 1, 2" ratings of / ŋ / in STD and CZE modes for the worst speakers.

3.3 Alveolar approximant /r/

In comparison to the previous speech sounds, the alveolar approximant proved to be generally much less problematic for the speakers to produce, which can be seen in Figure 15. In fact, there are overall more "0" rated tokens than those rated and "2". The Chi-Square test conducted to see whether the distribution of ratings was above chance level showed statistically significant results for the distribution in the STD mode [χ^2 (3, n=719) = 882,7, p< 0,001], distribution in the CZE mode [χ^2 (3, n=719) = 690, p< 0,001] and overall distribution of "0, 1, 2" ratings for both modes [χ^2 (3, n=1438) = 579,5, p< 0,001].

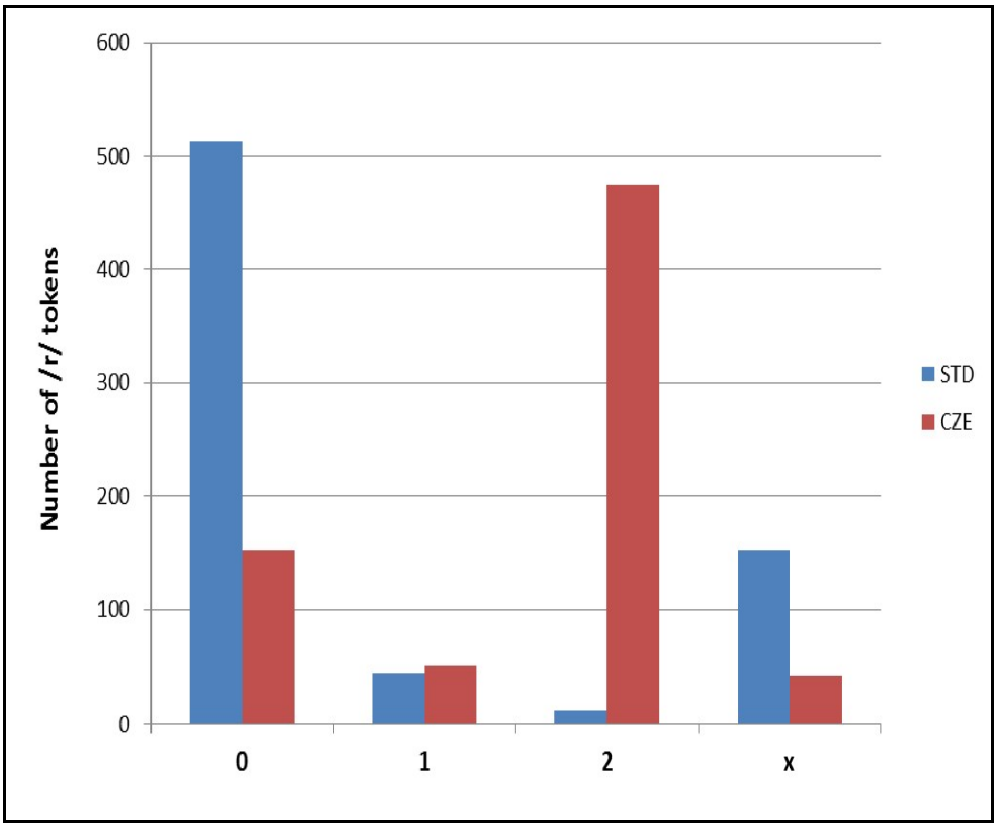

Figure 15: The distribution of "0, 1, 2, x" ratings for all / r / tokens in the STD and CZE modes where "x" represents an elided token.

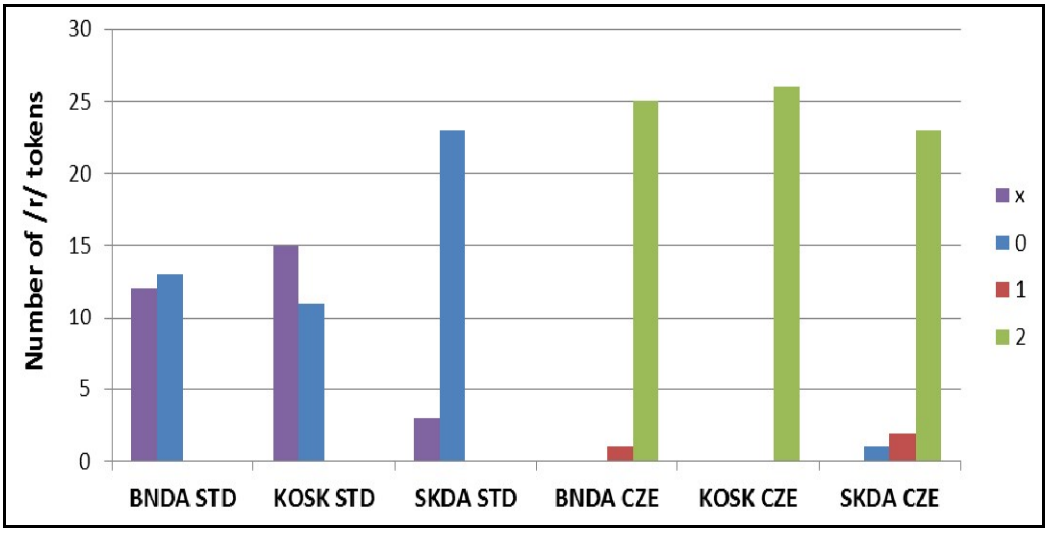

Figure 16: The distribution of "0, 1, 2, x" ratings of / r / in STD and CZE modes for the best speakers, where "x" represents an elided token

The three most successful speakers very reliably produced the alveolar approximant in the STD mode (excepting cases of legal elision), and created a vivid contrast with the alveolar

trill in the CZE mode (see Figure 16). As for the worst speakers, although their pronunciation was predominantly canonical, the result cannot be considered a successful completion of the given task, thus MCHA, MLCA, and HMNA are classified as least successful (refer to Figure 16). It's not the case of how well they pronounced the speech sound, but to what extent they were able to differentiate between the STD and CZE mode

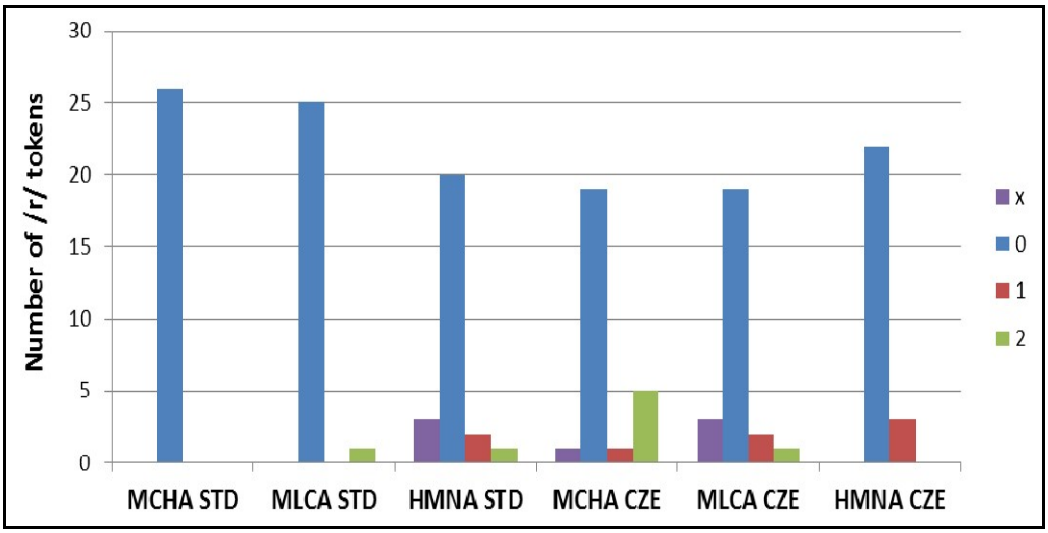

Figure 17: The distribution of "0, 1, 2, x" ratings of / r / in STD and CZE modes for the worst speakers, where "x" represents an elided token

3.4 Labio-velar approximant /w/

In the Czech pronunciation of English, /w/ is often confused with the labiodental fricative /v/. Soudek (1959, p. 78) observed this form of deviation from the British standard in 72,9% of his students. However, the data from the present study suggest that /w/ is one of the least problematic speech sounds. As shown below in Figure 18, there is a dominance of "0" ratings in the STD mode and "2" ratings in the CZE mode, which the Chi Square test showed to be well above chance level [χ^2 (2, n=562) = 107,2, p< 0,001].

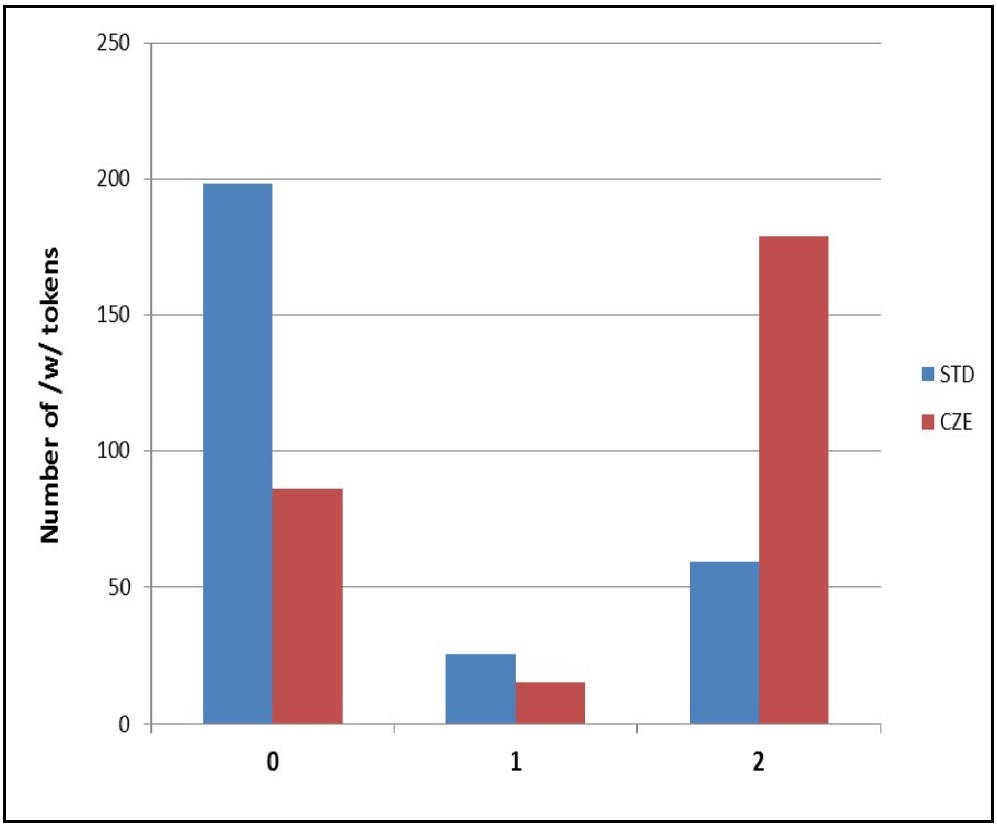

Figure 18: The distribution of "0, 1, 2" ratings for all /w/ tokens in the STD and CZE modes.

As can be seen in Figure 19 below, the most successful speakers, KOSK, DVRK and BCKA, managed to make a very perceptible distinction between the pronunciation of /w/ in the STD and CZE modes, the STD mode with a count of 30 canonically produced labio-velar approximants out of 31 total tokens. Alternatively, the three least successful speakers (below in Figure 20) had the same count of "0, 1, 2" ratings in both modes. HBTA was rated "2" for all /w/ tokens, which could suggest that this speaker was unable to produce a canonical /w/ speech sound. The tokens produced by speakers HSKR and MLCA, on the other hand, were rated "0" in 31 out of 40 cases, and yet there appears to be no difference in the production /w/ for the STD and CZE modes. One possible explanation could be that they do not consider /w/ to be an indicator of Czech accented speech.

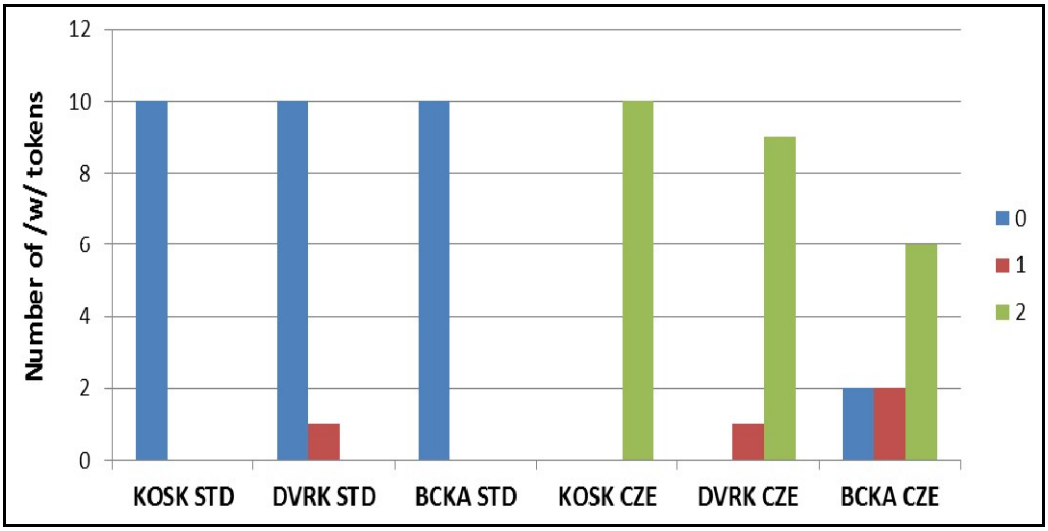

Figure 19: The distribution of "0, 1, 2" ratings of /w/ in STD and CZE modes for the best speakers

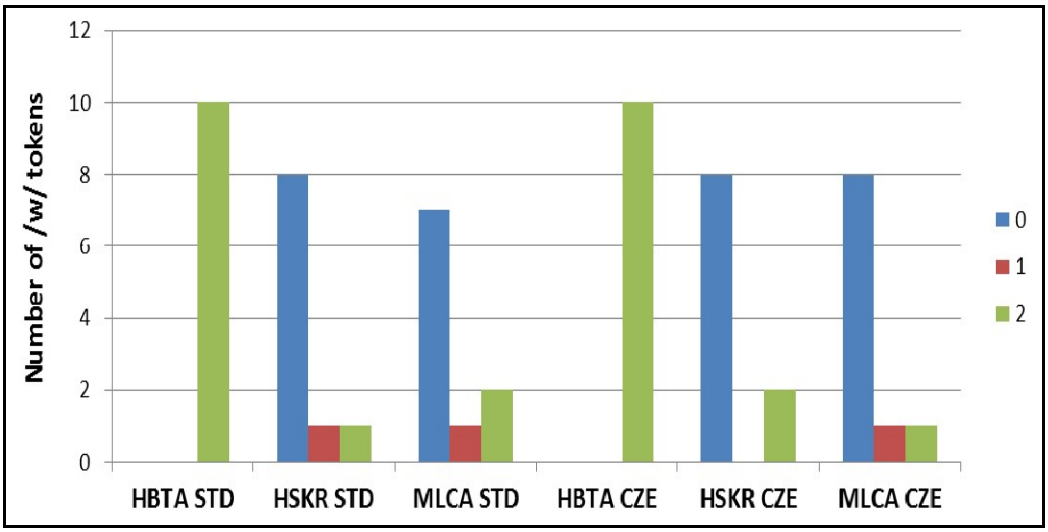

Figure 20: The distribution of "0, 1, 2" ratings of /w/ in STD and CZE modes for the worst speakers

3.5 Front open-mid unrounded vowel /æ/

For the vowel /æ/, a total of 222 tokens were analysed. Figure 21 (below) shows the global distribution of /æ/ in the STD and CZE modes. The results show that even in the STD mode the number of "0" ratings was lower than the "2" ratings. The Chi Square test showed the results to be well above chance level [χ^2 (2, n=222) = 16,2, p< 0,001]. It can therefore be supposed that the vowel /æ/ is difficult for Czech speakers of English to produce in its canonical form.

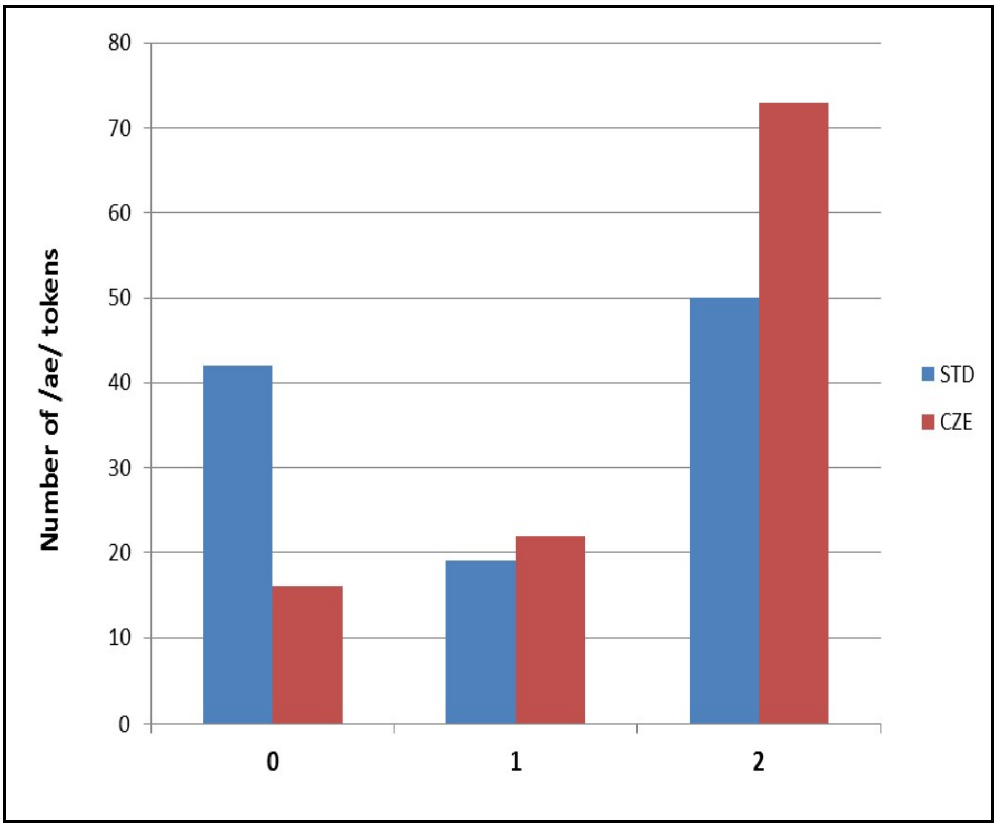

Figure 21: The distribution of "0, 1, 2" ratings of /æ/ in STD and CZE modes.

Due to the small number of /æ/ tokens per speaker, it is not easy to make definite conclusions. Below in Figure 22 it can be seen that the three most successful speakers had a full count of "0" ratings in the STD mode, and none in the CZE mode.

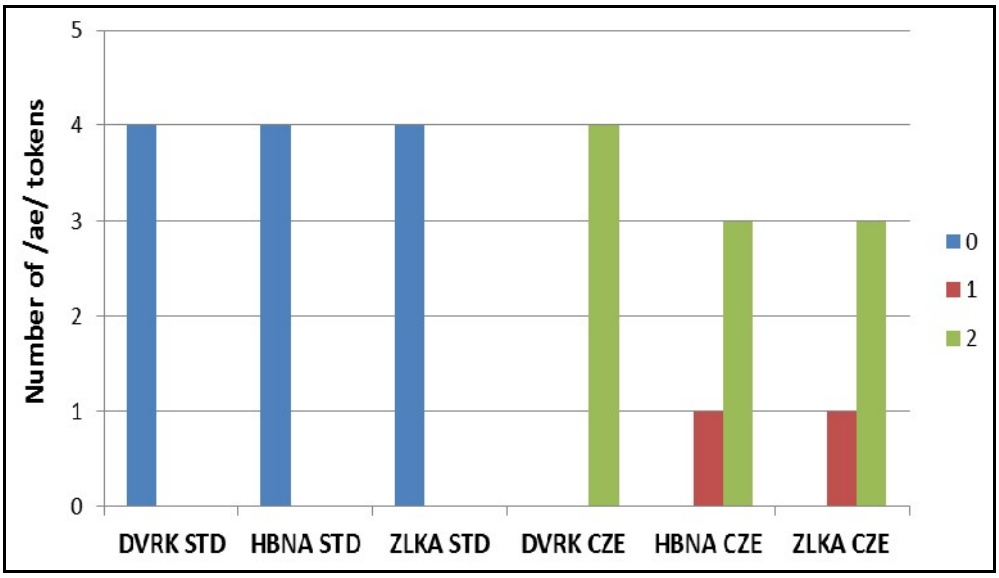

Figure 22: The distribution of "0, 1, 2" ratings of /æ/ in STD and CZE modes for the best speakers.

Below, Figure 23 shows results for the three speakers least successful in differentiating between the canonical and Czech accented pronunciation of /æ/. VLES had the same number of "0" ratings in both modes, whereas HMNA and KVTA had more counts of the "0" rating in the CZE mode than in the STD mode, suggesting that they may not have understood the task.

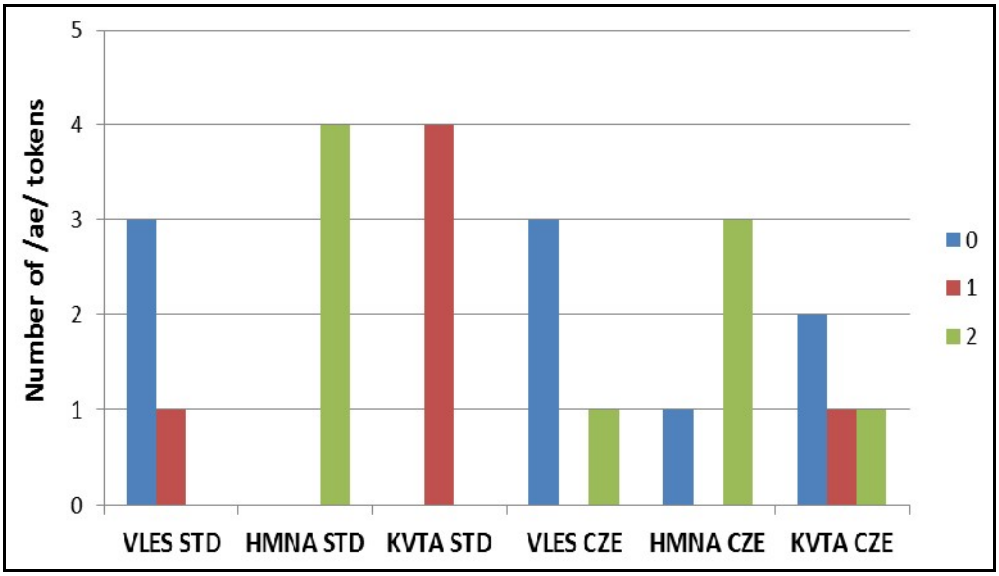

Figure 23: The distribution of "0, 1, 2" ratings of /æ/ in STD and CZE modes for the worst speakers.

3.6 Mid-central unrounded vowel /ɜː/

The speech sound with the smallest number of tokens was the vowel /ɜː/ with a total of 114 (labelled /er/ in graphs). Unlike /æ/ there are more "0" ratings than "2" ratings in the STD mode, indicating that this vowel is less likely to cause problems to Czech speakers of English. In fact, it can be seen in Figure 24 (below) that 40 of the 57 tokens in the STD mode were rated "0". These results are, according to the Chi Square test, well above chance level [χ^2 (2, n=114) = 17,1, p< 0,001]

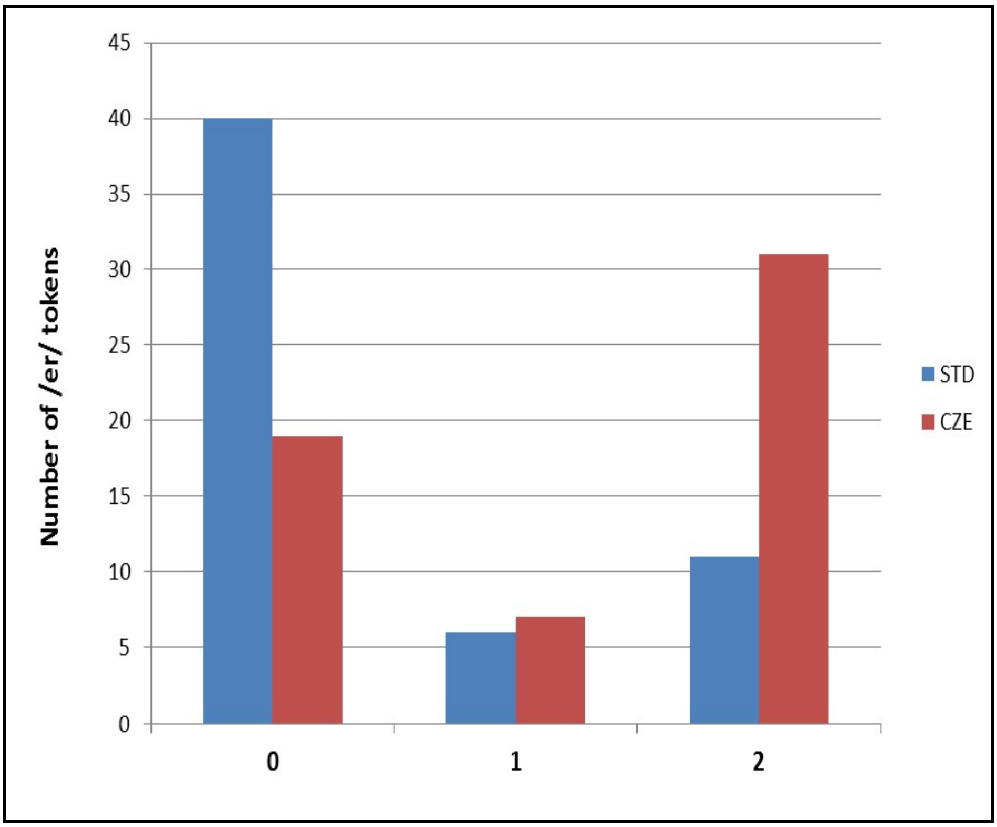

Figure 24: The distribution of "0, 1, 2" ratings of /ɜ:/ in STD and CZE modes.

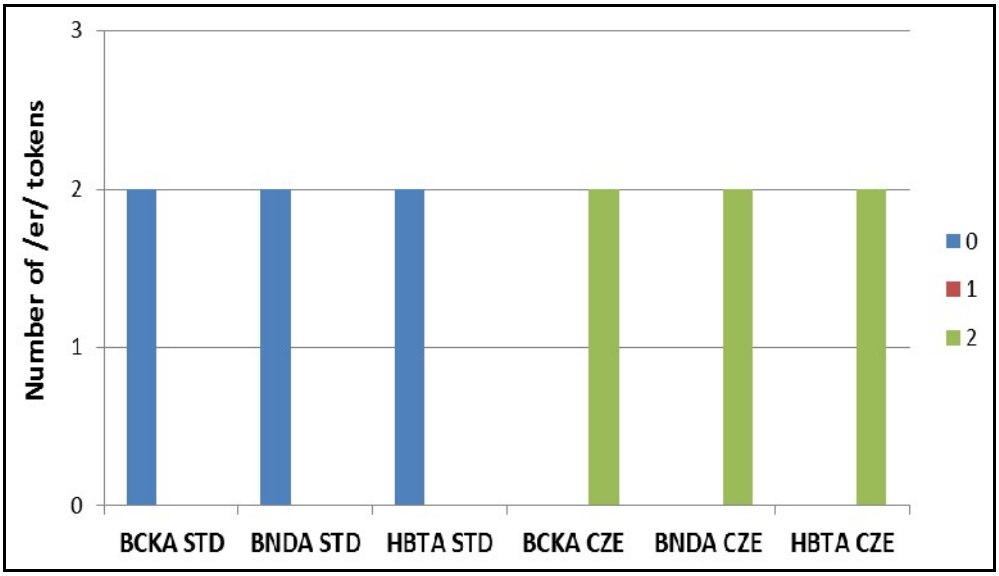

Figure 25: The distribution of "0, 1, 2" ratings of /ɜ:/ in STD and CZE modes for the best speakers.

Figure 25 above shows the three most successful speakers, who had a full count of "0" ratings in the STD mode and a full count of "2" ratings in the CZE mode (in actual fact 2 tokens per mode). With a mere two items per speaker per mode, even one count

mispronunciation in the STD mode was enough to put a speaker out of the running for a place in the successful category. Some speakers did not differentiate between the canonical and Czech accented pronunciation of /ɜ:/ at all, as can be observed for MLKA and ZLKA in Figure 26, who had a full score of "0" ratings in both modes.

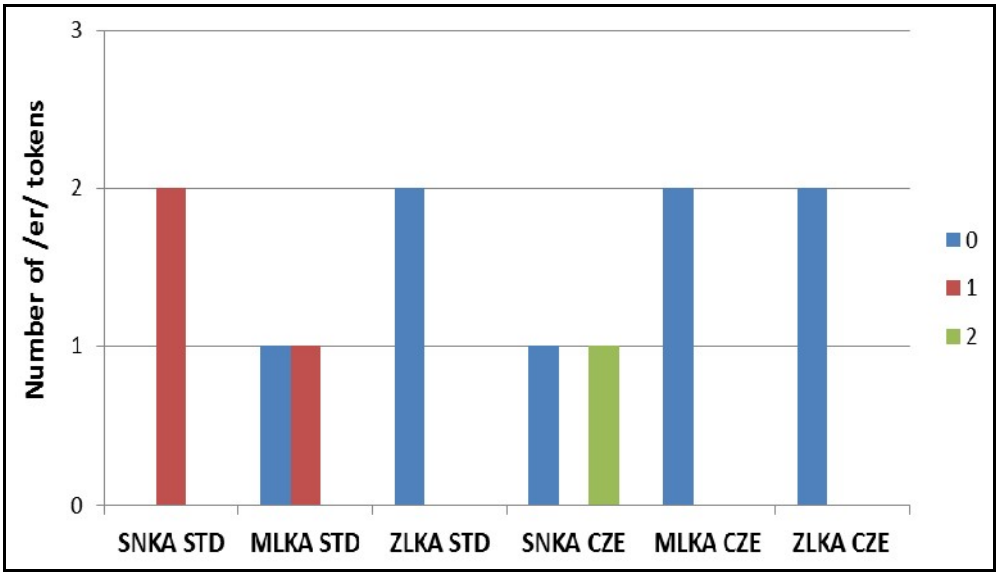

Figure 26: The distribution of "0, 1, 2" ratings of /ɜ:/ in STD and CZE modes for the worst speakers.

3.7 Overall success rate

As demonstrated in the preceding sections, the ability of the native speakers of Czech to successfully make an obvious distinction between the STD mode and the CZE mode varied enormously. The question remains as to which of the speech sounds has the highest success rate among the speakers. As has been previously stated, it was not possible to work simply with sums of occurrences due to the fact that not every speaker had produced the same number of tokens. The indicator was the percentage of canonical British standard realizations of the given speech sound in the standard mode and in the Czech mode. The difference in the percentage of canonical realizations has the potential to range from 100 to -100. The difference would be 100 if a speaker scored 100% of "0" ratings in the STD mode and 0% in CZE mode, and it would mean that the speaker was very successful at capturing the difference between the production of the given speech sound in the two modes. On the other hand, if the difference was -100, it would mean that the speaker had produced 0% of "0" speech sounds in the STD mode and 100% in CZE. In other words, that he did the exact opposite of what was

expected of him. For a full visualization of success rates for individual speech sounds refer to Appendix III.

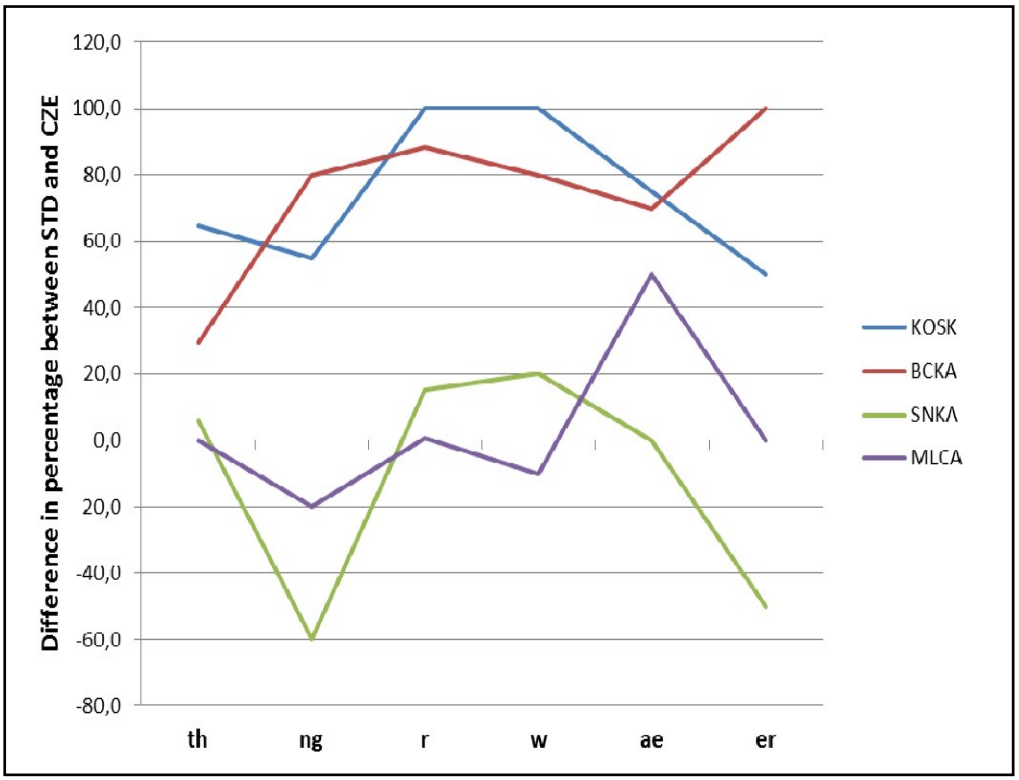

Figure 27: Success rates of the two best and two worst speakers for individual speech sounds.

Above in Figure 27 is a portrayal of the two speakers who were most successful at capturing the difference between the STD and CZE modes of pronunciation, and of the two speakers least successful at this task. The speakers were chosen based on the number of times they appeared among the 5 most successful or 5 least successful, speakers.

A success rate was calculated for every speaker in each of the speech sounds, i.e. /θ, ð, ŋ, r, w, æ, ɜ:/ and five categories were created for more effective visualization. The categories were based on the difference between the percentages of canonical realizations in the two modes: 100-75, 74-50, 49-25, 24-0 and < 0. The output is a comprehensive graph (refer to Figure 27) which will play a key role in the interpretation of the results.

As can be seen in Figure 28, the speech sound /r/ has the highest success rate, as the first category (100-75) encompasses the largest percentage of speakers than for other speech

sounds. The lenis dental fricative /ð/, on the other hand, has the lowest success rate. A number of conclusions will be drawn from these results in the following chapter.

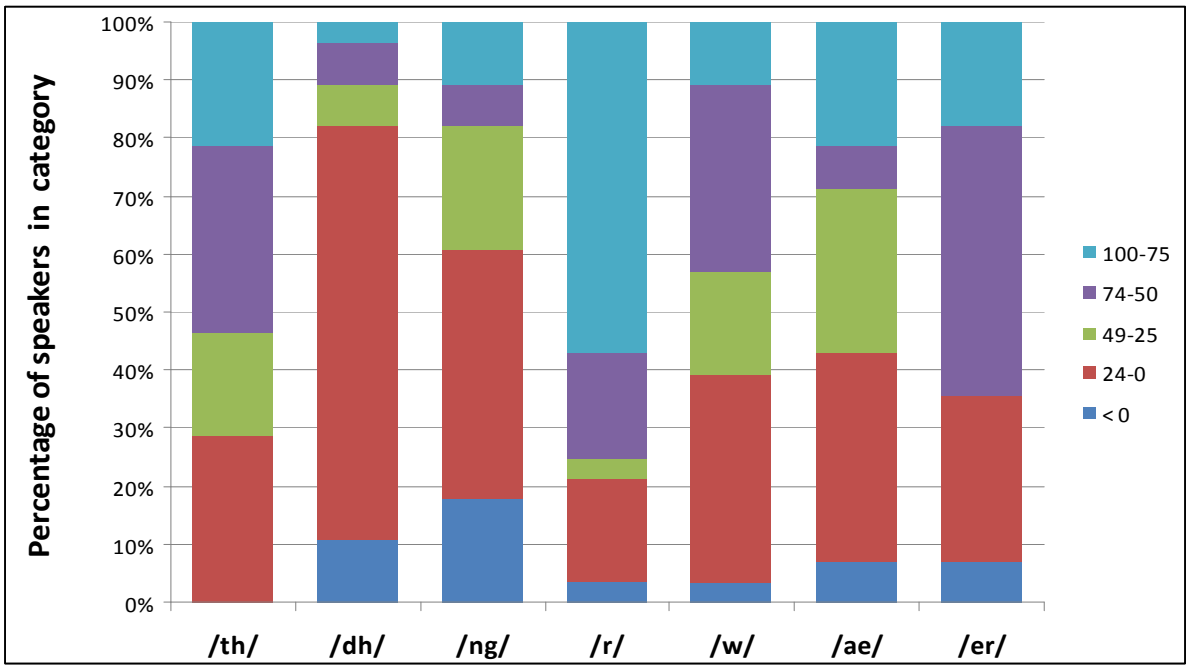

Figure 28: The percentage of speakers in given success category for individual speech sounds.

4 Discussion

The success rates of the speakers analyzed in the present study vary immensely. While imitating the Czech accent proved to be a difficult task for a large number of our participants, several were in fact very successful in making a clear distinction between the standard British accent and the Czech foreign accent. Furthermore, the success rates varied for individual speech sounds, as can be seen in Figure 28 (p. 48). Let us arbitrarily define a successful speaker as one whose success rate fell into one of the two categories, *100-75* or *74-50* for the given speech sound, meaning that in the STD mode, there were at least twice as many counts of "0" ratings than in the CZE mode.

The data in the graph show that the highest number of successful speakers was recorded for /r/. On the other hand the lowest number of successful speakers can be found for /ð/ and /ŋ/. Our preliminary assumption was that when asked to imitate the Czech accent in English, native Czech speakers would focus on what they considered to be the key features accounting for the most perceptible differences between the Czech accent from the British standard, just as in van Dommelen's experiment (2011) the speakers all had focused on the most characteristic difference (i.e. preaspiration) between the Stavanger and Trondheim dialects. With that in mind, the data in Figure 28 (p. 48) possibly indicate that /r/ as an alveolar trill is generally perceived as the key feature of the Czech accent, whereas the canonical pronunciation of /ð, ŋ/, or lack thereof, does not play an important role in the perception of the Czech foreign accent. However, this conclusion seems extremely unlikely to a trained listener who is familiar with the Czech accent in English.

There is an additional factor involved, which was encountered early in the study when rating the individual speech sound tokens. It is the fact that a large number of the speakers had a perceptible Czech accent even in the British standard mode, indicating that these speakers were either unaware of their Czech accent, or incapable of toning it down for the British standard mode. In either case, the result was that for these Czech accented speakers, the difference in pronunciation of the given speech sounds between the STD and CZE modes was either insignificant or non-existent, as can be clearly seen in Figures 3, 7, 10 and 20 (pp. 31, 34, 36, 42). This offers a plausible explanation for the unexpectedly low success rates for /ð/ and /ŋ/. For future studies it is highly recommended that the speakers are chosen based on the strength of their Czech accent, as judged by experienced phoneticians, ruling out the

possibility that a speaker would be incapable of producing a canonical realization of the given speech sound. This would ensure that failure to differentiate between the STD and CZE modes would automatically indicate that the given speech sound was not a determining factor in the perception of the Czech foreign accent.

In Zetterholm's study (2003), the imitators had all focused on the same characteristics of the target voice, a finding which served as a basis for our preliminary hypothesis; we had expected all of the participants to focus on the same features of the Czech foreign accent. However, the results clearly show that not even the most successful speakers all managed to differentiate between the STD and CZE modes for all speech sounds. For example KOSK was always among the top five speakers, with a success rate of 64,7 for /θ, ð/ and 55 for /ŋ/, whereas VIDR was very inconsistent with his scores; a high success rate of 70,6 for /θ, ð/ but -20 for /ŋ/. VIDR's low success rate for /ŋ/ can be attributed to his inability to consistently produce a canonical velar nasal word finally. Therefore it can only be the subject of speculations whether he in fact considers that particular phoneme to be an indicator of the Czech accent.

A point of interest is the distribution of "0" ratings for the speech sounds /θ/ and /ð/. As seen in Figure 4 (p. 32), the words *ethnic, thousand, thousands* and *within* were rated as "0" on considerably more counts than *the, that* and *with*. The first explanation which comes to mind is that lexical words, burdened with meaning, are likely to be articulated with greater care and precision. This could well be true of *ethnic, thousand,* and *thousands*. However, *within* is a preposition and therefore a functional word, and yet it has the highest percentage of canonical realizations (34%).

Another possibility is that words with /θ/ have a higher number of "0" ratings than words containing /ð/. As discussed earlier in Chapter 2, there is reason to believe that the two dental fricatives are not in a traditional phonemic relationship. In addition, Figure 11 (p. 36) clearly suggests that the lenis /ð/ is generally more difficult for students to pronounce in its canonical version than its fortis counterpart /θ/. These results confirm the observations recorded by Soudek (1959) who reported that deviations from the standard /ð/ were produced by 97,2% of students, while /θ/ was replaced by another speech sound by 91,8% of students. Once again, there is a problem with the classification of *within,* which contains the supposedly more problematic lenis /ð/, and yet has the highest percentage of correct

realizations. However, unlike the remaining words which contain /ð/, *within* is not monosyllabic.

In light of this, we may speculate that poor pronunciation concerns primarily monosyllabic grammatical words which, apart from the fact that they are semantically unburdened, are often in an unstressed position in the sentence, and so less care is taken with their pronunciation. Apart from being predominantly monosyllabic, grammatical words have a high frequency of occurrence, a fact that is mentioned by Smith (2007) as a possible explanation for greater variability in pronunciation of /ð/. In the spoken part of The British National Corpus, the number of hits for the given words was as follows; *the* 409 714, *that* 227 029, *with* 47 042, *within* 2 640, *thousand* 5 343, *thousands* 370 and *ethnic* 53. This frequency breakdown, together with the results of the present study, suggests that the pronunciation of a speech sound is to an extent dependent on the word containing the given speech sound. The possibility that a speaker's ability to produce canonical realizations of a given speech sound could be dependent on morphological and semantic variables is very intriguing, and it is definitely an aspect which should be explored further.

It was often the case with the speech sound /r/ that the speakers received high "0" and "x" ratings in both modes, as depicted in Figure 17 (p. 40). Although these speakers were judged as least successful for their failure to differentiate between the two modes, it is highly unlikely that the reason behind this result was their inability to pronounce the Czech alveolar trill. A more probably explanation would be that, unlike the majority of our speakers (21 out of 28 speakers had a success rate over 50 for the phoneme /r/), these particular individuals did not consider the discrepancy between the British alveolar approximant and the Czech alveolar trill to be one of the more salient features of the Czech accent.

In the present study, a total of 3568 speech sound tokens were analyzed. Unfortunately, the numbers of individual speech sounds varied significantly; every speaker produced 14 /ð/, but only 2 /ɜ:/ tokens in each mode (a total of 34 and 4 tokens respectively). With such limited data, it is highly probably that the final results will be distorted. If a speaker were to mispronounce 2 /ð/ tokens and 2 /ɜ:/ tokens in the STD mode, his success rating for the given phonemes would vary extremely. Therefore it would be desirable to dispose with comparable amounts of tokens for each speech sound in future experiments. The text read by the participants should be composed specifically for that purpose in order to achieve, if possible, an even distribution of target speech sounds.

Although the individual tokens were rated by an experienced listener, the final results would surely have more gravity if the speech sound realizations were assessed by a group of experienced listeners. Ideally, the group should contain native speakers of English as well as native speakers of Czech. Another possible expansion of the present study would include the addition of English native speakers as a second group of participants. It would be necessary for the speakers to have spent several years living in Prague and to therefore be to some extent familiar with the Czech language. The method would be the same as in the present study, and comparing the results with those of the present study could reveal whether the speech sounds affecting the perceived strength of the Czech accent are the same for Czech and English native speakers.

The results presented in this study suggest that some speakers are unable to imitate the Czech accent altogether, which can be deduced from their negative success rates. Others are capable of differentiating between the Czech accented and British standard pronunciation, but only provided that they are able to produce the speech sounds in the first place. In general, /r/ appears to be the most popular indicator of the difference between the Czech foreign accent and the British standard, whereas /ð/ was not utilized nearly as often. There are, however, many factors which must be taken into consideration before it can be patently said whether these results are definitive.

5 Conclusion

The aim of the present study was to identify the most prominent features of the Czech foreign accent in English, as perceived by its speakers. The results were based on the perceptual analysis of the speech sounds /θ, ð, ŋ, r, w, æ, ɜ:/, encompassing a total of 3568 tokens produced by 28 speakers. Two recordings were made for each speaker, one in the standard mode (STD) and one in the Czech accented mode (CZE), where the speakers were asked to imitate the Czech foreign accent. Individual speech sounds were rated "0" (zero deviation from the standard pronunciation), "1" or "2" (Czech accented pronunciation). The data were then analyzed in order to see where the greatest differences between the STD and CZE pronunciation could be found.

Results showed that the distinction between the alveolar approximant in STD mode and alveolar trill in CZE mode was utilized by the greatest percentage of speakers, while the pronunciation of the dental fricatives /θ, ð/ and the velar nasal /ŋ/ in word final position varied the least between modes. However, it was discovered that there are other important factors affecting the scores for individual speech sounds, namely the speaker's L2 proficiency and, perhaps more surprisingly, lexical features of the word in which the speech sound occurs.

Arguably the most notable finding is the fact that the pronunciation of a speech sound varies based on morphological and frequency factors. Specifically for /θ, ð/ the results suggest that words with a high frequency of occurrence (e.g. *the*) are more likely to contain a non-canonical realization of the target speech sound than low frequency words (e.g. *ethnic*). The fact is that high frequency words are generally grammatical and often monosyllabic, and so it cannot be said for certain whether the determining factor is frequency, syllable count or semantic weight. This is a matter definitely worth exploring further, ideally with a larger corpus and a greater selection of words containing the given speech sound.

The principal limitation of this study was the disparity in speaker proficiency. A number of speakers in the present study was unable to produce the canonical British pronunciation of individual speech sounds, and therefore also unable to differentiate between the Czech accent and British standard pronunciation. Limitations Levy and Strange (2008) suggested that previous L2 experience plays an important role in discrimination of vowel contrasts. This hypothesis is supported by the findings of this study, which showed that the

level of L2 proficiency has a significant effect on the speaker's ability to differentiate between Czech accented speech and British standard pronunciation.

The present study further demonstrates the complexity of the foreign accent phenomena, suggesting that there are many variables which work together and must be taken into account when studying the pronunciation of individual speech sounds. The conclusions drawn from the results suggest that a more careful selection of participants would be necessary for obtaining conclusive results. It is hoped that the findings regarding morphological and frequency variables could contribute to further discussion about factors influencing the perceived strength of the foreign accent.

References

Boersma, P. & Weenink, D. (2013). Praat: doing phonetics by computer [Computer program]. Version 5.3.56, retrieved 15 September 2013 from http://www.praat.org/

Derwing, T. & Munro, M. (2009). Putting Accent in its Place: Rethinking Obstacles to Communication. *Language Teaching – LANG TEACH 42 (4)*. 476–490.

Doughty, C. J., & Long, M. H. (2003) The Handbook of Second Language Acquisition. Oxford: Blackwell Publishing Ltd.

Flege, J. E. (1981). The Phonological Basis of Foreign Accent: A Hypothesis. *TESOL Quarterly 15 (4)*. 443-455

Flege, J. E. (1984). The detection of French accent by American listeners. *The Journal of the Acoustical Society of America*, *76*(3). 692-707.

Flege, J. E., Munro, M. J., & MacKay, I. R. (1995). Factors Affecting Strength of Perceived Foreign Accent in a Second Language. *Journal of the Acoustical Society of America, 97(5)*. 3125-3134.

Flege, J. E., Munro, M. J., & MacKay, I. R. (1995). Effects of age of second-language learning on the production of English consonants. *Speech Communication 16*. 1-26.

Flege, J. E. (2007). Language Contact in Bilingualism: Phonetic System Interactions. *Laboratory Phonology 9*. 353-381.

Gluszek, A & Dovidio, J. F. (2010). Speaking With a Nonnative Accent: Perceptions of Bias, Communication Difficulties, and Belonging in the United States. *Journal of Language and Social Psychology 29 (2)*. 224-234.

Gluszek, A & Dovidio, J. F. (2010). The Way *They* Speak: A Social Psychological Perspective on the Stigma of Nonnative Accents in Communication. *Personality and Social Psychology Review 14 (2)*. 214-237.

Haak, N. J., & Darling, R. M. (1993). Imitation of suprasegmental patterns by non-native speakers of English. *American Journal of Speech-Language Pathology*, 2(1). 47-50.

Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet. (1999). Cambridge: Cambridge University Press.

Iverson, P. & Evans, B. G. (2007). Learning English vowels with different first-language vowel systems: Perception of formant targets, formant movement, and duration. *The Journal of the Acoustical Society of America.* 2842–2854

Kamiyama, T. (2004). Perception of Foreign Accentedness in L2 Prosody and Segments: L1 Japanese Speakers Learning L2 French. *SP-2004.* 721-724.

Lagefoged, P. (2001). *Vowels and Consonants.* Massachusetts: Blackwell Publishing.

Lagefoged, P. (2001). *Course in Phonetics.* Boston: Heinle & Heinle.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: John Wiley and Sons.

Levy, E. S., & Strange, W. (2008). Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics*, *36*(1). 141-157.

Machač, P., & Skarnitzl, R. (2009). Principles of Phonetic Segmentation. Praha: Epocha.

Millward, C. M. & Hayes, M. (2012) *A Biography of the English Language 3rd Edition.* Boston: Wadsworth Cengage Learning.

Munro, M. J., Derwing, T. M., & Burgess, C. S. (2003). The detection of foreign accent in backwards speech. *Proceedings of the 15th International Congress of Phonetic Sciences*. 535-538.

Piske, T. et al. (2001). Factors Affecting Degree of Foreign accent in an L2: a review. *Journal of Phonetics 29*. 191-215.

Preacher, K. J. (2001). Calculation for the chi-square test: An interactive calculation tool for chi-square tests of goodness of fit and independence [Computer software]. Available from http://quantpsy.org.

Roach, P. (2009). *English Phonetics and Phonology.* Cambridge: Cambridge University Press.

Segalowitz, N., & Gatbonton, E. (1977). Studies of the nonfluent bilingual. *Bilingualism: Psychological and social implications.* 77-90.

Smith, B. (2007). Dental fricatives and stops in Germanic: Deriving diachronic processes from synchronic variation. *Historical Linguistics*. 6-11.

Soudek, L. (1956). On Some Essential Deviations from the Standard Pronunciation of English in Czechoslovakia. *Sborník Filozofickej fakulty Univerzity Komenského, XI-XII*. 73-84.

The British National Corpus, version 3 (BNC XML Edition). 2007. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: http://www.natcorp.ox.ac.uk/

Timmis, I. (2002). Native-speaker norms and International English: A classroom view. *ELT Journal 56*. 240–249.

van Dommelen, W. A., Holm, S., & Koreman, J. (2011). Dialectal feature imitation in Norwegian. *Proc. 17th ICPhS*. 599-602

Werker, J. F., Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child development*. 349-355.

Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *37*(2). 278-286.

Yuan, J. & Liberman, M. (2008). Speaker identification on the SCOTUS corpus. *Proceedings of Acoustics '08*.

Yuan, J., Jiang, Y., & Song, Z. (2010) Perception of foreign accent in spontaneous L2 English speech. *Proceedings of Speech Prosody Vol. 100884*. 1-4.

Zetterholm, E. (2003). The same but different–three impersonators imitate the same target voices. *Proceedings of the 15th International Congress of Phonetic Sciences: Barcelona, 3-9 August 2003. Vol. 1/3*. 2205-2208.

Appendix I

BBC News Bulletin

BBC News with Kyle Andrews.

Kenya has agreed to open a new refugee camp near its border with Somalia as thousands of people are fleeing the region's worst drought in sixty years. The prime minister of Kenya announced that a camp, which can fit up to eighty thousand people, would open within ten days. Some ministers had feared opening the camp would encourage more Somalis to cross the border. But the prime minister said that turning away the refugees was not an option. In his opinion, that would amount to ethnic cleansing.

You're listening to the news from the BBC in London.

Appendix II

	STD			CZE			Total tokens		Number of "0" Ratings		
Speaker	***0***	***1***	***2***	***0***	***1***	***2***	STD	CZE	% STD	% CZE	Success rate
BCKA	5	0	12	0	1	16	17	17	29,4	0,0	29,4
BNDA	13	1	3	0	1	16	17	17	76,5	0,0	76,5
BSTA	2	2	13	0	0	18	17	18	11,8	0,0	11,8
CMFA	9	3	5	2	0	15	17	17	52,9	11,8	41,2
DVRK	9	3	5	3	2	12	17	17	52,9	17,6	35,3
FJTA	6	2	10	4	0	13	18	17	33,3	23,5	9,8
HBNA	0	3	14	0	0	17	17	17	0,0	0,0	0,0
HBTA	2	0	15	0	0	17	17	17	11,8	0,0	11,8
HMNA	2	0	15	1	1	15	17	17	11,8	5,9	5,9
HSKA	4	1	12	1	4	12	17	17	23,5	5,9	17,6
HSKR	4	5	8	1	3	13	17	17	23,5	5,9	17,6
KACT	3	0	14	4	0	13	17	17	17,6	23,5	-5,9
KOSK	13	2	2	2	0	15	17	17	76,5	11,8	64,7
KVTA	1	0	16	0	0	17	17	17	5,9	0,0	5,9
MCHA	2	1	14	0	1	16	17	17	11,8	0,0	11,8
MLCA	0	0	17	0	0	17	17	17	0,0	0,0	0,0
MLKA	3	1	13	0	0	17	17	17	17,6	0,0	17,6
MRES	3	5	9	0	0	17	17	17	17,6	0,0	17,6
PVLK	2	2	13	2	0	15	17	17	11,8	11,8	0,0
RMSA	6	1	10	2	0	14	17	16	35,3	12,5	22,8
SKDA	0	0	17	0	0	17	17	17	0,0	0,0	0,0
SLAB	2	1	14	0	0	17	17	17	11,8	0,0	11,8
SNKA	2	2	13	1	1	15	17	17	11,8	5,9	5,9
TCHA	1	1	15	0	0	18	17	18	5,9	0,0	5,9
VIDR	13	2	2	1	2	14	17	17	76,5	5,9	70,6
VLES	1	0	16	1	0	16	17	17	5,9	5,9	0,0
VLKA	7	2	8	0	0	17	17	17	41,2	0,0	41,2
ZLKA	2	1	14	0	0	17	17	17	11,8	0,0	11,8

Table 1: Complete data for the speech sounds / θ, ð /, including sums of individual "0, 1, 2" token ratings for STD and CZE modes, a total number of tokens for each mode, the percentage of "0" ratings for STD and CZE, and finally the difference between the two percentual values (i.e. the success rate).

	STD			CZE			Total tokens		Number of "0" Ratings		
Speaker	***0***	***1***	***2***	***0***	***1***	***2***	STD	CZE	% STD	% CZE	Success rate
BCKA	3	0	0	0	0	3	3	3	100,0	0,0	100,0
BNDA	2	1	0	0	0	3	3	3	66,7	0,0	66,7
BSTA	1	1	1	0	0	3	3	3	33,3	0,0	33,3
CMFA	3	0	0	0	0	3	3	3	100,0	0,0	100,0
DVRK	3	0	0	0	0	3	3	3	100,0	0,0	100,0
FJTA	4	0	0	1	0	2	4	3	100,0	33,3	66,7
HBNA	0	2	1	0	0	3	3	3	0,0	0,0	0,0
HBTA	1	0	2	0	0	3	3	3	33,3	0,0	33,3
HMNA	2	0	2	1	0	2	4	3	50,0	33,3	16,7
HSKA	3	0	1	0	3	1	4	4	75,0	0,0	75,0
HSKR	2	0	2	0	1	3	4	4	50,0	0,0	50,0
KACT	2	0	2	2	0	2	4	4	50,0	50,0	0,0
KOSK	4	0	0	0	0	3	4	3	100,0	0,0	100,0
KVTA	1	0	3	0	0	3	4	3	25,0	0,0	25,0
MCHA	2	1	1	0	1	3	4	4	50,0	0,0	50,0
MLCA	0	0	3	0	0	3	3	3	0,0	0,0	0,0
MLKA	2	0	1	0	0	3	3	3	66,7	0,0	66,7
MRES	2	1	1	0	0	4	4	4	50,0	0,0	50,0
PVLK	2	1	0	2	0	1	3	3	66,7	66,7	0,0
RMSA	3	0	0	1	0	2	3	3	100,0	33,3	66,7
SKDA	0	0	3	0	0	3	3	3	0,0	0,0	0,0
SLAB	2	0	2	0	0	4	4	4	50,0	0,0	50,0
SNKA	0	0	3	0	0	3	3	3	0,0	0,0	0,0
TCHA	1	0	3	0	0	3	4	3	25,0	0,0	25,0
VIDR	4	0	0	1	1	2	4	4	100,0	25,0	75,0
VLES	1	0	3	0	0	3	4	3	25,0	0,0	25,0
VLKA	2	0	1	0	0	3	3	3	66,7	0,0	66,7
ZLKA	0	1	2	0	0	4	3	4	0,0	0,0	0,0

Table 2: Complete data for / θ /, including sums of individual "0, 1, 2" token ratings for STD and CZE modes, a total number of tokens for each mode, the percentage of "0" ratings for STD and CZE, and finally the difference between the two percentual values (i.e. the success rate).

	STD			CZE			Total tokens		Number of "0" Ratings		
Speaker	***0***	***1***	***2***	***0***	***1***	***2***	STD	CZE	% STD	% CZE	Success rate
BCKA	2	0	12	0	1	13	14	14	14,3	0,0	14,3
BNDA	11	0	3	0	1	13	14	14	78,6	0,0	78,6
BSTA	1	1	12	0	0	15	14	15	7,1	0,0	7,1
CMFA	6	3	5	2	0	12	14	14	42,9	14,3	28,6
DVRK	6	3	5	3	2	9	14	14	42,9	21,4	21,4
FJTA	2	2	10	3	0	11	14	14	14,3	21,4	-7,1
HBNA	0	1	13	0	0	14	14	14	0,0	0,0	0,0
HBTA	1	0	13	0	0	14	14	14	7,1	0,0	7,1
HMNA	0	0	13	0	1	13	13	14	0,0	0,0	0,0
HSKA	1	1	11	1	1	11	13	13	7,7	7,7	0,0
HSKR	2	5	6	1	2	10	13	13	15,4	7,7	7,7
KACT	1	0	12	2	0	11	13	13	7,7	15,4	-7,7
KOSK	9	2	2	2	0	12	13	14	69,2	14,3	54,9
KVTA	0	0	13	0	0	14	13	14	0,0	0,0	0,0
MCHA	0	0	13	0	0	13	13	13	0,0	0,0	0,0
MLCA	0	0	14	0	0	14	14	14	0,0	0,0	0,0
MLKA	1	1	12	0	0	14	14	14	7,1	0,0	7,1
MRES	1	4	8	0	0	13	13	13	7,7	0,0	7,7
PVLK	0	1	13	0	0	14	14	14	0,0	0,0	0,0
RMSA	3	1	10	1	0	12	14	13	21,4	7,7	13,7
SKDA	0	0	14	0	0	14	14	14	0,0	0,0	0,0
SLAB	0	1	12	0	0	13	13	13	0,0	0,0	0,0
SNKA	2	2	10	1	1	12	14	14	14,3	7,1	7,1
TCHA	0	1	12	0	0	15	13	15	0,0	0,0	0,0
VIDR	9	2	2	0	1	12	13	13	69,2	0,0	69,2
VLES	0	0	13	1	0	13	13	14	0,0	7,1	-7,1
VLKA	5	2	7	0	0	14	14	14	35,7	0,0	35,7
ZLKA	2	0	12	0	0	13	14	13	14,3	0,0	14,3

Table 3: Complete data for / ð /, including sums of individual “0, 1, 2” token ratings for STD and CZE modes, a total number of tokens for each mode, the percentage of “0” ratings for STD and CZE, and finally the difference between the two percentual values (i.e. the success rate).

	STD			CZE			Total tokens		Number of "0" Ratings		
Speaker	***0***	***1***	***2***	***0***	***1***	***2***	STD	CZE	% STD	% CZE	Success rate
BCKA	4	1	0	0	0	5	5	5	80,0	0,0	80,0
BNDA	2	2	1	0	0	5	5	5	40,0	0,0	40,0
BSTA	1	2	2	0	0	5	5	5	20,0	0,0	20,0
CMFA	2	1	2	0	1	4	5	5	40,0	0,0	40,0
DVRK	3	0	2	2	0	3	5	5	60,0	40,0	20,0
FJTA	2	0	3	1	0	4	5	5	40,0	20,0	20,0
HBNA	5	0	0	4	0	1	5	5	100,0	80,0	20,0
HBTA	1	0	4	0	0	5	5	5	20,0	0,0	20,0
HMNA	0	0	5	0	0	5	5	5	0,0	0,0	0,0
HSKA	1	0	4	1	0	4	5	5	20,0	20,0	0,0
HSKR	3	0	2	1	0	4	5	5	60,0	20,0	40,0
KACT	1	1	3	3	0	2	5	5	20,0	60,0	-40,0
KOSK	4	0	1	1	0	3	5	4	80,0	25,0	55,0
KVTA	1	0	4	1	0	4	5	5	20,0	20,0	0,0
MCHA	2	0	3	3	0	2	5	5	40,0	60,0	-20,0
MLCA	1	0	4	2	0	3	5	5	20,0	40,0	-20,0
MLKA	4	1	0	0	0	5	5	5	80,0	0,0	80,0
MRES	1	0	4	0	1	4	5	5	20,0	0,0	20,0
PVLK	3	0	2	1	0	4	5	5	60,0	20,0	40,0
RMSA	5	0	0	4	0	1	5	5	100,0	80,0	20,0
SKDA	2	1	2	0	0	5	5	5	40,0	0,0	40,0
SLAB	2	0	3	1	0	4	5	5	40,0	20,0	20,0
SNKA	1	2	2	4	0	1	5	5	20,0	80,0	-60,0
TCHA	3	0	1	1	0	4	4	5	75,0	20,0	55,0
VIDR	1	2	2	2	0	3	5	5	20,0	40,0	-20,0
VLES	2	2	1	0	2	3	5	5	40,0	0,0	40,0
VLKA	2	1	2	1	0	4	5	5	40,0	20,0	20,0
ZLKA	5	0	0	1	0	4	5	5	100,0	20,0	80,0

Table 4: Complete data for/ ŋ /, including sums of individual “0, 1, 2” token ratings for STD and CZE modes, a total number of tokens for each mode, the percentage of “0” ratings for STD and CZE, and finally the difference between the two percentual values (i.e. the success rate).

	STD				CZE				Total tokens		Number of "0" Ratings		
Speaker	***0***	***1***	***2***	***x***	***0***	***1***	***2***	***x***	STD	CZE	% STD	% CZE	Success rate
BCKA	13	0	0	12	0	1	25	0	25	26	100,0	0,0	100,0
BNDA	11	0	0	15	0	0	26	0	26	26	100,0	0,0	100,0
BSTA	23	0	0	3	1	2	23	0	26	26	100,0	3,8	96,2
CMFA	25	0	0	1	0	2	23	1	26	26	100,0	3,8	96,2
DVRK	14	0	1	9	0	0	25	0	24	25	95,8	0,0	95,8
FJTA	19	0	0	6	2	1	23	0	25	26	100,0	7,7	92,3
HBNA	17	1	0	7	0	0	24	1	25	25	96,0	4,0	92,0
HBTA	15	1	0	10	2	4	20	0	26	26	96,2	7,7	88,5
HMNA	15	2	1	8	0	1	24	0	26	25	88,5	0,0	88,5
HSKA	25	0	0	1	0	1	22	3	26	26	100,0	11,5	88,5
HSKR	24	1	0	1	0	0	23	2	26	25	96,2	8,0	88,2
KACT	19	1	0	6	2	1	22	1	26	26	96,2	11,5	84,6
KOSK	17	2	0	7	2	3	20	1	26	26	92,3	11,5	80,8
KVTA	9	3	0	14	1	1	22	2	26	26	88,5	11,5	76,9
MCHA	24	1	0	1	5	0	21	0	26	26	96,2	19,2	76,9
MLCA	18	5	0	2	1	0	25	0	25	26	80,0	3,8	76,2
MLKA	12	0	0	14	2	1	17	5	26	25	100,0	28,0	72,0
MRES	20	3	0	3	5	6	14	1	26	26	88,5	23,1	65,4
PVLK	20	1	1	4	7	1	16	1	26	25	92,3	32,0	60,3
RMSA	16	3	0	7	1	3	13	8	26	25	88,5	36,0	52,5
SKDA	26	0	0	0	9	3	10	4	26	26	100,0	50,0	50,0
SLAB	11	0	0	15	10	1	11	4	26	26	100,0	53,8	46,2
SNKA	26	0	0	0	19	1	5	1	26	26	100,0	76,9	23,1
TCHA	10	12	3	1	6	8	11	0	26	25	42,3	24,0	18,3
VIDR	18	4	4	0	13	3	9	1	26	26	69,2	53,8	15,4
VLES	25	0	1	0	19	2	1	3	26	25	96,2	88,0	8,2
VLKA	20	2	1	3	22	3	0	0	26	25	88,5	88,0	0,5
ZLKA	21	1	0	2	23	1	0	2	24	26	95,8	96,2	-0,3

Table 5: Complete data for/ r /, including sums of individual “0, 1, 2” token ratings for STD and CZE modes, a total number of tokens for each mode, the percentage of “0” ratings for STD and CZE, and finally the difference between the two percentual values (i.e. the success rate).

	STD			CZE			Total tokens		Number of "0" Ratings		
Speaker	***0***	***1***	***2***	***0***	***1***	***2***	STD	CZE	% STD	% CZE	Success rate
BCKA	10	0	0	2	2	6	10	10	100,0	20,0	80,0
BNDA	4	2	4	0	1	9	10	10	40,0	0,0	40,0
BSTA	10	0	0	3	1	6	10	10	100,0	30,0	70,0
CMFA	9	0	1	3	0	7	10	10	90,0	30,0	60,0
DVRK	10	1	0	0	1	9	11	10	90,9	0,0	90,9
FJTA	10	1	0	4	0	6	11	10	90,9	40,0	50,9
HBNA	5	3	2	3	1	6	10	10	50,0	30,0	20,0
HBTA	0	0	10	0	0	10	10	10	0,0	0,0	0,0
HMNA	7	1	2	6	1	3	10	10	70,0	60,0	10,0
HSKA	8	1	1	1	1	8	10	10	80,0	10,0	70,0
HSKR	8	1	1	8	0	2	10	10	80,0	80,0	0,0
KACT	2	2	6	1	1	8	10	10	20,0	10,0	10,0
KOSK	10	0	0	0	0	10	10	10	100,0	0,0	100,0
KVTA	6	1	3	2	0	8	10	10	60,0	20,0	40,0
MCHA	5	2	3	2	3	5	10	10	50,0	20,0	30,0
MLCA	7	1	2	8	1	1	10	10	70,0	80,0	-10,0
MLKA	7	2	1	6	0	4	10	10	70,0	60,0	10,0
MRES	5	1	4	3	1	6	10	10	50,0	30,0	20,0
PVLK	7	1	2	6	0	4	10	10	70,0	60,0	10,0
RMSA	4	3	3	0	0	10	10	10	40,0	0,0	40,0
SKDA	6	0	4	5	1	4	10	10	60,0	50,0	10,0
SLAB	9	1	0	4	0	6	10	10	90,0	40,0	50,0
SNKA	10	0	0	8	0	2	10	10	100,0	80,0	20,0
TCHA	7	0	3	0	0	10	10	10	70,0	0,0	70,0
VIDR	9	1	0	2	0	8	10	10	90,0	20,0	70,0
VLES	8	0	2	3	0	7	10	10	80,0	30,0	50,0
VLKA	9	0	1	5	0	5	10	10	90,0	50,0	40,0
ZLKA	6	0	4	1	0	9	10	10	60,0	10,0	50,0

Table 6: Complete data for/ w /, including sums of individual “0, 1, 2” token ratings for STD and CZE modes, a total number of tokens for each mode, the percentage of “0” ratings for STD and CZE, and finally the difference between the two percentual values (i.e. the success rate).

	STD			CZE			Total tokens		Number of "0" Ratings		
Speaker	***0***	***1***	***2***	***0***	***1***	***2***	STD	CZE	% STD	% CZE	Success rate
BCKA	3	0	1	0	2	2	4	4	75,0	0,0	75,0
BNDA	1	1	2	0	0	4	4	4	25,0	0,0	25,0
BSTA	0	1	3	0	0	4	4	4	0,0	0,0	0,0
CMFA	0	1	3	0	0	4	4	4	0,0	0,0	0,0
DVRK	4	0	0	0	0	4	4	4	100,0	0,0	100,0
FJTA	0	0	4	0	0	4	4	4	0,0	0,0	0,0
HBNA	4	0	0	0	1	3	4	4	100,0	0,0	100,0
HBTA	0	0	4	0	0	4	4	4	0,0	0,0	0,0
HMNA	0	0	4	1	0	3	4	4	0,0	25,0	-25,0
HSKA	0	0	4	0	0	4	4	4	0,0	0,0	0,0
HSKR	2	0	2	0	2	2	4	4	50,0	0,0	50,0
KACT	0	0	4	0	3	1	4	4	0,0	0,0	0,0
KOSK	4	0	0	1	1	2	4	4	100,0	25,0	75,0
KVTA	0	4	0	2	1	1	4	4	0,0	50,0	-50,0
MCHA	0	0	4	0	0	4	4	4	0,0	0,0	0,0
MLCA	2	1	1	0	2	2	4	4	50,0	0,0	50,0
MLKA	1	2	1	0	0	4	4	4	25,0	0,0	25,0
MRES	3	1	0	2	0	2	4	4	75,0	50,0	25,0
PVLK	2	1	1	1	0	3	4	4	50,0	25,0	25,0
RMSA	4	0	0	1	1	2	4	4	100,0	25,0	75,0
SKDA	1	2	1	1	2	1	4	4	25,0	25,0	0,0
SLAB	1	0	3	0	3	1	4	4	25,0	0,0	25,0
SNKA	0	0	4	0	1	3	4	4	0,0	0,0	0,0
TCHA	2	1	1	1	1	2	4	4	50,0	25,0	25,0
VIDR	4	0	0	3	0	1	4	4	100,0	75,0	25,0
VLES	3	1	0	3	0	1	4	4	75,0	75,0	0,0
VLKA	1	2	1	0	1	3	4	4	25,0	0,0	25,0
ZLKA	4	0	0	0	1	3	4	4	100,0	0,0	100,0

Table 7: Complete data for/ æ /, including sums of individual “0, 1, 2” token ratings for STD and CZE modes, a total number of tokens for each mode, the percentage of “0” ratings for STD and CZE, and finally the difference between the two percentual values (i.e. the success rate).

	STD			CZE			Total tokens		Number of "0" Ratings		
Speaker	***0***	***1***	***2***	***0***	***1***	***2***	STD	CZE	% STD	% CZE	Success rate
BCKA	2	0	0	0	0	2	2	2	100,0	0,0	100,0
BNDA	2	0	0	0	0	2	2	2	100,0	0,0	100,0
BSTA	1	1	0	1	1	0	2	2	50,0	50,0	0,0
CMFA	2	0	0	1	0	1	2	2	100,0	50,0	50,0
DVRK	2	0	0	1	0	1	2	2	100,0	50,0	50,0
FJTA	1	0	1	0	1	1	2	2	50,0	0,0	50,0
HBNA	2	0	0	1	1	0	2	2	100,0	50,0	50,0
HBTA	2	0	0	0	0	2	2	2	100,0	0,0	100,0
HMNA	2	0	0	2	0	0	2	2	100,0	100,0	0,0
HSKA	2	0	0	1	0	1	2	2	100,0	50,0	50,0
HSKR	2	0	0	1	0	1	2	2	100,0	50,0	50,0
KACT	1	0	1	0	0	2	2	2	50,0	0,0	50,0
KOSK	2	0	0	1	1	0	2	2	100,0	50,0	50,0
KVTA	2	0	0	1	0	1	2	2	100,0	50,0	50,0
MCHA	1	0	1	1	0	1	2	2	50,0	50,0	0,0
MLCA	1	0	1	1	0	1	2	2	50,0	50,0	0,0
MLKA	1	1	0	2	0	0	2	2	50,0	100,0	-50,0
MRES	2	0	0	1	0	1	2	2	100,0	50,0	50,0
PVLK	1	0	1	1	0	1	2	2	50,0	50,0	0,0
RMSA	2	0	0	1	1	0	2	2	100,0	50,0	50,0
SKDA	1	0	1	0	0	2	2	2	50,0	0,0	50,0
SLAB	0	0	2	0	0	2	2	2	0,0	0,0	0,0
SNKA	0	2	0	1	0	1	2	2	0,0	50,0	-50,0
TCHA	2	0	0	0	0	2	2	2	100,0	0,0	100,0
VIDR	2	0	0	1	0	1	2	2	100,0	50,0	50,0
VLES	2	0	0	0	0	2	2	2	100,0	0,0	100,0
VLKA	0	1	1	0	1	1	2	2	0,0	0,0	0,0
ZLKA	2	0	0	2	0	0	2	2	100,0	100,0	0,0

Table 8: Complete data for/ ɜː /, including sums of individual “0, 1, 2” token ratings for STD and CZE modes, a total number of tokens for each mode, the percentage of “0” ratings for STD and CZE, and finally the difference between the two percentual values (i.e. the success rate).

Appendix III

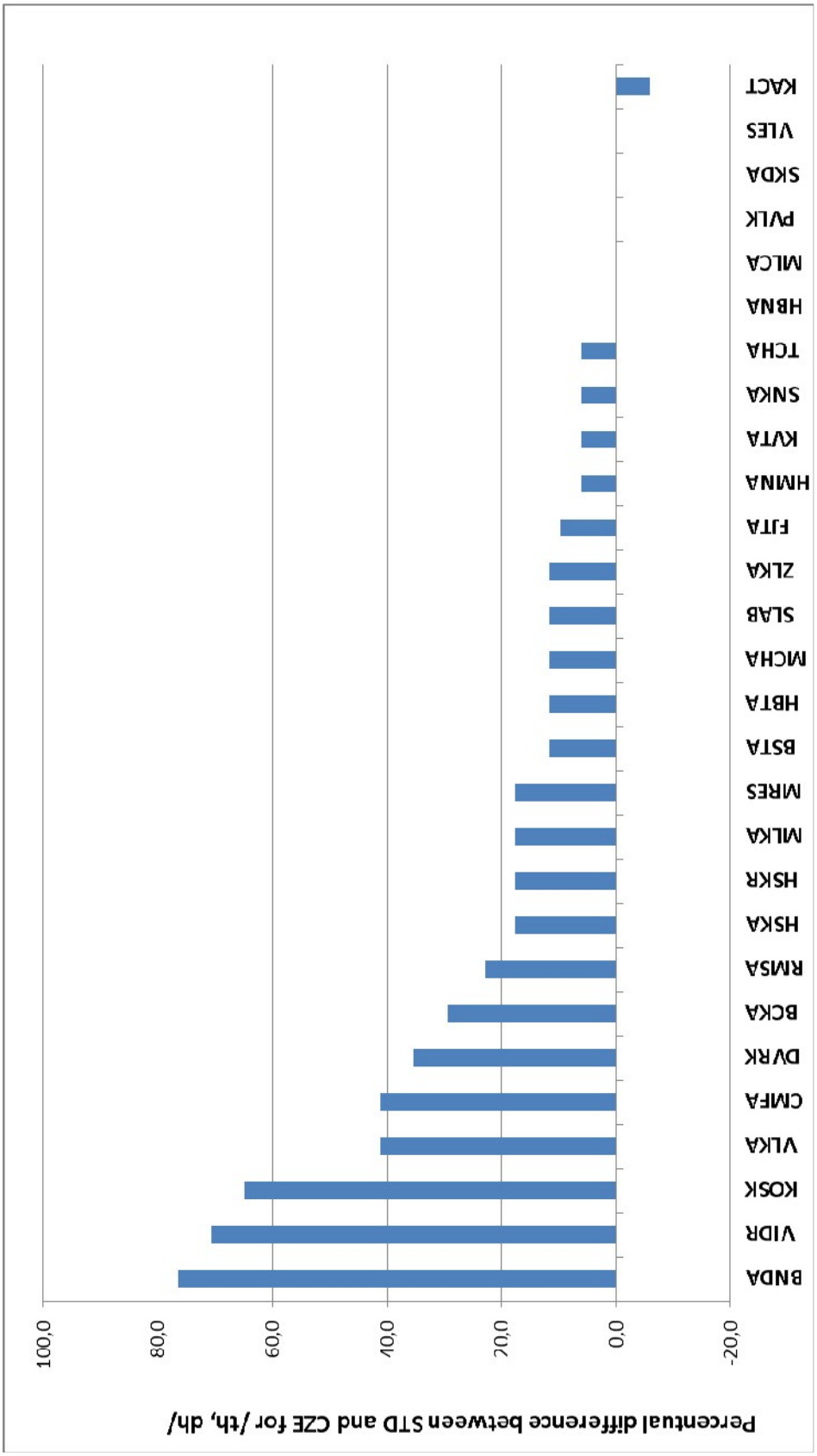

Figure 1: The difference between the percentage of "0" ratings in the STD mode and in the CZE mode for / θ, ð /.

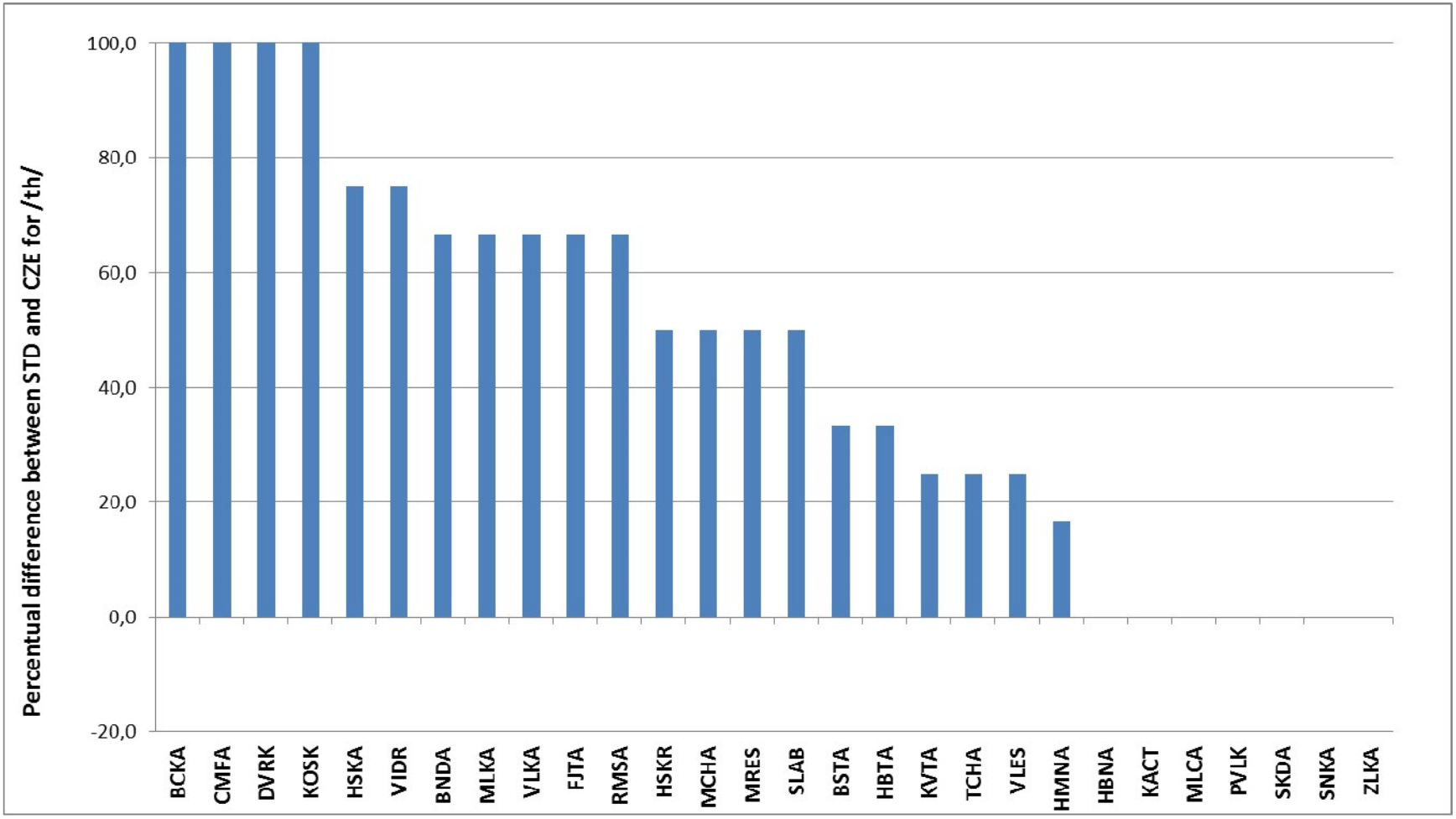

Figure 2: The difference between the percentage of "0" ratings in the STD mode and in the CZE mode for / θ./

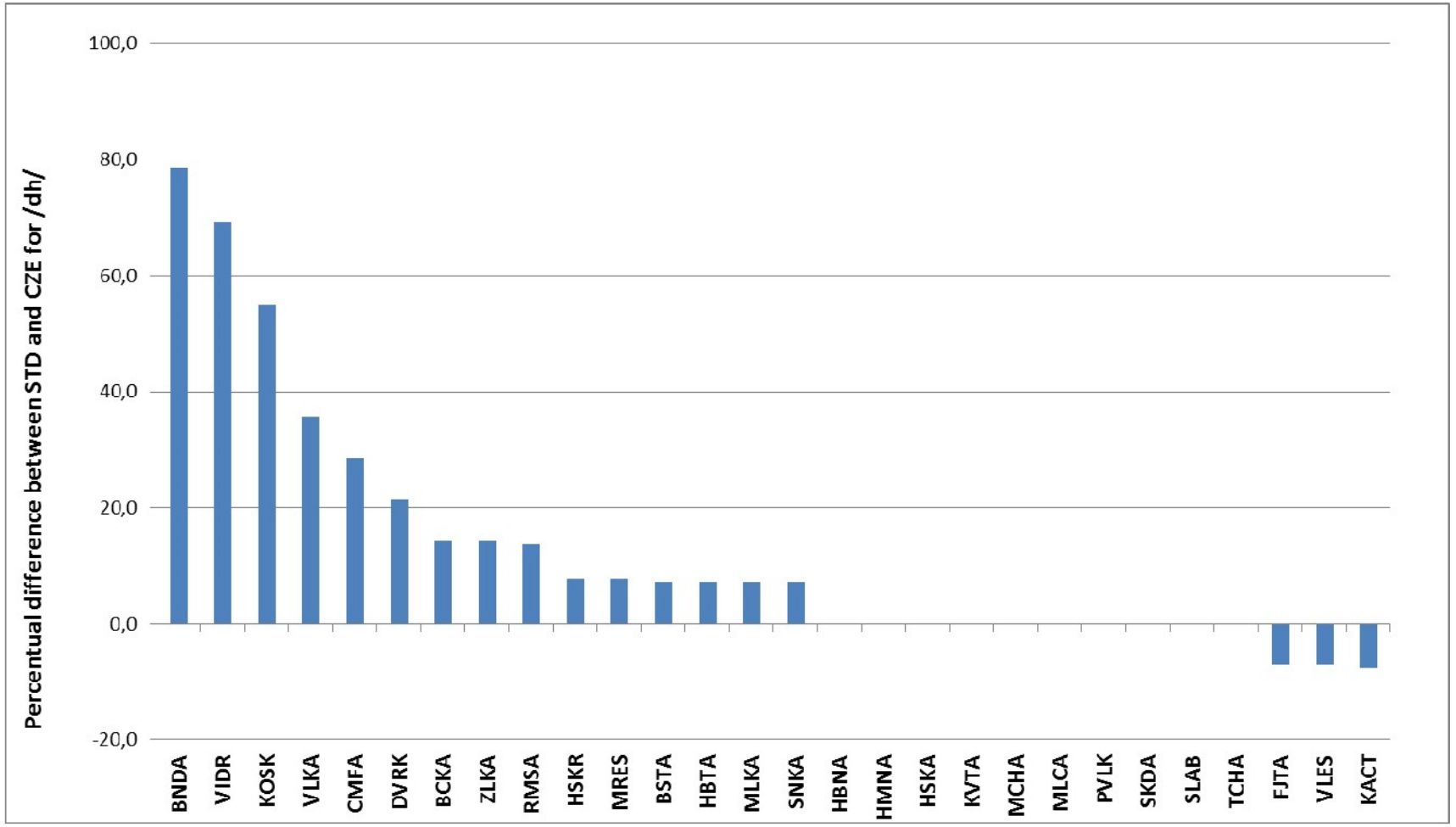

Figure 3: The difference between the percentage of "0" ratings in the STD mode and in the CZE mode for / ð /.

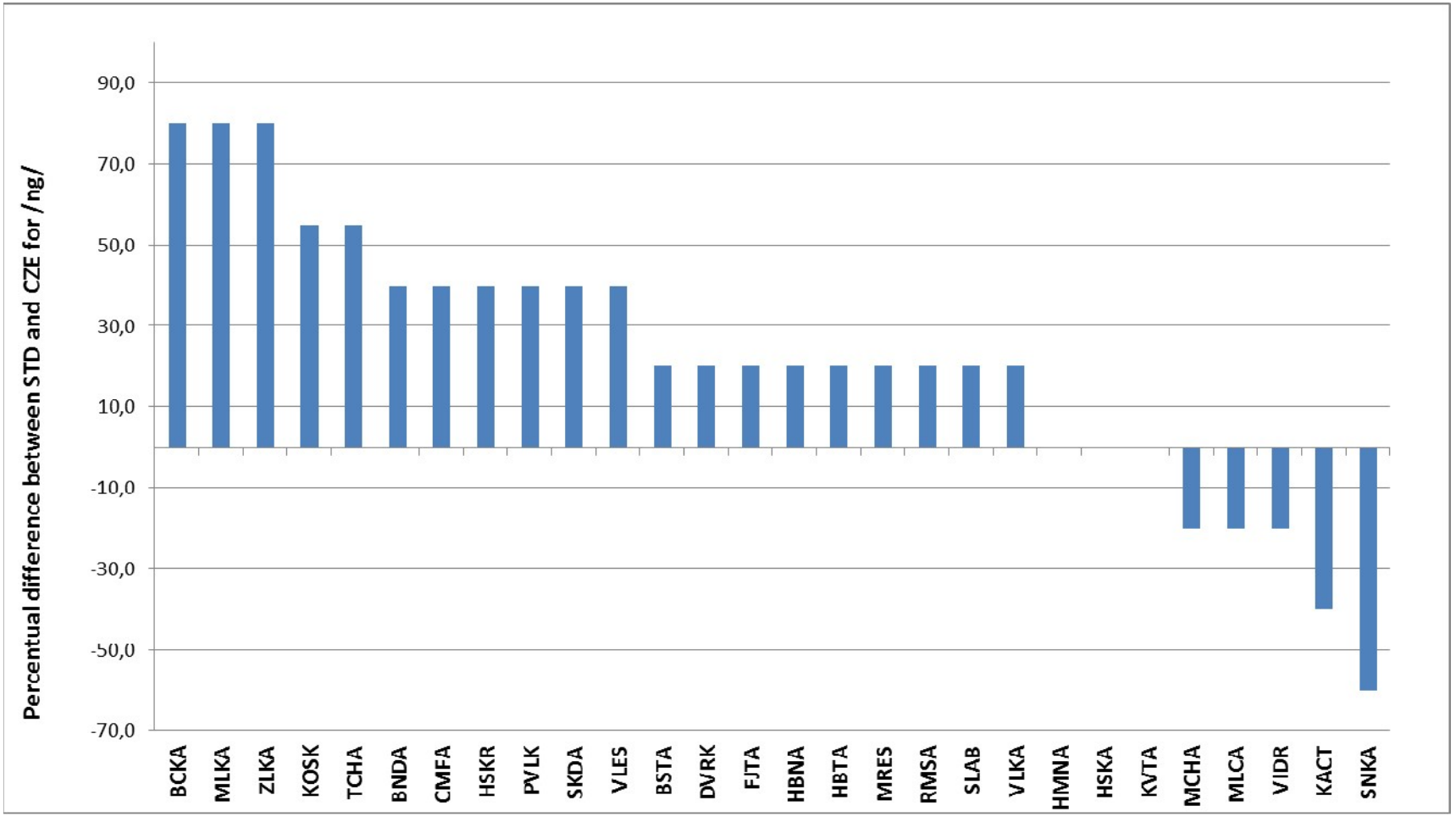

Figure 4: The difference between the percentage of "0" ratings in the STD mode and in the CZE mode for / ŋ /.

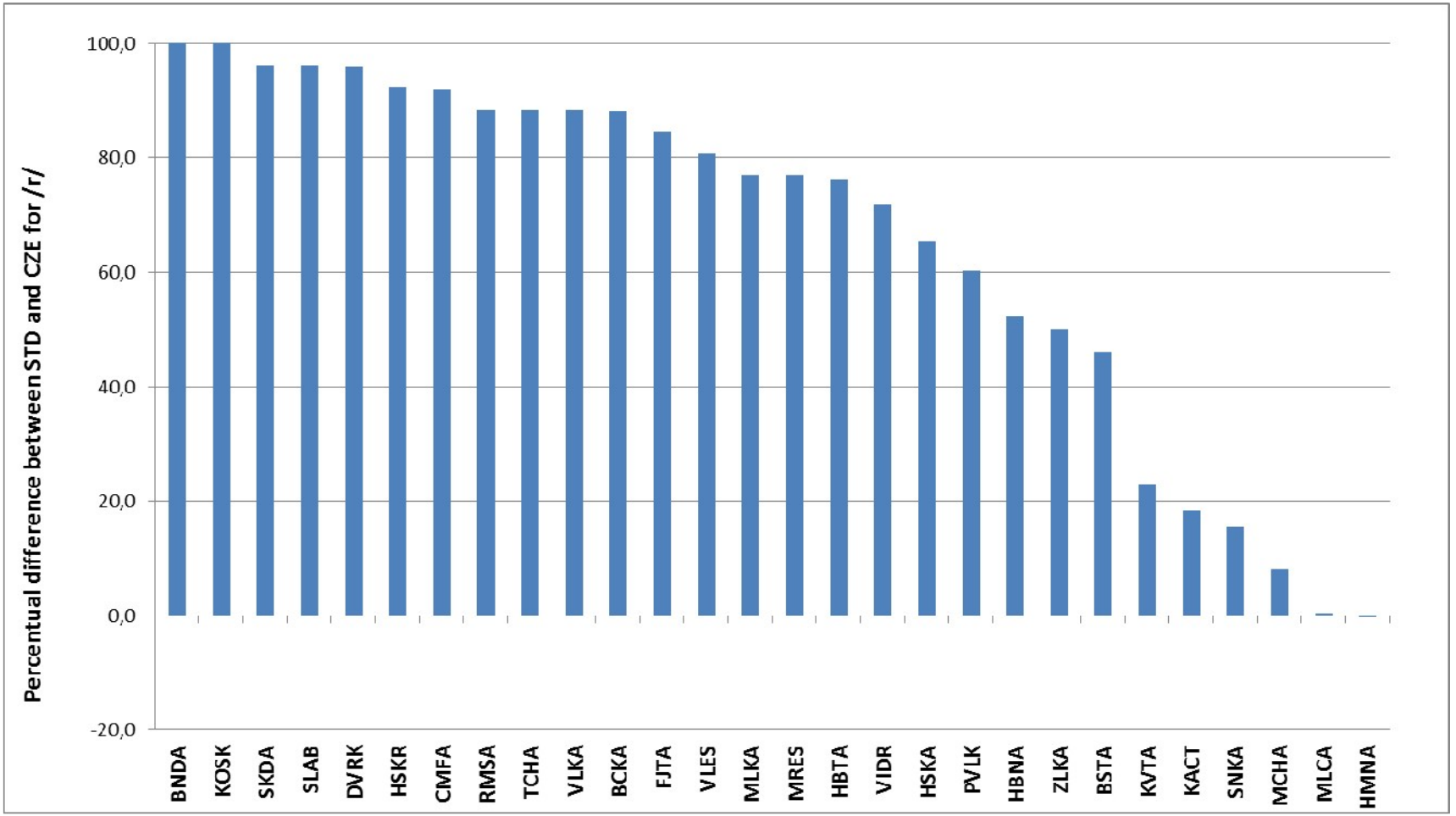

Figure 5: The difference between the percentage of "0" ratings in the STD mode and in the CZE mode for /r/.

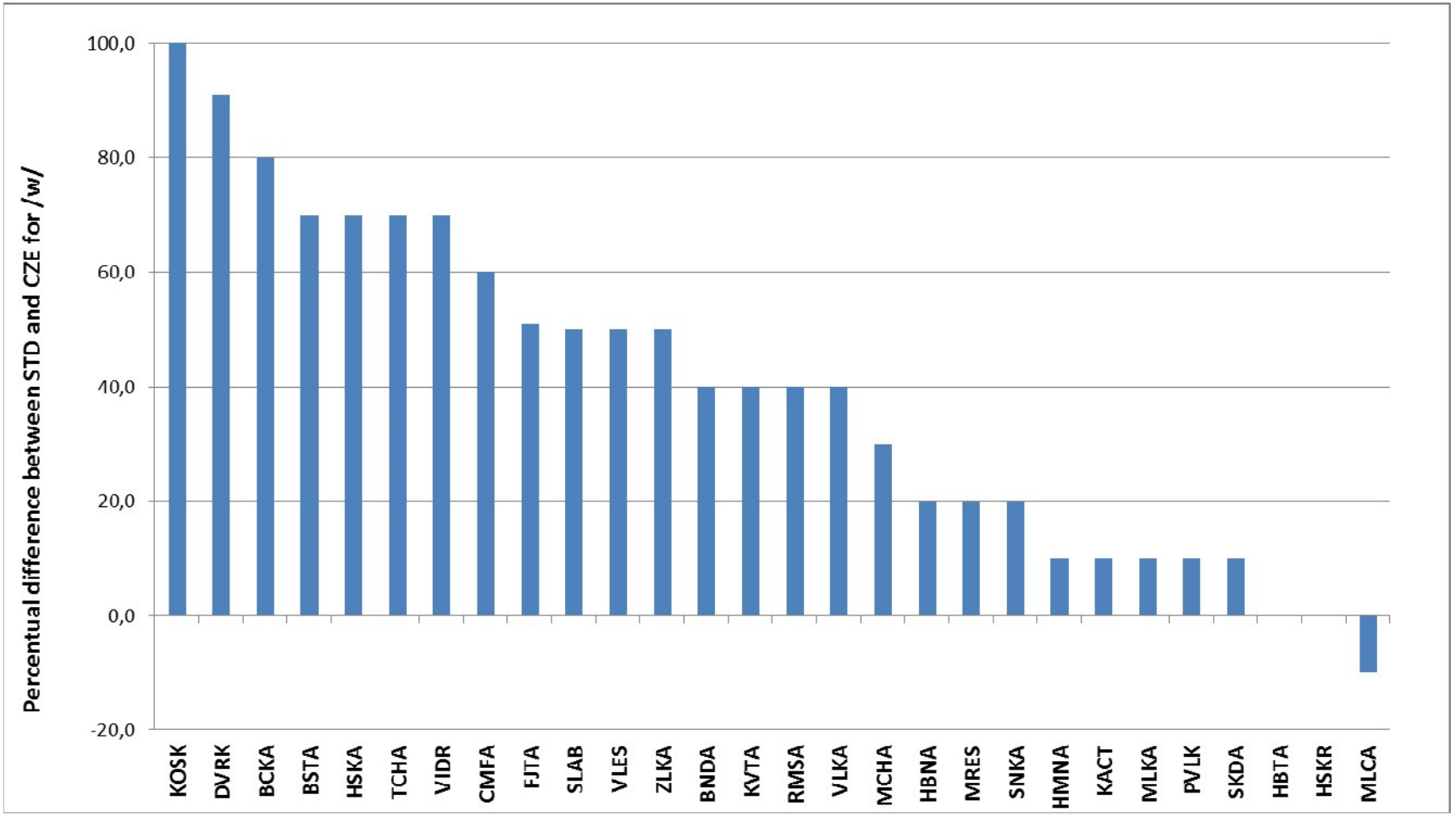

Figure 6: The difference between the percentage of "0" ratings in the STD mode and in the CZE mode for /w/.