Filozofická fakulta Univerzity Palackého

Short-term Phonetic Interference in Interpreting and Code-switching

(Bakalářská práce)

2016 **Tereza Šreková**

Filozofická fakulta Univerzity Palackého
Katedra anglistiky a amerikanistiky

Short-term Phonetic Interference in Interpreting and Code-switching

(Bakalářská práce)

Autor: **Tereza Šreková**

Studijní obor: Angličtina se zaměřením na komunitní tlumočení a překlad

Vedoucí práce: **Mgr. Šárka Šimáčková, Ph.D.**

Počet stran: 44

Počet znaků: 65 639

Olomouc 2016

Prohlašuji, že jsem tuto diplomovou práci vypracovala samostatně a uvedla jsem úplný seznam citované a použité literatury.

V Olomouci dne 1. 5. 2016 Tereza Šreková

I would like to express my gratitude to Mgr. Šárka Šimáčková, Ph.D. for valuable advice and guidance on elaborating the thesis as well as for her willing and helpful approach. Further, I am very grateful to the participants of the research without whose cooperation this paper would not have been possible.

Abstract

This bachelor's thesis focuses on manifestation of foreign accent in speech production of native Czech students of English with specialization in translation and interpreting. The speech production is explored under three experimental conditions: speaking L2 English, code-switching and interpreting into English. The aim of this experimental study is to investigate under which of these conditions the influence of L1 on L2 is the most eminent. A specific parameter used for comparison of the accent degree is the voice onset time of voiceless consonants *p* and *t*. The paper could be divided into two parts, theoretical and practical. The theoretical part summarizes previous research on the given topic, deals with possible causes of the non-native pronunciation and the factors which could further influence the degree of accent manifestation in particular situations. The practical part includes methodology of the experiment, information about the participants, analysis of the retrieved data and reflection on the results.

Key words

interference, speech production, foreign accent, voiceless stops, VOT, code-switching, interpreting

Abstrakt

Tato bakalářská práce se zabývá mírou cizího přízvuku studentů angličtiny se zaměřením na tlumočení a překlad, kteří jsou rodilými mluvčími češtiny. Jejich produkce řeči je zkoumána za třech experimentálních podmínek: mluvení anglicky, přepínání kódů a tlumočení do angličtiny. Daná experimentální studie má za cíl zjistit, za které z určených tří podmínek se vliv rodného jazyka na další získaný projeví nejvíce. Vliv je zkoumán na neznělých plozivních souhláskách *p* a *t*, konkrétně pomocí jejich doby nástupu hlasivkového tónu. Obsah práce lze rozdělit na dvě části, teoretickou a praktickou. Teoretická část shrnuje předchozí výzkum o jazykové interferenci, zabývá se možnými důvody, proč nerodilí mluvčí často nedosahují stejné výslovnosti jako ti rodilí a v neposlední řadě zkoumá faktory, které by dále mohly ovlivnit míru přízvuku v konkrétní situaci. Praktická část poté obsahuje metodologii experimentu, informace o účastnících, analýzu získaných dat a diskuzi nad výsledky.

Klíčová slova

interference, produkce řeči, cizí přízvuk, neznělá ploziva, doba nástupu hlasivkového tónu, přepínání kódů, tlumočení

Obsah

1 Introduction

The present experimental study investigates speech production of students of the 3rd year of English with a specialization in translation and interpreting. All are native speakers of Czech advanced in L2 English who have not been exposed to native English environment during their L2 acquisition. Second language learners are often known to speak with a certain degree of foreign accent (Grosjean and Li 2013, 7). A study conducted by Abrahamsson and Hyltenstam (2009) selected non-native but very experienced speakers of L2 Swedish and subjected them to a number of tests which aimed to reveal whether L2 learners can reach a native-like performance. In one of the tests, utterances produced by these participants were evaluated by native speakers who were supposed to identify, whether to them presented speech was uttered by a native or a non-native speaker. The results showed that many quite proficient learners (especially late learners) could not pass test as a native speaker. That indicates that even advanced learners such as the bilingual participants of the current research might speak with a distinguishable foreign accent.

As Grosjean and Li (Grosjean and Li 2013, 14–21) state, bilinguals have to face decisions concerning the everyday language use. Their often unconscious choices lead to a determination of language mode for situations they find themselves in. The language mode could be described as "[the] state of activation of the bilingual's languages and language processing mechanisms at a given point in time" (Grosjean and Li 2013, 15). The state of activation then might show an impact on the degree of language interference – "[the] influence of a bilingual's two languages either uni or bi-directional" (Simonet 2014, 26).

Recent studies of language influence have investigated speech production of a non-native group of speakers under various conditions requiring a different level of activation of the languages involved (Bullock et al. 2006, Antoniou et al. 2011, Simonet 2014) rather than comparing the native-likeness. One of these studies is an experiment conducted by Šimáčková and Podlipský (2015) that observed interference of L1 Czech on L2 English under three conditions requiring different activation of bilinguals' languages (speaking in L2 English, code-switching between L1 Czech and L2 English and interpreting from L1 Czech to L2 English). The results revealed that when the participants had to activate both of their languages (in code-switching and

interpreting), the interference was greater. It was hypothesised that in the interpreting, the influence of L1 will be manifested more than in code-switching because of the cognitive load imposed on the speaker. However, the results showed the same effect, with an insignificant trend of more native-like speech when interpreting.

The purpose of this experiment is to replicate the above illustrated experiment with a small but significant change in the procedure, and to further explore the observations made, especially the distinction between code-switching and interpreting. Identically as in the original study, the language influence will be investigated on the production of voiceless stops *p* and *t*, specifically on their VOT (voice onset time), defined as "[the] interval between release of a closure and the start of the voicing", since the typical values of VOT differ in individual languages (Ladefoged and Johnson 2010, 151).

2 Literature review

2.1 Language influence

It is quite common for non-native speakers to have a certain degree of an accent in the foreign language (Grosjean and Li 2013, 7). The issue of a foreign accent has been discussed in a wide range of studies, some dealing with aspects influencing its degree. For example, the study by Flege, Yeni-Komshian and Liu (1999) focuses on the impact of age on acquisition of second language by assessing participants' degree of foreign accent and their knowledge of grammar. Piske MacKay and Flege's article reviews several factors influencing the degree of foreign accent and in an experiment tests the relative impact of these factors (Piske, MacKay and Flege, 2001). These two studies use a global accent rating method to evaluate the degree of influence of one language on another by relying on native listeners. Other studies examine production of specific sounds in native and nonnative pronunciation (Fowler et al. 2008; MacLeod, Stoel-Gammon, and Wassink 2009).

One phonetic experiment which takes into consideration both perception and production is a study by Flege and Eefting (1987a) concerned with stop consonants. The experimenters investigated whether bilingual speakers of Dutch and English perceive stops differently under two different language conditions (English and Dutch) – so called "language sets". By these "language sets" the authors refer to the documents and instructions given during the experiment. The participants of the research were adult native speakers of Dutch, forty of them studied English and ten pursued education in technical fields. All of them started learning English at the age of 12 years, however, the technically oriented group was not in contact with English as much as the group of students of English. In the experiment, they were exposed to synthetized scale of syllable representations /*da*/ and /*ta*/. The scale began at VOT[1] value of −60ms and included values increased always by 10ms up to the 90ms. The subjects were led to believe that they were about to listen to two different scales of stimuli corresponding to the Dutch and English "language sets", which was not the

[1] Ladefoged and Johnson define VOT – voice onset time as"[the] interval between release of a closure and the start of the voicing" (Ladefoged and Johnson 2010, 151). As they explain, values of VOT form a scale showing how voiced or voiceless the sound is. Negative values mean that the voicing began before the release of a closure, indicating that the sound is voiced, on the other hand, positive values refer to voiceless sounds (Ladefoged and Johnson 2010, 151). For more information on VOT, see chapter 3.2.1.

case. Their task was to determine whether a presented stimulus is syllable /*da*/ or rather /*ta*/.

In addition, the experimenters also evaluated proficiency of the subjects based on three criteria. Firstly, 8 native speakers of English were asked to evaluate a foreign accent degree of 3 sentences produced by the bilinguals. The second criterion consisted in the self-assessment of their foreign accent in the form of filling in a questionnaire. Lastly, the ability to authentically pronounce specific sounds in English (in this case production of phoneme /*t*/ in terms of its VOT value) was taken into consideration. According to these criteria, subjects were divided into 3 proficiency groups.

The results confirmed that the participants perceived stops differently in the two distinct language environments (although the difference was not very prominent). The threshold for English environment shifted towards higher VOT values in comparison to the Dutch mode, but the extent of the shift did not grow together with the proficiency. Similarly as in the perception task, a VOT value of stops increased when participants spoke English, which corresponds to the fact that VOT values of stops are higher in English than in Dutch (Lisker and Abramson 1964). The more experienced the speakers of L2 English were, the greater was the increase. Finally, it was found that native Dutch speakers with a higher degree of proficiency in English pronounced /*t*/ the VOT which of reached smaller values than VOT of native Dutch speakers with a lesser English language competence. The authors interpreted the last finding as a possible L2 influence on L1.

Similar results in terms of production were reported in another study conducted by Flege and Eefting (1987b). The purpose of this study was to reveal whether the phonetic categories for phoneme realizations not-included in the native language are established in the process of learning second language. Participants formed 3 Spanish-English bilingual groups: a group of bilingual children and two groups of adult bilinguals who differed in the age at which they started learning their second language (English). To these 3 groups of bilinguals were assigned 2 corresponding groups of English monolinguals (children, adults) and Spanish monolinguals as well (children, adults). Their first task was to read words in English and Spanish starting with a stop consonant for the purpose of VOT analysis. In the second task, the participants were presented with synthetized continuum of /*da*/ and /*ta*/ syllables with different VOT

values (beginning at −60ms to 90ms, with an increase by 10ms) and were supposed to decide whether the stop consonant they hear is /*t*/ or /*d*/.

The data from the first part of the research showed that bilinguals produced English words with voiceless stops with shorter VOT than the native speakers, although longer in comparison with their production of voiceless stops in Spanish words. This corresponded with authors' expectations that phoneme realizations of voiceless stops in English and Spanish are considered as members of one phonetic category. However, in the view of further results, the authors considered that the bilinguals might have created new phonetic categories at least partially dependent on the L1 categories for voiceless stops, since in the production of Spanish words, bilinguals realized voiceless stops with even shorter VOT than their monolingual Spanish counterparts. The phoneme identification test revealed that the average threshold point for adult bilinguals was located in the area with lower values of the VOT continuum than the threshold point for native English speakers. Nevertheless, that did not apply to the children whose threshold (both for bilingual and native English children) occurred at approximately same point.

So far described studies were concerned rather with influence of L1 on L2, however the result of this research which showed decrease in voiceless stop VOT value produced by native Spanish bilinguals in comparison to native Spanish monolinguals indicate that L2 can affect L1 as well.

The mutual influence of languages in pronunciation was explored in the study of Flege and Hillenbrand (1984). The purpose of the study was to discover whether L2 speech sounds which sufficiently differ from all L1 realizations of phoneme categories are produced more authentically than the L2 speech sounds which might be taken as a part of some L1 category. The experiment regarded production of French phonemes /*y*/ and /*u*/ by native English bilinguals, /*y*/ being the phone without corresponding representation in English, /*u*/ having similar "counterpart" in English. In addition to that, the experimenters asked whether a greater language proficiency leads to more authentic production. They further investigated validity of Flege's hypothesis which states that foreign language learners establish "… superordinate acoustic models based on pairs of corresponding sounds or phonetic dimensions in two languages [that] serve

as input for phonetic learning in second language acquisition", leading to speech with a foreign accent (Flege 1981, 452).

In the first experiment three groups of speakers – one group of native French, two groups of native English speakers differing in their experience in L2 French – recorded French syllables /*tu*/ and /*ty*/. These syllables were then presented to subjects – native speakers of French – whose assignment was to identify whether they encountered /*tu*/ or /*ty*/. Outcome of the identification task showed that the subjects were more successful in recognition of syllables produced by native French speakers than those with French as an L2. As far as regards the comparison of the less and the more proficient group, the recognition rate of */ty/* was approximately the same, however /*tu*/ produced by more proficient speakers was recognized with a significantly greater success than /*tu*/ pronounced by a less proficient group. As regards the comparison within the groups, more proficient group produced more recognizable /*tu*/ than /*ty*/ syllable, however, in case of non-proficient speakers it was the other way around.

The second experiment further explored the influence of proficiency on the successful production of /*y*/ and /*u*/. The same recorded stimuli were employed, however this time they were presented to native English speakers, not French as in the first part of the experiment, since the authors presumed that it will be easier for native English speakers to differentiate between /*ty*/ and /*tu*/ produced by other native English speakers. The collected results were in compliance with the results of the first part. Again, the utterances were recognized correspondingly to the proficiency in French – the best recognition rate was reached in case of syllables produced by native French, slightly worse by the more proficient group of non-native speakers and worst by the less proficient group.

Finally, the recorded sounds were evaluated in terms of VOT and formant[2] values. As the authors expected, both native French bilinguals' and native English bilinguals' VOT values were considerably higher than the values normally produced by monolingual French speakers. In terms of formants of the vowels, both native English and native French reached quite similar values in terms of /*y*/ vowel

[2] According to Ladefoged and Johnson, formant is "[a] resonating frequency of the air in the vocal tract" (Ladefoged and Johnson 2010, 307).

production, however, quite different as regards /*u*/ (the less degree of proficiency, the higher F2 values). The authors also compared the F2 values to the values reported in previous studies on French monolinguals. In case of /*y*/ there was no significant difference, however in /*u*/ production, they observed increase in F2 of the native French bilingual speakers compared to monolingual values. The overall findings correspond to the Flege's hypothesis on merging the phonetic properties of "similar" phones (two corresponding phones) in two languages (Flege 1981). The authors also found that a greater proficiency led to more authentic pronunciation.

The study showed not only that non-native speakers of French retain a certain degree of accent, in other words, that their L1 influences L2, but also that French native speakers were affected by their L2 in production of French /*u*/.

Another study which is in favour of a statement that the language influence is mutual is Sancier and Fowler's study (1997), dealing with a so called "gestural drift" defined by the authors as "perceptually guided changes in speech production" (Sancier and Fowler 1997, 421). These changes were observed in speech production of a native speaker of Brazilian Portuguese with L2 English, whose speech was allegedly different after each stay in the USA. This claim was verified by means of 3 related experiments. Firstly, the subject was recorded translating sentences both from Brazilian Portuguese into English and in the opposite direction. The recording was conducted three times and it took place on a departure to Brazil after 4.5 month-stay in the USA, then after 2.5 month-stay in Brazil and again after 4 month-stay in the USA.

The first experiment was based on evaluating a degree of the manifested accent. Native speakers of Brazilian Portuguese were invited to participate and were presented with before recorded corresponding translations of the sentences into Brazilian Portuguese, one of which was produced after the subject's stay in Brazil and the other after the stay in the USA. Their task was to identify a sentence manifesting (more) sings of a foreign accent and alternatively to enlist particular features contributing to its foreign impression. The results showed that in 66% cases, the participants marked the sentences produced after the stay in USA as more accented over sentences produced after the stay in Brazil. Further, among features contributing to the foreign impression were placed changes in intonation and nasality and excessive articulation.

The second experiment was conducted in the same way with two differences – the concerned sentences were translations into English and they were presented to native speakers of English. The task was identical, that is, to establish which of the presented utterances has a higher degree of foreign accent. The results revealed that the participants found English sentences produced after the stay in Brazil as more accented in 48% of cases which according to the authors reflects "choice by chance". They believed that this occurrence could be explained in two ways. Either it is only Portuguese and not English that is influenced by the different language environment or distinguishing native speech and accented speech is easier than distinguishing between various degrees of a foreign accent.

The third experiment concerned the comparison of voiceless stops in sentences after stay in Brazil and the USA. Since English /*p*/ and /*t*/ is aspirated however, in Portuguese it is not, the experimenters expected that the VOT of these stops will adapt to VOT values of a language to which the subject was exposed to (Sancier and Fowler 1997). In general, the VOT of stops when speaking English (regardless the environment where the subject spent time) was longer than when speaking Portuguese. That matches the difference in aspiration of languages. Both consonants registered increase in VOT when the recording was conducted after the stay in the USA.

In conclusion, subject's accent was influenced by the stay in a different language environment, however only speakers with the same native language as the subject were able to distinguish this change.

3 Manifestation of a foreign accent in speech

As we have seen in the above described studies, not even a high level of proficiency prevents from the occurrence of a foreign accent in a non-native language. That leads to a question what causes this phenomena.

Chomsky stipulated that when studying language, it should be distinguished between language competence and language performance. The language competence refers to "the speaker-hearer's knowledge of his language", whereas the performance denotes "the actual use of language in concrete situations" (Chomsky 1969, 4). The reason why the accent is manifested in speech might lie within both of these areas.

3.1 Studies investigating accent in terms of the language competence

A range of studies, such as the above described experiments, provide a comparison of speech production of bilingual speakers to the production of monolingual native speakers. They focus on subjects' overall ability to speak a foreign language, their knowledge of a language i.e. the language competence. They investigate its possible lack leading to a non-native pronunciation.

One of the possible explanations of non-native pronunciation is that the L2 speakers fail to establish phonetic categories. As Kuhl (2004) explains, it is believed that as infants, humans are able to distinguish phonetic units, defined as "specific articulatory gestures that constitute vowels and consonants in a particular language" (Kuhl 2004, 832). This enables them to adapt to any language. However, as they grow older, they cease to distinguish phonetic units which are not relevant for their native language (the phonetic units which are a part of a one phonemic category in their native language). This change starts when they reach 6 months. After 12 months of their age, their ability to recognize phonetic units of other languages decreases, nevertheless, at the same time, they show progress in recognizing phonetic units of their native language. As a result, the perception of adult speakers is "categorical", e.g. the adults "classify the sounds used in their languages as one phoneme or another, showing no sensitivity to intermediate sounds" (Kuhl 2004, 833).

As an example of these phonetic units could serve velar /*ki*/ and uvular /*qi*/ of the Thompson language the perception which of was explored in an experiment conducted by Werker and Tees (2002). In Thompson language, these sounds serve as phonemes, however in English their distinction in place of articulation is not relevant.

The subjects i.e. English infants, English adults and Thompson language adults were presented with these syllables and their task was to distinguish them from one another. The results showed that Thompson language speakers distinguished these sounds in all cases. In accordance with the above stated information, English infants reached a comparable score to the Thompson speakers, whereas English adults obtained considerably worse results.

Exactly the imperfect perception of sounds due to categorization might lead to a foreign accent. Flege (1987) elaborated on the topic of phonetic categories and proposed a hypothesis regarding so called "new" and "similar" phones. He describes "new" phones as L2 sounds which do not have a corresponding phone in L1, whereas "similar" are perceived as a realization of a single category existing in L1. He argues that equivalence classification, "…a basic cognitive mechanism which permits humans to perceive constant categories in the face of the inherent sensory variability found in the many physical exemplars which may instantiate a category", hinders from establishing new categories for a "similar" phone in L2 (Flege 1987, 49). He claims that unless a category for L2 phone is established, speakers will not be able to produce these phones in a native-like manner. These assumptions were tested in an experiment dealing with production of French vowels /*u*/ and /*y*/ as a part of syllable with a stop at the beginning. Native speakers of English consider French /*y*/ to be a "new" phone, whereas /*u*/ is a phone with its counterpart in English. The majority of groups of English native speakers of French produced /*y*/ with values only slightly different from monolingual French speakers. However, in production of /*u*/ none of the groups acquired results approximately corresponding to French. That appears to be in compliance with above mentioned hypothesis that a new category (for /*y*/) can be established, however the authentic production of a "similar" phone is not possible because of the equivalence classification.

Other possible reason why the foreign accent occurs might be that the speaker creates a new phonetic category, however, he or she is not able to produce it. That was suggested for example by Flege and Hilllebrand (1984) in their above described research on production of French /*y*/ and /*u*/ as a possible explanation to quite accurate yet still not precisely nativelike French /*y*/ produced by non-native speakers. They argued that the non-native speakers might have failed to learn a proper way to

articulate this phone, although they have established a phonetic category for this phone.

3.2 Fluctuation of the accent degree – speakers′ performance

However, other studies, instead of dealing with speakers' competence i.e. exploring the overall accent of the bilingual speakers in comparison to monolingual speakers, focus on the degree of speakers' accent in particular conditions – on their linguistic performance.

As an example of these studies, known as code-switching[3] studies, could be enlisted an experiment conducted by Antoniou et al. (2011) which investigated production of stops in Greek and English environments. Instead of involving only monolingual language mode[4], it included a code-switching condition, i.e. the participants were required to activate both of the concerned languages at a single point (Grosjean and Li 2013, 15). The language influence was observed in speech production of L2 dominant[5] bilinguals. According to the self-evaluation, the bilinguals obtained higher level of proficiency in English (L2) than in their native Greek (L1). They were divided into two groups and were asked to read speech elements (syllables such as "pa", "aba", "anda", etc.) that included stops *b, d, p, t*. In the monolingual mode, which was part of the previous experiment (Antoniou et al. 2010), the elements were implemented into phrases in a language corresponding to the language in which the element was (e.g. English phrase and English element). In the bilingual mode, the element was in a different language than the phrase (e.g. Greek element within English phrase) so that the participants were obliged to switch from one language to the other. (Differentiating English and Greek elements was possible because of the distinction in their orthography.)

[3] Grosjean and Li define code-switching as „the alternate use of two languages, that is, the speaker makes a complete shift to the other language and then reverts back to the base language"(Grosjean and Li 2013, 18). For more information on code-switching, see the chapter 3.2.2.1.

[4] According to Grosjean and Li, language mode could be characterized as „the state of activation of the bilingual's languages and language processing mechanisms at a given point in time"(Grosjean and Li 2013, 15). The authors further clarify that there are two fundamental modes, monolingual and bilingual. Monolingual mode indicates that only one of the bilingual's languages is activated, bilingual mode signifies activation of two or more bilingual's languages. (Grosjean and Li 2013, 15). This issue is discussed in more detail in chapter 3.2.2.

[5] As Wei explains, dominant bilingualism refers to "[a] situation in which a bilingual has greater proficiency in one of his or her languages and uses it significantly more than other language(s)"(Wei 2000, 495).

The experimenters compared the values of VOT in code-switching condition to the monolingual mode. They evaluated stops in various locations within the element – initial and medial (either preceded by a vowel or a nasal). As regards the influence of code-switching on the production of initial stops, it manifested in case of English stops, however the VOT values of Greek stops remained the same. When the same participants were exposed only to the English environment (monolingual mode), their initial stop VOT values were same as for native speakers of English (Antoniou et al. 2010). Identically as in case of the initial position stops, the English medial position stops following a vowel were influenced by the Greek environment, however, so were Greek stops (voiced) influenced by the English environment. The production of stops in medial following a nasal was again influenced only in case of English stops (L1 on L2 influence).

In conclusion, the authors observed a mutual influence of L1 and L2 (more Greek-like VOT values) even though the participants were L2 dominant. The influence was asymmetric, since L1 affected the L2 almost in all cases, but L2 on L1 appear only in case of voiced stops in medial position. When comparing the two modes (code-switched and the monolingual mode), the impact was clearly greater in code-switching condition.

Bullock et al. (2006) also explored influence of languages using code switching condition. As in experiment by Antoniou (2011), they also found that the influence is asymmetric, however in the first experiment the effect occurred in L2 and in the second experiment in L1. The investigators examined production of initial voiceless stops by native Spanish bilinguals rather proficient in L2 English. The aim of the first experiment was to compare the degree of influence in the monolingual and bilingual (code-switching) mode. The task of the participants was to read monolingual (both Spanish and English) and bilingual code-switched (both from Spanish to English and English to Spanish) sentences. The results showed that participants' values of VOT in monolingual English mode and monolingual Spanish mode significantly differ (25ms for Spanish and 55ms for English). When in the bilingual mode, the Spanish values remained the same, however, the English values declined, approximating the Spanish values. The L1 influence was observed regardless of the direction of the switch (English-Spanish, Spanish-English). The effect of a code-switch did not last long, however the lowest English VOT values were observed before the switch to Spanish.

The second experiment was identical with one exception, the participants were not native speakers of Spanish but native speakers of English quite proficient in Spanish. As in the first experiment, the values for monolingual modes differed: 31ms for Spanish and 61ms for English. When comparing the code-switch and monolingual mode values, the same effect occurred, i.e. the Spanish VOT code-switched values did not go through a change in comparison to the monolingual mode, however the English did. The change therefore happened in L1, not L2. In conclusion, in both experiments the code-switching lead to asymmetric effect. The influenced language was English, but once as participants' L2 and then as their L1.

The aforementioned studies created an artificial code-switching situation, however Balukas and Koops (2014) investigated this phenomenon in common spontaneous conversation. They retrieved recorded speech of Spanish-English bilinguals (majority of them had Spanish as L1) from a specialized corpus and extracted initial voiceless stops in order to analyse their VOT. Subsequently, the experimenters found the points when the code-switch occurred and determined the distance of the words containing stops from the previous code-switch.

By analysing the VOT, the authors discovered that the general VOT values of speakers' Spanish voiceless stops corresponded to the values of monolingual Speakers of Spanish. Nevertheless, the English stop values were rather close to lower boundary of the monolingual speaker spectrum which appeared to be an overall influence of Spanish (L1 for the majority of speakers) on English. As regards the code-switching effect, the Spanish words following a code-switch remained unchanged, whereas English words displayed a more English-like VOT of the stops. As in the above described studies, there was an asymmetry in the effect of code-switching.

The authors suggested that the asymmetry might have been caused by the sequence in which the languages were acquired, which is partially in the conflict with the results obtained by Bullock et al. (2006) who found that English was affected even when it was participants' L1, whereas the L2 remained the same. Other explanation they proposed was that the reason of the asymmetry dwells in the difference of the compared languages. They argued that since the VOT extent of English stops is greater than of the Spanish stops, it is possible to lower the VOT and at the same time remain within the spectrum of native speakers, as it was suggested in the study of Bullock et

al (2006), or rather within a correct part of the border dividing aspirated and unaspirated stops. The experiment also confirmed that the effects of code-switching appear not only in artificial experimental conditions but also in common speech.

All of the above described papers on code-switching have used the same assessment feature – VOT. There are also studies employing different methods such as Simonet's experiment on production of Catalan vowel by Catalan-Spanish bilinguals (2014) or an experiment conducted by Schwartz, Balas and Rojczyk (2015) who observed language influence on occurrence of glottalization or linking of words beginning with a vowel, but generally, VOT is the most commonly investigated criterion for observing inter-language code-switching influence (Piccinini and Arvaniti 2015, López 2012, Olson 2013).

3.2.1 VOT

Voice onset time is a parameter used to describe oral stop consonants[6]. Ladefoged and Johnson define VOT as "the interval between the release of a closure and the start of the voicing" (Ladefoged and Johnson 2010, 151). As they state, when the vocal folds start vibrating before the release of the stop, the value of VOT is negative and the sound is marked as voiced. On the contrary, the start of voicing after the release indicates positive VOT values i.e. a voiceless sound. Three types of stops can be distinguished based on their VOT: voiced, voiceless unaspirated and voiceless aspirated (Ladefoged and Johnson 2010, 151–157). As Lisker and Abramson observed, the VOT values differ significantly in individual languages, which also a reason why it is so commonly used as an indicator of language interference. The values are measured in milliseconds and they vary for each consonant in relation to the place of articulation (Lisker and Abramson 1964). Further, the VOT can be affected by the speaker's speech tempo (Kessinger and Blumenstein 1998).

[6] As Ladefoged and Johnson explain, oral stop consonants are produced in way that articulators create a complete obstruction (closure) which prevents the air from flowing through neither oral nor nasal cavity. That leads to an increase of the pressure in the cavity followed by a release when the articulators are loosened. These consonants can be further described by means of voicing. Voicing refers to the state of vocal folds, whether they are vibrating or not. Based on the voicing, stops are further divided into voiced and voiceless depending on the fact whether the vocal chords start vibrating during the closure (Ladefoged and Johnson 2010, 14–18; 148–157).

3.2.2 Language mode

As it was already mentioned, another common indicator for code-switching studies is the employment of language modes. Language mode could be defined as "the state of activation of the bilingual's languages and language processing mechanisms at a given point of time" (Grosjean and Li 2013, 15). Grosjean and Li argue that when bilinguals find themselves in a situation requiring use of language such as (conversation, reading, writing, etc.) they face (usually subconsciously) two decisions. Firstly, they must determine which one of the languages they speak will be used as a base language – a language which will be fully activated during the situation. Further, they decide whether to involve also the other language[7]. Based on the latter decision, the other language is either activated to some extent as well, or it is not engaged. When there is only one active language, then the mode in question is called monolingual mode. In accordance, the activation of both languages (at least to some extent) denotes bilingual mode. It is necessary to point out that the state of activation forms rather a scale than representing only a choice between two possibilities (either only one active or two fully active). Both bilinguals' degree of activation of the language and their base language can change very quickly as a response to the development of the situation (Grosjean and Li 2013, 14–21).

3.2.2.1 Code-switching

Code-switching, as an example of bilingual mode situation, could be described as "shifting (switching) completely to the other language for a word, a phrase, a sentence, etc." (Grosjean 1995, 263). According to Grosjean and Li, there are various reasons why the bilinguals code-switch such as using a precise word (which might not exist in the other language), involving someone into conversation, or even making a good impression. In addition, switching might be also initiated by the so called "complementarity principle". This principle states that the more thematic areas a language covers, the more often it will be used and the more proficient the speaker will be in this language. On the other hand, if a language is used only to talk about a very few topics, there is a high chance that the speaker will not reach a high level of fluency. Therefore, the speaker might feel the need to switch to a language, in which he or she

[7] For the purpose of clarity, only two languages will be discussed, however, according to Grosjean and Li, the number of languages involved can be (arbitrarily) higher (Grosjean and Li 2013, 14–21).

obtains the necessary knowledge (vocabulary) on the concerned thematic area (Grosjean and Li 2013, 11–21).

3.2.2.2 Interpreting

Interpreting is an activity requiring a full activation of both speakers' languages (Grosjean and Li 2013, 16). There are two basic types of interpreting are distinguished, simultaneous and consecutive. In simultaneous interpreting, the utterance being heard is transferred by the interpreter at the same time, whereas in the consecutive interpreting, the interpreter first hears the utterance and then renders it. As Christoffels and De Groot state, both of these types pose demands on the interpreter, the simultaneous increases strain on cognitive processing and the consecutive on the long-term memory. (Christoffels and De Groot 2005, 454–479). Gile (2001) described interpreting in terms of a cognitive load which it imposes on the interpreter. He hypothesized that when simultaneously interpreting, the speaker has to exert 3 types of effort: listening and analysis effort (perception), production effort and short-term memory effort. The situation concerning consecutive interpreting slightly differs, the effort is exerted "merely" on production (taking notes and rendering them) and on the long-term memory. All of these efforts draw from a processing capacity, which is limited. Once it is spent, it leads to a deterioration of interpreter's performance. Gile (1999) further proposed that during the process, the capacity used by interpreters is almost the total capacity they possess altogether. He tested this so called "tightrope hypothesis" by asking professional interpreters to render a speech segment twice in a row. The speech was then assess as regards errors and omissions made in case of segments which do not pose any special transfer problems. Gile presupposed that if the problems at these places are made only a few of the interpreters, it would indicate that these mistakes might be caused by the cognitive load. He also argued that if the subjects make a mistake in the second interpretation, although its translation was correct the in the first instance, the result might be attributable to a cognitive load. Approximately half of the mistakes were made by only one or two interpreters in the first session. Also, the majority of subjects made a new mistake in the second interpretation. The obtained results supported the Gile's claim, however, as Gile adds, more elaborate methods need to be applied for further research on this subject.

3.2.2.2.1 Cognitive load

An idea of a cognitive load was further discussed by Yap et al. (2015), who defined this concept as "[a] mental demand experienced while performing a cognitive task" (Yap et al. 2015, 74). In their research, analogically as Gile (2001), they assumed that the working memory capacity is limited and when performing a cognitive task, the load on the memory increases. When the capacity is not sufficient, the task will deteriorate. In their own study, they observed the effects of the cognitive load on a speech production, namely on the voice source, with the aim to create a database serving as a measurement system of a cognitive load degree. They recorded participants during performing various cognitive tasks. For example, in one of the series of experiments, the participants were presented with words whose meaning referred to colors. The words also had a colored font. Participants were asked to say the color of the font. The words were presented at a different speed, the faster speed, the higher cognitive load. In another experiment, the participants determined, whether the sentence they were presented with was logical or not, while at the same time they were supposed to memorize letters. The overall results of the experiments were that a higher cognitive load leads to the narrower flow of the air through the vocal folds.

Since interpreting is an activity which places demands on cognitive processing – the interpreter needs to remember information, to put effort into comprehension, translate, it can be expected that the cognitive load will affect the speech production of interpreters (Christoffels 2004).

3.2.3 Bilingual speech production

As it was outlined above, when in the bilingual mode, the speakers have activated more than one of their languages (Grosjean and Li 2013, 14–21). The issue how this affect the production of the bilingual speakers has been discussed by De Bot (1992), who devised a comprehensive bilingual speech production model. This model has been construed by adapting Levelt's speech production model for monolingual speakers.

3.2.3.1 Levelt's speech production model for a monolingual speaker – a starting point

As Kormos (2006) and De Bot (1992) describe, Levelt's model consists of a range of processing components, namely a knowledge component (general knowledge of the world, discourse knowledge) a conceptualizer, a formulator, an articulator and

a speech-comprehension and an auditory system. Each of these components focuses on performing a particular task and is rather independent. The processing is parallel (components can process different inputs at the same time) and mostly automatic.

Firstly, information which speaker wishes to communicate is chosen in the conceptualizer. This is realized in two steps, macro- and microplanning. Macroplanning consists in choosing the communicative aim and accessing the information needed for its rendering, whereas microplanning leads to determining of the speaker's approach to the communicative aim (choices such as whether to use declarative/imperative mode, stipulating agent of the sentence). The result of these two processes is a preverbal message (plan) which does not have a linguistic form yet.

This message is then subjected to grammatical encoding which takes place in the formulator. In this phase the relevant lexical items are accessed from the mental lexicon of the speaker. This happens in two steps according to the division of lexical items into two constituents – a lemma and a lexeme. Firstly the lemmas (carrying syntactic and conceptual information) the meaning which of corresponds to the preverbal plan are activated, resulting in the formation of the surface structure. The phase which follows is called morpho-phonological encoding in which the lexemes (carrying morpho-phonological information) are gained. Then, by means of phonological encoding, the individual phonemes are chosen, leading to the creation of a phonological score. The last process taking place in the formulator is the phonetic encoding during which the articulatory patterns are retrieved and the phonetic plan (score) emerges. The plan is later transformed into speech in the articulator.

There is also a speech comprehension system (connected to an auditory system) sending information to a monitor (located in conceptualizer). By its means it is possible to receive information on the state of speech production. There are three loops connected to the monitor. First loop controls whether the preverbal plan and speakers aim are matching, second oversees the phonetic plan and the last monitors the production. (Kormos 2006, 7–11; De Bot 1992).

3.2.3.2 De Bot's adaptations of the model for a bilingual speaker

De Bot (1992) suggested changes to the Levelt's model, to make it suitable even for speech production of bilingual speakers. He investigates individual

components and argues whether they can remain as in the original model or need to be duplicated.

As stated above, the process starts in a conceptualizer which connected to a knowledge component. De Bot concluded that the information specific to the other language (L2) are simply added to the knowledge component, therefore there is no need for its (L2) copy. Identically, he assumes that there is a single conceptualizer, however, the second of the processes taking place within this component – microplanning – is language dependent. To summarize, in the first process called macroplanning, the language in which the speech unit will be articulated (base language) is determined. Subsequently, an encoding specific for each language occurs during a microplanning.

As regards storing lexical items, De Bot assumes that there is one mental lexicon containing items from all bilingual's languages. The individual items are inter-connected and clustered into subsets according to the strength of the links among them (subsets corresponding to individual languages). Another question De Bot asks is how the items in the lexicon are accessed. He sides with the idea that the lexical items have certain describing features and that the preverbal message is constantly examined for occurrence of these features. If a match is found, the item becomes active. The item with the highest number of corresponding features (the most active item) is then retrieved from the lexicon. As it was already said, in the conceptualizer, the language of the utterance is chosen. However, since bilinguals can activate more languages simultaneously (bilingual mode), De Bot presumes that two (possibly even more based on number of active languages) phonetic plans are created, although only the plan for language chosen in the macroplanning phase is articulated. Therefore, there are separate formulators for each language.

De Bot further adopted Levelt's model of the inner representation of sounds that is based on the perception of speaker's own speech. The bilingual then must have representations for all sounds of the languages he or she speaks, however, the sounds same in the two languages will be contained within a one set. In Levelt's model, the fundamental units of articulation process are however the syllables, for which there are so called syllable programs needed for their articulation. As regards the articulator, De Bot presumes that it is a single component common for all bilingual's languages,

since even in the speech of proficient speakers of L2 can be observed prosodic patterns of their native language.

3.2.3.3 Kormos' adaptations of the model for a bilingual speaker

Identically as De Bot (1992), Kormos (2006, 167–181) introduced a bilingual speech production model which is built on Levelt's model for a monolingual speaker.

Unlike in both the original model and De Bot's version where the knowledge stores are placed separately, Kormos argues that there is one information component called a long-term memory which comprises of four constituents: an episodic memory (experiences), a semantic memory (conceptual information, with a mental lexicon), syllabary (syllable programs) and a "store for declarative knowledge" containing non-automatized syntactic and phonological rules. This composite knowledge component is common for all bilingual's languages.

In accordance with De Bot's model, in the conceptualizer, the choice of the concepts takes place and the language of the utterance is determined. In Kormos' view, the concepts are "conglomerate[s] of interrelated memory traces consisting of information concerning word meaning" (Kormos 2006, 169). The language choice is carried out by assigning "language cues" to the individual concepts. In case that a concept overlaps in two languages (or more), only one concept in a chosen language is further processed. In the following phase, lexical encoding, concepts are matched with lexical items in an analogical way as described by De Bot (1992). According to Kormos, syntactic encoding proceeds in the same manner as in the monolingual speech production. The syntactic characteristics of a lexical item are activated. At this point, it might happen that less proficient L2 speakers will retrieve syntactic characteristics of a matching L1 lexical item. In the next step, a structure and order of the utterances is composed, which is for proficient speakers automatic, however less proficient learners might need to access the declarative knowledge store. The subsequent process is the phonological encoding. The author presumes that phonological information of lexical items of the other language are available as well and also that the activation in both direction from lemma to lexeme and from lexeme to lemma is possible.

The process of phonetic encoding is no different from the monolingual model (with the exception that bilinguals have articulatory patterns for syllables of all their languages). Only less proficient speakers might need to draw rules from declarative

knowledge store as in the syntactic processing. Finally, the model borrows Levelt's control mechanism, it however stipulates that this mechanism is not automatic when using the L2.

Both of these bilingual models account for code-switching. There is an agreement that the language choice is done in conceptualization phase. Kormos believes that it is executed by adding a language cue to the concept. On the other hand, De Bot suggests that two (or more according to the number of bilingual's activated languages) phonetic plans are created and processed, whereas only one only one is articulated.

4 Experiment

4.1 The foundation of the current study

The discussed experiment is an offshoot of an experiment conducted by Šimáčková and Podlipský (2015). The authors investigated speech production of voiceless stops by native Czech speakers of English. The participants were students of English with a focus on interpreting. All were rather proficient in L2 English, however half of them had significantly more experience with interpreting than the other. The experimenters observed their speech in three different situations: speaking English (in the monolingual mode), code-switching and interpreting. The aim of the research was to explore whether there will be a (more prominent) influence of L1 Czech on L2 English in the two bilingual situations in comparison to the monolingual. Assuming that the L1 affect appears, the further question was under which circumstance it will be the strongest. According to the experimenters' expectations, the results revealed that the VOT values of voiceless stops were lower in two bilingual sessions than in the monolingual session, indicating a greater L1 influence. The difference between the bilingual situations was rather insignificant, although the interpreting showed slightly higher VOT values (more English-like) than in the code-switching. Further, unlike in case of unexperienced students of interpreting, the speech production of experienced students under bilingual condition did not change in comparison to their monolingual mode production.

In addition to the VOT data analysis, as a part of this experiment was conducted also an analysis of formant (F2) height of English vowel /u/ by Lucia Šmelíková (2015). The collected data was common for both of these analyses, however this paper compared only the monolingual condition to the code-switching condition with identical aim – to investigate the extent of L1 influence on L2 speech production. That was possible due to differences between in the height of Czech and English /u/ vowel formants (namely F2). The results showed that when the participants were in the bilingual mode, their F2 values decreased in comparison to their monolingual mode production, showing a greater L1 interference.

4.2 Current experiment

The purpose of this study is to replicate experiment conducted by Šimáčková and Podlipský (2015) and further explore influence of L1 on L2 in the three stated language

modes. In the original research, the greater interference was observed in the two bilingual modes (code-switching and interpreting). Further, the distinction between these modes showed a small but rather insignificant increase of the VOT in the interpreting condition. This study aims to investigate the relevance of this increase by adopting a slight alternation of the procedure.

4.2.1 Participants

4.2.1.1 Bilingual speakers

Participants[8] of the experiment were 14 students of the 3rd year of English with a specialization in translation and interpreting at Palacký University in Olomouc. All were native speakers of Czech who had English as an L2. The group of 14 people consisted of 5 men and 9 women. Their age ranged between 21–24 years (the mean value in round numbers was 22 years). None of the participants spoke fluently more than one other foreign language apart from English. They started to learn English at school when their mean age was 9 years. The majority (roughly 64%) had not been taught English before, and those who had, had begun at 5-7 years of their age. Approximately 36 % have attended a grammar school with a special focus on language instruction or a bilingual school and a half received some kind of additional instruction in English. None of the participants spent a significant amount of time in English speaking country before or after the age of 15 (the only exception is a person who has spent a year in an English speaking country after the age of 15). The mean age when the participants started feeling comfortable using English is 16. The participants' evaluation of their accent in English was rather diverse, taken on average, it could be described as moderate. Apart from the university environment, the majority of participants reported not to communicate with native speakers of English more than two times a month, however the most of them claimed to be exposed to English speaking media on a daily basis. In addition, they all passed a practical English exam level C1 stipulated by CEFR (Verhelst et. al. 2009).

4.2.1.2 Native speakers

The stimuli presented to the bilingual participants were recorded by native speakers of Czech and English already for the purposes of the original experiment. Further, their

[8] The information about participants were collected by means of a questionnaire which was borrowed from the original research (Šimáčková and Podlipský 2015). Its answers are summarized in the tables presented in appendix.

VOT values were measured and used as a reference point for comparison with the values produced by the bilinguals. As Šimáčková and Podlipský (2015) state, the native speakers of Czech were 3 women and 2 men and their age ranged between 21 and 25 years. The group of native speakers of English was represented by 2 women and 3 men. Their age ranged between 24 and 54 years and both speakers of American and British English were included.

4.2.2 Procedure

The participants were invited to a sound studio for 2 recording sessions. For each session, the participants were asked to repeat 88 auditorily presented sentences. The stimuli contained 6 occurrences of words with initial voiceless stop/p/ and 6 occurrences of words with initial voiceless stop /t/ in both initial and final position in the utterance – i.e. the total of 24 sentences. (The rest of the sentences were collected for other experimental purposes and served only as fillers.). The stimuli were identical as in the original experiment conducted by Šimáčková and Podlipský (2015).

The first session which was the bilingual session took place within two days and included both code-switching and interpreting part. The bilingual session was recorded by a native Czech speaker of English, both Czech and English were used when giving instructions, although Czech prevailed. Between the two parts of the bilingual session, the participants had either 20 minute-long break outside the studio (meanwhile another participant was recorded) or 5 minute-long break in the studio where they were presented with a silent short movie. Six people firstly performed code-switching part and eight people started with the interpreting. Further, before the interpreting condition, participants had to complete a short priming exercise.

In the code-switching task, the participants heard an English sentence e.g. "Pubs were open all night" followed by a question in Czech asking about the sentence "*Co jsi slyšel?*" The participants then simply repeated the English sentence while using an introductory phrase in Czech "*Slyšel jsem…*" In the interpreting task, the participants were presented with a Czech sentence e.g. "Vepřové je nezdravé" followed by English question "*What should you say*?" The participants then answered with an introductory English phrase "*I should say*" and provided the translation of the Czech sentence "Pork is unhealthy". The interpreting part differs from the original experiment in a small but important detail. In Šimáčková and Podlipský (2015), the

translated English sentence was preceded by a Czech introductory phrase "*Teď řeknu*", whereas in the current experiment the code-switch was left out and replaced by an equivalent English introductory phrase "*I should say*".

The two-day long monolingual session took place approximately 3 months after the bilingual session. It was recorded by a native Czech bilingual speaker experienced in L2 (a teacher at Palacký University). The only language used during the recording process was English. The bilinguals were presented with English sentences, e.g. "Toys for boys and girls" followed by a question "*What should you say*". Then they said introductory phrase "*I should say*" and repeated the sentence.

4.2.3 Results

The recordings were acoustically analysed in the software Praat (Boersma and Weenink 2015) VOT of *p* and *t* was measured from the release of the closure to the start of periodicity of the following vowel (to the first "formant column"). Further, both the target word duration and sentence duration were annotated as well. Duration of sentences containing pauses was excluded from further analysis.

4.2.3.1 Bilingual speakers vs. native speakers

The results showed that when bilinguals spoke only English, their mean VOT value of *p* was 37.0 ms and in case of *t* reached 59.3 ms, whereas reference native speakers of English produced *p* and *t* with mean values 53.1 ms and 68.5 ms respectively. In addition, there was a great diversity in the *p* and *t* production between individual bilingual speakers. For illustration, Figures 1 and 2 show the mean values of both bilingual and native speakers.

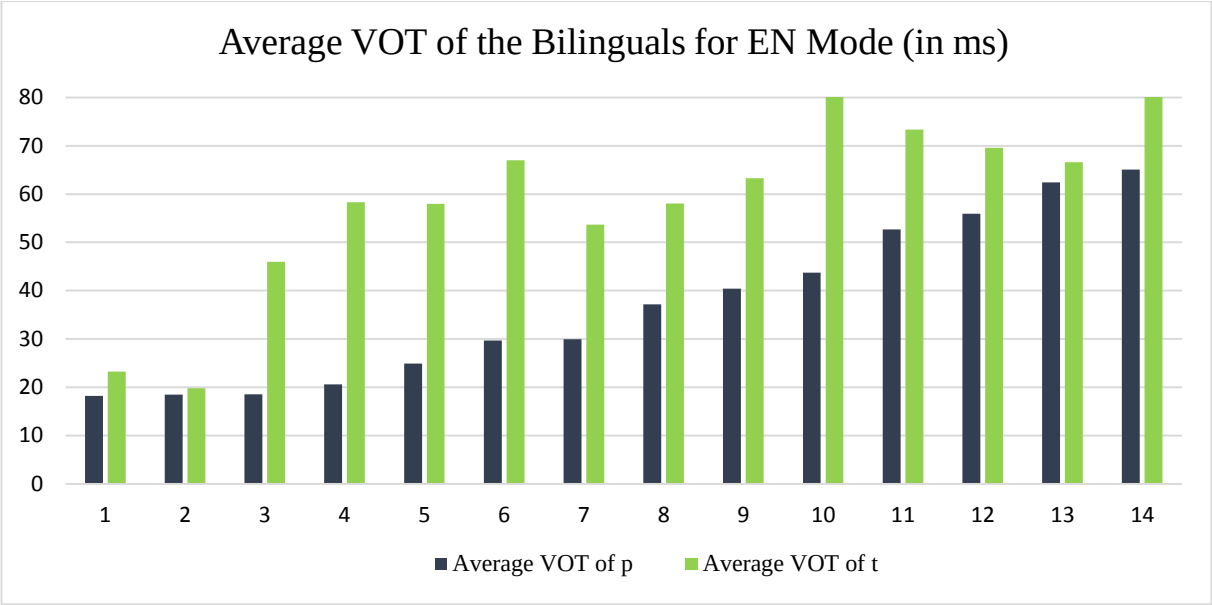

Figure 1. The mean VOT of individual bilinguals in the monolingual English mode. The numbers 1 –14 on the horizontal axis refer to individual participants (e.g. Participant 1). The vertical axis shows VOT values in milliseconds.

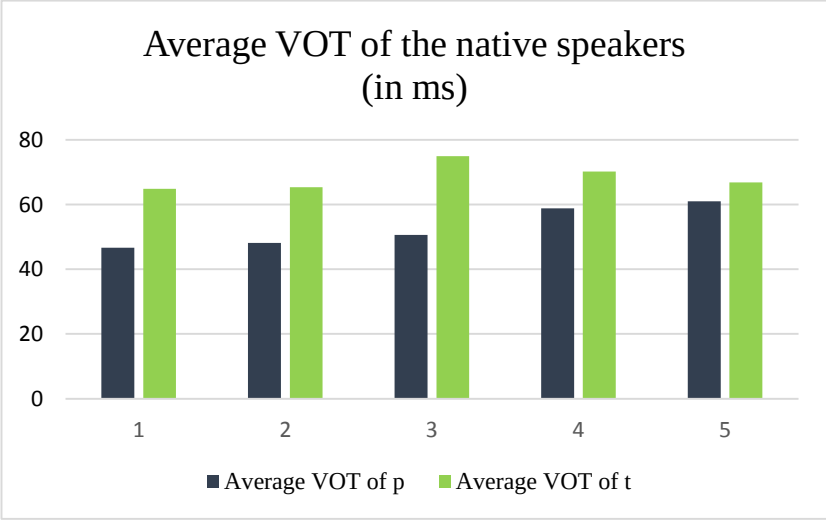

Figure 2. The mean VOT values of individual reference native speakers of English. The numbers 1–5 on the horizontal axis refer to individual speakers (e.g. Speaker 1). The vertical axis shows VOT values in milliseconds.

Both native and bilingual speakers produced *t* with higher VOT values than *p*, corresponding with a fact that VOT varies with place of articulation (Cho and Ladefoged 1999, Lisker and Abramson 1964). However, the above presented results do not account for the impact speaking rate. As Kessinger and Blumenstein state, the VOT can be further influenced by the rate of speech, moreover, in English, the VOT grows with slower rate (Kessinger and Blumenstein 1998).

Therefore the average sentence duration of both native speakers and bilinguals was calculated. It was revealed that the native speakers produced sentences with average length of 1196.54 ms, whereas the bilinguals with 1261.31 ms. Repeated Measures ANOVA showed that the effect of Speaker Group on sentence duration did not reach significance (F [1, 4] = .49307, *p*=.52127). Further, VOT used for the following analysis was so called "tempo-normalized", which can be defined as "the ratio of VOT duration to the duration of the given word" (Šimáčková and Podlipský 2016, 3). By applying this "tempo-normalized" VOT, a target word duration was also taken into consideration.

4.2.3.2 Bilingual speakers' production in the three modes

The Repeated Measures ANOVA was used to explore bilingual speakers' production in the three examined modes. Firstly, the effect of mode on sentence duration was investigated. The results showed that it did not reach significance (F [2, 26] = 2.6954, p = .08634), yet in comparison to code switching and English condition, the sentence duration in interpreting was to some extent longer.

The normalized VOT values were submitted to a Repeated Measures ANOVA with Place of Articulation (labial, alveolar) and Mode (English only, Interpreting, Code-switching). The test revealed a significant effect of Place (F [1, 14] = 23.672, p = .00025). It was observed, that the speakers produced the lowest mean VOT in the monolingual English mode and the highest in the interpreting condition, however, the effect of Mode only approached the level of significance &=.05 (F [2, 28] = 2.7640, p = .08026). The interaction of Place and Mode reached significance (F [2, 28] = 7.1358, p =.00313), therefore a Tukey's HSD post hoc test was conducted, which confirmed the overall statistically relevant interaction of Place and Mode ($p < 0.01$) with only two exceptions. The first exception was that the difference between *p* in the interpreting and *p* in the code-switching condition proved to be insignificant. Further, *t* in the English only and *t* in the code-switching condition showed no relevant variation. For illustration, see the Figure 3 below.

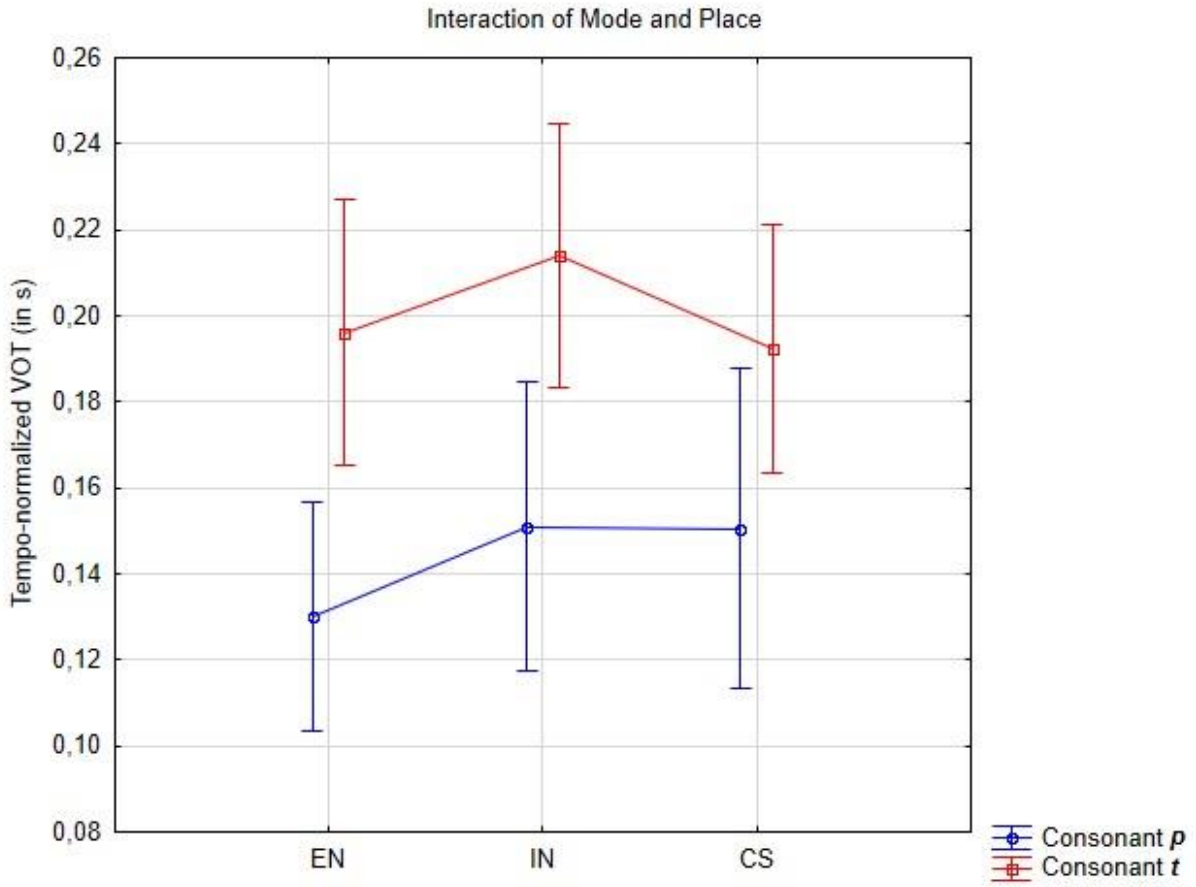

Figure 3. The following graph shows VOT of consonants p and t produced by the bilinguals in the individual conditions. The headings EN, IN and CS on the horizontal axis denote English-only, interpreting and code-switching condition, respectively. The vertical axis shows the "tempo-normalized" VOT in seconds. The bars denote 0.95 confidence intervals.

As it is apparent from the graph, the VOT in the interpreting condition was higher than in the monolingual English mode for both consonants. However, the increase did not concern all the participants, even the opposite trend could be observed, as illustrated in the Table 1.

Participant	t (EN-IN)	p (EN-IN)
P1	-8	19
P2	-16	-2
P3	-59	-2
P4	-33	-33
P5	-19	7
P6	-85	-79
P7	65	30
P8	-24	-48
P9	-42	-50
P10	-41	-68
P11	31	-12
P12	17	1
P13	-40	-37
P14	7	-18

Table 1. The table shows difference in milliseconds between "tempo-normalized" VOT in the English condition and interpreting condition for production of both consonants.

4.2.4 Discussion

Although the difference between the VOT of *t* in the English condition and *t* in the code-switching condition proved to be insignificant, the overall results show that the bilinguals in the monolingual English mode produced examined voiceless stops with shorter VOT than in the bilingual mode. Therefore, it might be said that the fact mode (monolingual or bilingual) influenced their speech production. However, the observed effect is in contradiction the original study conducted by Šimáčková and Podlipský (2015), where the bilinguals' VOT of voiceless stops declined in the bilingual mode. It might be hypothesized that the final effect of the mode depends on a particular group of speakers and is highly individual. The speakers could initially decide what approach to take in the conceptualizer, where according to both presented bilingual models, the language of the utterance is decided (Kormos 2006, De Bot 1992). Further, as regards the comparison of interpreting and code-switching condition, there is not a conclusive distinction, since *p* was produced with similar values in these two tasks, but *t* showed a longer VOT in the interpreting condition.

5 Conclusion

The thesis has dealt with the language influence manifested in speech production of bilingual speakers. Firstly, literature review provided insight into this issue. It showed that the influence can be observed in both L1 to L2 and L2 to L1 direction. The following chapter looked into the cause of foreign accent occurrence and its degree. It divided the experimental studies into two types – studies investigating accent in terms of the language competence and studies exploring speakers' performance. Further, a parameter commonly used for the evaluation of the language influence, VOT, was introduced. Then the Grosjean and Li's theory (Grosjean and Li 2013) based on that bilinguals' state of activation of languages they speak (language mode) affects their language performance was discussed. In the final paragraphs of the theoretical part, possible course of the speech production process was characterized. The practical part investigated speech production of native Czech students of English with specialization in translation and interpreting under three conditions. Out of these three conditions, one represented a monolingual language mode (speaking English) and the two remaining (code-switching and interpreting) were examples of a bilingual mode. The study replicated an experiment conducted by Šimáčková and Podlipský (2015) in which it was observed that bilingual speakers produced voiceless stops *p* and *t* with more native-like values in the monolingual mode than in the two examples of the bilingual mode. It was also revealed that the difference between the code-switching and interpreting was insignificant. However, the interpreting was preceded by a code-switched phrase which might have affected the result. By removing the code-switch, the current experiment aimed to further explore the behavior in these two conditions. Nevertheless, the results showed a contradictory trend. Although there was to some extent an effect of the mode on the speech production, instead of producing higher VOT values in the monolingual mode, the students VOT reached greater values in the two bilingual modes. When comparing the code-switching and interpreting condition, consonant *p* was produced in the same manner in both bilingual modes, whereas VOT of consonant *t* was higher in the interpreting task. It might be hypothesized that language mode indeed has an effect on the speech production, however, the direction of the effect is dependent on the particular group of participants.

6 Resumé

Tato bakalářská práce se zaměřila na produkci řeči studentů třetího ročníku Angličtiny se zaměřením na komunitní tlumočení a překlad. Jednalo se o rodilé mluvčí češtiny, kteří dosáhli pokročilé znalostní úrovně angličtiny. Avšak i zkušení nerodilí mluvčí často mluví s rozpoznatelným cizím přízvukem (Grosjean a Li 2013, 7; Abrahamsson a Hyltenstam 2009), proto se dalo očekávat, že ani tito studenti netvoří žádnou výjimku.

Jak Grosjean a Li (2013, 14–21) uvádějí, bilingvní mluvčí musí každodenně učinit (často podvědomě) několik rozhodnutí týkajících se užití jazyka. Na základě těchto voleb je pak zvolen jazykový mód, který představuje míru aktivace jednotlivých jazyků pro danou komunikační událost. Monolingvní mód značí, že pouze jeden jazyk je aktivován, bilingvní předpokládá aktivaci dvou či více jazyků.

Dopad míry aktivace jazyků na produkci řeči v současné době zkoumá řada studií (Bullock et al. 2006, Antoniou a kol. 2011, Simonet 2014). K těm lze zařadit i experiment, který podnikli Šimáčková a Podlipský (2015), v němž byl zkoumán vliv rodného jazyka češtiny na další osvojený jazyk, angličtinu, za třech podmínek lišících se již zmíněnou mírou aktivace jazyků (mluvení anglicky, přepínání kódů a tlumočení z češtiny do angličtiny).

Bylo zjištěno, že aktivace obou jazyků při tlumočení a přepínání kódů vedla k větší interferenci – vlivu jednoho jazyka na druhý (Simonet 2014). Autoři předpokládali, že při tlumočení se kvůli kognitivní zátěži daný vliv projeví ještě více než při přepínání kódů, avšak porovnání mezi těmito situacemi ukázalo jen nevýznamně větší interferenci při přepínání kódů.

Cílem této studie bylo zopakovat tento experiment za pozměnění metodologie, a tak více prozkoumat rozdíl mezi tlumočením a přepínáním kódů. V souladu s původním experimentem byl daný jev zkoumaný pomocí doby nástupu hlasivkového tónu neznělých ploziv *p* a *t*. Získané výsledky ukázaly, že studenti sice přizpůsobili produkci řeči podle jazykového módu, ale jejich výslovnost se přiblížila více rodilým mluvčím v bilingvních módech než v módu monolingvním, jednalo se tedy o trend zcela opačný než v původní studii. Možným vysvětlením je, že jazykový mód sice ovlivní produkci řeči, avšak to, jakým způsobem se bude odvíjet od konkrétní skupiny mluvčích. Dále bylo pozorováno, že produkce *p* zůstala doba hlasivkového tónu

přibližně stejná v obou bilingvních situacích, kdežto *t* zaznamenalo nárůst při tlumočení (a tím pádem se přiblížilo *t* produkovanému rodilými mluvčími).

Výzkum je podrobněji popsaný v praktické části práce, kterou předchází teoretický základ, v němž je představen dosavadní výzkum dané problematiky, možné příčiny cizího přízvuku a již zmíněné okamžité faktory ovlivňující jeho míru (komunikační situace). Dále teoretická část informuje o použitém parametru pro změření jazykového vlivu – doby nástupu hlasivkového tónu – a popisuje bilingvní proces produkce řeči.

7 References

Abrahamsson, Niclas, and Kenneth Hyltenstam. 2009. "Age of Onset and Nativelikeness in a Second Language: Listener Perception Versus Linguistic Scrutiny." *Language Learning* 59(2):249–306. doi:10.1111/j.1467-9922.2009.00507.x.

Antoniou, Mark, Catherine T. Best, Michael D. Tyler, and Christian Kroos. 2010. "Language context elicits native-like stop voicing in early bilinguals' productions in both L1 and L2." *Journal of Phonetics* 38(4):640–653. doi:10.1016/j.wocn.2010.09.005.

Antoniou, Mark, Catherine T. Best, Michael D. Tyler, and Christian Kroos. 2011. "Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching." *Journal of Phonetics* 39(4):558–570. doi:10.1016/j.wocn.2011.03.001.

Balukas, Coleen, and Christian Koops. 2014. "Spanish-English bilingual voice onset time in spontaneous code-switching." *International Journal of Bilingualism* 20(10):1–21. doi: 10.1177/1367006913516035.

Boersma, Paul, and David Weenink. 2015. "Praat." Version 6.0.04, retrieved November 15, 2015, http://www.fon.hum.uva.nl/praat/

Bullock, Barbara E., Almeida Jacqueline Toribio Verónica González, and Amanda Dalola. 2006. "Language Dominance and Performance Outcomes in Bilingual Pronunciation." In proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006), Banff, Canada, April 27–30.

Cho, Taehong, and Peter Ladefoged. 1999. "Variation and universals in VOT: Evidence from 18 languages." *Journal of Phonetics* 27(2):207 –229. doi:10.1006/jpho.1999.0094.

Chomsky, Noam. 1969. "Methodological Preliminaries." In *Aspects of the theory of syntax*, 3–63. Cambridge: MIT Press.

Christoffels, I. K. 2004. "Cognitive studies in simultaneous interpreting." PhD thesis, University of Amsterdam.

Christoffels, Ingrid. K., and Annette M. B. De Groot. 2005. "Simultaneous Interpreting: A Cognitive Perspective." In *Handbook of Bilingualism:*

Psycholinguistic Approach, edited by Judith F. Kroll and Annette M. B. De Groot, 454–479. Oxford: Oxford University Press.

De Bot, Kees. 1992. "A Bilingual Production Model: Levelt's Speaking Model Adapted." *Applied Linguistics* 13(1):1–24. doi: 10.1093/applin/13.1.1.

Flege, James Emil. 1981. "The Phonological Basis of Foreign Accent: A Hypothesis" *Tesol Quarterly* 15(4):443–455. doi:10.2307/3586485.

Flege, James Emil. 1987. "The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification." *Journal of Phonetics* 15: 47–65.

Flege, James Emil, and Wieke Eefting. 1987a. "Cross-language switching in stop consonant perception and production by Dutch speakers of English." *Speech Communication* 6:185–202. doi:10.1016/0167-6393(87)90025-2.

Flege, James Emil, and Wieke Eefting. 1987b. "Production and perception of English Stops by native Spanish speakers." *Journal of Phonetics* 15:67–83.

Flege, James Emil, and James Hillenbrand. 1984. "Limits on phonetic accuracy in foreign language speech production." *Journal of the Acoustical Society of America* 76(3):708–721. doi: 10.1121/1.391257.

Flege, James Emil, Grace H. Yeni-Komshian, and Serena Liu. 1999. "Age Constraints on a Second-Language Acquisition." *Journal of Memory and Language* 41(1):78–104. doi:10.1006/jmla.1999.2638.

Fowler, Carol A., Valery Sramko, David J. Ostry, Sarah A. Rowland, and Pierre Hallé. 2008. "Cross language phonetic influences on the speech of French-English bilinguals." *Journal of Phonetics* 36(4):649–663. doi:10.1016/j.wocn.2008.04.001.

Gile, Daniel. 1999. "Testing the Effort Models' tightrope hypothesis in simultaneous interpreting: A contribution." *Hermes* 23:153–172.

Gile, Daniel. 2001. "Consecutive vs. simultaneous: Which is more accurate?" *The Journal of the Japan Association for Interpretation Studies* 1:8–20.

Grosjean, François. 1995. "A psycholinguistic approach to code-switching: the recognition of guest words by bilinguals." In *One speaker, two languages: Cross-disciplinary perspectives on code-switching,* edited by Milroy, Lesley, and Pieter Muysken, 259–275. Cambridge: Cambridge University Press.

Grosjean, François, Ping Li, and Ellen Bialystok. 2013. *The psycholinguistics of bilingualism*. Chichester: Wiley-Blackwell.

Kessinger, Rachel H., and Sheila E. Blumenstein. 1998. “Effect of speaking rate on voice-onset time and vowel production: Some implications for perception studies.” *Journal of Phonetics* 26:117–128. doi:10.1006.jpho.1997.0069.

Kormos, Judit. 2006. *Speech production and second language acquisition*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Kuhl, Patricia K. 2004. “Early language acquisition: Cracking the speech code.” *Nature Reviews Neuroscience* 5(11): 831–843. doi:10.1038/nrn1533.

Ladefoged, Peter, and Keith Johnson. 2010. *A Course in Phonetics*. 6th ed. Canada: Wadsworth Cengage Learning.

Lisker, Leigh, and Arthur S. Abramson. 1964. “A Cross-Language Study of Voicing in Initial Stops: Acoustic Measurements.” *Word* 20 (3):384–422. doi: 0.1080/00437956.1964.11659830

López, Verónica Gonzáles. 2012. “Spanish and English word-initial voiceless stop production in code-switched vs. monolingual structures.” *Second Language Researc*h 28(2):243–263. doi:10.1177/0267658312439821.

MacLeod, Andrea A. N., Carol Stoel-Gammon, and Alicia B. Wassink. 2009. “Production of high vowels in Canadian English and Canadian French: A comparison of early bilingual and monolingual speakers.” *Journal of Phonetics* 37(4):374–387. doi: 10.1016/j.wocn.2009.07.001.

Olson, Daniel. 2013. “Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production.” *Journal of Phonetics* 41(6):407–420. doi:10.1016/j.wocn.2013.07.005.

Piccinini, Page, and Amalia Arvaniti. 2015. “Voice onset time in Spanish-English spontaneous code-switching.” *Journal of Phonetics* 52:121–137. doi:10.1016/j.wocn.2015.07.004

Piske, Thorsten, Ian R. A. MacKay and James E. Flege. 2001. “Factors affecting degree of foreign accent in an L2: a review.” *Journal of Phonetics* 29(2):191–215. doi:10.006/jpho.2001.0134.

Sancier, Michele L., and Carol A. Fowler. 1997. “Gestural drift in a bilingual speaker of Brazilian Portuguese and English.” *Journal of Phonetics* 25(4):421–436. doi: 10.1006/jpho.1997.0051.

Simonet, Miquel. 2014. “Phonetic consequences of dynamic cross-linguistic interference in proficient bilinguals.” *Journal of Phonetics* 43(1):26–37. doi:10.1016/j.wocn.2014.01.004.

Schwartz, Geoffrey, Anna Balas, and Arkadiusz Rojczyk. 2015. “Phonological factors affecting L1 phonetic realization of proficient Polish users of English.” *Research in Language* 13(2):181–198. ” doi: 10.1515/rela-2015-0014.

Šimáčková, Šárka, and Václav Jonáš Podlipský. 2015. “Immediate phonetic interference in code-switching and interpreting.” Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, UK: the University of Glasgow, August 10–14.

Šmelíková, Lucia. 2015. “Variation in Cross-Linguistic Interference in Monolingual and Bilingual Performance.” Bachelor’s thesis, Palacký University.

Verhelst, N., Van Avermaet, P., Takala, S., Figueras, N., and B. North. 2009. Common European Framework of Reference for Languages: learning, teaching, assessment. Cambridge: Cambridge University Press.

Wei, Li. 2000. *The bilingualism reader*. London: Routledge

Werker, Janet F., and Richard C. Tees. 2002. “Cross-language speech perception: Evidence for perceptual reorganization during the first year of life.” *Infant Behavior and Development* 25(1):121–133. doi:10.1016/S0163-6383(02)00093-0.

Yap, Tet Fei, Julien Epps, Eliathamby Ambikairajah, and Eric H. C. Choi. 2015. “Voice source under cognitive load: Effects and classification.” *Speech Communication* 72:74–95. doi: 10.1016/j.specom.2015.05.007

8 Appendix

	Age of onset of learning English at school	**Age of onset of learning English before school education**	**Attending a grammar school with extended language instruction**	**Additional instruction in an English language**
P1	9 years	7 years	attended	-
P2	15 years	-	-	individual lessons
P3	9 years	-	attended	-
P4	10 years	-	-	an US highschool (year), a language school (year)
P5	10 years	-	attended (a bilingual school)	a summer language school in Canada
P6	8 years	-	-	-
P7	10 years	7 years	attended (a bilingual school)	a language school, a private teacher
P8	10 years	5 years	-	a language school, a summer language school in Ireland
P9	7 years	5 years	-	a language school
P10	8 years	-	-	-
P11	6-7 years	5 years	attended	a language school
P12	10 years	-	-	-
P13	9 years	-	-	-
P14	9 years	-	-	-

(1) A table summarizing answers of the participants to questions concerning their language acquisition and experience. “P” stands for a Participant.

	Amount of time spent in an English speaking country before the age of 15 (months)	**Amount of time spent in an English speaking country after the age of 15 (months)**	**Frequency of exposure to the English speaking media (films/music/news/radio/others)**
P1	-	-	daily – 1 to 2 hours
P2	-	-	several hours per month
P3	-	-	1-2 hours per week
P4	-	12 months	daily – up to 1 hour
P5	2 months	-	daily – up to 1 hour
P6	-	-	more than 2 hours per day
P7	-	-	more than 2 hours per day
P8	1 month	3 months	daily – 1 to 2 hours
P9	-	3 months	more than 2 hours per day
P10	-	-	more than 2 hours per day
P11	-	-	several hours per month
P12	-	-	daily – 1 to 2 hours
P13	-	-	more than 2 hours per day
P14	-	-	several hours per week

(2) A table indicating exposure of the participants to English spoken by native speakers. “P” stands for a Participant.

	Age when the participants started feeling comfortable using English	**Degree of a foreign accent in English (based on self-evaluation)**	**Importance of having speech free of foreign accent for the interpreter according the participants (on a scale 1 to 9)**	**Importance of improving own pronunciation during studies according to participants (on a scale 1 to 9)**
P1	15	6. considerable	3	4
P2	21	5. moderate	5	6
P3	15	4. some	5	7
P4	17	5. moderate	7	6
P5	17	3. light	7	5
P6	16	5. moderate	2	3
P7	15	7. heavy	4	8
P8	14	4. some	7	7
P9	15	4. some	7	9
P10	11	6. considerable	7	7
P11	15	3. light	6	9
P12	17	2. almost none	7	8
P13	17	3. light	6	7
P14	16	7. heavy	7	9

(3) A table showing participants' self-evaluation of their accent and motivation to improve