

Univerzita Karlova v Praze  
Filozofická fakulta  
Ústav anglického jazyka a didaktiky

## **BAKALÁŘSKÁ PRÁCE**

Dominika Ortutayová

**Asimilace znělosti u slovenských mluvčích angličtiny**

**Assimilation of voicing in Slovak speakers of English**

## **Acknowledgments**

I would like to express my very great appreciation to doc. Mgr. Radek Skarnitzl, Ph.D., my research supervisor, for patient guidance and for providing me with constructive and useful suggestions concerning the progress of my research work. His effort and advice have been immensely appreciated.

Simultaneously, I am particularly grateful to all members of the English-Studies Department who have supervised my previous written assignments and thus not only provided me with invaluable amount of experience but also encouraged me to grow stylistically.

## **Declaration of authorship**

Hereby I declare that the following BA thesis is my own work for which I used only the sources and literature mentioned.

In Prague: 23 May 2014

Signature

## **Abstract**

The aim of this thesis is to examine the influence of the mother tongue, more specifically Slovak, in the speech of the Slovak speakers of English and subsequent appearance of the assimilation of voicing in the environments in which it is expected in both Slovak and English and then in those which are typical only for Slovak - and the production of the assimilation of voicing in those environments in English is considered an error. In the first part we introduce and specify the process of voicing in speech in general and then in the second part we proceed with the presentation of the concept of the second language acquisition and language interference. The empirical part is based on the analysis of the material spoken by 18 students, both males and females, reading the BBC bulletins. Our results show that the assimilation of voicing is an integral part of the speech of Slovaks in English, proving that they assimilate extensively in the environments typical for the Slovak language. Those are only partially typical for English as well – and thus they simultaneously produce assimilation errors. This knowledge could be taken in account during the learning process in order to eliminate such errors as much as possible.

**Key words:** voicing, assimilation, Slovak, English, interference

## **Abstrakt**

Cieľom tejto bakalárskej práce je skúmať vplyv materinského jazyka u Slovákov hovoriacich anglicky na angličtinu. Sledujeme výskyt asimilácie znelosti ako produkt jazykovej interferencie, ktorá sa objavuje v miestach typických pre oba jazyky, avšak zaznamenávame aj výskyt asimilácie znelosti v miestach, ktoré sú typické iba pre slovenčinu a prispôsobovanie znelosti v týchto okoliach je v angličtine považované za chybu. V prvej časti práce sa venujeme všeobecnému predstaveniu javu znelosti a jeho korelátov, v časti druhej sa zaoberáme javom jazykovej interferencie a konceptom osvojovania si cudzieho jazyka. V praktickej časti analyzujeme zvukový materiál poskytnutý osemnástimi študentmi, mužmi aj ženami, ktorí mali za úlohu prečítať správy BBC. Výsledky nám ukazujú, že asimilácia znelosti je naozaj základnou súčasťou reči Slovákov, ktorí v nahrávkach výrazne asimilujú tam, kde je to správne pre oba jazyky, ale aj tam, kde je to v angličtine považované za chybu. Znalosť takejto chyby nám umožňuje ju vziať do úvahy počas procesu osvojovania si angličtiny a tak ju do najvyššej možnej miery eliminovať.

**Kľúčové slová:** znelosť, asimilácia, slovenčina, angličtina, interferencia

## List of tables

<b>Figure 2</b>	“Glottal views at the phonatory position immediately before a) and after b) the onset of vocal fold vibration.” (Sawashima, 1983: 15).....	<b>13</b>
<b>Figure 2.1</b>	“Voicing profiles of the closure phases of German stop consonants, pooled across all left and right segmental contexts.” (Möbius, 2004: 12).....	<b>17</b>
<b>Figure 2.2</b>	“Voicing profiles of German fricatives, pooled across all left and right segmental contexts.” (Möbius, 2004: 14).....	<b>19</b>
<b>Figure 5</b>	Labelling the target environment in Praat.....	<b>33</b>
<b>Figure 6</b>	The distribution of voicing ratio for both voiced and voiceless following segments.....	<b>36</b>
<b>Figure 6.1</b>	The distribution of voicing when voiced and voiceless segments follow for male speakers.....	<b>37</b>
<b>Figure 6.2</b>	The distribution of voicing for the following voiced and voiceless segments for the female speakers.....	<b>38</b>
<b>Figure 6.3</b>	Voicing of final [t] together with the glottalization.....	<b>38</b>
<b>Figure 6.4</b>	The assimilation of voicing tendencies according to the type of the following word for both voiced and voiceless following segments.....	<b>39</b>
<b>Figure 6.5</b>	The assimilation of voicing tendencies according to the following type of the word for both voiced and voiceless segments for male speakers.....	<b>40</b>
<b>Figure 6.6</b>	The assimilation of voicing tendencies according to the following type of the word for both voiced and voiceless segments for female speakers.....	<b>41</b>
<b>Figure 6.7</b>	The average voicing profiles according to the presence of $F_0$ in the duration of the word-final segments.....	<b>42</b>
<b>Figure 6.8</b>	The voicing profiles and variation of the $F_0$ presence for male speakers.....	<b>43</b>
<b>Figure 6.9</b>	The voicing profiles and the variation of the $F_0$ presence for female speakers.....	<b>43</b>
<b>Figure 6.10</b>	The variation of the $F_0$ presence for the stop consonants.....	<b>44</b>
<b>Figure 6.11</b>	The variation of the $F_0$ presence for fricatives.....	<b>45</b>
<b>Figure 6.12</b>	The variation of the $F_0$ presence for affricates.....	<b>46</b>
<b>Figure 6.13</b>	The variation of the presence of voicing according to the type of the following consonant.....	<b>46</b>
<b>Figure 6.14</b>	The voicing ratios for the individual speakers for both voiced and voiceless following segments.....	<b>47</b>

**Figure 6.15** The average value of the assimilation taking place in the pair-consonant environments..... **48**

## Table of contents

<b>1</b>	<b>Introduction</b> .....	<b>10</b>
<b>2</b>	<b>Theoretical part</b> .....	<b>11</b>
	2.1 Speech and speech sound production.....	11
	2.2 Voiced and voiceless sounds distinction.....	12
	2.3 Voicing influences.....	14
	2.4 Terminology problem.....	15
	2.5 Devoicing.....	15
	2.6 Voicing in specific consonant groups.....	16
<b>3</b>	<b>Second language acquisition</b> .....	<b>21</b>
	3.1 Factors influencing the second language acquisition process.....	21
	3.2 Impact of mother tongue and interference theory.....	22
	3.3 Acquisition of phonemic segments.....	23
	3.4 Perception and production of sounds.....	24
	3.5 Error types.....	26
	3.6 Differences between Slovak and English with respect to the appearance of errors...	26
	3.7 Assimilation.....	28
<b>4</b>	<b>Hypotheses</b> .....	<b>31</b>
<b>5</b>	<b>Method</b> .....	<b>32</b>
	5.1 Speakers.....	32
	5.2 Method.....	32
	5.3 Statistical analysis.....	34
<b>6</b>	<b>Results</b> .....	<b>35</b>
	6.1 Overall results.....	35
	6.2 Gender differences.....	36
	6.3 Lexical and grammatical words.....	39
	6.4 Presence of F <sub>0</sub> variation.....	42



6.5 Variation of $F_0$ presence for separate consonant types.....	44
6.6 Individual speakers.....	47
6.7 Homorganic pair consonants.....	48
<b>7 General discussion and conclusion.....</b>	<b>49</b>
<b>Bibliography.....</b>	<b>53</b>
<b>Zhrnutie.....</b>	<b>55</b>

## 1 Introduction

The presence of voicing and its assimilation are the features that are found in most of the languages, including English and Slovak. The conditions of their presence can, however, mark the possible differences in the two languages, and therefore such properties and their distribution can mark the second language interference when concerning the second language acquisition process. Voicing can be described as the consequence of the movement of the air in the vocal folds and their subsequent slackening and vibration. The assimilation of voicing of a segment is caused either by its left or right voicing context and in this process the segment adjusts in voicing according to the voicing of the neighbouring sound. This process appears in both the languages, but its distribution in the native speeches is likely to differ due to the diverse rules the native speakers employ in their languages. This difference can also be the source of the interference errors in the speech of Slovaks in English.

The theoretical part of this work consists in providing the background on the voicing and on the second language interference with respect to the assimilation of voicing and its properties. In the first part we are concerned with the definition of voicing, its correlates and the distribution of voicing for the specific consonant groups. In the second part we discuss the second language acquisition process and the factors which can influence it. We are also concerned with the most influential theories describing the interference together with the most common mistakes emerging, taking in account the differences in the Slovak and English consonantal systems. Then the assimilation process is described in more detail.

The empirical part is based on the analysis of the recordings of the 18 speakers reading the BBC bulletins. We focused on the examination of the word-boundary environments in which the assimilation of voicing was likely to occur with the aim to see the distribution of the assimilation in the environments typical for both Slovak and English and in those typical only for Slovak to mark the possible error appearances. The semantic properties of the words following the analysed segments were considered too. In order to determine the significance of our results, we conveyed the specific statistical tests.

## **2 Theoretical part**

### **2.1 Speech and speech sounds production**

“Speech is, physiologically, an overlaid function, or to be more precise, a group of overlaid functions. It gets what service it can out of organs and functions, nervous and muscular, that come into being and are maintained for very different ends than its own.” (Vaissiere 1997 from Sapir 1923: 115) The speech-sounds production consists in “a presence of (quasi-) periodic excitation signal. Periodicity of the speech signal, a harmonic spectrum, and the presence of low-frequency energy have been identified as acoustic consequences of voicing. Accordingly the feature has been identified in articulatory terms [...] and in acoustic and auditory terms.” (Möbius, 2004: 5) The production of sounds takes place in larynx, to be more specific in its significant part - the vocal folds. Larynx fulfills several essential functions in human body, its primary purpose being that of protection, as “it prevents the entrance of foreign materials into the lungs,” (Vaissiere, 1997: 115) while simultaneously securing the breathing process and last but not least, providing the possibility of the speech-sounds creation. (Sawashima, 1983: 14) The generation of sounds is not simple, requiring the assistance of numerous parts of larynx and vocal folds which are interconnected and supplement their simultaneous activities. Larynx as one of the main articulators in speech production can be divided into several parts, namely the tongue root, the tongue body, the tongue blade, the velum and the lips; and at the same time it consists of four different cartilages: the epiglottis, thyroid, cricoid and arytenoid. The thyroid and cricoid cartilages are connected by the cricothyroid joint, the movements of which change the length of the vocal folds and at the same time contribute to their opening and closing. Vocal folds, on the other hand consist of the cover, transition and body, all of them in their movements substantially influencing the quality of the outcoming speech sound. (Sawashima, 1983: 11,12)

Opening (anatomically abduction) and closing (anatomically adduction) of the vocal folds is the process which allows us to distinguish between the phonation and respiration, by changing the posture of the vocal folds according to their required function in a particular moment. This reciprocal activity is caused by the movements of the arytenoid cartilage on the surface of the cricoarytenoid joint. “For phonation, the vocal folds are in adducted position and are set vibrating by transglottal airflow and in this position a narrow spindle-shaped gap is usually seen in the membranous position of the glottis before the vocal folds start to

vibrate.” (Sawashima, 1983: 14) Apart from mechanical procedures taking place in the larynx, in order to create a speech sound, not only is a specific configuration of the vocal folds required, yet also a considerable amount of energy, usually in the form of airflow, must be present. “All speech sounds are made with some movement of air, and the egressive pulmonic is by far the most commonly found in the languages all over the world.” (Roach, 1991: 26) The reason why the egressive pulmonic air stream is the most effective and most used energy source is that it allows us to spare some air in the lungs during breathing and immediately employ it in speech. “The excessive subglottal pressure forces the closed vocal folds to go apart and the elastic recoil and the Bernoulli Effect suck them again together without any muscle action. The vibrations divide the continuous stream of expired air coming from the lungs into puffs of air that will excite the supraglottic cavities. The quasiperiodic modulation of the respiratory airstream by the vibration of the vocal folds provides the primary source of energy for the production of voiced sounds.” (Vaissiere, 1997: 115)

## **2.2 Voiced and voiceless sounds distinction**

Nonetheless, the overall course of speech-sound production is not stable and the abduction-adduction of the vocal folds together with the “constriction of the false vocal folds, change in the length and thickness of the vocal folds and up and down movements of the larynx,” have got the capacity to influence it radically. To provide an example of a potential influence on the process, it is possible to examine the change of the thickness of the vocal folds which mediates the control of the vocal pitch and fundamental frequency ( $F_0$ ) during voicing, and which “is achieved mainly by controlling the effective mass and stiffness of the vocal folds.” (Sawashima, 1983: 21) Such a fundamental frequency control is often employed in order to produce falsetto tones, but even more importantly, to distinguish between voiced and voiceless consonants. (Sawashima, 1983 from Halle and Stevens, 1971: 22)

Halle and Stevens (1971: 22) present a theory that for the voiceless consonants the vocal folds exhibit the tendency to stiffen, while for the voiced consonants they tend to slacken. (Sawashima, 1983 from Halle and Stevens, 1971: 22) The vocal folds adjustment, however, does not represent the only way of distinguishing between voiced and voiceless consonants. From technical point of view, voicing is caused by the vocal fold vibration which occurs throughout certain interval of time during the process of the pronunciation of a consonant. (Stevens, 1992: 2979) When the vocal fold vibration is not present during the pronunciation, the consonant is voiceless. Apart from the distinction methods described above, one of the most important determiners of voicing and voicelessness is the function of

glottis, more specifically the appearance and disappearance of the gap in the glottis. Sawashima (1983: 14) states that “the general picture of the glottal condition in the abduction-adduction dimension during speech is that the glottis is closed or nearly closed for voiced sounds whereas it is open for voiceless sounds, the extent of the glottal opening varying with different phonemes and phonological environments” – which we can see in the figure 2 below. The functioning of glottis is connected with the function and activity of the respective muscles, thus allowing the researchers to observe a precise muscle activity pattern which closely determines the dimensions of the sounds in terms of voicing. “There is a clear reciprocal pattern in the activity of two muscles – suppression of the interarytenoid (INT) and activation of the posterior cricoarytenoid (PCA) muscle – for glottal opening corresponding to the voiceless sounds. Contrary to the voiceless sounds, there is an exactly reverse pattern – activation of the INT and suppression of the PCA – for the voiced segments when the glottis is closed.” (Sawashima, 1983: 15,17)

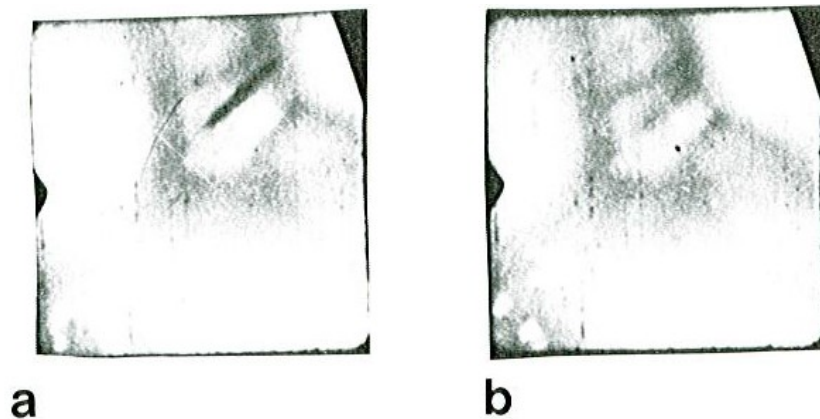


Figure 2 “Glottal views at the phonatory position immediately before a) and after b) the onset of vocal fold vibration.” (from Sawashima, 1983: 15)

When summarizing the previously discussed ways of distinguishing between the sounds in terms of voicing, a conclusion can be drawn that from the articulatory point of view, the distinction between the voiced and voiceless sounds “implies that the voiced sounds belong to a category sharing a common feature [voice], which in English corresponds to a specific vocal tract configuration: an articulatory setting for vocal fold adduction and an aerodynamic setting for a pressure drop across the glottis.” (Smith, 1997: 471)

The above analysed laryngeal adjustments are, however, not the only features which have a significant impact on the phonation. While Vaissiere (1997: 123) claims that the changes of the vocal fold length together with the variation of subglottic pressure and glottis

adjustment are the quintessential examples of the linguistically used controls of the larynx, she simultaneously implies that the phonation is variable according to the immediate attitudes and even psychological state of the speakers who subconsciously adjust the laryngeal mechanisms in order to produce the most accurate sound.

### **2.3 Voicing influences**

Voicing as an unstable feature is highly prone to be influenced by various factors and therefore one ought to examine those possible influences when discussing the voicing phenomenon. The most common correlates of the voicing according to Möbius (2004: 5,6) are “aspiration duration, closure voicing, fundamental frequency onset, first formant onset, closure duration, preceding vowel duration, following vowel duration, and the difference between the amplitude values of the first and second harmonics.” It is therefore possible to claim that the immediate environment of the sound, together with the position of the phonemes in the word or phrase, is not to be disregarded as it determines further characteristics and realization of the sound. Those features prove their significance as they shape the character of the sound and determine whether the voicing during the pronunciation of the consonant will persist throughout the whole process of pronunciation, or whether it will be much less observable and the occurrence will be detected only in certain phases of the sound, or as the last possibility, whether the voicing probability will be completely repressed and the consonant will be recorded as voiceless. What enables the listeners to distinguish whether the particular speech sound is voiced or voiceless, is the question that arises inevitably, the answer to which is provided by Shih and Möbius (1998: 1) that it is “a number of acoustic cues, in particular duration, fundamental frequency  $F_0$  and formant trajectories that have been shown to enable the listeners to make the voiced/voiceless distinction, provided that they can reliably perform the task without lexical information.” In terms of the amount of energy needed to produce the sounds, it is possible to claim that due to the bigger amount of energy needed to be exerted, the voiced sounds are considered to be more difficult to produce, (Smith, 1992: 471) which is connected with the length of the voiced sounds. “Compared to its voiceless counterpart in otherwise identical context, a voiced consonant typically has shorter duration.” (Shih, Möbius, 1998:1) The examples in English can be found particularly in the plosive and fricative consonant groups, such as [p,b] or [s,z], the latter ones of the couples being voiced and thus requiring bigger amount of energy for correct pronunciation – which is the reason why the speakers often seek ways of simplifying them in speech.

## 2.4 Terminology problem

When considering the voiced/voiceless terminology, it is important to notice that there exists an ambiguity, as sometimes even systemically voiceless sounds can be partially voiced in certain part of their pronunciation. Simultaneously, such a distinction of the sounds is likely not to correspond in diverse languages. Separation of the two sets of consonants is thus not to be determined by the voiced/voiceless contrast only, mostly because this feature is particularly unstable and prone to be influenced by the position of the sound within a word. “Although in medial position English [b,d,g] are voiced and [p,t,k] voiceless [...], in initial position both sets are commonly produced with silent closure intervals and should therefore be classed as voiceless according to the definitions.” (The mechanisms of the individual consonants pronunciation are to be discussed in the separate paragraph.) (Lisker, Abramson, 1964: 384) In order to resolve this issue, the force of articulation is often taken in account, “although the assessment of articulatory force appears ultimately to be a matter of proprioceptive judgment, this judgment is said to depend directly on the audible features” of the sounds. (Lisker, Abramson, 1964: 385) Subsequently, in order to erase the issue, the terms fortis/lenis or tense/lax were introduced. This category, however ought to be accepted with care as well, as even though it is widely said that it is the only distinctive feature separating the sets of consonants, it has not been proved yet, and in the languages which invoke the use of fortis/lenis the fact that many sounds are recognized as fortis and at the same time voiceless and vice versa, is only accidental. (Lisker, Abramson, 1964: 386)

## 2.5 Devoicing

Relative instability of voicing is, besides its tendency to be affected by its environment, often demonstrated by another feature as well, which is the devoicing process. Devoicing can be understood in several ways, one of which being a simplification of the sound creation as a devoiced sound is viewed to be easier for pronunciation, similarly as voiceless sounds. The devoicing effect is literally the loss of voicing from an originally voiced sound. The loss is usually conditioned by the specific environment, “where it could be viewed as an assimilation to an adjacent voiceless context, and where articulatory and aerodynamic effort tends to be reduced.” (Smith, 1997: 471)

The voiced/voiceless, fortis/lenis and devoicing effects are not recognized as compact homogeneous categories; on the contrary it is necessary to highlight the distinctions in the voicing profiles primarily between vowels and consonants and then in between the specific consonant groups, in which when elaborately analyzing, the various essential differences can

be found. The extent of the voicing therefore depends on the form of the obstruction created and simultaneously on the place of articulation of the particular sound.

## **2.6 Voicing in specific consonant groups**

Two fields of sounds are to be looked upon: single consonant sounds and consonant clusters. Both groups exhibit the tendencies to behave differently according to their environment and therefore it is significant to focus on them equally. The elaborate research of this particular field has already been done, according to which we separate consonants into three major categories depending on the form of obstacle.

The first major category discusses the stop consonants, or the plosives which share four essential characteristic features (Roach, 1991: 30):

- One or two articulators are moved against each other in order to form a stricture that prevents the air from escaping
- The air compressed behind the stricture is then released
- If the air is still behind the stricture in the process of releasing, it is likely to produce a plosion
- There may be voicing during part or all of the plosive articulation

The voicing is thus to be expected, yet one needs to distinguish between those stops commonly considered to be voiceless, where the voicing is not likely to occur if not in specific phonemic environment, and those generally acknowledged as voiced where the voicing is likely to emerge, yet its magnitude is not the same and exhibits the tendency to be changeable in diverse speech situations. Möbius (2004: 11,12) summarizes the idea as follows: “The probability of voicing in [b,d,g] series is consistent across the duration of the closure phase and stays within a narrow range (approximately 60-70%). [...] The [p,t,k] series, on the other hand, shows a considerable degree of voicing early in the closure, but the probability of voicing falls under 10% by the temporal midpoint of the closure.” It is also necessary to point out that even when concerning the voiced stops, the voicing, according to the Möbius’s research, never reaches the 100% level. On the contrary, there are two effects observable, “the devoicing effect on voiced stops and a closure-initial voicing effect on phonemically voiceless stops.” (Möbius, 2004: 12) Furthermore, the effect of the surroundings plays an important role in shaping the extent of voicing of stops as well. The



position of the stops within the word is therefore necessary to be analysed. Both contexts, left and right, have a considerable effect on voicing, but for the voiced stops the left context is more significant than the right one. When the left context is either sonorant or vowel, the voicing has the tendency to rise almost to 90%. For voiceless obstruents left contexts, the possibility of voicing stays below 10% during the whole closure phase, although vocalic or sonorant left context may raise the possibility of voicing by 10 or 15% at the beginning of the closure phase. Despite being less significant, the right context may have certain influence on voicing as well. As the voicing of [b,d,g] series is rising towards the end of the closure phase, the effect of prevoicing can happen there when followed by vocalic or sonorant context. (Möbius, 2004: 12,13) The illustration of the voicing profiles for the plosive series can be seen in the figure 2.1:

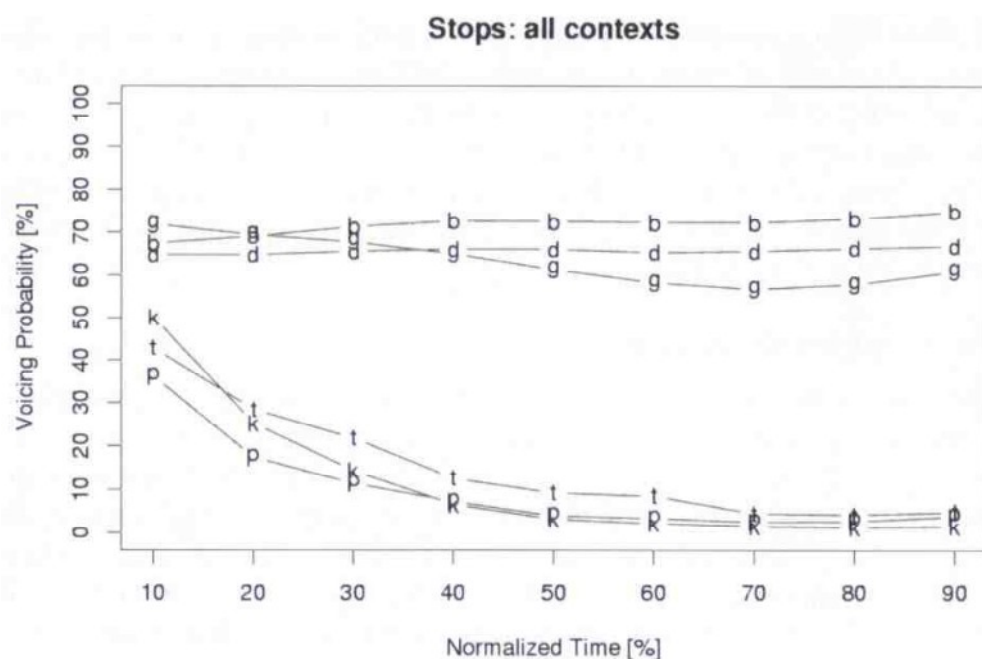


Figure 2.1 “Voicing profiles of the closure phases of German stop consonants, pooled across all left and right segmental contexts.” (from Möbius, 2004: 12)

The situation is quite different in the second group of the examined consonants, the fricatives. “Fricatives are consonants with the characteristic that when they are produced, air escapes through a small passage and makes a hissing sound. [...] They are continuant consonants, which means that you can continue making them without interruption as long as you have enough air in your lungs.” (Roach, 1991: 47) Stevens et al. (1992: 2980) provides

even more detailed view on their production and states that they are formed due to the narrow constriction in the vocal tract above the glottis. The airflow thus becomes turbulent, which is the process called frication. The problem which emerges here is the fact that both phonation and frication are characteristic for the air turbulence, and therefore it is rather complicated to distinguish between the frication turbulence and phonation, as they tend to overlap. Due to the increased requirements needed for the pronunciation of the voiced fricative, it is possible to assume it a sound far more difficult to produce than a stop consonant (from the energy usage point). “Voiced fricatives require higher subglottal pressure than oral pressure - oral pressure must be relatively low when compared to subglottal pressure. Speakers usually allow oral pressure to rise and subglottal pressure to fall in order to make the pronunciation easier.” (Smith, 1997: 472,473) Thus one ought to expect several issues emerging when analyzing the maintenance of the voicing during the period of frication. In order to produce noise at the supraglottal constriction together with voicing, both the glottal configuration and the supraglottal constriction must be adjusted one to another; otherwise it would cause a pressure drop. (Stevens et al., 1992: 2980) Individual fricatives are, similarly as plosives, different in terms of creating and maintaining voicing and they too exhibit a different behaviour when in consonant clusters. “The data show that singleton intervocalic voiced fricatives are most likely to exhibit glottal vibration throughout their duration. A voiced fricative in a cluster with another voiced fricative shows continuous glottal vibration almost as often. The most striking result is that when a voiced fricative is in a cluster with voiceless fricative, it exhibits glottal vibration throughout only 24% of time.” (Stevens et al., 1992: 2990) The result drawn from this can be that the voicing is lessened directly due to the presence of the voiceless fricative in a cluster. (Stevens et al., 1992: 2990) Thus it is possible to observe that while voiced plosives tend to keep voicing during the whole closure phase, the fricatives are excessively prone to be influenced by their surroundings, together with the fact that the subglottal vibration necessary for the frication has a significant impact on the voicing in such kind of obstruents. Möbius (2004: 14) adds that “the voicing probability stays within the narrow range of (60-80%), similar to what was observed for the phonemically voiced stops.” The individual sounds are much more distinguishable according to their voicing, contrary to the stops which have got the tendency to act similarly within a word or a clause. They all however show a considerable initial voicing, yet for the most of them it falls after one third of their respective duration. (Möbius , 2004: 15,16,17) The voicing profiles of the fricatives discussed above are displayed in figure 2.2:

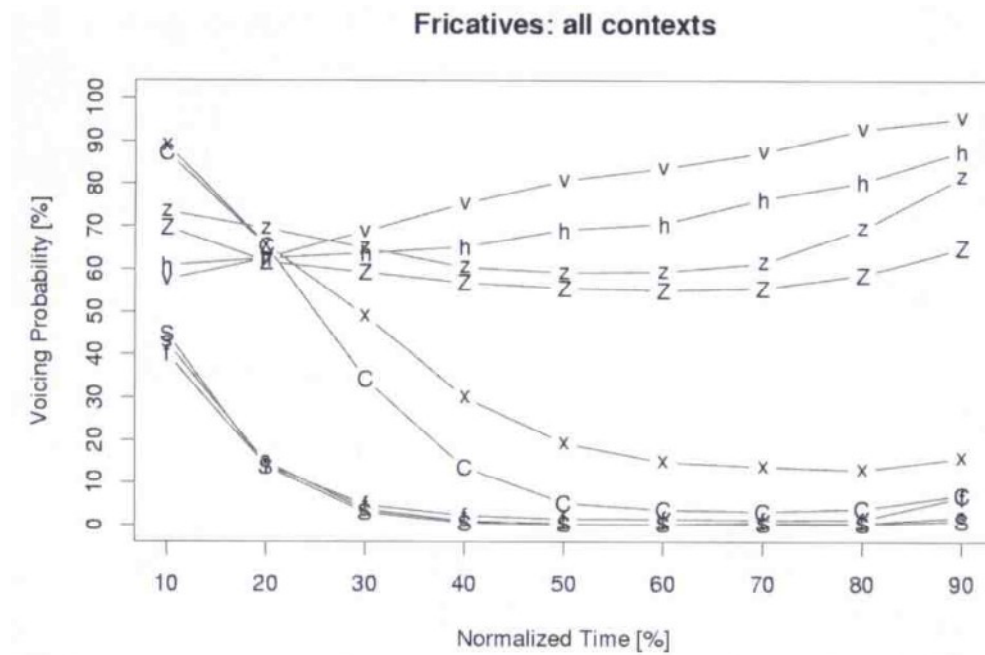


Figure 2.2 “Voicing profiles of German fricatives, pooled across all left and right segmental contexts.” (from Möbius, 2004: 14)

The next type of the consonants requiring further examination is the sonorants, which “are continuants and usually have no frication noise.” (Roach, 1983: 56) They seem to have the similar tendency as fricatives, that is to say, they are easily influenced by their left context, due to which “all sonorants are practically fully voiced throughout their duration,” (Möbius, 2004: 17) if there is another sonorant or a vowel. When considering the results for the consonant groups provided by Möbius, it is however important to take in account the fact that he was examining the profiles of German consonants and therefore there is the possibility that English plosives, fricatives and sonorants might behave in a different way, yet the difference ought not to be substantial.

Finally, in order to create a unified picture of the diverse consonant groups and their voicing tendencies in English, it is important to mention the affricates. “Affricates begin as plosives and end as fricatives, [...] so the plosive is followed immediately by fricative noise – and they must be homorganic or made with the same articulators.” (Roach, 1983: 47) “Since affricates involve precise temporal control over two constriction regions, as well as changes in the coronal configuration of the tongue, one can hypothesize that they are phonetically more complex than singleton stops or voiced fricatives. [...] They can also be longer in duration.” (M. Żygis, S. Fuchs, and L. L. Koenig, 2012: 312, 313) Thus what we can expect in terms of voicing are the tendencies similar for both plosives and fricatives: the big amount of voicing

at the beginning of the closure phase of and subsequent transfer of voicing to the beginning of the frication noise with its gradual fall towards the end of the pronunciation phase.

As we have seen, the voicing probability very closely depends not only on the characteristic features of the individual phonemes (whether they are fricatives or plosives,...), but simultaneously there is their environment playing a significant role in their voicing characteristics. The significance lies in the fact that the environment can cause either the increase or fall in the voicing, therefore the conclusion can be made that even though certain phonemes are always considered to be voiced, we ought to keep in mind that the degree of their voicing may differ in diverse phonemic contexts. When concerning the position of the consonant in the context, the distinction between the divergent types of the environment is important as they behave differently in clusters and for instance in the word boundaries, but for the purposes of this work we will disregard the word medial-position, and focus only on the word-boundary environment.

The analysis of the process of phonation together with the parameters of the consonants in which the voicing is observable in certain situations and contexts is significant for the further examination of the capacity of the people to learn a second language. Phonation will be observed in the context of the common errors of the Slovak speakers of English concentrating on the voicing of the final consonants in the environment inducing the use of the Slovak voicing principles, which is not common for the native English speakers. Fathoming what lies beneath the voicing process together with the second language acquisition capacity and its intricacies is thus essential in order to be able to provide a proper analysis of the Slovak-speech errors in English.

### **3 Second language acquisition**

The second language acquisition is a modern field in phonetics to which a considerable attention has been paid, owing to the increased tendency of the people all over the world to master a second language apart from their mother tongue, which is caused mostly by the “demographic trends that have brought the sociolinguistic and psycholinguistic dimensions of multiculturalism to the attention of a wider research community.” (Leather, 1996: 269) Such stimulation of the interest can therefore be considered to be very closely connected with the propagation of both the multiculturalism and a continuous expansion of English which, during the last decade, has become unarguably the world’s lingua franca and thus the necessity of being able to communicate in English has become imperative. This suggests that the target group of the people striving to learn it as their second language are often “those from essentially monolingual societies.” (Leather, 1996: 269) The learners of foreign language encounter several critical problems in their educational process, some of which being unavoidable.

#### **3.1 Factors influencing the second language acquisition process**

The second language acquisition process is not the same for everybody, the individual factors influencing the capacity either to speak or to perceive a foreign language properly thus must be taken in consideration. “Among the constraints on the acquisition of the second language speech determined by the learner are the maturational factors on one hand and individual and social constraint on the other hand. “ (Leather, 1996: 270) Thus what plays a significant role in the matter is the gender, social acceptance and distance, although it is primarily the motivation serving as one of the most serious extralinguistic impacts upon the learning process as despite having been proved that “most people can acquire minimally adequate second language pronunciation without systemic or formal training of any kind,“ the actual concern of the learner to acquire a native-like knowledge of the language is what to a great extent matters. (Leather 1996: 270) Apart from motivation and the other factors already listed, what we need to consider in the subsequent speech analysis, are the age and the way in which the language is acquired. (Kráľová, 2005: 30) From the two aspects it is necessary to highlight the role of the age in the second language acquisition process, as it is very closely connected with the learner’s ability to perceive and imitate the foreign sounds. Unlike the native speakers who were imitating the sounds in their early childhood, the grown-up learners of the second language have already developed their speech habits of their own language and during the learning process they are required to perceive and produce the segments only

according to their auditory and speech capacities. “A learner receiving explicit training in L2 [second language] articulation must adjust the configurations and movements of his articulators according to verbally formulated instructions, and the accuracy with which he or she is able to do this will ultimately be limited by tactile and proprioceptive feedback. [...] The capacity for stereognosis [capacity for accurate perception of spatial configurations within the mouth] apparently increases with age until the midteens and remains high into adulthood, declining with advancing age.” (Leather, 1996: 273) When the perceptiveness is limited, the learner is immensely handicapped and it takes them much longer to adjust their oral cavity in the pronunciation as required. This capacity is adjacent to the one of the imitation, “or the capacity to switch between the two languages, verbal abilities and verbal intelligence in both primary and secondary language.” (Kráľová, 2005: 29) Concerning the significance of correct pronunciation, for ordinary speakers it is not imperative to master it, yet the intelligibility and the clarity of speech is a condition in order to be respected and appreciated by the native speakers.

### **3.2 Impact of mother tongue and interference theory**

Apart from the sources of possible effects discussed above, the major factor affecting the acquisition of the language is the speaker’s mother tongue. Leather (1996: 273) asserts that “there is abundant evidence that the beginning learner seeking to impose phonetic structure on the second language speech to which he is exposed makes perceptual reference to the phonetic categories of the first language.” This phenomenon is called the interference and can be defined as “the mutual influence of the languages on several diverse levels. In the foreign language educational process the term ‘positive’ and ‘negative’ transfer is differentiated and the interference denotes only the negative influence of the mother tongue on the foreign language acquisition.” (Kráľová, 2005: 9) The language interference is not, however a process thoroughly automatic and there exist the stimulators and the inhibitors of the process itself.

The two basic types are recognized:

1. Structural factors, being a consequence of the structural differences of the languages,
2. Non-structural factors, usually divided into the subgroups:
  - Individual
  - Extraindividual (Kráľová, 2005: 19)

The structural factors operate on a syntactic, morphological and phonetic level while those non-structural can be identified as those which do not result from the structure of the language, but mostly from the capacities of the speaker or the influence of the environment.

Another distinction concerns the phonetic level in more detail. The influencing factors thus can be:

- Phonic
- Extraphonic (lexical-semantic)
- Extralinguistic (psycho and sociolinguistic)
- Random (Kráľová, 2005: 19)

The factors penetrate each other in several layers, therefore the final impact on the speech sound is not a consequence of the one factor's effect, but most probably it is a result of their co-operation. The interference theory gave birth to new hypotheses, one of the introduced having been called Markedness Differential Hypothesis which endeavours to provide an explanation why certain segments are much more difficult to learn than the others. It has been established that the phenomena which are perceived as difficult by the learners of the second language are only those more phonetically marked (unusual or different from the already established norm for the speaker) than the phenomena of the first language. (Kráľová, 2005: 10) Phonetic markedness enables the learners to identify the sounds they are not familiar with and this unfamiliarity often leads to presumptions that those sounds will require more effort to learn properly, which is primarily caused by "the tendency to correlate the sounds we know on primary language with those we hear in the second language." Leather (1996: 273)

Therefore, the process of learning the second language involves many diverse aspects which significantly determine its pace and efficiency with respect either to the specific needs and capacities of the students or the extent of interference of the languages. Having introduced the theory concerning the complexity of learning the sound segments system in the languages, we need to look upon their actual role in the whole process.

### **3.3 Acquisition of phonemic segments**

The segments of a second language from the phonetic point of view represent the core of the words, thus being of an immense importance for the learner to fathom for the sake of intelligibility. They can be of diverse nature but Kráľová (2005: 20) differentiates only three fundamental types:

- Identical,
- Similar,
- New

Those phonemes which are identical or similar for both languages usually do not signify a problem for the learner, even though resulting sound of the particular phonemes in the new language may subtly differ. More importantly, such phonemes are not the constraints hindering the comprehensibility of the speaker despite his possibly imprecise pronunciation of the phenomena. On the contrary, the phonemes with the tendency to occur only in the second language and without any counterpart in the speaker's mother tongue are often the stimulators of the errors in their pronunciation. There exists a relationship between the two languages' phonetic components as "in the processing and developmental realities of the second language speech acquisition it is likely that different properties of the second language sound input are picked out and associated with the mother tongue or gauged for their similarity or difference to the first language sound elements at different stages, with reference to different units (segments, syllables) of identification." (Leather, 1996: 287) The author thus claims that as soon as the speaker encounters a situation in which they are supposed to imitate a phoneme they are not familiar with, they automatically assume that there ought to be certain similarity between the particular phoneme and those they know from their mother tongue. Subsequently, they strive to categorize it according to the already established principles in their first language and employ their knowledge of the production of other sounds in order to convey the new sound in as appropriate way as possible. The influence of the first language on a new one is therefore undeniable, however the actual authenticity of the produced sound might be questionable.

### **3.4 Perception and production of sounds**

In terms of learning correct pronunciation of the second language sound repertoire, where the complications are easily presumed is both the correct perception and pronunciation of the new phonemes. The perception and production create an adjacent pair in the oral representation of the language and their relationship is more complicated than it seems at the first sight. Concerning the perception-production issue, several theories have been presented in order to provide an explanation concerning their relationship. "Trubetzkoy was the first one to assume that what causes the wrong production of a sound is primarily its being perceived in a wrong manner by the listener." Král'ová (2005: 20) Another theory presented by Král'ová



(2005: 20) dealing with the issue consists in the “incapacity of the learner to transmit sensorial perception of the second language phonemes to stabilized articulatory habits.” The capability of the listener to perceive certain phonemes however does not suggest that they will be able to recognize them in a fluent speech as there emerge the natural factors influencing the communication, such as loud environment, fast speech and even unintelligible pronunciation of the native speaker and thus the process of the new phoneme recognition might be rather complex. “A speech signal or phonetic interpretation of interlingual identifications must make a connection with a phonological interpretation of that part of a second language speech learner’s ‘mental grammar’ as much as the latter must connect with the former as a specification of part of a learner’s speech-processing arsenal.” (Leather, 1996: 289) The problem which Leather (1996: 289) identifies in the learning process of the new phoneme production consists in finding the phonetic target of the phoneme and the means of acquiring it. He opposes the previously presented theories by stressing that the “perceptual and productive mechanism are independent.” (Leather, 1996: 282) This statement is based on the theory introduced by Flege et al., called Speech Learning Model, which contradicts the previous hypotheses about the influence of the perception on the production of the new language phonemes. “It is concerned primarily with the ultimate attainment of L2 (second language) pronunciation, so work carried out within its framework focuses on bilinguals who have spoken their L2 for many years, not beginners. During L1 acquisition, speech perception becomes attuned to the contrastive phonic elements of the L1. Learners of an L2 may fail to discern the phonetic differences between pairs of sounds in the L2, or between L2 and L1 sounds, either because phonetically distinct sounds in the L2 are ‘assimilated’ to a single category because the L1 phonology filters out features (or properties) of L2 sounds that are important phonetically but not phonologically, or both. The model claims that without accurate perceptual “targets” to guide the sensorimotor learning of L2 sounds, production of the L2 sounds will be inaccurate. The model does *not* claim, however, that all L2 production errors are perceptually motivated.” (Flege, 1995: 238) Resulting from this theory thus “it seems likely that the adult’s L2 phonetic learning task is harder for a sound classified as an equivalent to one found in L1 than for one for which a phonetic category must be constructed from a scratch – because the influence of L1 category may cause learners to develop inaccurate perceptual targets for L2.” (Leather: 1996:276 from Flege 1987)

### **3.5 Error types**

The intricacies of the interconnectedness of perception and subsequent production of the sounds, significantly contribute to the amount of the emerging errors which, as the ultimate consequence of the language divergences, are unavoidable. There are several ways to interpret the reasons for the appearance of errors and Gas and Selinker (2008:102) suggest that we consider them “not just to be seen as something to be eradicated [...] as second language errors are not a reflection of faulty imitation. Rather, they are to be viewed as indications of a learner’s attempt to figure out some system, that is, to impose regularity on the language the learner is exposed to.” Kráľová (1996: 14) believes that there are several types of the errors, yet there are only two types relevant for this analysis:

- Phonetic errors
- Phonological errors

The errors are closely connected with the language interference resulting from the possible subtypes of the factors influencing the smoothness of the communication, which are as follows:

- Under-differentiation of the phonemes
- Over-differentiation of the phonemes
- Reinterpretation of distinctions
- Phone substitution (Kráľová, 1996: 63)

The errors and communication complications are not of the same magnitude in influencing the intelligibility and when Kráľová (2005: 18) compares the two types of errors, she asserts that those phonological (omitting or adding new phonemes) are more serious and cause more constraints than those phonetic (first language articulation implementation).

### **3.6 Differences between Slovak and English with respect to the appearance of errors**

The source of most of the errors emerging in the speech of Slovak learners of English is the diverse nature of the two languages. Kráľová (2005: 25) emphasizes several types of the differences between Slovak and English, starting with the contrast between the vocalic and consonantal system. More attention ought to be paid to consonants, as the core of this work is the realization of the consonants in Slovak English. Slovak language has at its disposal 27

consonants, while the English system is provided with only 24. This suggests that the Slovak learners of English should not have problems with the pronunciation of most of the English consonants. The English system involves, however, certain consonants which cannot be found in the Slovak one, namely the dental fricatives [ð,θ], the realisation of which tends to be the essential source for the pronunciation problems. Another example is the velar [ŋ] sound, which in spite of occurring in both the languages, causes pronunciation problems for the Slovak speakers, mostly due to the fact that while in English it is rather commonly spread, in Slovak it can be found only in specific environments, that is when preceding the plosives [k] and [g] – and therefore it is considered to be only the allophone of [n] and is produced mostly subconsciously. (Kráľová, 2005: 25) The errors in the pronunciation can be also based on the diverse character of the accent which in Slovak is stable and placed on the first syllable, while in English it is moving and distributed according to either morphological or structural type of a word. When concerning the intensity of the stressed syllables, it is necessary to point out that there is a sharp distinction between the quality and realization of those syllables in Slovak and English. In English the distinction is more marking and there is a considerable reduction of the syllables with no stress which does not happen in such an extent in Slovak. Consequently the pronunciation and accentual habits can have an immense impact on the production of the English sounds, despite the repertoire of the consonants being similar in both languages. (Kráľová, 2005: 25)

Having compared the two consonantal and suprasegmental systems one is capable of discussing the possible environments in which mistakes take place in the actual speech. First of all it is important to claim that the errors originate not only from the above-mentioned sources of differences but generally it is an intersection and permanent penetration of all the language subgroups. Mistakes which are loosely connected with the subject of this work will be briefly presented as their presence in speech can be the source of mutual influence with the other, relevant aspects. Apart from mispronouncing the velar [ŋ] and the dental fricatives, what causes the complications as well is the aspiration of the fortis plosives in the initial word position. Kráľová (2005: 25) The suprasegmental level of the speech realization cannot be omitted in the analysis, as the errors emerging there are very often considered to be the significant ones when concerning the intelligibility. The learners have the tendency not to separate the words according to the English accentual system but they employ their knowledge of the Slovak one, which is the effect of having incorrectly placed the accent on the first syllable of the English words (as the accent in Slovak is fixed on the first syllable). (Kráľová, 2005: 26) Despite having stated that the phonological errors are of greater

significance than the phonetic ones, the latter ones are, however, more intensive and immediately recognizable by the native speaker.

### 3.7 Assimilation

Král'ová (2005) notices that the pronunciation of the final lenis consonants in Slovak English is incorrect; the realization of the phonemes is either completely voiced or replaced by the adjacent fortis consonants. This interference or interlingual error is called the assimilation of voicing. According to Král' (1988: 111) “the assimilation is the process in which one or more phonemes changes in a certain manner in order to resemble the neighbouring phoneme. In this process the phoneme loses some of its characteristics or acquires new ones. The lost characteristic is the one which was different from the neighbouring phoneme and the acquired one resembles it acoustically and articulatory. Assimilation can take place only in the voice pairs, otherwise it is not possible.” The assimilation causes three types of differences in the consonants, namely:

- in place of articulation
- in manner of articulation
- in voicing. (Roach 1991: 124)

Regarding the direction in which the phonemes influence each other we can generally consider the process from two points of view: assimilation progressive (the preceding phoneme voicing characteristics influence the following phoneme, for example: begs [begz] vs. kicks [kɪks]) and regressive (the following phoneme voicing characteristics influence the preceding phoneme, for example: make believe). (Roach: 1991, 125) In English, “the assimilation processes are prevailingly regressive.” (Volín, 2002: 67)

As it is the change of voicing which is relevant for the further purposes in this work, we need to focus on it in detail. There exist two types of the assimilation of voicing:

- Voiced consonant changes into voiceless
- Voiceless consonant changes into voiced (Sičáková, 2002: 73,74)

In Slovak the voiced consonant changes into voiceless in several cases:

- in the word boundaries when the following word starts with voiceless stop or fricative,
- in the prepositional phrases,
- at the end of the word before the pause,
- in the compounds,
- in the medial position when a voiceless stop or fricative follows.

The voiceless consonant changes into voiced:

- in the word boundaries when the word starts with voiced stop or fricative,
- in the prepositional phrases, compounds, before the pause and in the medial position when a voiceless stop or fricative follows.

The quintessential example of the assimilation in Slovak is the pronunciation of the Slovak voiced fricative [v] which is in certain environments (word-final position) pronounced as back high vowel [u]. (Sičáková, 2002: 76) In English it means that in the words which end in [v], such as ‘live’, the Slovak speakers have the tendency to pronounce it as [liu]. “The assertion of the assimilation of voicing in English means that the pair consonants are assimilated even though the native speakers assimilate only the single consonants. The pair consonants in English always maintain their characteristics even in the neighbouring position.” (Kráľová, 2005: 55) On the other hand, in English we often find the assimilation of [l], [w], [r], [j] when preceded by the fortis pair consonant. (Kráľová, 2005: 25) In both cases a strong presence of Slovak can be observable, as while they produce assimilation in the environment unusual for the English native speaker, they do not assimilate in places where the natives generally do. Kráľová (2005: 39) provides a research in which she focuses on the errors emerging in the speech of the Slovak students of English with the result that the Slovak rules of the assimilation of voicing were implemented by fourteen out of sixty students. “In 73.68% the native speakers noticed the assimilation of the voiced consonant to the voiceless and vice versa in 26.32% (mostly before voiced consonant).” (Kráľová, 2005: 55) She later adds that according to the classification made during her research, this error can be classified as interference or caused by the native language influence, which is the prevailing type of errors recorded by her. Simultaneously when concerning the seriousness of this particular error in terms of intelligibility, it cannot be marked as a communication hindrance, as it does not influence the direct meaning of the words, yet it is noticeable by the native speakers. (Kráľová, 2005: 65,66) In terms of frequency of occurrence of the assimilation error, Kráľová

(2005: 71) states that it emerges in 3.76% cases of all the potential places recorded by her, in which it could appear.

## 4 Hypotheses

In the experimental part we will analyse the appearance or potential lack of the assimilation of voicing in the speech of Slovak students. The focus will be placed upon specific word-boundary environments in which the assimilation is expected either in both languages, or just in the one of them, following the rules of its distribution according to the information provided earlier in this work. Two kinds of environments are to be looked upon: the word-boundaries in which the voiceless consonant meets the voiced consonant or vowel, both for the word-final voiceless – word-initial voiced environments, and for the word-final voiced – word-initial voiceless environments. The concentration will be placed upon the environments which follow the Slovak rules of assimilation according to which the voiceless consonant changes into voiced in the word boundaries and thus we can formulate our first hypothesis.

H1:

Slovaks are expected to assimilate the voicing in the word-boundary environment when voiceless and voiced segments meet.

There is, however, a specific variant of the above described environment where the pair consonants meet, in which the assimilation of voicing is said not to take place in the speech of English native speakers as they always maintain the quality of the respective consonants. Thus we can formulate our second hypothesis.

H2: Slovak speakers are expected to assimilate the voicing of the pair consonants in the word-boundary environment.

The examination of the hypotheses can be conveyed with help of the research questions:

1. Do Slovaks assimilate when the following segment is voiced consonant?
2. Do Slovaks assimilate when the following segment is voiceless consonant?
3. Do Slovaks assimilate when the following segment is a vowel?
4. How do the following lexical and grammatical words influence the distribution of voicing in the target word-final consonant?
5. Is there a difference between males and females in terms of assimilation?

## **5 Method**

### **5.1 Speakers**

For the purposes of this work eighteen Slovak speakers were chosen from the database of the recordings of the Institute of Phonetics at the Faculty of Arts of the Charles University in Prague. The recordings were obtained in the soundproof room of the Institute and the sound was recorded using the AKG C4500 B-BC condenser studio microphone with the sampling rate of 32-kHz. All of them were the students of the Charles University in Prague, their age thus varying from approximately eighteen until twenty-six. Before the recording, the speakers were required to master at least intermediate level of English. Concerning the gender of the speakers, both males and females were selected to create two equal groups, which means that there were nine male and nine female students. The students were reading the BBC news bulletins with the additional time to prepare themselves. The texts were not continuous but rather the selections of short news divided into seven/eight passages. All of the speakers were reading the texts originally by the broadcaster Jackie Leonards (JLA) and then the text of another broadcaster was chosen for everyone so that we acquired two recordings from each of the speakers. The overall number of the texts considered was eight. The recordings were then divided into smaller sections according to the individual breath pauses, the longitude of the separate sections varying from 2 to 15 seconds.

### **5.2 Method**

The recordings were processed in the phonetic computer programme Praat, version 5.3.63. for the Windows XP Professional. (Boersma and Weenink, 2014) First we chose the specific environments from the respective bulletins, in which the assimilation was expected to occur. The environments we focused on consisted of the word-final voiceless segment followed by word-initial voiced segment and vice versa. When the word-initial segment was a vowel, we had to take in account the extensive glottalization of the Slovak speakers and the presence of the glottal stop in the majority of this type of environments. The glottal stop was then marked and processed as word-initial voiceless phoneme. The boundaries of the phonemes concerned required manual correction so that they were precise and possibly uninterrupted by neighbouring sounds and for the purposes of the correct boundary identification, Machač and Skarnitzl (2009) was consulted. Here it is necessary to keep in mind that in many cases the boundary distinction was unclear and the sounds often fluidly overlapped (which was mostly the case of sonorants and vowels in the neighbouring position).



The boundaries were then adjusted to the nearest zero crossing – “a point in which the waveform crosses the amplitude axis.” (Machač and Skarnitzl, 2009: 24) All the environments chosen from the texts and designed for further examination were identified by ‘x’ symbol in the point tier of Praat, irrespective of the presence of the assimilation of voicing. This symbol did not carry any specific information about the character of voicing in the target environment; it only marked the area in which presence/non presence of the assimilation was predicted, as we can see in Figure 5 below. The environments in which the assimilation was predicted, but not realized due to the breath pause were not included in the selection. The target environments were then extracted from the recordings using the Praat script which identified the duration of the word-final/initial phonemes, together with the presence of F<sub>0</sub>, eleven times for each of both word-final and word-initial phonemes, the aim being the identification of the voicing profiles of the individual speakers. Before the analysis we, however, excluded the voiced-voiceless environments due to their insufficient amount (only 30). The total number of the target environments further processed was 1101.

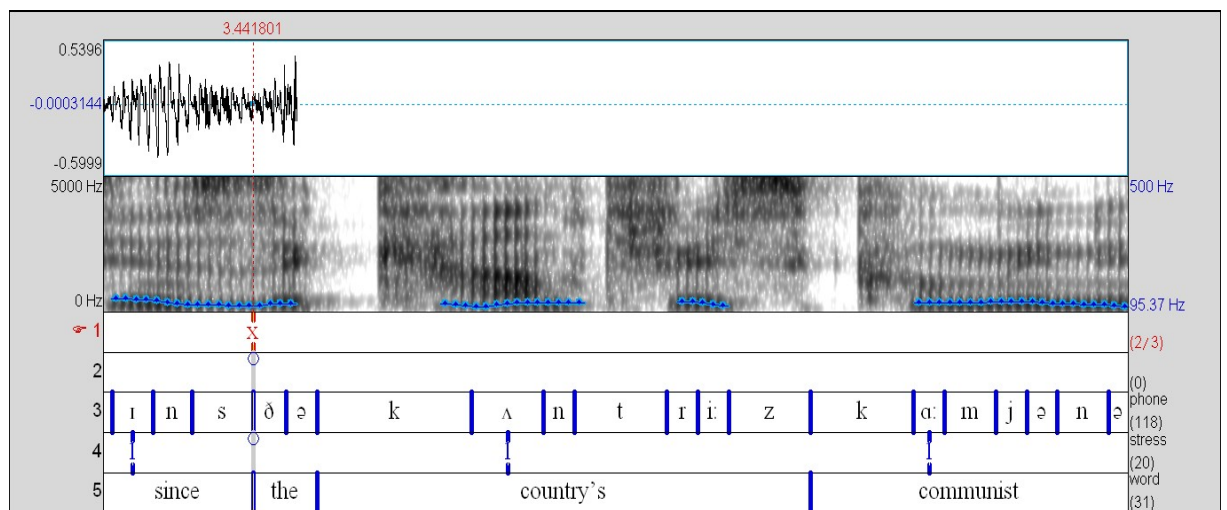


Figure 5 Labelling the target environment in Praat

The information obtained from the script analysis was then further processed in the Microsoft Excel. In order to acquire precise results we had to consider the semantic character of the words and thus we divided them into grammatical and lexical ones. The calculations and statistical analyses were then conveyed.

### 5.3 Statistical analysis

For this work, the statistical examination of the material provided had to be conveyed in order to determine whether the results were statistically significant or not. This was achieved using the statistical parametric method called “student’s t-test”. By means of the t-test we could compare two sets of data and determine whether they were the two parts of one population, or whether they were two separate groups. The basic component of the t-test is the so called standard error calculation which can be acquired from the formula,

$$s_x = \frac{s}{\sqrt{n}}$$

from which we determine the standard error by dividing the standard deviation by the rooted number of the data samples. The fundamental principle of the t-test lies in the measurement of the difference analysed (the null hypothesis) which shows the data variability – and this variability has always the nature of the standard error. (Volín, 2007: 108) The results of the t-test are always formulated as follows:

$$t(24) = 5.17; p < 0.001$$

This would mean that there was total of 26 samples as the degrees of freedom are not involved in calculation: as we have two sets of data, we excluded two degrees of freedom; and the number 5.17 tells us that there exists a difference between the two sets as the difference between their means is 5.17 bigger than the standard error, the overall result being statistically highly significant (the value of  $p$ ). (Volín, 2007: 108, 111) This value can differ, marking thus the diverse degrees of statistic significance. If the value of  $p$  is less than 0.001 the result is statistically highly significant, if it is less than 0.05 it is significant and if it is less than 0.08 it is only marginally significant. (Volín, 2007: 36, 37) The second type of the statistical test that we used for our purposes is the so-called ANOVA test which is based on the same principle as the “student’s t-test”, but it is used in the instances when three or more sets of data require comparison.

## 6 Results

### 6.1 Overall results

First of all the overall results of our analysis are to be presented, the specific measurements will be presented in subsequent passages.

Having completed the data marking according to our parameters, the calculations of the percentage rate of voicing of the target segments were conveyed. More specifically, we determined the ratio of the voicing present in the voiceless segments followed by the phonetically voiced and voiceless ones. Thus we determined that for the phonetically voiceless segments, on the average 59% of their duration is voiced, which leads to the conclusion that Slovak speakers do extensively assimilate in the majority of the target environments – and thus we confirmed our first hypothesis (H1). The examples of the assimilated sounds can be found in the following areas: British woman, took her, since then, but it's (speaker CIBK). There is, however, the abundance of the instances in which the assimilation did not take place as expected, for example: communist leader (speaker CIBK), governments around (speaker KANA) or rights group (speaker KIRA). As we can rule out the influence of the pause between the words, the explanation that we can draw from this sample is that the phoneme immediately preceding the word-final segment might have certain influence upon it as well. The results show us that the tendency not to assimilate increases when the word in which it is expected ends in more than one consonant, the instances can be found in the environments with the words such as: best, outbreaks, must, against or crashed. When analyzing the presence of voicing in the word-final voiceless and word-initial voiceless segments environments we calculated that even there the voicing is present in, on average, 10.1% of the word-final segments duration. For example: sent a (speaker CIBK), which included (speaker KANA), or statement also (speaker SPOA). In this case it is important to note that the glottalization is present in each of the examples and therefore the possible explanation for the voicing present in the word-final segments is the interference of the preceding voiced segment and subsequent transfer of voicing. The distribution of voicing for both following voiced and voiceless segments can be seen in the boxplot figure 6 below.

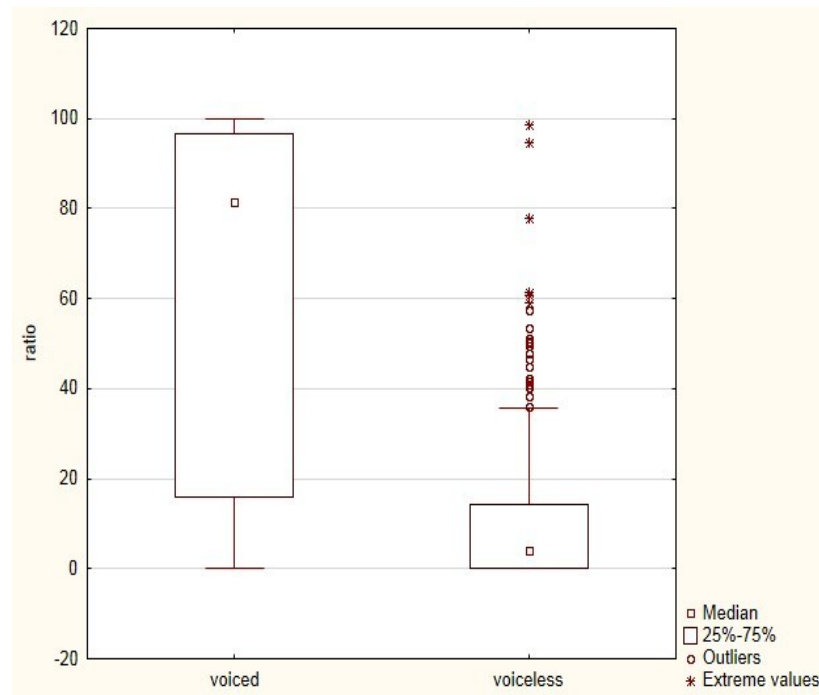


Figure 6 The distribution of voicing ratio for both voiced and voiceless following segments.

The box in both cases represents the middle 50% of the data sample, whiskers stretch to minimum and maximum values. As we can see, when the following segment is voiced, the voicing distribution varies from zero to 100%. For the following voiceless segments the voicing stretches from zero to almost 40%, the outliers and extreme values (the values that immensely differ from the majority of the samples) are present as well. The statistical t-test for the above described analyses proved the difference between the group of the word-final voiceless segments and their assimilation of voicing highly significant:

$$t(1099) = 21.0; p < 0.001$$

## 6.2 Gender differences

As we have seen according to our previous results, Slovaks tend to assimilate extensively. Now we can look upon the possible differences between the males and females. Our analysis of 565 environments read by male students showed that in the voiceless-voiced areas, on average 59.6% of the respective voiceless segments duration is voiced. When we examine the voiceless-voiceless areas we can see that at an average it is only 6.2% of the duration that is voiced. The t-test proved the result to be statistically highly significant.

$$t(563) = 16.7; p < 0.001$$

The distribution of voicing was very similar to that we saw with the overall results which we can see in the boxplot figure 6.1:

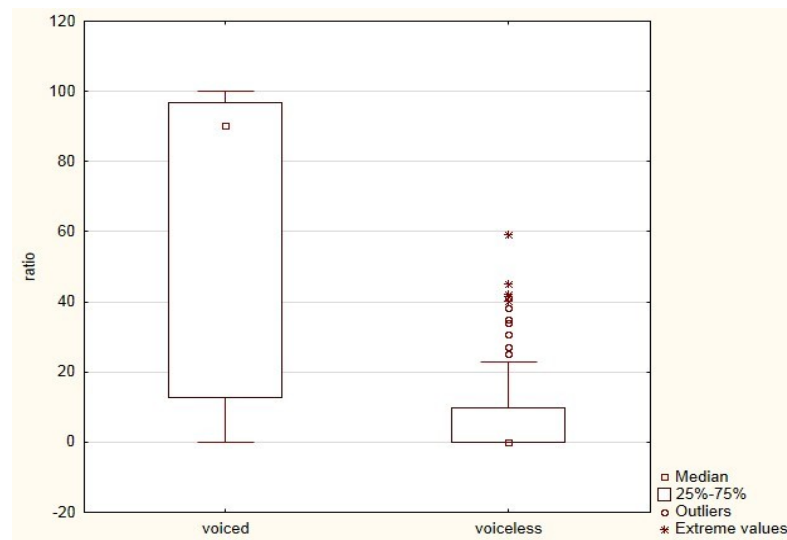


Figure 6.1 The distribution of voicing when voiced and voiceless segments follow for male speakers.

When the following segment is voiced the range of voicing for the preceding segments stretches from zero to 100% and the whiskers of the first box show us that there were more segments the duration of which was voiced from 95% to 100% than those with approximately 16% and less. For the following voiceless segments we can again see that the range of voicing is much smaller and that the outliers and extreme values are present as well.

For the female speakers the results were very similar. We analysed 536 samples and found out that for the voiced following segment, on average 58.4% of the duration of the preceding segments was voiced and when the following phoneme was voiceless the number was again much smaller – only 14.7% of the duration was recorded as voiced. According to the t-test this result is also highly significant:

$$t(534) = 13.0; p < 0.001$$

The boxplot figure 6.2 shows us the distribution of voicing for the female speakers:

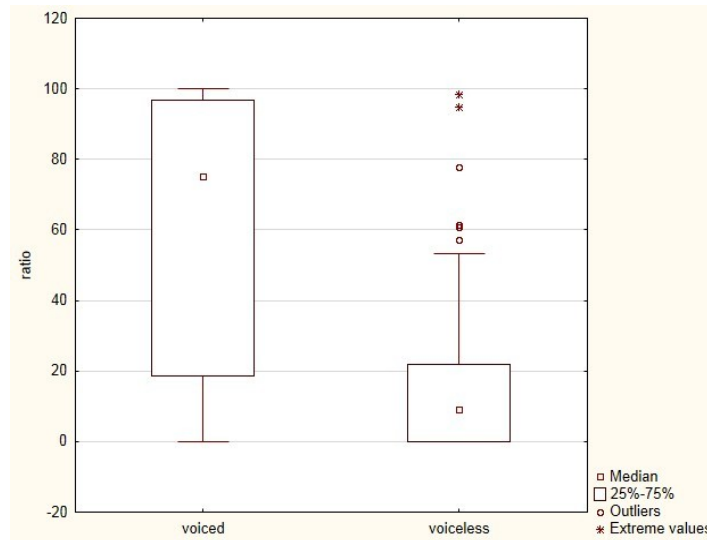


Figure 6.2 The distribution of voicing for the following voiced and voiceless segments for the female speakers.

While the box for the following voiced phonemes looks very similar to what we have seen in the figure 6.1, what is interesting is the box for the following voiceless phonemes. The ratio here ranges from zero to approximately 55% with several outliers and extreme values. The example of the extreme value is the pronunciation of [t] in ‘court of’ by the speaker KLIA who glottalizes at the beginning of the word ‘of’ and despite this pronounces this final [t] as voiced for 98.5% of its duration, which we can see in figure 6.3:

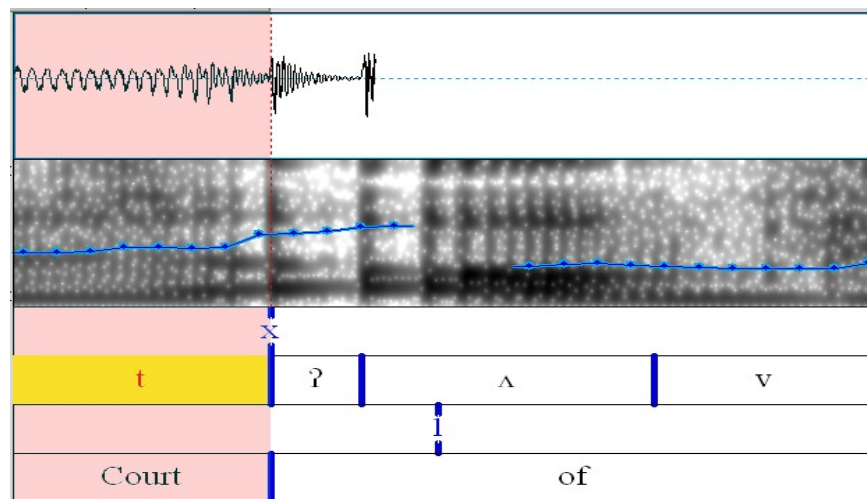


Figure 6.3 Voicing of final [t] together with the glottalization.

### 6.3 Lexical and grammatical words

The next aspect we examined concerned the possible differences in the assimilation of voicing according to whether the following word was lexical or grammatical. This analysis was conveyed twice, first for the whole set of data and second for males and females in order to notify the possible differences. The significance of our results was tested by ANOVA test which, as already described in the part 5.3 Statistical analysis, is used when we need to compare three and more groups of data and prove that they are not simply the parts of the same population. For the whole set of samples the ANOVA test proves the statistical significance of the results as well as the differences between the groups:

$$F(1, 1097) = 21.7; p < 0.001$$

The figure 6.4 shows us the development of the tendency to assimilate the voicing according to the type of the following word:

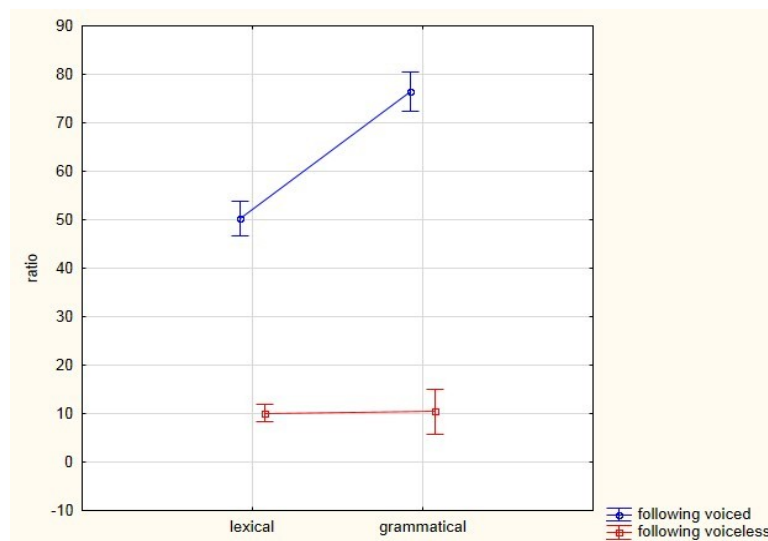


Figure 6.4 The assimilation of voicing tendencies according to the type of the following word for both voiced and voiceless following segments.

We can see that when the following segment is voiced rather than voiceless, there is a clear tendency to assimilate much more when the voiced segment is the part of the grammatical word.

In the figure 6.5 we can notice the development of the assimilation for the male speakers:

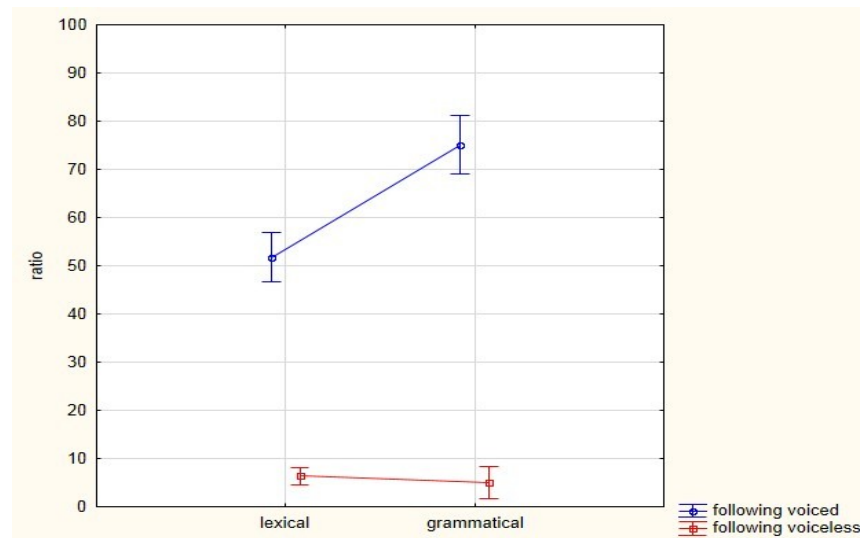


Figure 6.5 The assimilation of voicing tendencies according to the following type of the word for both voiced and voiceless segments for male speakers.

For the male speakers it is possible to point out that although for the following voiced segments the tendency is rather predictable from what we have already seen in the previous figure 6.4, the following voiceless segments development is rather striking. The overall voicing ratio is less than 10% but instead of rising for the grammatical words, we can see that the tendency to assimilate decreases. In terms of statistical significance we can say that although the results show that when the following words start with the voiceless segment the differences in voicing of the word-final phonemes are insignificant, the overall statistical significance is high due to the big differences in the voicing ratio of the word-final segments when followed by the lexical or grammatical words starting with voiced phonemes:

$$F(1, 561) = 10.3; p < 0.005$$

A comparison with female speakers can be made according to the figure 6.6:



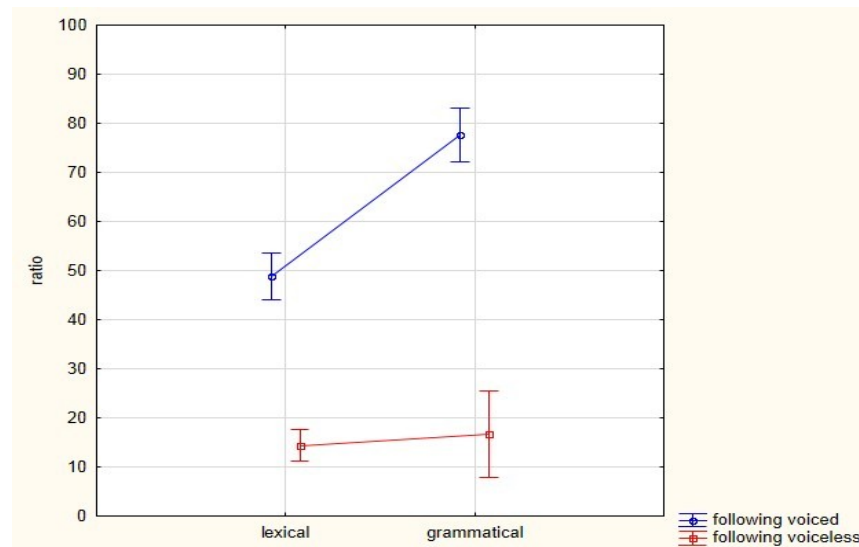


Figure 6.6 The assimilation of voicing tendencies according to the following type of the word for both voiced and voiceless segments for female speakers.

When compared with the male speakers we clearly see that the distinction lies mostly in the development of the voicing tendency when the following segment is voiceless.

The statistical significance of the result is high, again due to the differences in the voicing ratio of the word-final phonemes according to the following words being lexical or grammatical, both starting with voiced phonemes:

$$F(1, 532) = 11.189; p < 0.001$$

The overall differences between males, females and the whole set of samples are not big. Thus we can claim that there is general inclination of the Slovak speakers to assimilate more when the following word is grammatical, starting with voiced phoneme – the reason for this trend may be that we do not pay much attention to the pronunciation when the grammatical word follows as it is considerably bleached in terms of meaning and thus we do not feel the urge to emphasize it.

## 6.4 Presence of $F_0$ variation

What we examined next was the variability of the presence of the fundamental frequency, as its onset is one of the correlates of voicing, in order to create two average voicing profiles which would reflect this variability according to the nature of the following segment (voiced/voiceless). The average profiles can be seen in the figure 6.7:

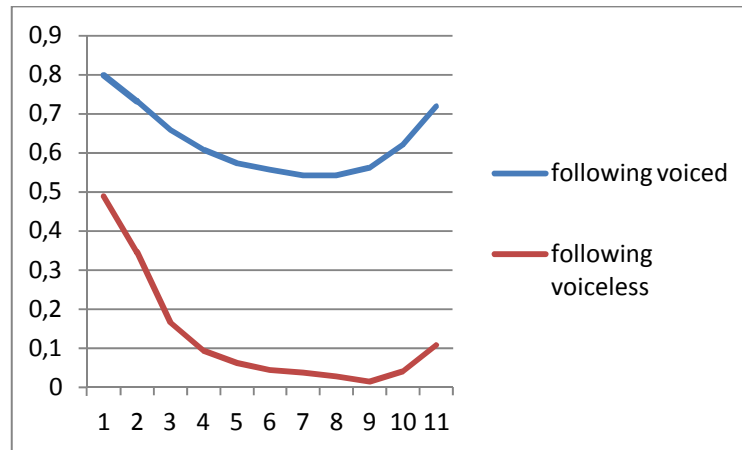


Figure 6.7 The average voicing profiles according to the presence of  $F_0$  in the duration of the word-final segments.

The figure 6.7 shows us that the distribution of  $F_0$  for the environments when the following segment is voiced (780 environments) varies and while almost 80% of the environments show its presence at the beginning of the duration, by the half of the overall duration the number of the environments which maintain the  $F_0$  presence decreases to almost 55% and towards the end it again increases to over 70%. As all the types of consonants were included in this measure, this result is not surprising as the voicing itself is of diverse nature for plosives and for example for fricatives and thus the presence of  $F_0$  varies according to the type of the consonant (we will look at the variation of  $F_0$  according to the type of consonants in the subsequent passage). On the other hand, for the environments in which the following segment is voiceless (overall number is 321), the result is somewhat surprising. We have already noticed that despite our expectations, some of these environments are voiced. From the figure 6.7 we can see that almost 50% of the word-final environments in this case show the presence of  $F_0$  at the beginning of the pronunciation phase, which is a result of preceding segment interference. The tendency to lower almost to 0% is however followed by repeated increase in the  $F_0$  occurrences; although their number is substantially lower than when the following segment is voiced – only something over 11%. While following voiced segment supports the reoccurrence of the  $F_0$  presence, the following voiceless segment does not and therefore there

might be other aspects to cause it – for example the incorrect pronunciation of the target word.

In order to be able to create a unified picture of the presence of  $F_0$  in the word-final consonants, we conveyed the analysis of the differences according to genders, which we can see in the figures 6.8 and 6.9:

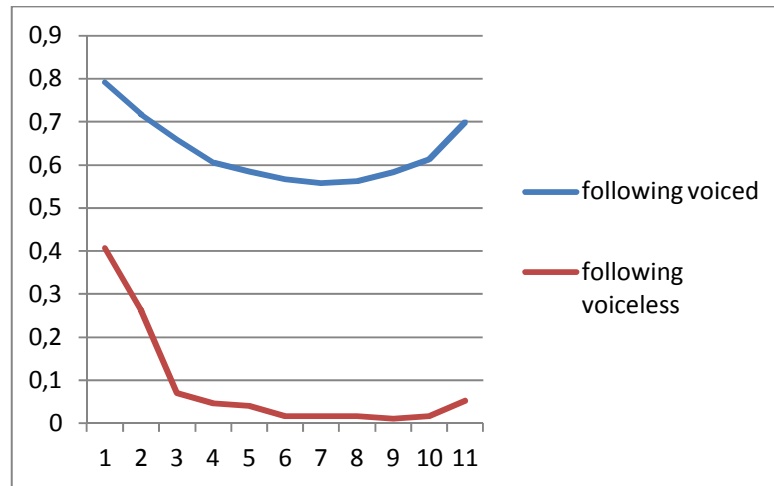


Figure 6.8 The voicing profiles and variation of the  $F_0$  presence for male speakers.

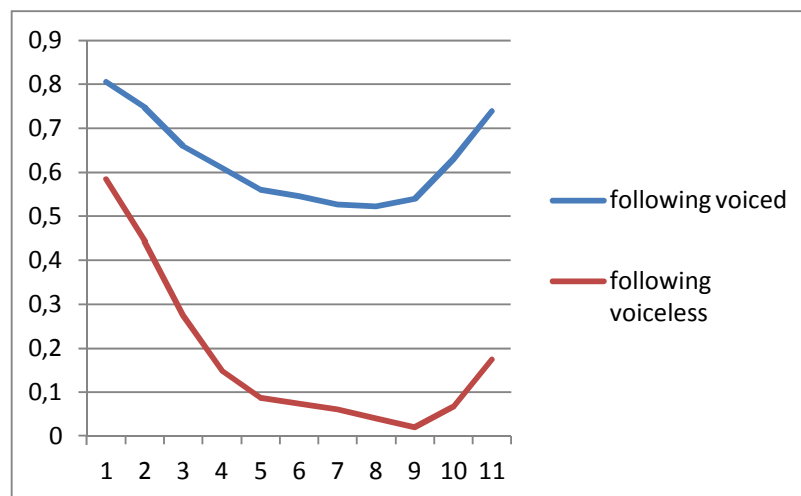


Figure 6.9 The voicing profiles and the variation of the  $F_0$  presence for female speakers.

What is interesting to notice is the difference between males and females in the  $F_0$  presence and variation for the following voiceless segments. We can notice that almost 60% of the environments when the following segment is voiceless pronounced by females show the presence of the fundamental frequency at the beginning of the pronunciation phase. Simultaneously, while the tendency to keep the voicing by the end of the pronunciation is very slight for males, we can see its presence in almost 20% of the environments pronounced by females. Re-listening the target words in Praat again, we noticed that there were several

female speakers who were reading quickly and inattentively which may have caused the mispronunciation of several target words, and thus the voicing occurred in the atypical areas.

### 6.5 Variation of $F_0$ presence for separate consonant types

This analysis was conveyed to show the differences in the  $F_0$  presence in the pronunciation of the consonants, due to the unstable nature of voicing which causes that it behaves differently for diverse types of the consonants. The variation of  $F_0$  throughout the pronunciation of all the consonants in general could be seen in the figure 6.7 in the preceding section. In this section we analysed both consonants undergoing and causing assimilation. The examination of the consonants undergoing assimilation was conveyed first. We examined the voicing variation for the three consonant groups, namely the plosives, fricatives and affricates and our results are presented in the figures 6.10, 6.11 and 6.12 below:

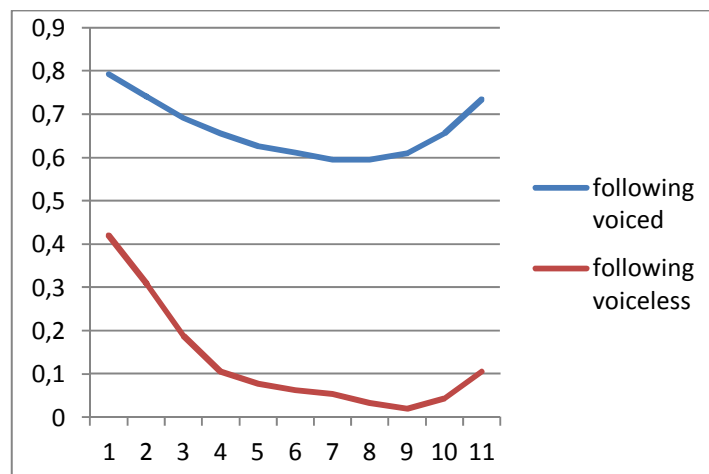


Figure 6.10 The variation of the  $F_0$  presence for the stop consonants.

Our results that we obtained for the plosives in our set of data correspond with what we learn about the nature of voicing for plosives in Möbius (2004). We are working with originally [p,t,k] series which are supposed to show the considerable amount of voicing at the beginning of their closure interval, then it is supposed to decrease substantially. This is what we can see for the following voiceless segments, with the exception of the appearance of the  $F_0$  by the end of the pronunciation phase of the consonants, already discussed in the preceding section.

What is worth noticing is, however, the presence of voicing when the following segment is voiced. This tendency much more resembles the one expected for the voiced [b,d,g] stops, which highlights the influence of the nature of the following phonemes and subsequent process of assimilation.

When analysing the fricatives we can see that our tendencies again correspond with the one's of Möbius' (2004) research. Here it is important to comment on the tendency of voicing to reappear by the end of the pronunciation phase in almost 70% of the recorded environments for the following voiced context, which we can see in the figure 6.11. Möbius (2004) claims that it is typical for the fricatives to exhibit high amount of voicing at the beginning of the pronunciation phase, but it is very likely to fall by the temporal mid-point, mostly due to the fact that it is difficult to produce voicing and frication at the same time. Our figure 6.11 shows us, however, that the voicing by the end raises substantially, the effect which is likely to be caused by the voiced nature of the following segment.

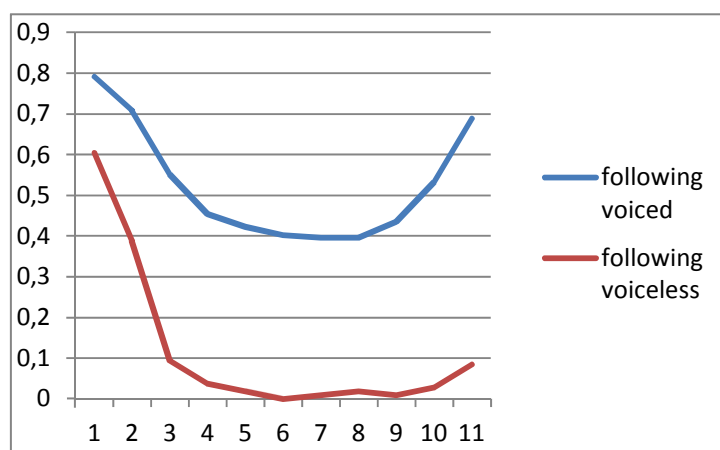


Figure 6.11 The variation of the F<sub>0</sub> presence for fricatives.

The last type of the consonants we were looking at in our analysis were the affricates, as one of the most frequent words under examination in our texts was 'which' (figure 6.12). As the affricates are composed of a plosive and of a fricative (they must be homorganic), we would expect that [tʃ] would behave half as a plosive and half as a fricative in terms of voicing. In the environments with the right voiced context we do not find the tendency displayed in the figure very surprising. What is very interesting, however, is the tendency exhibited for the following voiceless context. We can see that half of the environments show the presence of F<sub>0</sub> at the end of the pronunciation phase of the final [tʃ]. In the second half of the pronunciation we see that the phoneme is pronounced completely without voicing, but towards the end the voicing occurrence increases – and such a rising was not noticed for the fricatives alone with the voiceless right context. This sudden increase may be caused primarily by the interference of the voiced left context and possibly the mental preparation for the pronunciation of the subsequent voiced sound, despite glottalization anticipating it.

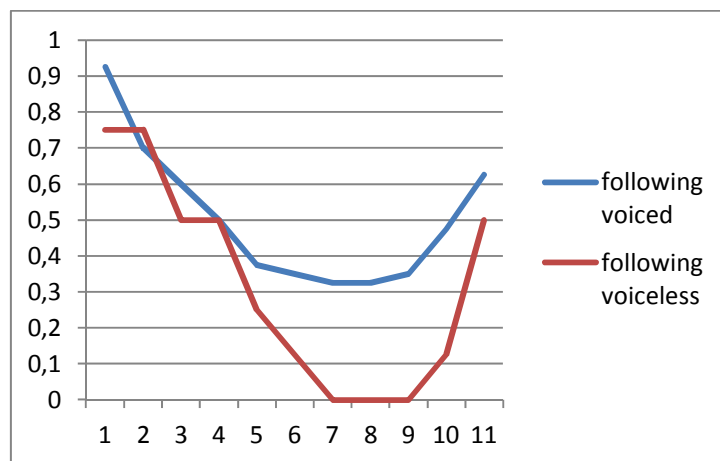


Figure 6.12 The variation of the F<sub>0</sub> presence for affricates.

The last voicing profiles analysis was conveyed to see the variation in the presence of F<sub>0</sub> according to the type of the consonant causing the assimilation, but we only divided them into two rather general groups: voiced obstruents and sonorants. Our results are visible in the figure 6.13:

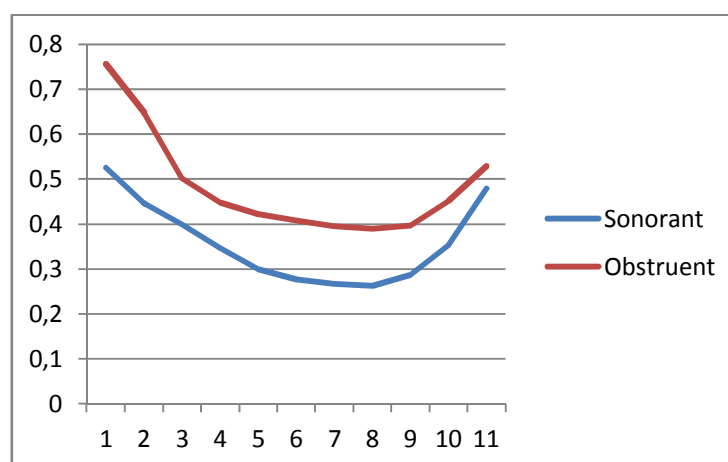


Figure 6.13 The variation of the presence of voicing according to the type of the following consonant.

We can see that the most essential difference lies in the actual influence of the two consonant groups on the preceding segment. Sonorants seem to be less prominent in causing the assimilation than the obstruents in general as the voicing is found in only a bit over 50% of the respective environments with sonorant right context, while for the obstruent right context it is almost 80%. The tendency of the voicing to reoccur towards the end of the pronunciation phase in most of the environments is similar for both groups, it is however more prominent for the sonorants as in the last phase its presence is marked in almost all the environments it was marked at the beginning. The effect of the sonorants therefore again gains prominence

and shows that despite having caused the assimilation of voicing in less environments than the obstruents, the actual assimilation caused by the sonorants seems more stable in comparison with the one caused by the obstruents and the voicing is less likely to disappear in the last phases of the pronunciation of the consonants concerned, on the contrary, it is present in almost all of them.

## 6.6 Individual speakers

In order to acquire a unified picture of the assimilation of voicing tendencies we proceeded with the analysis of the individual speakers profiles. Our results are presented in the figure 6.14 below:

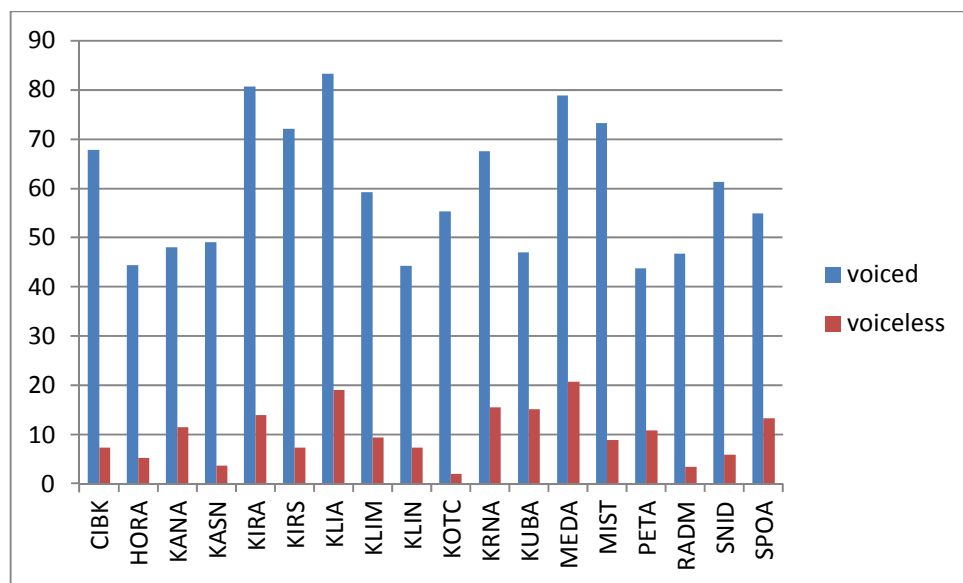


Figure 6.14 The voicing ratios for the individual speakers for both voiced and voiceless following segments.

Concerning the statistical significance, for all of them the t-test proves it highly significant, except for the speakers HORA, KUBA and PETA for whom the results are only statistically significant. As we can see, the individual values differ and what is striking is the difference in the voicing ratio between the speaker KLIA, for whom the assimilation is present in on average 83.3% of the duration of the voiceless segment when followed by voiced, and the speaker PETA, for whom it is only 43.7%. The reasons why there is such a difference might be multiple. Although both speakers have the tendency to glottalize and therefore not to assimilate when the following segment is a vowel, PETA seems to pay much more attention to the pronunciation of the words. Thus, although she binds them correctly, she keeps them in certain isolation and preserves their prescribed pronunciation, mostly without the extensive influence of the following segments. KLIA, on the other hand reads less attentively, tends to

make many mistakes and repeat herself, which increases the possibility of the assimilation emerging.

### 6.7 Homorganic pair consonants

Our second hypothesis concerned the assimilation errors, more specifically the presence of the phenomenon in the environments in which it does not occur in the speech of the native speakers – when the pair consonants meet. Those environments are rather rare and therefore we were able to select only 22 areas in which we found t-d, p-b, θ-ð or tʃ-dʒ. Figure 6.15 shows us that the assimilation in those areas is not uncommon in the speech of the Slovaks:

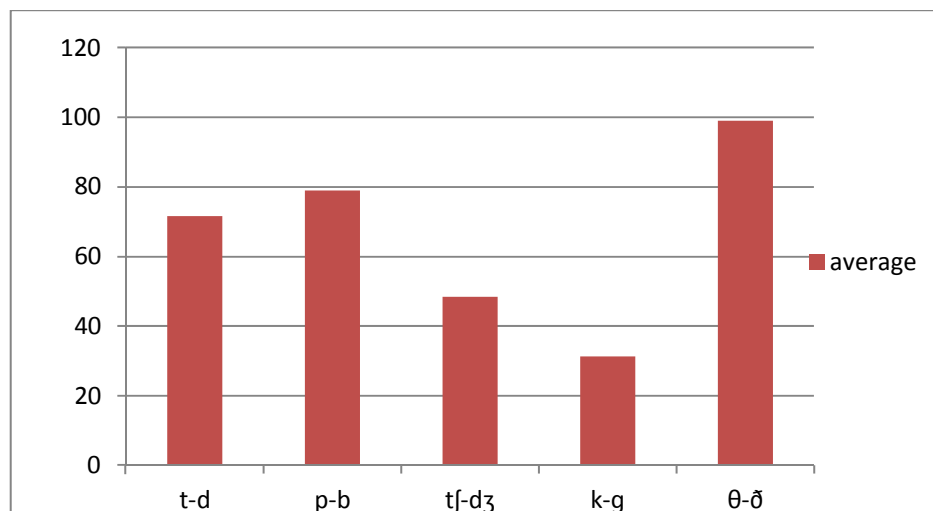


Figure 6.15 The average value of the assimilation taking place in the pair-consonant environments

The least present assimilation percentage is found in k-g environments, while the most noticeable one is in the θ-ð, which may be caused by the dental character of the sounds. For all the pair environments considered, the assimilation emerges on average in 65.78% of the pronunciation of the respective voiceless sounds which supports our hypothesis (H2) and thus we can claim that Slovaks do have the inclination to assimilate extensively, not only in the areas typical for the fluent speech, but also in the areas where it is not required or even natural for the native speakers.



## 7 General discussion and conclusion

The aim of our thesis was to examine the presence of the assimilation of voicing in the speech of the Slovak learners of English as their second language. We concentrated upon the word boundaries environments which we selected from the recordings of eighteen speakers, both males and females, who were reading the BBC bulletins. The two recordings for each of the speakers were further analysed. Both regressive and progressive assimilation environments were selected, however, due to the small amount of the environments in which the progressive assimilation was anticipated, we focused only on those with the regressive type of change. What was considered as well was the ratio of the assimilation of voicing for both male and female groups, together with the possible distinctions caused by the specific type of words – whether they were lexical or grammatical. In order to be able to determine the significance of our results, the statistical t-test and ANOVA test were conveyed.

Our thesis was based on two hypotheses. The first one presumed, according to what we studied about the principles of assimilation of voicing both in Slovak and English that Slovak would excessively assimilate in the word-boundary environments when word-final voiceless and word-initial voiced segments meet. This hypothesis was confirmed. The second hypothesis was built on the presumption that Slovaks would ignore the specific rules of the assimilation in English and would employ it in the uncommon environments for the native speakers, namely when the pair consonants meet in the word-boundary position. This hypothesis was confirmed as well, proving thus that Slovak speakers tend to produce assimilation errors in English.

The overall number of the environments considered in the analysis was 1101. Before the analysis we had to take in account the fact that although we expected to examine only the voiceless-voiced environments, it was necessary to take in account the fact that the Slovaks have the tendency to glottalize at the beginning of the words starting with a vowel. Thus we had two sets of environments: voiceless-voiced and voiceless-voiceless. We determined that the assimilation of voicing in the voiceless-voiced environments took place on the average in 59% of the respective duration of the word-final consonants. What was surprising was the result for the voiceless-voiceless environments, in which we determined that the assimilation of voicing was present on the average for 10.1% of the duration of the word-final segments. The possible explanation is that the segment immediately preceding the word-final phoneme has certain influence upon its voicing as well. In this context we also looked upon the possible differences between male and female speakers which showed no significant distinctions

between them and therefore we can claim that in general the assimilation of voicing is not conditioned by the gender of the speakers.

The next aspect we analysed was the influence of the type of the following word – whether it was lexical or grammatical. The obvious tendency is that the speakers assimilate much more when the following word is grammatical, starting with the voiced phoneme. The reason for this can be that the speakers seem to pay less attention to the grammatical words as they do not carry semantic load as the lexical ones and therefore do not require precise pronunciation as their meaning and distribution are partially predictable from our mental knowledge of the syntactical structure of the sentences in the language.

The creation of the voicing profiles was then conveyed in order to see the development of the voicing in the course of the pronunciation of the respective consonant. For the following voiced context the presence of  $F_0$  varied but it was present in almost 80% of all the environments at the beginning of the pronunciation phase, then the ratio lowered to 60% and towards the end of the pronunciation phase it again occurred in over 70% of the areas concerned. For the following voiceless segments the tendency is similar, but the fundamental frequency at the beginning of the pronunciation phase was present in around 50% of the environments and towards the end raised from almost zero to over 10%. Again this result can be caused by the previous phoneme interference. In order to see the further differences in the individual voicing profiles we conveyed the analysis of the gender distinction which showed one significant difference – the tendency of the voicing to appear towards the end of the pronunciation phase in almost 20% of the environments with following voiceless context, caused primarily by the quick and inattentive reading of the female speakers. As the general picture of the voicing profiles displays the tendencies for all types of the consonants present in this study, the analysis for the individual groups was conveyed in order to see the differences in the distribution of the  $F_0$ . The results proved what we learned from Möbius (2004) for the plosives and fricatives, the result was however rather interesting for the affricates, more specifically for their right voiceless context. We saw the unexpected rise in the presence of the fundamental frequency at the end of the pronunciation phase, where half of the tokens analysed exhibited this tendency. What might have caused it is again the impact of the quick reading and therefore possibly incorrect pronunciation, but here we can also take in account the possibility that the speakers could have anticipated the oncoming voiced vowel and thus certain amount of voicing was present in the final phase of the pronunciation of the word-final consonant, despite having inserted the glottal stop before the vowel as well. The analysis of the voicing profiles according to whether the following consonant is voiced

obstruent or sonorant was then conveyed as well. Here we saw that despite sonorants having caused the assimilation in fewer environments than the obstruents, the number of the environments in which  $F_0$  appeared at the end of the pronunciation phase almost equalled the number at the beginning of the phase. With the right obstruent context we noticed the significant decrease of the environments which were voiced at the end of the pronunciation phase as much as at its beginning. This might suggest that the voicing caused by the sonorants is more stable.

The measures for the individual speakers were conveyed as well, showing a variation of the assimilation tendencies. The speaker who assimilated the most was KLIA (83.3% of the environments with following voiced segment were assimilated) and the one who assimilated the least was PETA (only 43.7% of the environments with right voiced context assimilated). The most probable reason for such a difference is the effort of PETA to keep the distinct pronunciation for the words she read and thus she avoids to a great extent the influence of the surrounding sounds.

The last aspect analysed was the pronunciation and subsequent assimilation of voicing in the environments in which the homorganic pair consonants meet, as we have learned that it is unacceptable to assimilate in those areas for the native speakers of English. The result proved that the Slovaks assimilate excessively here as well – on the average 65.78% of the pronunciation of the respective consonants were voiced. This means that Slovak speakers do not distinguish between the rules of assimilation for Slovak and English and subconsciously employ the rules they know from their mother tongue also in the second language.

The reason why Slovak speakers of English apply what they know from their mother tongue is most probably the result of the language interference, as the Slovaks clearly exhibit the tendency to assimilate in the environments typical for the Slovak language. As for both the languages the places in which the assimilation of voicing occurs are the same, the problem does not have to arise necessarily, although there might be differences in the pronunciation of the consonants concerned as the quality of those sounds does not have to be identical. The problem however arises, when the assimilation takes place where it is not common and thus the error emerges. As the native speakers are likely to recognize it, it is desirable for the speakers who want to achieve higher level in their knowledge of English to eliminate as many pronunciation errors, including the incorrect assimilation, as possible. This could be taken in consideration during the learning process so that the students were aware of such a distinction and therefore would be able to avoid making mistakes in their future interactions with the native speakers.

The research of the assimilation of voicing in the Slovak English is not extensive and our analysis covers only the regressive type. In the future researches it would be desirable to analyse the progressive type as well in order to be able to see the possible differences between the two types, as the sources accessible to us provide no details concerning such an analysis.

## Bibliography

Boersma, P. & Weenink, D. (2014). Praat: doing phonetics by computer, version 5.3.63. [www.praat.org](http://www.praat.org)

Flege et al. (1995). "Second Language Speech Learning. Theory, Findings and Problems." In Winifred Strange (ed) *Speech Perception and Linguistic Experience* (233-277). Timonium, MD: York Press.

Gass, S. M. & Selinker, L. (2008). *Second Language Acquisition: An Introductory Course*. New York and London: Routledge. Third edition.

Kráľová, Z. (2005). *Slovensko-anglická zvuková interferencia*. Žilina: Žilinská univerzita, Fakulta prírodných vied.

Kráľ, Á. (1988). *Pravidlá slovenskej výslovnosti*. Bratislava: Slovenské pedagogické nakladateľstvo. Druhé vydanie.

Leather, J. & James, A. R. (1996). "Second Language Speech." In: Ritchie, W. C. & Bhatia, T. K. (eds.), *Handbook of Second Language Acquisition* (269-316). San Diego: Academic Press.

Lisker, L., Abramson, A.S. (1964). "A cross-language study of voicing in initial stops: acoustical measurements." *Word*, 20, pg. 384-422.

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

Möbius, B. (2004). "Corpus-based Investigations on the Phonetics of Consonant Voicing." *Folia Linguistica*, 38, pg. 5-26.

Roach, P. (1991). *English Phonetics and Phonology*. Cambridge: Cambridge University Press. Second edition.

Sawashima, M., Hirose, H. (1983). "Laryngeal Gestures in Speech Production." In Peter F. MacNeilage (ed) *The Production of Speech*. (11-38). New York: Springer-Verlag.

Shih, C., Möbius, B. (1998). "Contextual Effects on Voicing Profiles of German and Mandarin Consonants." (1-4). In: Proceedings of the Third International Workshop on Speech Synthesis. Jenolan Caves Australia. [www.isca-speech.org/archive](http://www.isca-speech.org/archive)

Sičáková, Ľ. (2002). *Fonetika a Fonológia pre Elementaristov*. Prešov: Náuka. <http://www.unipo.sk/public/media/11339/fonetikaIIIa.pdf>

Smith, C. L. (1997). "The devoicing of /z/ in American English: effects of local and prosodic context." *Journal of Phonetics*, 25, pg. 471-500.

Stevens et al. (1992). "Acoustic and Perceptual Characteristics of Voicing in Fricatives and Fricative Clusters." *Journal of the Acoustical Society of America*, 91, pg. 2979-3000.

Vaissière, J. (1997). "Phonological Use of the Larynx." *Larynx* 1997, Marseille, pg. 115-126.

Volín, J. (2002). *IPA-Based Transcription for Czech Students of English*. Praha: Karolinum.

Volín, J. (2007). *Statistické metody ve fonetickém výzkumu*. Praha: Epoque.

Zygis, et al. (2012). "Phonetic explanations for the infrequency of voiced sibilant affricates across languages." (299-336). Berlin: Centre for General Linguistics. <http://www.haskins.yale.edu/Reprints/HL1715.pdf>

## Zhrnutie

Cieľom tejto bakalárskej práce bolo zamerať sa na výskyt asimilácie znelosti v reči Slovákov hovoriacich anglicky. Náš výskum vychádzal z teoretického základu, v ktorom sme sa venovali popisu dvoch kľúčových termínov pre našu prácu, a to konkrétne znelosti a asimilácie.

V prvej časti teoretického základu sme rozoberali pojem znelosť, jeho definíciu, koreláty a vlastnosti v jednotlivých spoluhláskových skupinách. Znelosť ako taká je súčasťou reči samotnej a preto vzniká v hrtane na základe pôsobenia pohybu vzduchu v hlasivkách. Aby bolo možné odlišiť dýchanie (takisto spôsobené pohybom vzduchu) od tvorenia znelosti, je možné vziať do úvahy tvar hlasiviek v danom momente: ak sa v hlasivkovom aparáte tvorí znelosť, sú uzavreté, naopak ak ide iba o činnosť dýchaciu, hlasivky sú otvorené. Zároveň počas znelosti dochádza k ich vibrácii a podľa Sawashimu (1983) je možné poukázať aj na to, že hlasivky majú tendenciu sa počas tvorenia znelosti uvoľniť. Iný spôsob, ktorý spomína Sawashima (1983), čo nám umožní rozoznať na základe artikulačných parametrov či sa jedná o znelý, alebo neznelý zvuk, je otvorenie/uzavretie hlasivkovej štrbiny (glottis) – konkrétne pre znelé zvuky je hlasivková štrbina uzavretá. Dôležitý aspekt pri určovaní znelosti daného segmentu sú jej koreláty, ako napríklad prítomnosť a kontrola základnej frekvencie, aspirácia, alebo dĺžka predchádzajúceho vokálu. So znelosťou je úzko prepojená funkcia straty znelosti, ktorá sa môže udiť z viacerých dôvodov, avšak Smith (1997) uvádza najmä dva: strata znelosti môže byť spôsobená pôsobením asimilácie na základe susedného fonému, alebo ju môže spôsobovať fakt, že v danom artikulačnom momente je to pre hovoriaceho jednoduchšia možnosť, čo sa týka hlavne frikatív, keďže je nutné, aby bola znelosť vyprodukovaná spolu s frikačným hlukom. V závere prvej časti bolo poukázané na rozdiely v distribúcii znelosti medzi jednotlivými skupinami konsonantov.

V druhej časti teoretického základu sme rozoberali pojem akvizície druhého jazyka (Second language acquisition) a prípadné problémy s tým spojené. Tento proces nie je jednoduchý, naopak, ovplyvňuje ho veľké množstvo faktorov nevynímajúc pri tom individuálne schopnosti každého z nás naučiť sa nový jazyk, pohlavie alebo dokonca naše spoločenské postavenie. Jednu z najdôležitejších úloh pri tomto procese zohráva vek, pretože ako tvrdí Kráľová (2005), schopnosť naučiť sa jazyk na základe sluchového vnímania je najviac rozvinutá v detskej fáze nášho života, a potom postupne klesá. Zároveň deti, ktoré sa útlom veku učia angličtinu ako materinský jazyk, postupne získavajú zvukový repertoár, ktorý sa stáva prvotným a jediným, ktorý ovládajú. Dospelí hovoriaci slovenčiny počas výuky

angličtiny už jeden takýto repertoár ovládajú, a keďže medzi týmito dvoma systémami existujú rôzne odlišnosti, proces osvojovania si nového jazyka sa komplikuje.

Kráľová (2005) v tomto smere pracuje s pojmom *jazyková interferencia*, čo je vzájomný vplyv jazykov jeden na druhý, a zároveň odlišuje pozitívnu a negatívnu interferenciu. Zmieňuje sa o tom, že keď sa jedná o interferenciu materinského jazyka do procesu osvojovania si nového jazyka, hovoríme o negatívnej interferencii. Pri osvojovaní si je potrebné uvedomiť si, že zvukový repertoár nového jazyka sa môže do veľkej miery odlišovať od toho, čo už poznáme. Kráľová (2005) odlišuje konkrétne tri typy zvukov, ktoré môžeme v novom jazyku nájsť: identické s materinskými, zvuky podobné a nové. Dozvedáme sa, že študenti považujú za jednoduchšie naučiť sa zvuky, ktoré sú podobné s tými, čo už poznajú zo svojho materinského jazyka, to však nevylučuje možnosť, že práve u týchto zvukov vznikajú výslovnostné chyby, čo je spôsobené práve nesprávnou kategorizáciou daného zvuku na základe interferencie zo slovenčiny.

Aby sme lepšie pochopili, akým spôsobom si osvojujeme nové zvuky, je treba zaviesť pojmy percepcia a produkcia. Doterajší výskum ukazuje, že teórie zaoberajúce sa týmito dvoma javmi sa líšia v zásadnej veci – a to, či percepciu a produkciu zvukov môžeme vnímať ako prepojený a navzájom súvisiaci proces, alebo či je treba ich vnímať samostatne. Kým Kráľová (2005) vo svojej štúdií tvrdí, že na základe vnímania hlásky, by sme po určitom čase mali byť schopní ju rozpoznať a vysloviť, Leather (1996) nám predstavuje Flegeho et al.(1995) teóriu, ktorá túto schopnosť popiera. Naopak, na to, aby sme boli schopní vysloviť správne novú hlásku jazyka, ktorý sa učíme, je nutné, aby sme boli schopní identifikovať takzvaný fonetický cieľ tejto hlásky, pričom zdôrazňuje, že percepcia a produkcia sú v tomto smere dva nezávislé procesy. Z tejto teórie vychádza teda predpoklad, že napriek tomu, že zvuky podobné s tými, ktoré už hovoriaci pozná z materinského jazyka sa zdajú byť jednoduché pri učení, z hľadiska fonetického môžu byť zdrojom mnohých produkčných chýb, kvôli ich nesprávnemu zaradeniu v systéme.

Takto sa dostávame k chybám, ktoré pri osvojovaní si nového jazyka nutne vznikajú. Chyby vznikajú na základe pôsobenia rozdielov fonetických systémov dvoch jazykov, konkrétne u slovenčiny a angličtiny sa jedná o rozdiely v sadách hlások týchto jazykov. Napríklad slovenčina nepozná dentálne konsonanty [ð], [θ] a problém vzniká aj pri velárnom [ŋ], ktoré v slovenčine neexistuje ako samostatná hláska, ale iba ako alofón [ŋ].

Na úrovni suprasegmentálnej existuje rozdiel medzi charaktermi prízvukov týchto jazykov. Kým v slovenčine je prízvuk pevne ukotvený na prvej slabike slov, v angličtine je



prízvuk pohyblivý, čo častokrát spôsobuje jeho nesprávne umiestnenie zo strany slovenských študentov, a teda nesprávne vyslovenie daného slova.

Jednou z častých chýb u Slovákov hovoriacich anglicky je aj asimilácia znelosti foném, čím sa dostávame k jadrú našej práce. Asimilácia je proces, v ktorom jedna hláska prispôsobí svoj znelostný charakter hláske druhej, najčastejšie susednej. Existujú dva typy asimilácie: progresívna, pri ktorej je hláska ovplyvnená tou predchádzajúcou, a regresívna, pri ktorej je hláska ovplyvnená tou nasledujúcou. Asimilácia veľmi často vzniká práve na hraniciach slov, avšak je bežná aj vo vnútri slova. Základný rozdiel medzi slovenčinou a angličtinou spočíva v tom, že kým slovenčina bez problémov asimiluje v okoliach, kde sa stretne znelá a neznelá hláska, vrátane okolí, kde sa stretnú homorganické párové spoluhlásky, angličtina asimiluje len v tom prvom prípade, kým v tom druhom je asimilácia neakceptovateľná. Naopak angličtina asimiluje [l], [w], [r] a [j] po fortisovej párovej spoluhláske, kým slovenčina asimiláciu v tomto prípade nepozná. Z toho nám teda vyplýva, že Slováci by mali asimilovať v prípadoch, ktoré angličtina neuznáva, a zároveň neasimilovať v prípadoch, kde je to v angličtine potrebné.

Na základe týchto teoretických znalostí sme sformulovali dve hypotézy, na ktorých sme postavili náš ďalší výskum.

H1: Slováci budú asimilovať v okoliach na hranici slov, kde sa stretne neznelý a znelý konsonant, alebo vokál.

H2: Slováci budú asimilovať v okoliach, kde sa stretnú neznelý a znelý párový konsonant.

Metóda spočívala v tom, že sme analyzovali nahrávky 18 Slovákov, z toho bolo 9 chlapcov a 9 dievčat, ktorí čítali BBC správy, rozdelené do malých úsekov. Od každého hovoriaceho sme spracovali dve nahrávky. Potom bolo nutné vysegmentovať v Praate okolia, v ktorých sme očakávali asimiláciu znelosti, ktoré sme potom za pomoci skriptu izolovali a rozanalyzovali v Exceli. Ešte pred analýzou sme však museli vylúčiť okolia, kde sa nachádzal koncový znelý konsonant a začiatkový neznelý na hranici slov, kvôli nízkemu počtu dát, takže sme spracovávali iba okolia s neznelým koncovým konsonantom a nasledujúcim znelým konsonantom alebo vokálom. Pri spracovávaní našich výsledkov sme určovali ich význam štatistickým t-testom a testom ANOVA.

Všeobecné výsledky nám ukázali, že Slováci naozaj do veľkej miery asimilujú v nami vybraných okoliach, čím sme potvrdili našu prvú hypotézu (H1). Celkový počet spracovaných okolí bol 1101. Tu sme museli vziať do úvahy fakt, že Slováci do veľkej miery glotalizujú

v slovách, ktoré sa začínajú na vokál, a preto sme pracovali nie len s okoliami, kde sa stretli neznelý a znelý segment, ale aj s takými kde sa stretli neznelý konsonant a ráz. Výsledky nám ukázali, že znelostná asimilácia prebieha v priemere v 59% trvania koncových konsonantov v okoliach, kde nasledoval znelý segment. Prekvapivo, asimilácia sa objavila aj v neznelých okoliach, a to v priemere v 10,1% celkového trvania segmentov, čo mohlo byť spôsobené preznievaním predchádzajúcej hlásky. Zároveň sme tu skúmali možné rozdiely medzi pohlaviami, no tie sa neukázali byť významné, a preto je možné skonštatovať, že pohlavie na výslovnosť znelostnej asimilácie nemá veľký vplyv.

V ďalšej sekcii sme analyzovali, aký vplyv na asimiláciu môže mať typ nasledujúceho slova, konkrétne či sa jedná o slovo plno alebo neplnovýznamové. Výsledky nám ukázali, že ak sa jedná o slovo gramatické, začínajúce na znelú hlásku, je tendencia asimilovať väčšia, než v prípade, že sa jedná o slovo lexikálne, alebo začínajúce rázom – to pravdepodobne spôsobilo, že vo všeobecnosti máme tendenciu klásť väčší dôraz na slová nesúce lexikálny význam, kým slová gramatické vyslovujeme menej dôsledne.

Aby sme zistili vývoj znelosti v jednotlivých cieľových hláskach, vytvorili sme znelostné profily, v ktorých sme sledovali prítomnosť základnej frekvencie ako jedného z korelátov znelosti. Znovu sme analyzovali celkové dáta pre nasledujúce znelé a neznelé segmenty, ďalej sme sa venovali znelostným profilom rozdeleným podľa pohlavia, a vypracovali sme jednotlivé profily aj pre samostatné skupiny konsonantov. Videli sme, že všeobecne ak nasleduje znelý segment, základná frekvencia sa objavuje u 80% okolí, ale iba na začiatku výslovnostnej fázy. Počet okolí, ktoré  $F_0$  obsahujú potom klesne, a znova stúpne ku koncu výslovnostnej fázy, približne na 70%. Podobný vývoj sledujeme aj pri nasledujúcom neznelom segmente. Čo sa týka rozdielu medzi pohlaviami, hlavný rozdiel pozorujeme pri nasledujúcom neznelom segmente, kde u žien stúpne prítomnosť  $F_0$  ku koncu výslovnostnej fázy viac ako u mužov. Ďalšia analýza sledujúca vývoj prítomnosti  $F_0$  v jednotlivých skupinách konsonantov neponúka prekvapujúce výsledky u plozív a frikatív, naopak sledujeme prudký nárast  $F_0$  ku koncu výslovnostnej fázy u afrikátov pre nasledujúci neznelý kontext. Poslednú analýzu znelostných profilov sme vypracovali na základe rozdielu typu nasledujúceho konsonantu, v tomto prípade sme ich však rozdelili iba na znelé obštruenty a sonóry. Ukázalo sa, že napriek tomu, že sonóry spôsobujú asimiláciu iba u približne polovice skúmaných okolí (narozdiel od obštruentov, u ktorých sa číslo pohybuje okolo 80%), zaznamenali sme, že výskyt základnej frekvencie v daných okoliach je pri nasledujúcej sonóre takmer zhodný na začiatku a na konci výslovnostnej fázy, s poklesom výskytu v strednej časti výslovnosti. U obštruentov sme naopak zaznamenali prudký pokles počtu okolí, u ktorých sa

F<sub>0</sub> objavila aj na konci výslovnostnej fázy. Z toho vyplýva, že asimilácia znelosti spôsobená sonórou je stabilnejšia, než tá spôsobená obštruentom.

Ďalšia analýza sa zaoberala rozdielmi medzi jednotlivými hovoriacimi. Ukázal sa veľký rozdiel medzi hovoriacou, ktorá asimilovala najviac a tou, ktorá asimilovala najmenej: 83,3% ku 43,7%. Pravdepodobným dôvodom, ktorý tento rozdiel spôsobil bola nesprávna výslovnosť hovoriacej, ktorá asimilovala najmenej, vzhľadom na to, že pri čítaní slová izolovala (aj keď nerobila pauzy).

Posledná analýza bola urobená, aby sme poukázali na asimilačnú chybu u Slovákov v okoliach, kde sa stretnú homorganické párové konsonanty. Potvrdili sme, že Slováci asimilujú do veľkej miery aj v týchto okoliach (v priemere 65,78% celkového trvania cieľových konsonantov je asimilovaných), čím sme potvrdili aj našu druhú hypotézu (H2).

Z nášho výskumu sme zistili, že asimilácia znelosti je pre slovenských hovoriacich v angličtine bežným javom, ktorý však uplatňujú aj v okoliach, ktoré angličtina z asimilačného hľadiska nepozná. Takto vzniká chyba, ktorá je pravdepodobne spôsobená práve jazykovou interferenciou. V tomto prípade sa však nejedná o nesprávne vyslovenie fonémy kvôli nesprávnemu určeniu fonetického cieľa danej hlásky, ale vo veľkej miere je táto chyba výsledkom nevedomosti slovenských študentov o nesprávnosti takejto výslovnosti. Vzhľadom na to, že je možné, že rodení hovoriaci túto chybu v reči cudzincov počujú, je vhodné, aby bolo na ňu poukázané počas edukačného procesu študentov angličtiny, ktorí chcú dosiahnuť vysokú úroveň v znalosti tohto jazyka a tak sa priblížiť reči rodených hovoriacich. Náš výskum však pokrýva iba regresívny typ asimilácie, a na to, aby sme sa o tomto procese a zároveň aj vznikajúcej chybe dozvedeli viac, je treba preskúmať aj progresívny typ, aby boli vytvorené čo najvhodnejšie podmienky pre ďalšiu štúdiu týchto javov.