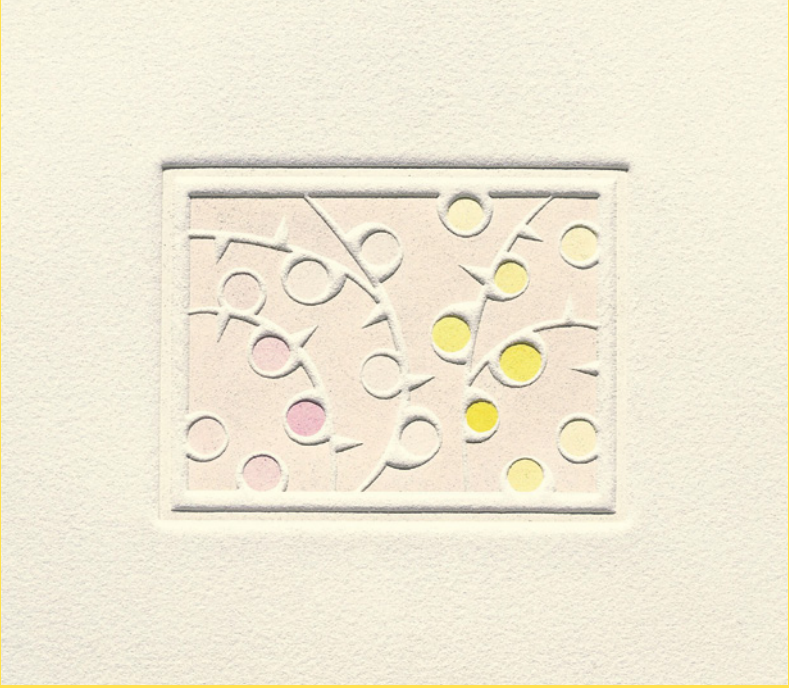
ELOPE
English Language Overseas Perspectives and Enquiries



Vol. 16, No. 1 (2019)

The Sounds of English

Guest Editors: IVO FABIJANIĆ and ANDREJ STOPAR

Journal Editors: SMILJANA KOMAR and MOJCA KREVEL

Ljubljana University Press, Faculty of Arts
Znanstvena založba Filozofske fakultete Univerze v Ljubljani

Ljubljana, 2019
ISSN 1581-8918

English Language Overseas Perspectives and Enquiries

Vol. 16, No. 1 (2019)

The Sounds of English

Guest Editors: IVO FABIJANIĆ and ANDREJ STOPAR

Journal Editors: SMILJANA KOMAR and MOJCA KREVEL

University of Ljubljana
FACULTY OF ARTS

Ljubljana University Press, Faculty of Arts
Znanstvena založba Filozofske fakultete Univerze v Ljubljani

Ljubljana, 2019

CIP - Kataložni zapis o publikaciji
Narodna in univerzitetna knjižnica, Ljubljana

811.111'243(082)
37.091.3:81'243(082)

The SOUNDS of English / guest editors Ivo Fabijanić and Andrej Stopar. - Ljubljana : University Press, Faculty of Arts = Znanstvena založba Filozofske fakultete, 2019. - (ELOPE : English language overseas perspectives and enquiries, ISSN 1581-8918 ; vol. 16, no. 1)

ISBN 978-961-06-0223-1
1. Fabijanić, Ivo
COBISS.SI-ID 300676096

Contents

Part I

Introduction

The Sounds of English: Introduction

This volume of ELOPE is dedicated to phonetics and phonology. The seven papers contributed by nine authors present a selection of topics that reflect the developments in these traditional fields of linguistics that have been (re)gaining momentum in recent years.

The idea for the issue stems from the 14th conference of the European Society for the Study of English (Brno, Czech Republic, 29 August–2 September 2018), where the guest editors of this issue convened the seminar group "The Sounds of English: Cross-Linguistic and Cross-Cultural Aspects". The seminar participants who contributed their articles are joined by other researchers working in the field.

The papers in this issue are organized alphabetically, by authors.

Biljana Čubrović (University of Belgrade, Serbia) presents a study that examines how non-native speakers of American English with Serbian as their mother tongue produce L2 vowels. She identifies vowel duration as an important feature that non-native speakers rely on in their production of American English vowels.

Nataša Hirci (University of Ljubljana, Slovenia) writes about the role of pronunciation in speech technologies used by translation professionals. Her findings show that translation students see good pronunciation as an important feature of their skillset that also enables them to use speech recognition tools more easily.

Saša Jošt and Andrej Stopar (University of Ljubljana, Slovenia) assess the perception of foreign language phonemes by devising an experiment that tests several phonemic contrasts. Their findings identify some of the most problematic pairs for Slovene L1 speakers, while showing that in the process of acquiring non-native phonemes, Slovene students tend to prioritize vowels over consonants.

Oleksandr Kapranov (Western Norway University of Applied Sciences, Norway) presents a mixed-method study that explores the sounds of English that pre-service and in-service teachers of English as a foreign language identify as challenging. While the results of self-assessment in the two groups are similar, they do not necessarily match the participants' errors in transcription tasks. The study also lists some problematic phoneme contrasts relevant for Norwegian speakers of English.

Smiljana Komar (University of Ljubljana, Slovenia) studies the oral production of General British phonemes in connected speech as read by Slovene EFL students. Her paper compares the performance of the students before and after their systematic training in English pronunciation. The findings confirm the influence of their L1 and the positive effects of explicit instruction in phonetics.

Emilija Mustapić and Frane Malenica (University of Zadar, Croatia) discuss the relationship between sign languages, spoken languages, and co-speech gestures. They examine the phonology of spoken and sign languages, while providing an insight into the basic features of co-speech gestures in order to explore how these means of communication convey or complement the meaning.

Serkan Şen (Başkent University, Ankara, Turkey) provides a comparative analysis of Turkish and English with respect to their vowel systems and orthography. He identifies the most problematic non-native vowels for Turkish speakers of English and devises a set of activities that address their production and perception.

The guest editors would like to thank the authors for their contributions, and the reviewers for their insightful comments and suggestions.

Ivo Fabijanić, University of Zadar, Croatia

Andrej Stopar, University of Ljubljana, Slovenia

Guest Editors of ELOPE Vol. 16, No. 1 (2019)

Part II

Articles

Biljana Čubrović
University of Belgrade, Serbia

2019, Vol. 16 (1), 15–28(164)
revije.ff.uni-lj.si/elope
doi: 10.4312/elope.16.1.15–28
UDC: 811.111(73)'243:81'342.41

Duration as a Phonetic Cue in Native and Non-Native American English

ABSTRACT

This vowel study looks at the intricate relationship between spectral characteristics and vowel duration in the context of American English vowels, both from a native speaker (NS) and non-native speaker (NNS) perspective. The non-native speaker cohort is homogeneous in the sense that all speakers have Serbian as their mother tongue, but have been long-time residents of the US. The phonetic context investigated in this study is /bVt/, where V is one of the American English monophthongs /i ɪ u ʊ ɛ æ ʌ ɔ ɑ/. The results of the acoustic analysis show that the NNS vowels are generally longer than the NS vowels. Furthermore, NNSs neutralise the vowel quality of two tense and lax pairs of vowels, /i ɪ/ and /u ʊ/, and rely more heavily on the phonetic duration when prononuncing them.

Keywords: vowel; duration; vowel quality; native; non-native; American English

Trajanje kot razločevalna fonetična lastnost v ameriški angleščini rojenih in tujih govorcev

POVZETEK

Pričujoča študija samoglasnikov obravnava zapleteno razmerje med spektralnimi značilnostmi in trajanjem samoglasnikov v ameriški angleščini z vidika rojenih in tujih govorcev. Skupina obravnavanih tujih govorcev je homogena, saj so vsi rojeni govorci srbščine, ki že daljše obdobje živijo v ZDA. Študija proučuje fonetični kontekst /bVt/, kjer je V eden od enoglasnikov v ameriški angleščini /i ɪ u ʊ ɛ æ ʌ ɔ ɑ/. Izsledki akustične analize pokažejo, da so samoglasniki tujih govorcev običajno daljši od samoglasnikov domačih govorcev. Sledi ugotovitev, da tuji govorci nevtralizirajo kakovostno razliko med pari samoglasnikov /i ɪ/ in /u ʊ/, in da se pri izgovarjavi bolj zanašajo na fonetično trajanje.

Ključne besede: samoglasnik; trajanje; kakovost samoglasnikov; tuji govorci; rojeni govorci; ameriška angleščina

1 Introduction

The vowels of American English (AE) differ in their durations, and the different durations are often said to be phonetically realised as tense or lax. Even though vowel duration in American English has been thoroughly explored, it still offers a fine terrain for further research in several different niches. First, there is intrinsic duration that is studied by Black (1949) and Lehiste and Peterson (1961), who found that open vowels are longer than the vowels produced with a more close jaw opening. To illustrate this point, we can say that the American English vowel /ɑ/, that is more open, is intrinsically longer than the close vowel /i/. Another important and well-investigated phonetic characteristic of English vowels pertains to the duration that is brought into connection with the voicing of the following consonant (House and Fairbanks 1953; House 1961; Chen 1970). The vowel of *bead* is realised as longer than the vowel of *beat*, for instance. The lack of voicing in the following consonant shortens the vowel duration, all else being equal.

Second, some extralinguistic factors may influence vowel duration: vowels produced by female speakers are generally longer than the vowels produced by male speakers (Hillenbrand et al. 1995; Holt, Jacewicz, and Fox 2015). Sociolinguistic influences may also have impact on vowel duration in English. In a recent study of Southern African-American English compared to White American English, Holt, Jacewicz and Fox (2015) claim that African-American speakers in the same geographical area as White Americans produce longer vowels.

2 Earlier Research on the Duration of American English Vowels

One of the early studies that aims at establishing the relationship between frequency, intensity and duration of vowels characteristic of the reading style in American English was carried out by Black (1949). We will here only point to the results that Black (1949) offered in relation to the phonetic duration. A sample of 16 male speakers was analysed in a voiceless phonetic environment, where 11 English vowels were recorded in isolation at a specific pace in the context where the vowel is preceded by /t/ and followed by a voiceless /p/, in short /tVp/. This article provides a good comparison ground against the corpus specifically designed for the purposes of the current study that utilises a similar phonetic context, namely /bVt/. Table 1 provides the averages of the phonetic duration for 9[1] vowels, which are read from a magnetic tape.

Table 1. Mean values of vowel duration produced by 16 male speakers (adapted and taken from Black (1949, 218)).

Vowel	i	ɪ	ɛ	æ	ʌ	u	ʊ	ɔ	ɑ
Mean duration in ms	159	135	153	208	154	200	153	209	192

[1] We only give data for the 9 AE vowels that are the subject of the current study.

Peterson and Lehiste (1960) studied the influence of the consonants that precede and follow an AE vowel (a monophthong or diphthong). They used two sets of data, and here we will present the measurements only of the larger set that involves 30 monosyllabic minimal pairs and 10 additional disyllabic minimal pairs (Peterson and Lehiste 1960, 693–94). Table 2 systematises data for the vowels that are the subject of the current study.

Table 2. Mean values of vowel duration produced by five speakers (adapted and taken from Peterson and Lehiste (1960, 702)).

Vowel	**i**	**ɪ**	**ɛ**	**æ**	**ʌ**	**u**	**ʊ**	**ɔ**	**ɑ**
Mean duration in ms	240	180	200	330	N/A	260	200	310	260

Peterson and Lehiste (1960) delved further into every individual consonantal context looking at the magnitude of influence of the preceding and following consonants. They arrived at the conclusion that the impact of the preceding consonant on the duration of the syllable nucleus is negligible, but the following consonant has a significant impact on the duration of the vowel. Table 3 provides the mean durations of short and long nuclei in different consonantal contexts. We have singled out only the stop consonants that close the syllable in question. It is evident that the ratio of the duration of vowels in front of voiceless and voiced consonants is approximately 2:3 for American English in favour of voiced segments. The importance of Peterson and Lehiste's research study lies in the fact that they systematically examined all phonetic environments and came to a robust conclusion that voicing is an important cue in distinguishing long and short vowels of AE.

Table 3. Mean values of short and long nuclei as a function of English stops (adapted and taken from Peterson and Lehiste (1960, 702)).

	/p/	**/b/**	**/t/**	**/d/**	**/k/**	**/g/**
Short	138	203	147	206	145	243
Long	188	307	210	318	200	314

House's study (1961) contributes to the overall phonetic research in the sense that he investigated vowel durations across phonetic contexts, similar to Peterson and Lehiste (1960). Each of the 12 AE vowels is followed either by a voiced or voiceless consonant (three stops, one affricate and three fricatives) in the speech of three male talkers. The vowel speech sound occurs in a stressed syllable of a disyllabic nonsense word. House (1961) provides duration ratios for different phonetic contexts (voiced vs. voiceless consonants that affect vowel durations), different characters of vowels (tense vs. lax), and compares cumulative durations for groups of vowels (close/mid/open tense or lax) in different consonantal contexts (stop, affricate, fricative). House (1961, 1176) finds

that there is a systematic progression in vowel durations depending on vowel features and phonetic environment. Table 4 gives an overview of values resulting from his study.

Table 4. Mean duration in ms (rounded to the nearest 10 ms) after voiceless consonants (taken and adapted from House (1961, 1176)).

	Tense	**Lax**
Close	150	120
Mid	170	
Open	220	150

The longest vowel duration is observed in open tense vowels followed by voiced fricatives – 400 ms (stops shorten the vowel duration most, all else being equal, and affricates are positioned in between these two classes of consonants). However, close lax vowels followed by a voiceless stop are characterised by the shortest duration that is only one quarter of that mentioned above, at 100 ms.

Table 5. Mean duration in ms (rounded to the nearest 10 ms) after voiced consonants (taken and adapted from House (1961, 1176)).

	Tense	**Lax**
Close	320	220
Mid	350	
Open	360	250

Hillenbrand, Clark and Houde (2000) came to interesting findings in their study of the effects of duration on vowel recognition in American English. The pairs of vowels including /i–ɪ/, /u–ʊ/ and /ɪ–e–ɛ/ are minimally affected by duration because their spectral features are different enough to distinguish between them, unlike /ɑ–ɔ–ʌ/ and /e–æ/ that are significantly affected by variable duration.

This paper looks into the vowel duration as a cue that plays a role in distinguishing between American English vowels as produced by two groups of speakers, native and non-native. To the best of our knowledge, this is the first study of its kind that compares the American English spoken by Serbian expatriates with native speakers of this variety. The Serbian vowel inventory is traditionally described as one of the commonest vowel systems: a five-vowel system that comprises /i e a o u/. These five vowels are combined with four pitch accents (long and falling, long and rising, short and falling, and short and rising). Some more recent approaches, starting with Jakobson (1937 [1962]), propose a novel approach to the Serbian vocalic system, where quantity and pitch are factored out as two distinct dimensions. Such a view was adopted in several other studies on Serbian pitch accent (Browne and McCawley 1965, Inkelas and Zec 1988). We will adopt the latter approach in this vowel study and regard Serbian as a quantity

language following Lehiste (1970), who claims that short and long vowels may also differ in their spectral characteristics (see also Čubrović 2016, 26–29).

3 Experiment

3.1 Participants

Ten native speakers of Serbian who live in the United States and five native speakers of American English took part in the experiment. All ten participants are male.

At the beginning of the recording session, each participant was required to fill in a questionnaire. The Serbian participants were asked to report the length of residence (LOR) in the United States and language(s) spoken at home. The Serbian participants were also asked to rate their own English fluency on a scale (1–5, 5 being the highest) at the time of relocation from Serbia and at the time of the recording in the States. All ten participants were born in Belgrade, Serbia (except for one participant who was born in the south of Serbia, but lived in Belgrade for 27 years prior to moving to the US), and continued to live in Belgrade until they moved to the States. They all live in Atlanta, GA, and their age ranges from 35–44. Nine of them had lived in Atlanta for more than 12 years at the time of the recording. Seven out of ten speakers mostly speak Serbian at home, and all participants use exclusively English at work.

Native speakers of English were asked to report on their place of residence and languages spoken. All five lived in the North-East of the United States at the time of the recording. Three were undergraduate students at Cornell University, Ithaca, NY, and two were employees (former and present) of the same university. Table 6 summarises this information.

Table 6. Background information on participants.

Subject	Sex	Age	L2 fluency (self-evaluated) then/now	Place of residence	Length of residence	Language(s) spoken at home
NNS						
GV	M	40	4/5	Atlanta, GA	12	Mostly Serbian
SG	M	41	3/5	Atlanta, GA	23	Serbian/English
MR	M	40	2/5	Atlanta, GA	14	Mostly Serbian
MS	M	40	1/4	Atlanta, GA	15	Mostly English
IS	M	44	2/4	Atlanta, GA	15	Mostly Serbian
NC	M	37	2/4	Atlanta, GA	16	Mostly Serbian
VG	M	42	2/3	Atlanta, GA	8	Mostly Serbian
NN	M	36	2/4	Atlanta, GA	13	Mostly Serbian
MP	M	35	N/A	Atlanta, GA	14	English/Spanish
UZ	M	45	3/4	Atlanta, GA	16	Mostly Serbian

NS						
MB	M	19	/	New York City, NY	/	English
TC	M	70	/	Ithaca, NY	/	English, some French
PI	M	19	/	Pittsburgh, PA	/	English
MI	M	73	/	Ithaca, NY	/	English
NI	M	20	/	Cortland, NY	/	English

3.2 Materials and Recording Procedures

The acoustic experiment targets nine vowels of American English (AE) in the following monosyllabic words: *beat*, *bit*, *bet*, *bat*, *but*, *boot*, *put*, *bought* and *pot*. The words were all embedded in the frame sentence "Say ___ again", and repeated three times in a random order, giving a total of 270 (10 speakers x 3 repetitions x 9 vowels) tokens for Serbian NSs and 135 (5 speakers x 3 repetitions x 9 vowels) tokens for English NSs, totalling 405 repetitions.

All NNS recordings were made using Sennheiser noise-cancelling headphones and a Sony laptop computer running Praat, Version 5.3.51 (Boersma and Weenink 2013). The NSs of American English were recorded in a sound-attenuated booth in the Cornell University Phonetics Lab. Participants were given sets of sentences in a Power Point presentation, and only one sentence was presented on a slide at a time. They were also given the opportunity to familiarise themselves with the sentences before the recording started. After they had got acquainted with the materials, the participants were instructed to read the sentences "as naturally as possible".

3.3 Analysis and Discussion

The recordings were digitised at 22,000 Hz and analyzed using the Praat software for acoustic analysis of speech (Boersma and Weenink 2013). All elicited materials were first manually labelled and vowel segmental acoustic features measured with the use of a script written by DiCanio (2013). This script generated eight acoustic measures: vowel duration, F1, F2, F3, centre of gravity, standard deviation, skewness, and kurtosis. Even though duration is in the focus of the present study, F1 and F2 will also be shown in order for vowels to be fully analysed.

The vowels will be analysed in groups in the following sections – /i ɪ/, /u ʊ/, /ɛ æ ʌ/, and /ɔ ɑ/.

3.3.1 High Vowels /i ɪ/

The first pair of vowels are those of *beat* and *bit*. In AE, they are most often described as tense and lax, respectively. NSs clearly differentiate them by vowel quality, which is shown in Figure 1.

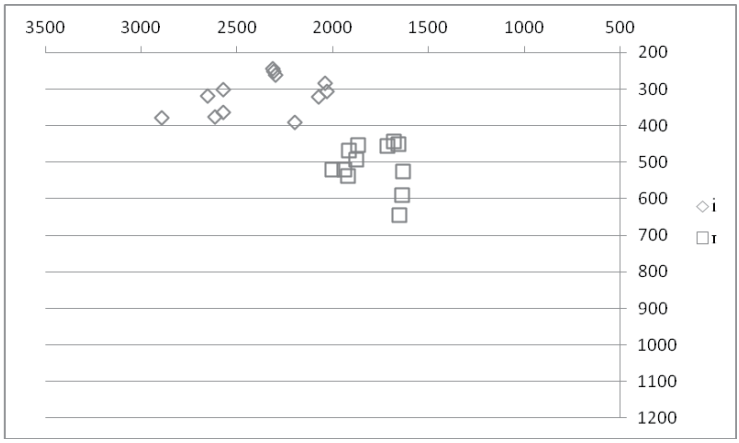


Figure 1. NS /i ɪ/.

NNSs have a tendency to merge /i/ and /ɪ/. This phenomenon may be accounted for by the fact that Serbian language background speakers rely heavily on the phonetic duration when distinguishing between these two vowels. They transpose this phonetic property from L1 (Serbian) into L2. The merger is not observed in speaker SG, who has the longest LOR in the States (23 years). Similarly, the acoustic characteristics of /ɪ/ of speaker MS, whose LOR is 15 years, approximate the NSs production of this vowel.

The duration measurements for the native and non-native participants are looked at next. The mean value for the tense vowel /i/ in the native speaker group is 112.5 ms (SD 15.1 ms). Its lax counterpart has a mean value of 88.6 ms (SD 14 ms) in the same group of participants.

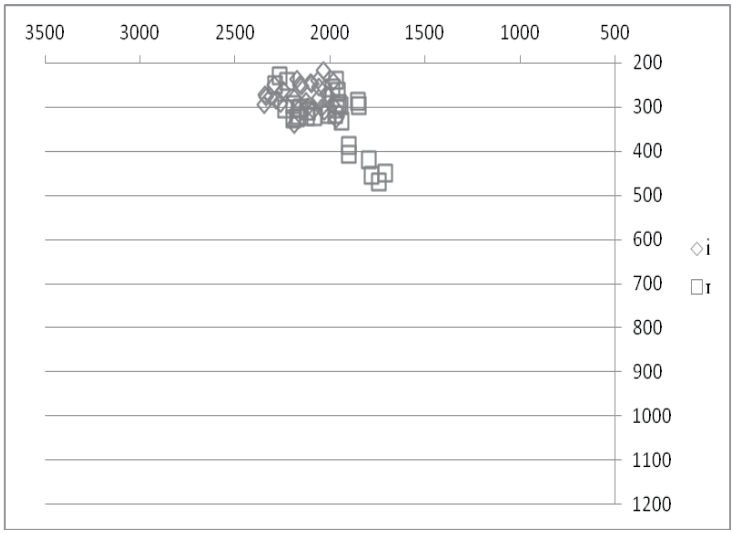


Figure 2. NNS /i ɪ/.

The non-native speakers' tense vowel is significantly longer at 132.51 ms and with a larger SD (30.05 ms). The non-native speakers' lax vowel duration is only marginally different from the native speakers' average duration measurement at 87.6 ms (SD 20.74 ms).

It seems worth noting the standard deviations that are significantly higher in the non-native speaker group, which points to vowel duration instability in this participant group.

3.3.2 High Vowels /u ʊ/

The next pair of vowels are the vowels of *boot* and *put*. In AE, /u/ is tense and /ʊ/ is lax. NSs clearly separate the two vowels in the vowel area, which is shown in Figure 3. For native speakers of English there is no overlapping of /u/ and /ʊ/ when vowel quality is concerned.

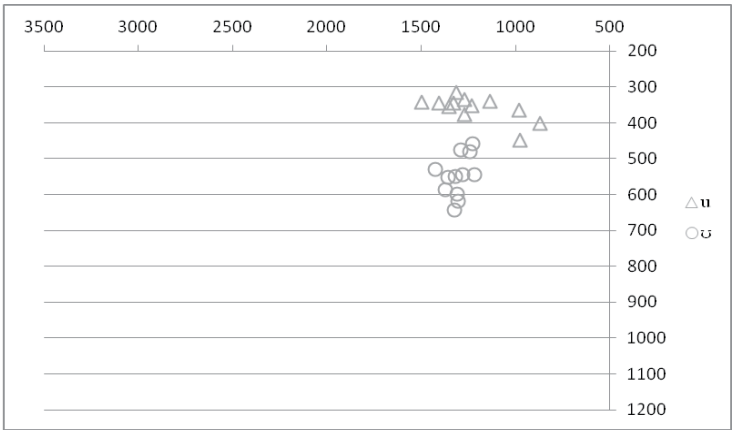


Figure 3. NS /u ʊ/.

NNS vowels are realised differently and their tendency to merge the two vowels is reported in this vowel study. The nonnative speaker participants consistently produce F1 of /ʊ/ with a lower frequency. This makes the lax /ʊ/ a higher vowel and closer to /u/ in the NNS group. NNSs obviously disregard the quality difference between the two English vowels and rely more on the phonetic duration, similar to the /i/–/ɪ/ pair, which is displayed in Figure 4 below.

As for the duration measurements, NS tense /u/ is on average 125 ms long (SD 15.8), while its lax counterpart is 73.8 ms long (SD 9.9 ms). NNSs tend to produce the tense vowel as even longer. The mean value of /u/ in the NNS group amounts to 154.5 ms (SD 35.9 ms). The lax /ʊ/ is realised as significantly shorter, and its average value in the NNS group is 80.3 ms (SD 12). To conclude, NNSs seem to rely on the phonetic duration when producing the AE tense /u/. The SD for the

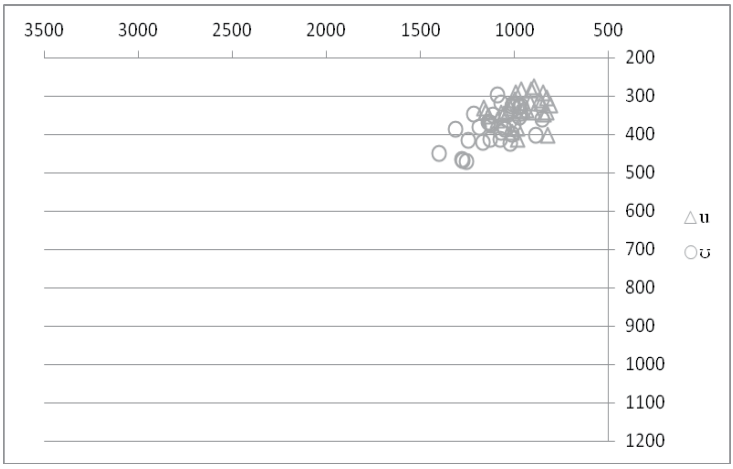


Figure 4. NNS /u ʊ/.

NNS vowel is also larger when compared to the NS realisation of this vowel which points to the vowel duration instability and hesitation on the part of the NNSs.

3.3.3 The Vowels /ɛ æ ʌ/

In the next group of vowels, we first investigate the phonetic characteristics of two vowels /ɛ/ and /æ/. It has been noted that there is more variation in the acoustic vowel space for /ɛ/ and /æ/ even in the group of NSs. There is a general tendency, though, for /ɛ/ to be produced with a lower F1, which makes it a higher vowel than /æ/ in AE native speech. The three tokens of all three vowels that have consistently lower values of F1 are all produced by speaker MB. This may be due to his vocal tract length, which can be longer in tall people. This reduction in F1 values in one speaker may be seen as his idiosyncratic characteristic.

In order to provide a comprehensive overview of AE monophthongs in this paper, the vowel /ʌ/ is also displayed on Figure 5, but it clearly does not overlap with /ɛ æ/ in the native speaker group.

The NNS participants in this study tend to merge /ɛ/ and /æ/, i.e. they do not clearly differentiate between the two. The vowel /æ/ is a new sound to Serbian language speakers. A similar finding can be observed in other Slavic languages. Slovenian learners of English, for example, find this vowel contrast in BE the most challenging (Stopar 2015, 89; Komar 2017, 163). However, there are nine tokens of the *bat* vowel in Figure 6 that have higher F1 values and they are repetitions of three speakers, SG, NN and UZ. Two of these participants have a relatively long LOR in the States, of 23 and 16 years. The speaker NN has lived in the States for 13 years, which is slightly under the mean value for LOR in the study (13.5 years). These nine tokens of *bat* approximate the NS spectral characteristics.

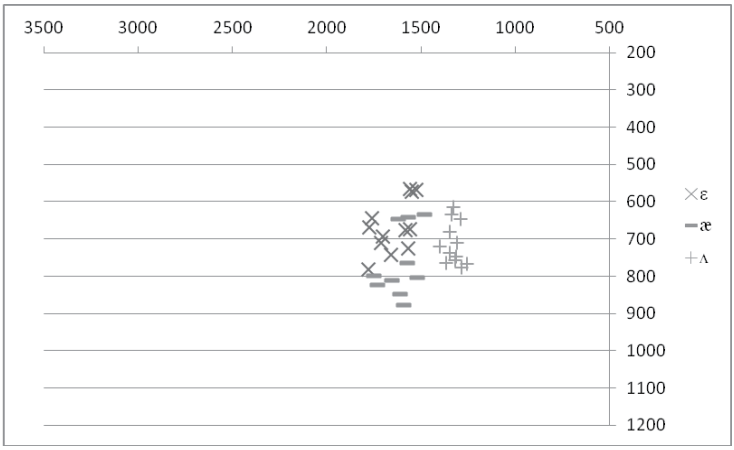


Figure 5. NS /ɛ æ ʌ/.

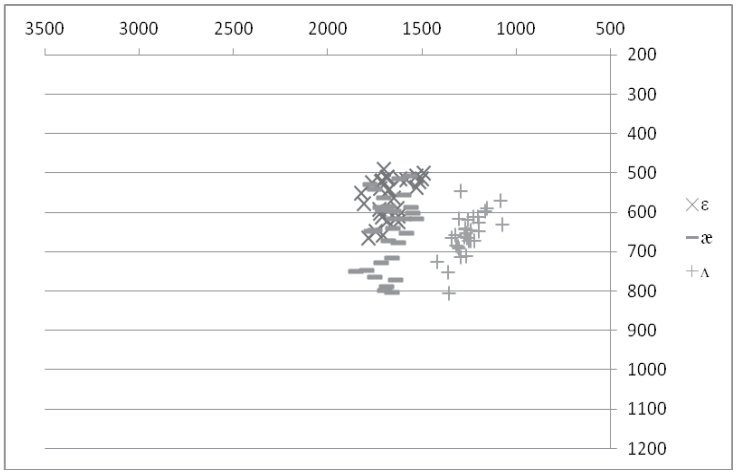


Figure 6. NNS /ɛ æ ʌ/

As for vowel duration, the three English vowels are given here in ascending order, from the shortest to the longest: /ɛ æ ʌ/. The average values for the NS vowels are consistently shorter than for the NNSs' ones. As expected, the standard deviations are larger in the NNS group. All duration measurements for the three vowels are given in Table 7.

Table 7. Duration measurements for /ɛ æ ʌ/.

NS	**ɛ**	**æ**	**ʌ**
Mean	110.4	161.6	100.6
SD	10.7	13.6	11.7
NNS	**ɛ**	**æ**	**ʌ**
Mean	120.8	173.1	116.2
SD	16.3	31.1	21.8

3.3.4 The Vowels /ɔ ɑ/

Both /ɔ/ and /ɑ/ are described as back vowels in AE. Neutralisation of these two vowels is observed in many regions of the US and Canada, and is known as *cot-caught* merger. As a result of this phonological process, the two vowels become one (see Čubrović 2017; 2018).

The NS participants of this study mostly differentiate the vowels in question, i.e. the vowels in the words *bought* and *pot*. The NNSs utilise the same speech habit as NSs. However, the /ɔ/ vowel is articulated with a lower F2 by most subjects in the NNS cohort. This implies that the NNS /ɔ/ is a more back vowel than in native AE speech.

The vowel of *pot* shares the same phonetic characteristics as /ɑ/ in the two groups of participants of this study. The values for both formants are lower in NNSs. This vowel is, therefore, produced as a higher vowel and with a greater degree of backness, see Figures 7 and 8.

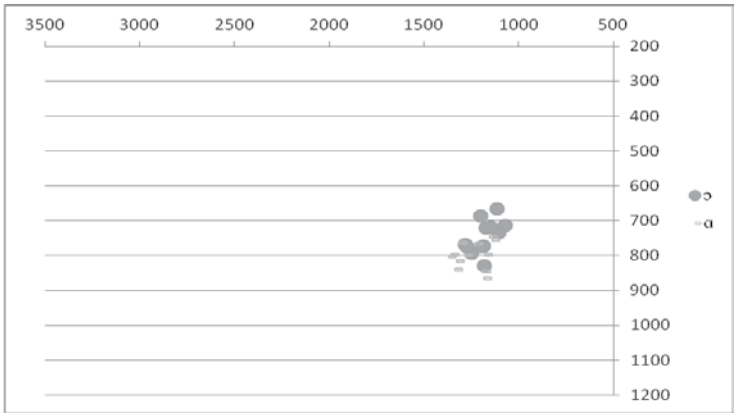


Figure 7. NS /ɔ ɑ/.

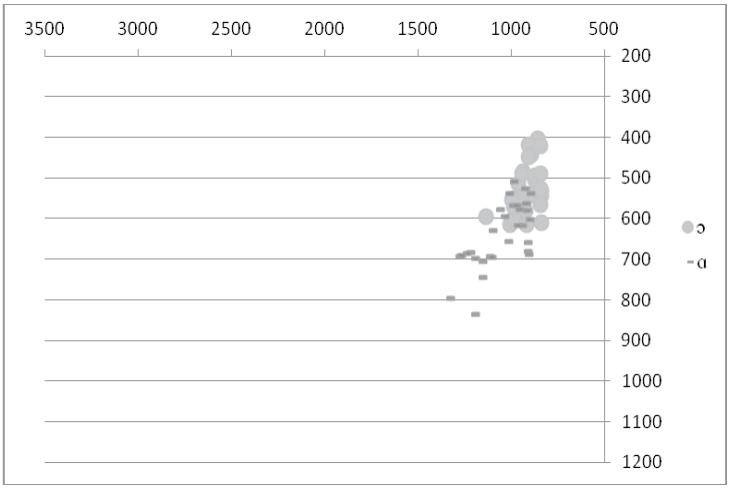


Figure 8. NNS /ɔ ɑ/.

A brief look at the duration measurements provides an interesting finding: the vowel of *bought* is significantly longer in the NNS group at the mean value of 199.5 ms (SD 56.9 ms) compared to the NSs' mean value of 156 ms (SD 14 ms). This is consistent with the findings for the tense /i u/, where NNSs also rely heavily on vowel length.

The vowel of *pot* is shorter in the NNS group (117.3, SD 23) than in the NS group of participants (123.2 ms, SD 13.7), which runs counter to the general rule that NS vowels are consistently longer compared to NNS vowels.

4 General Discussion and Conclusions

The tables that follow summarise the mean values of the vowel duration for all nine vowels investigated in this research study in both NS and NNS groups. The non-native speaker group data is given first, followed by the average duration measurements for the native speaker group. The third row displays the duration ratio for the two groups of participants (NNS vowel duration/NS vowel duration).

Table 8. Duration ratio.

	i	ɪ	ɛ	æ	ʌ	u	ʊ	ɔ	ɑ
NNS	132.5	87.6	120.8	173.1	116.2	154.5	80.3	199.5	117.3
NS	112.5	88.6	110.4	161.6	100.6	125	73.8	156.2	123.2
	1.17	0.98	1.09	1.07	1.15	1.23	1.08	1.27	0.95

The vowel duration ratio analysis shows that NNS vowels are consistently longer in duration than their NS counterparts, with the exception of /ɪ/ and /ɑ/. This is also shown in Figures 9 and 10.

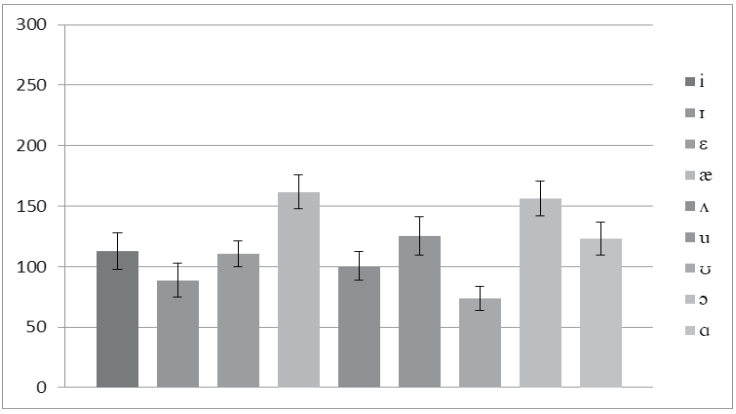


Figure 9. NS vowel duration with SD.

One of the significant findings of this small-scale vowel study is the neutralisation of vowel-quality of several vowels in the NNS production, which leads to a heavy

reliance on phonetic duration as a single, most important phonetic cue. The vowels /i/ and /u/ are reported to undergo a sort of a spectral merger in the NNS production, therefore NNS have to rely more heavily on the phonetic duration. It is a matter of debate why NNS articulate /ɔ/ as a very long vowel in AE. One of the reasons might be the influence of British English that these speakers were taught at school, where /ɔ:/ and /ɒ/ form a vowel pair, similar to /i: ɪ/ and /u: ʊ/.

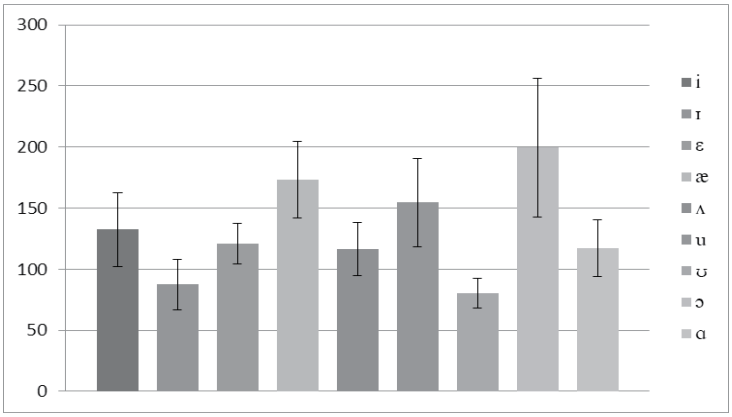


Figure 10. NNS vowel duration with SD.

Last but not least, the open vowel /æ/ is intrinsically long so both native and non-native speakers produce it this way. As a new sound to the NNS group, its vowel quality is somewhat more difficult to acquire, but its universal vowel duration is a good place to start.

References

Black, John W. 1949. "Natural Frequency, Duration, and Intensity of Vowels in Reading." *Journal of Speech and Hearing Disorders* 14: 216–21. https://doi.org/10.1044/jshd.1403.216.

Boersma, Paul, and David Weenink. 2013. *Praat: Doing Phonetics by Computer*. Version 5.3.51. Accessed June 20, 2013. http://www.praat.org/.

Browne, E. Wayles, and James D. McCawley. 1965. "Srpskohrvatski akcenat." *Zbornik za filologiju i lingvistiku* 8: 147–51.

Chen, Matthew. 1970. "Vowel Length Variation as a Function of the Voicing of the Consonant Environment. *Phonetica* 22: 129–59. https://doi.org/10.1159/000259312.

Čubrović, Biljana. 2016. *Acoustic Investigations of Serbian and American English Vowel Inventories*. Belgrade: Faculty of Philology.

—. 2017. "Low Back Merger in Native and Non-Native Speakers of American English." *The Linguistics Journal* 11 (1): 222–31.

—. 2018. "The Acquisition of Low Back Merger in Nonnative Speakers of American English." *Nasleđe* 41: 17–26.

Dicanio, Christian. 2013. "Script." Accessed February 16, 2019. http://www.acsu.buffalo.edu/~cdicanio/scripts/Vowel_Acoustics_for_corpus_data.praat.

Hillenbrand, James M., Laura A. Getty, Michael J. Clark, and Kimberly Wheeler. 1995. "Acoustic Characteristics of American English Vowels." *The Journal of the Acoustical Society of America* 97 (5): 3099–3111. https://doi.org/10.1121/1.409456.

Hillenbrand, James M., Michael J. Clark, and Robert A. Houde. 2000. "Some Effects of Duration on Vowel Recognition." *The Journal of the Acoustical Society of America* 108: 3013–22. https://doi.org/10.1121/1.1323463.

Holt, Yolanda Feimster, Ewa Jacewicz, and Robert Allen Fox. 2015. "Variation in Vowel Duration Among Southern African American English Speakers." *American Journal of Speech-Language Pathology* 24: 460–69. https://doi.org/10.1044/2015_AJSLP-14-0186.

House, Arthur S. 1961. "On Vowel Duration in English." *The Journal of the Acoustical Society of America* 33 (9): 1174–78. https://doi.org/10.1121/1.1908941.

House, Arthur S., and Grant Fairbanks. 1953. "The Influence of Consonant Environment upon the Secondary Acoustical Characteristics of Vowels." *The Journal of the Acoustical Society of America* 25 (1): 105–13. https://doi.org/10.1121/1.1906982.

Inkelas, Sharon, and Draga Zec. 1988. "Serbo-Croatian Pitch Accent: The Interaction of Tone, Stress, and Intonation." *Language* 64 (2): 227–48.

Jakobson, Roman. (1937) 1962. "On the Identification of Phonemic Entities." In *Selected Writings I*, 418–25. The Hague: Mouton.

Komar, Smiljana. 2017. "The Relationship between the Perception and Production of Four General British Vowels by Slovene University Students of English." *Linguistica* 57 (1): 161–70. https://doi.org/10.4312/linguistica.57.1.161-170.

Lehiste, Ilse, and Gordon E. Peterson. 1961. "Transitions, Glides and Diphthongs." *The Journal of the Acoustical Society of America* 33 (3): 268–77. https://doi.org/10.1121/1.1908638.

Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, MA: MIT Press.

Stopar, Andrej. 2015. "Perception of Four General British Vowels by Slovenian University Students of English as a Foreign Language." In *PTLC2015: Proceedings of the Phonetics Teaching and Learning Conference, London, 5–7 August 2015*, 87–90. London: Phonetics Teaching and Learning Conference.

Nataša Hirci
University of Ljubljana, Slovenia

2019, Vol. 16 (1), 29–45(164)
revije.ff.uni-lj.si/elope
doi: 10.4312/elope.16.1.29-45
UDC: 811.111'355:81'25

Trainee Translators' Perceptions of the Role of Pronunciation and Speech Technologies in the Technology-Driven Translation Profession

ABSTRACT

We live in a world of rapid technological advances which constantly affect the work of professional translators. Suitable training is therefore required for future translators to be able to compete on the translation market. With the rise of translation technologies, new ideas have been put forward on how to make translators faster and more efficient. Among the technologies that future translators may not be adequately familiar with are speech recognition tools; these enable translators to dictate their sight translation and have it typed out, allowing more time to focus on the content. However, as with all digital tools, the quality of input is important; a question thus arises on the role pronunciation assumes in such work. The present study aimed to establish how much awareness there is amongst the trainee translators of the possibilities afforded by speech technologies and to explore their perceptions of the role played by pronunciation.

Keywords: translator training; pronunciation; speech recognition tools; trainee translators' perceptions; the future of translation work

Bodoči prevajalci o vlogi izgovarjave in govornih tehnologij v sodobnem prevajalskem poklicu

POVZETEK

Živimo v času vse hitrejšega tehnološkega razvoja, v kar je nenehno vpeto tudi delo profesionalnih prevajalcev. V luči tega je nujno sprotno prilagajanje izobraževanja bodočih prevajalcev, da bodo primerno usposobljeni in bodo konkurenčni na prevajalskem trgu. S porastom sodobnih prevajalskih tehnologij se pojavljajo ideje o tem, kako bi lahko bili prevajalci pri svojem delu hitrejši in učinkovitejši. Eden od tehnoloških pripomočkov, ki bi k temu lahko pripomogel, a ga bodoči prevajalci premalo poznajo, so govorne tehnologije. S pomočjo prevajanja na vpogled prevajalec lahko besedilo narekuje: s tem se izogne tipkanju, in se bolj osredotoča na vsebino. A kot pri vseh digitalnih orodjih je pomembna kakovost vnosa podatkov, zato se poraja vprašanje, kakšno vlogo igra pri tem izgovarjava. V pričujoči študiji smo želeli raziskati, v kolikšni meri se bodoči prevajalci zavedajo možnosti, ki jih ponujajo govorne tehnologije, in ali imajo predstavo o vlogi, ki jo pri tem igra izgovarjava.

Ključne besede: poučevanje prevajalcev; izgovarjava; govorne tehnologije za razpoznavo govora; zavedanje bodočih prevajalcev; prevajalsko delo v prihodnosti

1 Introduction

The impact of new technologies on translation work over the last few decades has significantly changed the way people perceive the work of professional translators. The usual translator's workstation or translator's workbench no longer involves working only with computers and computer-assisted (CAT) tools, but may, under certain conditions, also involve working with machine translation (MT) and speech recognition technologies. According to a Stanford study (cf. Weiner 2016[1]) speaking is much faster than typing on a touchscreen, while typing on a computer keyboard is seemingly easier and faster. However, even a few years ago speech recognition software was criticised due to its error-prone performance which inevitably lead to spending too much time correcting the mistakes. It therefore seemed reasonable to assume that professionals who use a keyboard as part of their daily routine, translators included, would not be inclined to integrate into their work technologies which actually slow them down. However, a lot has changed since then: Nuance has produced Dragon Speech Recognition software, one of the leading speech recognition technologies, and claims that it is now able to transcribe up to 160 words per minute, which is also about three times faster than typing, with an enviable 99% recognition accuracy (cf. Dragon NaturallySpeaking[2]). This suggests speech technologies are now much more effective, and can perhaps make translation work more efficient. Moreover, any technological advantage is worth exploring to ensure that professional translators remain competitive on the translation market.

With the swift rise of digital innovations and artificial intelligence (AI), significant endeavours will constantly, and increasingly so, be put into speech technologies for translation undertakings, at least for fairly basic communication purposes and simple translation tasks, with the aim to establish basic contact and ease communication for those who do not speak a particular language. Students might already be aware of the possibilities afforded by virtual AI speech assistants such as Amazon's Alexa, Microsoft's Cortana, Google's Assistant or Apple's Siri, and might have tried using such services. Large brands are all investing heavily into voice technologies, and they are associated with a growing number of applications (cf. for more details on virtual assistants see Moren 2018). Armour (2018) reports on the data provided by Adobe Analytics, which indicates that "71% of owners of smart speakers like Amazon Echo and Google Home use voice assistants at least daily" [...] with "44% using them multiple times a day" while "[o]ver 76% of smart speaker owners increased their usage of voice assistants in the last year". Armour (2018) also quotes Steve Rabuchin, VP of Amazon Alexa, who stated that the vision they have for their customers is to "be able to access Alexa whenever and wherever they want. This means customers may

[1] Cf. https://www.popularmechanics.com/technology/a22684/phone-dictation-typing-speed/.

[2] Compare with data provided by Nuance at https://www.nuance.com/dragon/industry/education-solutions.html.

be able to talk to their cars, refrigerators, thermostats, lamps and all kinds of devices in and outside their homes". Armour (2018) believes that "voice is the future of how brands will interact with their customers". These virtual assistants are all monolingual, however, and do not engage in multi-lingual communication. Even so, "[t]o build a robust speech recognition experience, the artificial intelligence behind it has to become better at handling challenges such as accents and background noise. And as consumers are becoming increasingly more comfortable and reliant upon using voice to talk to their phones, cars, smart home devices, etc., voice will become a primary interface to the digital world and with it" (Armour 2018).

Virtual assistants no longer work only with English[3]; Cortana, for example, is currently also available in Chinese, French, German, Italian, Japanese, Portuguese and Spanish versions, making these voice technologies increasingly accessible to a much wider audience[4]. Even regular dictation services available to Windows and Mac users have the option of choosing between language varieties, with American, Australian, British or Canadian English, for example, already embedded while, depending on the tool, other varieties can easily be downloaded from the Internet. However, more time may be required to have languages of lesser diffusion[5] successfully integrated into existing systems. Slovene is a language spoken by only about two million people, and thus is less likely to be automatically added to other major language options. However, there are some speech recognition tools available, such as Voice Notepad, which already have Slovene embedded, and the dictation performance is relatively accurate. This is in contrast to the Google Translate dictation option, as the quality of translation is often still highly questionable and the final output more frequently than not inadequate and unusable. There is even a virtual AI assistant SecondEGO, designed by Amebis[6], and several other systems available for Slovene, which were originally created on the basis of large corpora and other language resources[7], such as the speech-to-speech communicator VoiceTRAN[8] or eBralec[9] (eReader): the direction, however, is speech-to-speech or written to spoken rather than spoken to written, which would be most suitable for translators. Moreover, these technologies are only available commercially or for research purposes (cf. Sepesy Maučec et al. 2009; Donaj and Kačič 2012; Žgank and Sepesy Maučec 2010; Žgank, Verdonik, and Sepesy Maučec 2016, to name just a few), while their non-commercial availability is still a matter for the future.

---

3 Other languages are also gaining ground on the Internet (cf. Internet World Stats 2017).

4 For more on English and its relative share online see Holly Young's article available at http://labs.theguardian.com/digital-language-divide/ and Laura Gonzales' article available at http://uxpamagazine.org/improving-digital-translation/.

5 Slovene included (cf. Pokorn 2005; Hirci 2012).

6 Cf. https://www.amebis.si/novice/npi-2015.

7 For more on Slovene in the digital age see Rehm and Uszkoreit (2012).

8 Cf. http://www.alpineon.si/voicetran/slovensko/html/index.html.

9 Cf. https://ebralec.si/?jezik=sl.

Still, none of these technologies are directly applicable to regular translation work as they are aimed at the general public to ease their daily routines. None of the virtual assistants are applicable to ease the tedious task of typing which has to be regularly undertaken by translators; translators thus need more specialised translation tools to facilitate their work (cf. Cronin 2013). One option that could possibly aid their daily routines and reduce the need for constant typing is dictation. Combined with sight translation it could change the way translation is habitually performed. It might thus be worth investigating the usability of speech-to-text technologies in translator training, foregrounding the time-efficiency ratio in particular. The awareness of trainee translators of the role of pronunciation and their familiarity with speech recognition technologies deserve research attention, in order to establish whether the application of such technologies could be motivating and beneficial for future translators.

2 Literature Review

Professional translation work is usually associated with the written output. However, the spoken modality should not be neglected in today's information society and its digital world, so heavily imbued with multimodality. It is therefore worth exploring the issues in translator training that address these modalities, spoken included, especially since – within the scope of interpreter training – Shlesinger (1995, 193–214) already maintained that "one modality can teach us about the constraints, conventions and norms of the other". This suggests that sight translation, a bridge between the oral and written mode of translation (cf. Agrifoglio 2004), should perhaps play a more prominent role not only in professional translation, but also in translation pedagogy.

So far, sight translation has been recognised as relevant in interpreting studies and interpreting pedagogy (cf. Agrifoglio 2004; Angelelli 1999; Li 2014; Gile [1995] 2009; Gonzalez, Vásquez, and Mikkelson 2012; Jimenez Ivars 2008; Lambert 2004; Mikkelson 1994; Moser-Mercer 1995; Pöchhacker 2004, 2010; Riccardi 2002; Schlesinger 1995; Song 2010; Viaggio 1995; Viezzi 1990; Weber 1990). Although there is still a fairly small body of literature focusing on the advantages of sight translation for written translation (cf. Baxter 2016; Dragsted and Hansen 2009; Dragsted, Hansen, and Sørensen 2009; Dragsted, Mees, and Hansen 2011; Gorszczyńska 2010; Mees et al. 2013), a recent study has shown (cf. Hirci, Mikolič Južnič, and Pisanski Peterlin forthcoming) that engaging in sight translation for the purposes of written translation can result in creative, novel translation solutions, which gives an added value to the translation process and can make the entire process of translating much faster and more efficient. Some scholars have already explored the application of dictation in sight translation and foregrounded its benefits for translation work in terms of time efficiency (cf. Biela-Wolonciej 2007). Possible advantages were also reported by Dragsted, Mees, and Hansen (2011), who compared written and sight translation output with and without speech recognition software.

They concluded that with additional training and better familiarity with speech recognition tools, "greater time savings and higher quality are likely to be achieved as technical obstacles are either reduced or overcome" (Dragsted, Mees, and Hansen 2011, 26). Baxter (2016) also investigated the application of sight translation skills to written translation combined with speech recognition; although there were no considerable time differences for the two studied groups, idiomaticity was enhanced, suggesting that combining sight translation with speech recognition "improves the spontaneity of the final text, thereby producing a more natural-sounding translation than the traditional W2W[10] method" (Baxter, 2016, 14). However, the most interdisciplinary approach was adopted in a study by Mees et al. (2013) where close collaboration among phoneticians, translators and interpreters yielded sound grounds for further interdisciplinary cooperation, proving that speech recognition technologies[11] can be successfully applied in translator training.

In Slovenia, no study has been carried out on having speech recognition technology fully integrated into translation work, focusing on a hybrid which "involves crossing borders between translation and interpreting since the translation is produced orally, as in interpreting, but is visible on the screen, as in translation" (Mees et al. 2013, 141). There is an introductory course on English phonetics and phonology for translators offered in year one of the undergraduate programme at the Department of Translation Studies in the University of Ljubljana to help students improve their pronunciation. As the advances in speech-to-text technology are relatively recent, students enrolled in the course may not be familiar with the relevance of pronunciation skills in technological applications, and may perceive pronunciation to be more important for interpreters than translators. Yet this issue is particularly relevant for those who may wish to use software which is heavily reliant on one's pronunciation. As Nuance is claiming a 99% accuracy for its software, it needs to be acknowledged that such accuracy is only possible if one's pronunciation is also highly accurate, otherwise the success rate of speech recognition is much lower. Near-native and intelligible pronunciation is required for the dictation systems to work well, at least for the time being, otherwise the rate of mistakes due to mispronunciation is too great to have such tools considered effective. However, so far the potential relevance of pronunciation skills for the trainee translators' work in the translation modules offered later as part of the graduate programme in Translation/Interpreting has not yet been addressed, as none of the specialised translation courses involve working with speech recognition technologies. As there are built-in dictation options available on computers (both for Windows and Mac users) that enable working with English, translation modules focusing on translation from L1 to L2 could possibly benefit

[10] W2W means written to written translation.

[11] For more details on speech recognition technology see Jurafsky and Martin (2000).

from integrating this technology into their regular translation instruction. In their study, Mees et al. (2013) also report on working into L2 (cf. studies by Dragsted and Hansen 2009; Dragsted, Mees, and Hansen 2011). In Denmark, and the rest of Scandinavia, where, according to Phillipson (2003, 96) there are "good grounds for referring to English as *a second language* rather than a foreign language", working into L2 is not perceived as unusual. Both Danish as well as Slovene are comparable in this respect, as they can both be considered as languages of lesser diffusion, so L1 to L2 translation (cf. Pokorn 2005; Hirci 2012) is not uncommon in Slovenia either. In fact, children in Slovenia start learning English as part of their primary school curriculum at the age of six. Films and TV shows are regularly subtitled rather than dubbed, and Slovene translators work into both directions, L2 to L1 as well as L1 to L2. Many professional translators in Slovenia find themselves in a position where they are required to undertake translation into L2, English in particular, on a regular basis, since there is a serious shortage of native English speakers working with Slovene. Thus training is necessary in the L2 direction and is offered as part of the translator training curriculum at the Department of Translation Studies in the University of Ljubljana.

2.1 Future Prospects – More Work with Speech Recognition Systems?

So far, no research has been undertaken in Slovenia to explore working with speech recognition systems focusing on time efficiency in translation. However, a study was carried out on the possible benefits of applying speech recognition technologies in the pronunciation training of non-native speakers of English. Šuštaršič (2005, 87) investigated some software packages to explore their "usability within an English phonetics curriculum for EFL learners at the university level" that can be applied to pronunciation training. Šuštaršič (2005, 93–97) suggested that "speech recognition can be applied in phonetics (or more precisely, in pronunciation) teaching, and that a number of aspects of articulatory and auditory phonetic principles can be observed in the way that speech recognition programs transfer (or fail to transfer) the received speech signals into written form." He pointed out that "using any speech recognition program with English pronunciation students has several other justifications. Firstly, the program needs to be trained to one's voice, which requires a great deal of loud reading. […] The basic rule is: the more you train the program (i.e. the more you read), the higher will be the accuracy of recognition, and thus the usefulness of the program for any practical task." Šuštaršič (2005, 98) also suggested that students can be encouraged to record their own speech and apply a speech recognition programme to convert it into a written text, an idea which in itself is closely related to sight translation from Slovene into English. Šuštaršič (2005) reported working with commercial speech recognition technologies such as Via Voice and Dragon's NaturallySpeaking, which, however, are not freely available. A cost-free option

nowadays is to simply activate the automatically built-in dictation option on the computer (either for Windows or Mac users), as it comes at no additional price, and explore its usability before obtaining some more sophisticated commercial software.

Drawing on Mees et al. (2013) and Šuštaršič (2005), a study was thus conceived to explore the possible benefits of using speech technologies in translator training for two reasons:

- to improve trainee translators' pronunciation,
- to use speech instead of typing to speed up the process of translation.

3 Study Design and Methodology

The present study was designed to explore the trainee translators' perceptions of the role of English pronunciation, as well as their familiarity with speech recognition tools, to establish whether or not it might be viable to introduce such technologies into translator training at the University of Ljubljana.

3.1 Methodology and Participants

An online questionnaire was designed for the purposes of the present study to foreground the perceptions of both undergraduate and graduate trainee translators studying at the Department of Translation Studies at the Faculty of Arts, University of Ljubljana, in the academic year of 2018/2019.

3.2 Data Collection

The questionnaire was made available online for 18 days, between 4 January 2019 and 22 January 2019, with a total of 94 participants taking part in the study. The questionnaire, designed using the online Google Forms survey mode, consists of 18 questions. The first part of the survey aims to collect general information about the participants, eliciting data on their age, gender and year of study. The second part of the questionnaire explores the participants' perceptions and self-awareness of their own pronunciation and their familiarity with the existing speech-to-text technologies that might prove to be useful in their future profession.

The trainee translators were asked to respond to several statements referring to their perceptions of the role pronunciation in English and their aspirations to improve it (i.e. a total of nine questions corresponding to yes/no answers, and four statements using a five-point Likert-type scale, ranging from "totally unmotivated" = 1 to "extremely motivated" = 5 related to the participants' motivation to have good pronunciation of English, from "the least important" = 1 to "the most important" = 5 on how important they find pronunciation in relation to other language skills, from "extremely poor" = 1 to "excellent" = 5 on how they would rate their own pronunciation at the time of

filling out the questionnaire, and finally from "do not aspire to this at all" = 1 to "aspire to this 100%" = 5 on how much they aspire to have a near-native pronunciation of English).

Additional information on the existing speech recognition tools and the students' experience with the application of these technologies to their work was elicited using a number of multiple choice questions. The participants were also encouraged to provide additional comments on the possible benefits of using speech-to-text technologies in the final section of the questionnaire.

4 Results and Discussion

This section reports on the results of the questionnaire completed by the participants of the study. First, general demographic information on the participants is provided, followed by the data related to their pronunciation and awareness of speech recognition technologies. Due to the limited scope of this paper only those results that directly address the topic are discussed in detail.

4.1 General Information on the Participants

The study involved 94 participants, of whom all completed the questionnaire in full. All of the participants are either undergraduate BA students of Interlingual Mediation, or graduate MA students of Translation/Interpreting in the University of Ljubljana (cf. Figure 1). Of the 94 participants, 76 were female and 18 were male, and all were aged between 17 and 26 (average 21).

Most participants (41, i.e. 43.6%) are enrolled in year 1 of the BA in Interlingual Mediation, with 14 (14.9%) respondents from year 2 of the BA in Interlingual Mediation, and 16 (17%) respondents from year 3 of the BA in Interlingual Mediation (cf. Figure 1). At the graduate level, there were 15 (16%) participants from MA I in Translation, three (3.2%) from MA I in Interpreting, and five (5.3%) from MA II in Translation (there is no MA II in Interpreting available for this academic year).

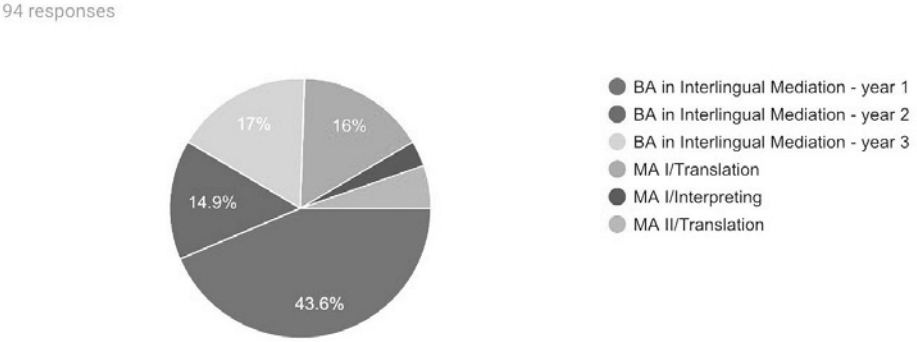


Figure 1. Participants in the study (N=94).

4.2 Specific Information on Pronunciation

Importance to speak English well

As evident from the results of the questionnaire, all of the participants believe that it is important to speak English well to make a good impression on their clients and employers, and all but one believe the same is important to be a successful interpreter, while 88 out of 94 participants (i.e. 93.6%) were of the opinion that this is also important for translators (cf. Hirci 2017). In addition, 90 (95.7%) respondents think that it is important to speak well to sound professional, and 83 (88.3%) to be able to use speech recognition tools more easily.

Significance of speaking English well

The participants seem to have rather diverse views on what speaking English well actually means. Most of the participants, i.e. 89 (94.7%), agreed that this meant having pronunciation which is intelligible and easy-to-understand, with 65 (69.1%) believing it meant speaking with an accent which is close to standard varieties of English. Fewer than half of the respondents in all (45 or 47.9%) believe that this meant having a native-like pronunciation.

Motivation to have a good pronunciation of English

The questionnaire yielded an insight into the participants' motivation with regard to having good pronunciation: the results show that over half of the participants (52 or 55.3%) are extremely motivated and an additional 30 (31.9%) are very motivated to have a good pronunciation of English (a mean score[12] of 4.4, cf. Table 1), which confirms that the respondents regard having good pronunciation in English as essential for their future profession.

Table 1. Mean scores for pronunciation.

Perceptions about pronunciation	Mean score
Motivation to have a good pronunciation of English	4.4
Importance of pronunciation compared to other language skills	3.8
Assessment of own pronunciation	3.5
Aspirations to improve their pronunciation	4.3

When asked about how important they find pronunciation compared to other language skills, the participants showed considerable agreement that pronunciation skills are quite important (a mean score of 3.8).

[12] The central tendency for each Likert-type statement was summarised using the mean score.

The participants' replies furthermore revealed that they tend to aspire to have English pronunciation which is intelligible yet close to one of the English standards. They deemed their own pronunciation at the time of filling out the questionnaire as only "good" or "fairly good", while only two participants considered it "excellent". Three participants even believed their pronunciation was "extremely poor" or "rather poor" (mean score 3.5). The responses revealed that over half (54.3%) of the participants have extremely high aspirations to improve their pronunciation, and an additional 29.8% of the participants have high aspirations to improve it (mean score 4.3).

These results are quite valuable, as they reveal that most participants are aware of the significance of having a good pronunciation of English. Whether they see a correlation with speech-to-text technologies, however, is yet to be explored. As clear, accurate and intelligible pronunciation is required to have speech recognition systems work well, at least for the time being, improving non-native English pronunciation is undoubtedly worth investing time and effort into if we also wish to gain from the advantages afforded by such technologies.

4.3 Specific Information on Speech Recognition Technologies

We wished to establish if the respondents were aware of the differences in speed as related to speech and typing. According to Nuance's Dragon speech recognition software, speaking is three times faster than typing. Most respondents of this study, i.e. 45 (47.9%), believed that speech was two times faster than typing, while 37 (39.4%) participants in fact responded that it was actually three times faster. Only two participants were of the opinion that speaking was slower than typing, three assumed that it was four times faster, while another four responded that these two activities were both of equal speed (cf. Figure 2, where responses are provided as option Other, after the option 4x faster).

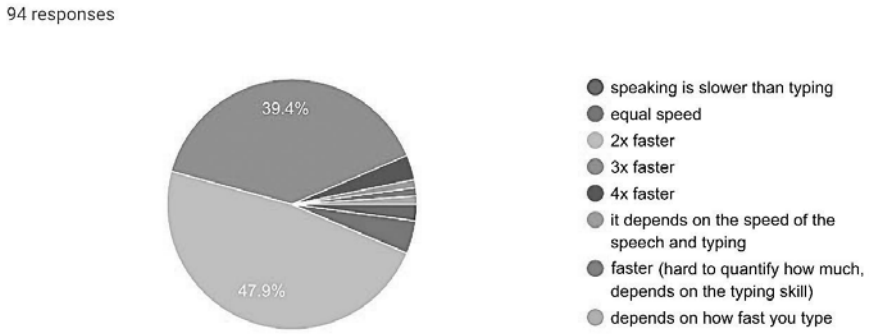


Figure 2. Speed of speaking v typing (N=94).

It was no surprise to see that almost half of the participants (44.7%) responded that they have already used the built-in dictation software on their smartphones; nevertheless, the number is much lower for computers, where only 11 of participants out of 94 reported using this technology. Interestingly enough, 28 of the participants reported that their dictation was successful, or at least sometimes or to some extent. It is fair to assume that with more accurate pronunciation of English the perception of the success rate would most likely be even higher. Some participants also pointed out that they used the dictation option only on their smartphones, without ever realising that this was also possible on their computers.

In all, 70 (74.5%) of the participants responded that they would consider using dictation in their translation work; even more, i.e. 84 (89.4%) believed that it would be useful to work with speech recognition tools as part of their translator training at the university. In additional, individual comments, the participants provided a number of reasons why they assumed it would be useful to work with speech recognition tools as part of translator training (cf. Figure 3).

P14: "It could improve the student's pronunciation skills and, more importantly, the proper flow of speech."
P8: "Speech recognition tools are great for improving ones pronunciation and I think we should focuse on that and phonetics in general more thoroughly."
P4: "I believe that students should be familiar with any translation- or language-related technology. This can be useful in their careers."
P53: "I think that such thing as a speech recognition tool would help me a lot with my poor pronunciation."
P16: "Working with these tools would improve our pronounciation."
P15: "I think we would be able to translate everything faster. And we would also practice our pronunciation and expand our vocabulary, because when we say something outloud, we remember it faster."
P18: "So that we learn different approaches to translating and figure out for ourselves which best suits us. Also I think it is less time consuming than typing and prevents you from making spelling mistakes"
P23: "It's a tool that is becoming increasingly popular and it could potentially make future work easier."
P29: "Knowledge of new technologies is always useful, the more you know the more you can learn, new skills can easily improve our employability, variation of skills is important for adapting to the market"
P19: "The more education we get - conected to our studies and technology connected to languages – the better."
P20: "Speech recognition tools are developing and becoming a bigger part of our everyday life".

P42: "It would improve out studying and it would be a variation of "teaching" that is not often used."
P49: "Because any aspect of the translation work that we are presented is welcome and useful. Anything that we learn might come in handy and we are better because of each of those experiences."
P56: "So that we learn different techniques and figure out which approach best suits us."
P59: "The advancement of technology will impose these tools sooner or later and it would be best if the new generations of translators and interpreters had mandatory training with them."

Figure 3. Comments provided by the participants.[13]

These comments show that there is already some degree of awareness amongst the trainee translator population of the possible advantages associated with the integration of speech recognition technologies into translator training.

The results of the questionnaire related to the various types of speech recognition software that the participants might have heard of are specified in Figure 4. The most frequently recognised speech recognition technologies were Windows Speech Recognition (60), Apple's dictation (49) and Google Docs Voice Typing (48), followed by IBM's Speech to Text (38), Amazon's Transcribe (27) and Speechnotes (25). The other speech-to-text tools (such as Via Voice, Dragon NaturallySpeaking or Voice Finger) were much less frequently recognised, while only one participant in this study had heard of Braina Pro.

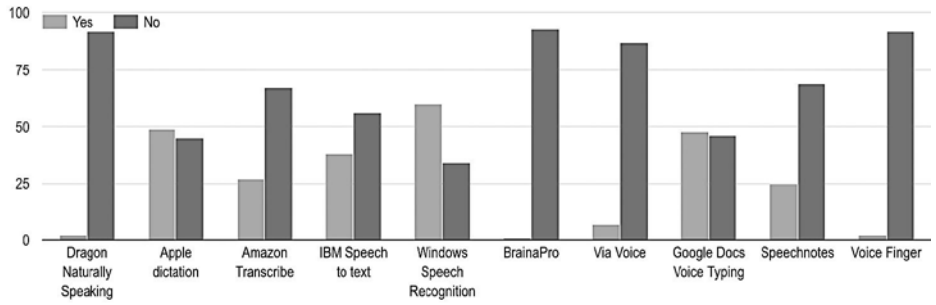


Figure 4. Familiarity with speech recognition technologies (N=94).

In addition, only three other online speech recognition tools were mentioned by the participants who were offered an option to list any other speech recognition technologies of which they might be aware: one of the participants noted using Google Keep, while another participant had not only heard of but has tried Voice Notepad for Slovene (they reported, however, that their dictation work was not

[13] All comments by the participants are provided in their original form, verbatim, with spelling mistakes and other errors left unchanged.

highly successful). It is interesting to note that 32 (i.e. 34%) of the participants responded that they have already tried using some of the tools mentioned, selecting mainly Apple Dictation, Google Docs Voice Typing and Windows Speech Recognition (only two participants selected Speechnotes, while only one mentioned IBM's Speech to text and another one Via Voice, cf. Figure 4). Most of the participants (62, or 66%) learnt about these speech tools online, by themselves, and only five (i.e. 5.3%) at the university.

P13: "It's a faster way of writing down what you need to translate and possibly a more fun and/or interesting way of translation."
P15: "Good speech recognition tools could help us learn proper pronunciation."
P20: "Because they could save quite a lot of time and work for translators (there would be also be no typos in the text etc)."
P23: "An extra aid one might find useful (like a dictionary or a thesaurus)."
P29: "They could be useful for translating things that need to be transcribed anyway, like speeches, directly, or just as an alternative to typing."
P40: "For transcribing and general text formation – the limits of one's typing skill can cause the occurence of getting lost in thought while typing and forgetting what you were about to say. In speech it happens less often"
P44: "It could be helpful if one has to translate videos or with subtitling."
P48: "These tools can facilitate the translation of audio documents"
P58: "They could replace typing, which can be time-consuming and tiring."
P65: "We could see where the problem with our speech is."
P67: "Because it is useful knowing tools that can make the translation work easier. This presents us with what the translation work is like and prepares us for it."
P68: "It is faster, so they can earn more money in a shorter period of time and thus have more free time. :)"
P70: "These tools could mean that translators would finish their work faster. Some may speak faster than they type so it could improve their working conditions."
P75: "To facilitate transcribing spoken language, could be useful for making subtitles"
P81: "Since speech recognition tools are very accurate nowadays, I believe it would save a lot of time."

Figure 5. Participants' comments on the usefulness of speech recognition tools for translators (also verbatim).

Judging from the comments provided in the questionnaire, some participants are also aware of the drawbacks of the current speech recognition technologies and their reliability:

P7: "Translators could work faster, but the speech recognition tools would need to be very good, especially when it comes to punctuation. Going over a text two or more times to correct punctuation that was wrongly placed by speech recognition tools is very time consuming."
P37: "They might be useful for cases, when translators have to write subtitles e.g. a speech or movie, and are not sure about what a person is saying. However the speech recognition tools are not yet reliable enough to be completely sure of whether their result is correct."

As it can be observed from the participants' comments, the predominant idea revolved around the opinion that speech is "faster than typing", and that the application of speech technologies could make translators more efficient. Some students are well aware of the current situation in the ever-evolving digital world, recognizing that "the use of speech recognition today is growing and many people use it on their phones (Siri) or have devices (Alexa) that help them with everyday tasks." (P34)

All this suggests it might be worth raising the awareness of the trainee translator population about the existence of such tools, and possibly even integrate speech recognition technologies into translator training. This could be achieved in several ways: either by implementing information on speech recognition technologies into the already existing technology-related courses, or by introducing it as part of a new course focusing on this particular topic with hands-on training within L1 to L2 translation modules.

Some studies have already shown (cf. Mees et al. 2013; Désilets et al. 2008) that the implementation of speech technologies into translation work is something that could possibly be better addressed in the future. There are also interesting pedagogical implications of this: if dictation may soon become an increasingly dominant mode of communication, it is important to gain an in-depth insight into the aspects of pronunciation that would be particularly relevant in translator training.

5 Conclusion

The present study explored the perceptions of trainee translators studying at the Department of Translation Studies in the University of Ljubljana on pronunciation and speech technologies. The results of the study offer good grounds for a more prominent role to be assigned to both pronunciation instruction and speech technologies in translator training. The study yielded results showing that an overwhelming majority of trainee translators (just under 90%) believe that having good pronunciation of English is important for their profession (cf. Hirci 2017), while over 80% also have aspirations to improve their pronunciation. In addition, the results show that all the participants believe it is important to speak English well to make a good impression on clients and employers; all but one find this important for interpreters, while 93.6% also find it important for translators. Moreover,

95.7% respondents stated that it is important to speak well to sound professional, and 88.3% believe this is important to be able to use speech recognition tools more easily.

These results suggest that equipping trainee translators with pronunciation skills for speech recognition technologies is of relevance and would most likely be embraced by the students. This is in line with the study by Mees et al. (2013, 149), whose retrospective interviews revealed that "a number of students feel that they have become more aware of their pronunciation problems in the course of training the SR [speech recognition] program". Their study also revealed that speech recognition "provides a potentially useful supplement to written translation, or indeed an alternative to it" (Mees et al. 2013, 140–42). The immediate time-efficiency aspect is therefore yet another reason why speech recognition technologies could be applied in translator training: a new modality could also enhance the learning experience in the translation classroom. As some participants of this study have observed, "Time is valuable. Every second saved from sitting in front of a screen and keyboard is warmly welcome" (P41) or "It is faster, so they can earn more money in a shorter period of time and thus have more free time. :) (P68)."

With the increasingly rapid advances in voice activated technologies, translator trainers should seize the opportunity to retain tech-savvy students' interest and channel it into their regular coursework. Staying ahead is vital to remaining competitive; having that special 'edge' might be a deciding factor in having trainee translators turn into successful players on the professional translation market. Thus aiming to have good pronunciation and speak English well enough to be able to work with speech recognition technologies could prove to have added value for translators' professional careers.

References

Agrifoglio, Marjorie. 2004. "Sight Translation and Interpreting: A Comparative Analysis of Constraints and Failure." *Interpreting. International Journal of Research and Practice in Interpreting* 6 (1): 43–67. http://dx.doi.org/10.1075/intp.6.1.05agr.

Angelelli, Claudia. V. 1999. "The Role of Reading in Sight Translation." *The ATA Chronicle (Translation Journal of the American Association of Translators)* 28 (5): 27–30.

Armour, Britt. 2018. "7 Key Predictions for the Future of Voice Assistants and AI." Accessed January 15, 2019. https://clearbridgemobile.com/author/britt_clrbridge/.

Baxter, Neal. R. 2016. "Exploring the Effects of Computerised Sight Translation on Written Translation Speed and Quality." *Perspectives* 1: 1–18. https://doi.org/10.1080/0907676X.2016.1241287.

Biela-Wolonciej, Aleksandra. 2007. "A-vista: New Challenges for Tailor-Made Translation Types on the Example of Recorded Sight Translation." *Kalbotyra* 57 (3): 30–39.

Cronin, Michael. 2013. *Translation in the Digital Age.* London: Routledge.

Désilets, Alain, Marta Stojanovic, Jean-François Lapointe, Rick Rose, and Aarthi Reddy. 2008. "Evaluating Productivity Gains of Hybrid ASR-MT Systems for Translation Dictation." In *IWSLT 2008, International Workshop on Spoken Language Translation, 20–21 October 2008, Waikiki, Hawai'i, USA*, Waikiki, Hawai'i, 158–65.

Donaj, Gregor, and Zdravko Kačič. 2012. "Širjenje slovarja in dvoprehodni algoritem v razpoznavalniku tekočega govora UMB Broadcast News." In *Proceedings of the Eighth Language Technologies Conference, October 8th-12th, 2012,* Ljubljana, Slovenija, 48–51. Ljubljana: Institut Jožef Stefan.

Dragsted, Barbara, and Inge G. Hansen. 2009. "Exploring Translation and Interpreting Hybrids. The Case of Sight Translation." *Meta: Journal des traducteurs/Meta: Translators' Journal* 54 (3): 588–604. https://doi.org/10.7202/038317ar.

Dragsted, Barbara, Inge G. Hansen, and Henrik S. Sørensen. 2009. "Experts Exposed." In *Methodology, Technology and Innovation in Translation Process Research* (Copenhagen Studies in Language 38), edited by Inger M. Mees, Fabio Alves and Susanne Göpferich, 293–317. Copenhagen: Samfundslitteratur.

Dragsted, Barbara, Inger M. Mees, and Inge G. Hansen. 2011. "Speaking Your Translation: Students' First Encounter with Speech Recognition Technology." *Translation & Interpreting* 3 (1): 10–43.

Gile, Daniel. [1995] 2009. *Basic Concepts and Models for Interpreter and Translator Training.* Amsterdam/Philadelphia: John Benjamins Publishing.

Gonzalez, Roseann D., Victoria F. Vásquez, and Holly Mikkelson. 2012. *Fundamentals of Court Interpretation: Theory, Policy and Practice.* 2nd ed. Durham, North Carolina: Carolina Academic Press.

Gonzales, Laura. 2017. "Improving Digital Translating: Research Findings from Multilingual Communicators." *User Experience Magazine* 17 (5). http://uxpamagazine.org/improving-digital-translation/.

Gorszczyńska, Paula. 2010. "The Potential of Sight Translation to Optimize Written Translation: The Example of the English-Polish Language Pair." In *Translation Effects. Selected Papers of the CETRA Research Seminar in Translation Studies 2009,* edited by Omid Azadibougar, 1–12. Leuven: KU Leuven. https://www.arts.kuleuven.be/cetra/papers/files/paula-gorszczynska-the-potential-of-sight.pdf.

Jimenez Ivars, Maria. 2008. "Sight Translation and Written Translation. A Comparative Analysis of Causes of Problems, Strategies and Translation Errors within the PACTE Translation Competence Model." *FORUM. International Journal of Interpretation and Translation* 6 (2): 79–104.

Hirci, Nataša. 2012. "Electronic Reference Resources for Translators. Implications for Productivity and Translation Quality." *The Interpreter and Translator Trainer* 6 (2): 219–35. https://doi.org/10.1080/13556509.2012.10798837.

—. 2017. "Investigating Trainee Translators' Views on the Pronunciation of English: a Slovene Perspective," *Linguistica*: *Sounds and Melodies Unheard: Essays in Memory of Rastislav Šuštaršič* 57 (1): 93–106. https://doi.org/10.4312/linguistica.57.1.93-106.

Hirci, Nataša, Tamara Mikolič Južnič, and Agnes Pisanski Peterlin. (forthcoming). "Enriching Translator Training with Interpreting Tasks: Bringing Sight Translation into the Translation Classroom." *Convergence, Contact and Interaction in Translation and Interpreting Studies*, edited by Eugenia Dal Fovo and Paola Gentile. Berlin: Peter Lang.

Internet World Stats. 2017. Accessed January 22, 2019. https://www.internetworldstats.com/stats7.htm

Jurafsky, Dan, and James H. Martin. 2000. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. New Jersey: Prentice Hall.

Lambert, Sylvie. 2004. "Shared Attention During Sight Translation, Sight Interpretation and Simultaneous Interpretation." *Meta: journal des traducteurs/Meta: Translators' Journal* 49 (2): 294–306. https://doi.org/10.7202/009352ar.

Li, Xiangdong. 2014. "Sight Translation as a Topic in Interpreting Research: Progress, Problems and Prospects." *Across Languages and Cultures* 15 (1): 67–89. https://doi.org/10.1556/Acr.15.2014.1.4

Mees, Inger M., Barbara Dragsted, Inge G. Hansen, and Arnt Lykke Jakobsen. 2013. "Sound Effects in Translation." *Target* 25 (1): 140–54. https://doi.org/10.1075/target.25.1.11mme.

Mikkelson, Holly. 1994. "Text Analysis Exercises for Sight Translation." In *Vistas: Proceedings of the 31st Annual Conference of ATA*, edited by Peter W. Krawutschke, 381–90. NJ: Learned Information.

Moser-Mercer, Barbara. 1995. "Sight Translation and Human Information Processing." *Basic Issues in Translation Studies. Proceedings of the Fifth International Conference* 2: 159–66.

Moren, Dan. 2018. "Alexa vs. Google Assistant vs. Siri: Google Widens Its Lead." Accessed January 15, 2019. https://www.tomsguide.com/us/alexa-vs-siri-vs-google,review-4772.html

Phillipson, Robert. 2003. *English-only Europe: Challenging Language Policy?* London: Routledge.

Pokorn, K. Nike. 2005. *Challenging the Traditional Axioms: Translation into a Non-Mother Tongue.* Amsterdam/Philadelphia: John Benjamins Publishing.

Pöchhacker, Franz. 2004. *Introducing Interpreting Studies*. London: Routledge.

—. 2010. "The Role of Research in Interpreter Education." *Translation & Interpreting* 2 (1): 1–10.

Rehm, Georg, and Hans Uszkoreit, eds. 2012. *The Slovene Language in the Digital Age.* Berlin, Heidelberg*:* Springer. https://doi.org/10.1007/978-3-642-30636-5.

Riccardi, Alessandra. 2002. "Interpreting Research: Descriptive Aspects and Methodological Proposals." In *Interpreting in the 21st Century: Challenges and Opportunities*, edited by Guiliana Garzone and Maurizio Viezzi, 73–82. Amsterdam/Philadelphia: John Benjamins Publishing.

Sepesy Maučec, Mirjam, Tomaž Rotovnik, Zdravko Kačič, and Janez Brest. 2009. "Using Data-Driven Subword Units in Language Model of Highly Inflective Slovenian Language." *International Journal of Pattern Recognition and Artificial Intelligence* 23 (2): 287–312. https://doi.org/10.1142/S0218001409007119.

Shlesinger, Miriam. 1995. "Shifts in Cohesion in Simultaneous Interpreting." *The Translator* 1 (2): 193–214. https://doi.org/10.1080/13556509.1995.10798957.

Šuštaršič, Rastislav. 2005. *English-Slovene Contrastive Phonetic and Phonemic Analysis and Its Application in Teaching English Phonetics and Phonology.* Ljubljana: Filozofska fakulteta.

Viaggio, Sergio. 1995. "The Praise of Sight Translation (and Squeezing the Last Drop Thereout of)." *The Interpreters' Newsletter* 6: 33–42.

Viezzi, Maurizio. 1990. "Sight Translation, Simultaneous Interpretation and Information Retention." In *Aspects of Applied and Experimental Research on Conference Interpretation*, edited by Laura Gran and Christopher Taylor, 54–60. Udine: Campanatto.

Weber, Wilhelm K. 1990. "The Importance of Sight Translation in an Interpreter Training Program." In *Interpreting: Yesterday, Today, and Tomorrow*, edited by David Bowen and Margareta Bowen, 44–52. Amsterdam: John Benjamins.

Weiner, Sophie. 2016. "Study Says Speech-to-Text Is 3 Times Faster Than Typing On Your Phone." Accessed January 15, 2019. https://www.popularmechanics.com/technology/a22684/phone-dictation-typing-speed/

Young, Holly. 2018. "The Digital Language Divide. How Does the Language You Speak Shape Your Experience of the Internet?" Accessed January 15, 2019. http://labs.theguardian.com/digital-language-divide/

Žgank, Andrej, and Mirjam Sepesy Maučec. 2010. "Razpoznavalnik tekočega govora UMB Broadcast News 2010: nadgradnja akustičnih in jezikovnih modelov." In *Proceedings of the Seventh Language Technologies Conference,* Ljubljana, Slovenia, 28–31. Ljubljana: Institut Jožef Stefan.

Žgank, Andrej, Darinka Verdonik, and Mirjam Sepesy Maučec. 2016. "Razpoznavanje tekočega govora v slovenščini z bazo predavanj SI TEDx-UM." In *Proceedings of the Conference on Language Technologies & Digital Humanities,* Ljubljana, Slovenija, 186–89. Ljubljana: Ljubljana University Press.

Saša Jošt,
Andrej Stopar
University of Ljubljana, Slovenia

2019, Vol. 16 (1), 47–76(164)
revije.ff.uni-lj.si/elope
doi: 10.4312/elope.16.1.47–76
UDC: 811.111'243'342:37.091.3

Perception of Foreign Phonemes: The Case of Slovene Students of English

ABSTRACT

The study focuses on assessing the state of foreign phoneme acquisition by foreign language (FL) students at the end of their undergraduate studies. To determine whether they prioritise vowels over consonants, a perception experiment was devised that focuses on the phonemes in Standard Slovene and General British reported as most problematic for Slovene learners of English. Thirty-three Slovene students of English were tested, along with a positive and a negative control group (CG). A set of stimuli was tested using the AX discrimination method; the participants listened to 60 phonemic contrasts, played in a pseudo-randomised order. The results foregrounded the most problematic phonemes which act as perceptual magnets. Analysis shows that the students can discriminate foreign phonemes well, and that they consistently perform better in discriminating vocalic contrasts.

Keywords: foreign language learning; phonological category; discrimination; perceptual magnet; vocalic prioritisation

Percepcija tujih fonemov: primer slovenskih študentov angleščine

POVZETEK

S to raziskavo smo želeli oceniti, kako uspešno študenti tujih jezikov usvojijo tuje foneme do konca študija prve bolonjske stopnje. Da bi ugotovili, ali v procesu učenja tujega jezika dajejo prednost samoglasnikom pred soglasniki, smo zasnovali percepcijski eksperiment, ki se osredinja na foneme standardne slovenščine in angleščine, ki veljajo za najbolj težavne za Slovence, ki se učijo angleščine. Testirali smo 33 slovenskih študentov anglistike ter pozitivno in negativno kontrolno skupino. V eksperimentu smo uporabili metodo AX; udeleženci so poslušali 60 fonemskih kontrastov, predvajanih v psevdonaključnem vrstnem redu. Z analizo rezultatov smo izpostavili najbolj problematične foneme, ki se vedejo kot percepcijski magneti. Analiza je pokazala, da imajo študenti ob zaključku prve stopnje študija dobro sposobnost razločevanja tujih fonemov in da so se bolje odrezali pri razločevanju samoglasniških kontrastov.

Ključne besede: učenje tujih jezikov; fonemska kategorija; razločevanje; percepcijski magnet; prioritizacija samoglasnikov

1 Introduction

Second/foreign language acquisition is a process in which a second language (L2) is acquired/learned in addition to the first language (L1). Some authors distinguish between the different contexts in which this occurs and choose their terminology accordingly – a second language may be described as 'acquired' when the acquisition occurs as part of everyday communicative situations, while a foreign language (FL) may be referred to as 'learned' when the learning happens through guided instruction in a classroom (Ellis 2015, 6). The term second language acquisition (SLA) is frequently used to cover both these contexts.

The role of L1 in the process of SLA has been a focus of numerous studies. Researchers now mostly believe that L1 language transfer, be it positive or negative, is "just one of several factors" in L2 acquisition, and that "learners follow a similar order and sequence of acquisition irrespective of their first language" (Ellis 2015, 11). When it comes to FL sounds, language transfer is most noticeable on the level of production – L2 speakers often have easily detectable and recognisable accents. Studies have shown that explicit instruction of pronunciation can be beneficial, so teachers should work on it by considering a variety of factors, the L1 background of their students being one of them (Saito 2012; Saito and Lyster 2012).

The difficulties observed in learners' production are often described as being closely linked to perception (Watkins, Strafella, and Paus 2003). The first studies looking at perception of FL sounds can be dated to the beginning of the 1970s, with prominent papers by authors such as Abramson and Lisker (1970) and Stevens et al. (1969). These studies investigate the ability to discriminate between similar FL sounds – an ability which has been described as almost non-existent for people unfamiliar with the FL in question (Escudero 2009). Most researchers in the field agree that it is possible to acquire the ability to discriminate these contrasts (Escudero 2002), and their studies aim to identify the factors affecting these processes, such as the listener's age and native language (Bohn 1995).

While the context of acquisition has been thoroughly explored, the intrinsic difficulty of certain phonemic contrasts in relation to others in SLA remains largely uncharted territory. In this context, the distinction between consonants and vowels becomes especially relevant. Thus, the main aim of the experimental study presented here is to identify General British (GB) phonemes that are problematic for speakers of Slovene, and to determine if the acquisition rate of FL consonantal contrasts is different to the one measured for FL vocalic contrasts. The results of the study can be used to determine the causes of perception difficulties, and to provide suitable feedback to students and language teachers alike. The paper presents an overview of the relevant literature (Sections 1 and 2), the methodology and research questions (Section 3), the

results of the experiment (Section 4), a discussion of the results (Section 5) and the conclusion (Sections 6).

2 Literature Overview

2.1 Perception of Phonemes

Speech perception always involves cognitive processing of the incoming sound wave. The listener analyses its properties before assigning it a symbolic representation used in the subsequent linguistic processing. Tatham and Morton (2011, 152–68) explain that according to some theories the acoustic information is analysed on the level of sound segments and the parameters that define them. The consequent categorisation of phonemes differs from language to language, which means that certain categories may coincide, while others differ significantly.

Vowels possess comparatively simple acoustic structures, which makes their recognition relatively easy. They have a near-periodic waveform and can mostly be distinguished from one another by analysing two of their formants, F1 and F2 (Carlson, Fant and Granstrom 1970, 1975), and their transitions (Lindblom and Studdert-Kennedy 1967). They can be described with a small set of phonetic features, being restricted to the dorsal place of articulation. The distinction is made according to where (front, back) and by how much (high, low) the tongue is raised, and the shape of the lips (rounded, spread). Consonants, on the other hand, are aperiodic and have a quasi-chaotic waveform, making it difficult to determine their boundaries. They exhibit a wider spectrum of possible places of articulation, and, correspondingly, a wider array of possible phonetic features.

Since utterances of sounds in natural speech seldom occur in a vacuum, a transition over the phoneme boundaries can result in an audible effect on an adjacent vowel, affecting its length and sometimes even quality (e.g. nasality). As some consonants tend to be very briefly articulated, the cues for their discrimination are held in the adjacent vowel (Cruttenden 2014, 21–22). For example, the length of the preceding vowel can aid in the discrimination between /t/ and /d/ in words such as 'foot' and 'food' when the two plosives are inaudibly released in the word-final position.

In the process of learning English as L1, vowels are generally mastered by the age of 3, whereas fricatives, for example, are challenging for some children until the age of 5 or 6, especially due to the five possible places of articulation which require "delicate adjustments" of the tongue (Cruttenden 2014, 6, 195). At first, the fricatives in initial positions are generally replaced by the nearest plosives (e.g. /f/>/p/), and only after the voiceless fricatives are successfully acquired, can the voiced ones be expected to appear (Cruttenden 2014, 195–6).

Research has shown that across all relevant ages vowels contribute more to the intelligibility and comprehensibility of sentences than consonants. The study on normal-hearing and hearing-impaired listeners by Kewley-Port, Burkle and Lee (2007) suggests that vowels are also more easily recognised in poor listening conditions due to their lower frequency and stronger intensity of articulation. The finding may indicate that FL learners should prioritise new (unfamiliar) vowels over consonants, as this would likely lead to their faster proficiency in the new language.

In the process of second language acquisition, learners create categories in which related data are combined into larger units with their own internal structures. Such a categorisation of sounds starts at an early age – babies are able to successfully discriminate more sounds than there are in their L1. But even in the first year of their lives, they lose some of this ability, especially when it comes to sounds to which they are not exposed regularly, i.e., the sounds outside the repertoire of their L1 (Werker et al. as quoted in Kuhl 1991, 104). By this time, each L1 sound starts to form an individual phonetic category, and phonologically related phonetic categories (allophones) start forming phonological categories. Jurančič (2014, 48) observes that the fact that children establish their phonological inventory so early and that this development is greatly influenced by their L1 can "provide a basis on which to predict which features are potential sources of L1 interference in second or foreign language learning".

Each category has a range of what we still consider to be the same sound; hence, phonologically different segments can be assigned the same categories. According to Kuhl (1991, 93), there exist prototypes or "best versions" of each sound. They are recognised faster and remembered at a better rate. A prototype is the basis of a category, and as such acts as a 'perceptual magnet'. Kuhl (1991, 99) explains that the "[s]urrounding members of the category are perceptually assimilated to it to a greater degree than would be expected on the basis of real psychophysical distance. […] The perceptual space appears to be 'warped', effectively shrunk around the prototype."

As can be expected, the compacting of categories observed in L1 affects the perception of L2. Best (1995) proposes that the assimilation of new phonemes occurs in various ways. A pair of L2 segments can be assimilated to different L1 categories, to the same L1 category or outside any particular L1 category. Flege asserts that the more different an L2 sound is from an L1 sound, the more likely it is that it will be successfully discerned (Flege 1995, 240), and that "the greater the perceived difference of an L2 sound from the closest L1 sound, the more likely that a separate category will be established for the L2 sound" (Flege 1995, 264).

The L1 sound system shapes our perception of other languages, functioning as a sieve to filter the cues which are phonologically relevant in our L1 (Trubetzkoy as cited in

Flege 1995, 237). Thus, a sound that is a phoneme in L2, but only an allophone in L1, is likely to be assimilated into the L1 phonological category. However, phonetic (and, by extension, phonological) categories are not set in stone and can be somewhat moulded with sufficient practice:

> The phonetic systems used in the production and perception of vowels and consonants remain adaptive over the life span, and […] reorganise in response to sounds encountered in an L2 through the addition of new phonetic categories, or through the modification of old ones. (Flege 1995, 233)

When an L2 phoneme (or position-sensitive allophone) is successfully acquired, it forms a new phonological category which is no longer related to the existing L1 phoneme (Flege 1995, 263). Flege's observations also led him to the conclusion that although we possess the motoric abilities to pronounce foreign sounds, we fail to produce them accurately due to insufficient perception input (1995, 236).

Flege's and other studies (see Flege 1988 1992a, 1992b, 1995; Rochet 1995) have shown that the cause of mispronunciation responsible for foreign accents is of a perceptive nature. The reasons include inaccurate perception of L2 sounds, inadequate phonetic input, insufficient motivation, a wish to retain a foreign accent and incorrect habits established in the early stages of L2 learning.

2.2 Sound Systems

In Sections 2.2.1 and 2.2.2 the Standard Slovene (StS) and the General British (GB) sound systems are presented and compared. The discussion is mostly limited to the phonemes that have been previously identified as problematic for Slovene speakers (Collins, Šuštaršič, and Komar 2014; Šuštaršič 2005; Komar 2017; Stopar 2015, 2017, 2019); they represent the focus of the present study.

One of the more notable differences between StS and GB is in the number of phonemes. While the number of consonants is similar (24 and 21, respectively), StS has a much smaller vowel system than GB (8 and 21, respectively), which often leads to neutralisation of GB vowels (Šuštaršič 2005, 9). As Šuštaršič points out (2005, 5–6), problems for Slovene speakers mainly occur with the front GB vowels /e/ and /æ/.

The consonants that have been identified as problematic for Slovene speakers involve the type of articulation that can be described as a marker of a foreign accent (Collins, Šuštaršič, and Komar 2014). Šuštaršič (2005, 24) categorises them in three groups: 1) the equivalent consonantal phonemes that do not present problems for FL learners, as they are already found in L1; 2) the partially equivalent consonantal phonemes that differ either in place (/t, d, h/) or in manner (/ɹ/) of articulation, which makes them more difficult to acquire; and 3) the non-equivalent consonantal phonemes

that exist in L1 only as allophones (/ŋ, w/) or do not exist in L1 at all (/θ, ð/), which makes them the most challenging for Slovene speakers.

2.2.1 Standard Slovene

The relatively small set of StS vowels (8)[1] is presented in Figure 1 (cf. Šuštaršič, Komar and Petek 1999, 137).

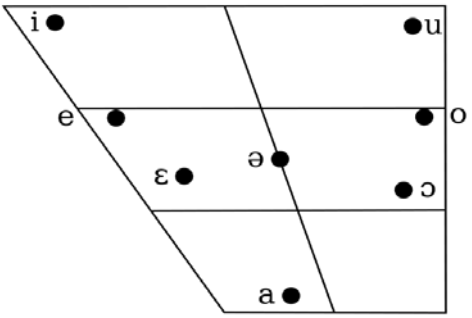


Figure 1. StS vowels.

Of the eight vowels depicted in the above vowel quadrilateral, two, /e/ and /ɛ/, are studied in this paper.

- **/e/:** Toporišič (2000, 47–49) describes /e/ as an unrounded close-mid front vowel. The tongue is pressed forward and raised somewhat lower than for /i/, the high point is approximately in line with the upper teeth. The rims of the tongue slightly touch the upper molars. The lips are spread. It corresponds to C2. Šuštaršič (2005, 10) states that StS /e/ is very close compared to the nearby GB vowels.[2] Most Slovene speakers correctly identify it as too close to replace the GB /e/.

- **/ɛ/:** In line with Toporišič (2000), **/ɛ/** is an unrounded open-mid front vowel. The tongue is pressed forward and raised slightly above the lower teeth. The rims of the tongue barely touch the upper molars. The lips are spread. It is somewhat above C3 and slightly more centralised. StS **/ɛ/** is identified by most Slovene speakers as closest to GB /e/ and /æ/, which is why it tends to be neutralised in production and perception (Šuštaršič 2005, 10; Collins, Šuštaršič, and Komar 2014, 53; Stopar 2015, 89; 2019).

[1] It should be noted that the StS vowel system distinguishes two phonological systems: the non-tonemic, based on stress and vowel-length, and the tonemic, based on stress, vowel length and pitch (Toporišič 2000, 63–64). In this paper, the issue of pitch is not addressed as the speaker involved in the experiment was instructed to produce each stimulus with the same, falling pitch.

[2] In some contexts (e.g. before /ɾ/) and especially in dialects of certain regions (e.g. Upper Carniola, Central Slovenia), it is pronounced even more close, so that it may at times be identified as /ɪ/ by speakers from other regions, cf. Tivadar (2004).

The set of StS consonants comprises 21 phonemes. This study focuses on the following consonants (cf. Šuštaršič, Komar, and Petek 1999, 135–36; Toporišič 2000, 73–82).

- **/d/** is a voiced dental plosive; it differs from GB /d/ in the place of articulation (dental in StS and alveolar in GB). When compared to GB /ð/, Šuštaršič (2005, 26, 29) observes that it has a different manner of articulation, but is nevertheless often used as a substitute for GB /ð/ even by advanced Slovene speakers of English.
- **/t/** is a voiceless dental plosive; it differs from GB /t/ in the place of articulation (StS dental and GB alveolar). StS /t/ is frequently used as a substitute for GB /θ/ (Šuštaršič 2005, 26, 29); the two phonemes differ in the manner of articulation.
- **/f/** is a voiceless labiodental fricative. Although the phoneme /f/ has the same VPM labels in StS and GB, we include it here because Slovene speakers sometimes use it as a substitute for GB /θ/ (the two sounds differ only in the place of articulation).
- **/x/** is a voiceless velar fricative; it differs from GB /h/ in the place of articulation (velar and glottal).
- **/ɾ/** is a (voiced) alveolar tap, which differs from the GB approximant **/ɹ/** in the manner of articulation.

2.2.2 General British

The inventory of GB phonemes lists 21 vowels (14 monophthongs and seven diphthongs) and 24 consonants. Their descriptions in this section are summarised from Cruttenden (2014, 96–237).

GB vowels are divided into monophthongs and diphthongs; this study focuses on the former only, namely on the KIT-vowel /ɪ/, the DRESS-vowel /e/, the SQUARE-vowel /ɛː/, and the TRAP-vowel /æ/. It is noteworthy that the list includes the SQUARE-vowel /ɛː/, which had been described as a diphthong /eə/ until the last, 8th edition of *Gimson's Pronunciation of English*. The current edition establishes /ɛː/ as the GB standard and considers the diphthong /eə/ to be a marked pronunciation (Cruttenden 2014, 78, 84).

GB has a rich vowel system – crowding in the vowel chart is most noticeable on the front axis. The presented phonemes are limited to the range from just above C2 to C4 in order to keep the quantity of data manageable and the experiment practical. The vowel chart below follows Wells (2008, xxxiii–xxiv) but is adapted to include the latest changes discussed above, namely /ɛː/ and /a/.

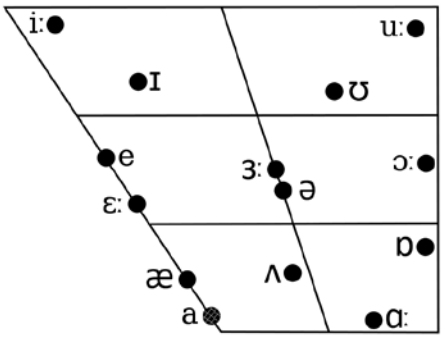


Figure 2. GB vowels.

- **/ɪ/** is an unrounded near-close near-front vowel. The part of the tongue nearer to the centre is pressed slightly forward and raised just above the close-mid position. The side rims make slight contact with the upper molars. The lips are loosely spread. It is closest to C2, but more centralised. Compared to StS /e/, it is more centralised (Collins, Šuštaršič, and Komar 2014, 47). Because Slovene does not have an intermittent vowel between StS /i/ and StS /e/, GB /ɪ/ can be problematic for Slovene FL learners. The results of a contrastive analysis by Šuštaršič (2005, 12) show that the degree of opening is approximately the same as for StS /e/.
- **/e/** is an unrounded mid front vowel. The front of the tongue is pressed forward and approximately in line with the lower teeth. The side rims make slight contact with the upper molars. The lips are loosely spread. It falls between C2 and C3, a bit closer to the latter. /e/ may present difficulties for learners if L1 has phonemes that correspond to C2 and C3 (Cruttenden 2014). Šuštaršič's contrastive analysis also shows the great proximity of GB /e/ and StS /ɛ/ (Šuštaršič 2005, 12). GB /e/ is often produced too open, and so replaced by StS /ɛ/ (Collins, Šuštaršič, and Komar 2014, 116).
- **/ɛ:/** is an unrounded open-mid front vowel. The front of the tongue is pressed forward and dropped somewhat below the lower teeth. The side rims make no contact with the upper molars. The lips are neutrally spread. It is most similar to C3.
- **/æ/** is an unrounded open-mid to near-open front vowel. The mouth is more open than for /ɛ:/. The tongue is raised between the open-mid and fully open positions. The side rims make slight contact with the back upper molars. The lips are neutrally open. It is most similar to C3. This phoneme has dropped recently to a position closer to C4, which is shown by the placement of /a/. The higher of the two positions explains the confusion with Slovene /ɛ/: in line with Šuštaršič (2005, 71), /æ/ is often pronounced too close.

With regard to the GB consonants, the study focuses on seven phonemes that have been reported as problematic in the literature (Šuštaršič 2005; Collins, Šuštaršič, and Komar 2014). They are as follows:

- **/t/** is a voiceless alveolar plosive;
- **/d/** is a voiced alveolar plosive;
- **/f/** is a voiceless labiodental fricative; in some regional English accents (e.g. Cockney), it replaces the voiceless dental fricative /θ/;
- **/θ/** is a voiceless dental fricative;
- **/ð/** is a voiced dental fricative;
- **/h/** is a voiceless glottal fricative;
- **/ɹ/** is a (voiced) alveolar approximant.

3 Methodology

The main goal of the experimental study was to assess the state of accurate discrimination of non-native phonemes in Slovene students at the end of their third year at the Department of English in Ljubljana. We hypothesised that the students would be better at discriminating vocalic contrasts rather than consonantal ones, as according to some linguistic theories vowels contribute more to the understanding of speech and carry more information on the surrounding sounds (Kewley-Port, Burkle, and Lee 2007; Cruttenden 2014). To confirm this, an experiment was designed that examined and compared the performance of Slovene third-year students of English.[3]

3.1 Participants

The participants were third-year BA or first-year MA students at the Faculty of Arts, University of Ljubljana, in the academic year 2016–17. They were invited to participate in the experiment a week prior to the experiment, and the volunteers were assigned individual time slots. The student participants formed the focus group.

To accurately place the results of the focus group in the range of ability to discriminate non-native phonemes, positive and negative control groups (hereafter CG) were also formed: one comprising advanced/native speakers and the other consisting of beginners. Since it proved impossible to find a group of participants with little to no exposure to English, the negative CG included participants who do not use English in their daily lives, never or seldom listen to English music, and who claim that they have great difficulties in speaking and understanding the language. The negative CG allowed us to assess whether the experiment is too easy. The positive CG included participants who are in contact with English daily. Their results were used to evaluate the phonemic quality of the recorded stimuli, i.e. whether the recordings sound genuine to a native speaker's ear.

[3] The experiment was part of S. Jošt's MA project (2018, supervised by A. Stopar and M. Šekli), which evolved from a previous experiment conducted by S. Jošt and D. Krassnig at the University of Konstanz (2015, supervised by M. Pohl).

Table 1. Participants in the focus and control groups.

	Focus group	Positive CG	Negative CG
Total no. of participants	33	5	5
Sex	22 female, 10 male	2 female, 3 male	3 female, 2 male
Semester of study	6th (22 part.), 8th (11 part.)	NA	NA
Age of participants	21–28, avg. 23	26–32, avg. 29	18–24, avg. 22
Native language	Slovene	Slovene	Slovene
Regular contact with English	81%	100%	40%
Type of contact	travel to English-speaking countries, media, books, audio books, work (tourism, translating, tutoring, teaching English), native-speaker friends	native-speaker friends, work (teaching English, interpreting…), media, study or work in England	music, travel

3.2 Speaker and Recording Procedure

The stimuli were recorded by a female native speaker of Slovene with professional training in English phonetics and with experience teaching this subject at the Faculty of Arts, University of Ljubljana. A bilingual native speaker of Slovene and English was not available during the preparation of the experiment; however, the speaker's pronunciation is at the level of a native speaker. The speaker also has experience with public speaking on radio, which simplified the recording procedure.

The speaker was given a list of monosyllabic words, grouped by categories and divided by languages, and asked to produce all items with the same pitch. The words were written in phonemic transcription. The speaker recorded a string of words for each phoneme separately, so that the quality of the tested phoneme (or category) remained the same. The string recordings were later cut into individual recordings (one for each word) with the program Audacity.

The digital voice recorder SONY ICD-SX1000 was used to record the stimuli. The recordings were processed in Audacity: noise reduction was applied, the pitch was equalised, and the recordings were also amplified to increase the volume of the audio.

3.3 Stimuli

The choice of stimuli was based on the 11 GB phonemes that were identified as

absent from StS, and their possible substitutes in StS (see the discussion in 2.2). They were categorised as shown in Table 2.

Table 2. Categorisation of stimuli.[4]

Label	Description	Label	Description
S1	GB vowel DRESS	S6	StS vowel DEBLO
S2	GB vowel TRAP	S7	GB voiceless consonant with a different place of articulation
S3	GB vowel SQUARE	S8	GB voiced consonant with a different place of articulation
S4	GB vowel KIT	S9	StS voiceless consonant with a different place of articulation
S5	StS vowel LES	S10	StS voiced consonant with a different place of articulation

The aim of the experiment was to determine whether the participants are able to discriminate between two similar FL phonemes (GB–GB), and a combination of an L1 phoneme and a similar FL phoneme (StS–GB). Table 3 presents the contrasts observed in the experiment.

Table 3. Categorisation of contrasts.

Code	Label	Relation	Categories	Contrast
1	DRESS VS. TRAP	GB–GB	S1–S2	/e/–/æ/
2	DRESS VS. SQUARE	GB–GB	S1–S3	/e/–/ɛ:/
3	TRAP VS. SQUARE	GB–GB	S2–S3	/æ/–/ɛ:/
4	DRESS VS. DEBLO	GB–StS	S1–S6	/e/–/ɛ/
5	TRAP VS. DEBLO	GB–StS	S2–S6	/æ/–/ɛ/
6	SQUARE VS. DEBLO	GB–StS	S3–S6	/ɛ:/–/ɛ/
7	DRESS VS. LES	GB–StS	S1–S5	/e/–/é/[5]
8	KIT VS. LES	GB–StS	S4–S5	/ɪ/–/é/
9	Place of articulation	GB–StS	S7–S9, S8–S10	/θ/–/t/; /θ/–/f/; /h/–/x/; /ɹ/–/ɾ/; /ð/–/d/
10	Voicing	GB–GB; StS–StS	S7–S8; S9–S10	/θ/–/ð/; /t/–/d/
11	Control	identical items	SX–SX	

When possible, we opted for nonsensical monosyllables, which follow the same pattern (e.g. /θaʊ/–/ðaʊ/).

[4] For GB vowels, the lexical sets created by Wells (1982) were used. In this system each vowel is represented by a prototypical monosyllabic word. Parallel lexical sets were created for StS: *les* 'wood' for StS /e/ and *deblo* 'tree trunk' for StS /ɛ/.

[5] For StS /e/ the notation with the acute accent is used henceforth /é/ to distinguish it from the GB /e/.

3.3.1 Vowels

The tested vowels appeared between two consonants (C–V–C). The initial phoneme varied to test the vocalic phoneme in different environments and to avoid the influence of the initial phoneme. The final phoneme was the voiceless plosive /t/, as it naturally appears in this context quite often in both tested languages. Voiced consonants would be unsuitable, as in word-final position in English they are devoiced (but lenis), while in Slovene they revert to their voiceless (but fortis) pairs.

Since the crowding of vocalic phonemes is most noticeable on the front axis of the vowel quadrilateral, the experiment focused on four GB monophthongs that range from just above C2 to C4 and two StS monophthongs in the same vowel space. Table 4 presents the relevant stimulus words.

Table 4. List of monosyllabic words with vocalic phonemes.

GB DRESS /e/	/let/	/jet/	/bet/	/pet/	/det/	/fet/
GB TRAP /æ/	/læt/	/jæt/	/bæt/	/pæt/	/dæt/	/fæt/
GB SQUARE /ɛː/	/lɛːt/	/jɛːt/	/bɛːt/	/pɛːt/	/dɛːt/	/fɛːt/
GB KIT /ɪ/	/lɪt/	/jɪt/	/bɪt/	/pɪt/	/dɪt/	/fɪt/
StS LES /é/	/lét/	/jét/	/bét/	/pét/	/dét/	/fét/
StS DEBLO /ɛ/	/lɛt/	/jɛt/	/bɛt/	/pɛt/	/dɛt/	/fɛt/

3.3.2 Consonants

The tested consonantal phonemes appeared in the word-initial position. In the GB set, they were followed by the diphthong /aʊ/, which is a relaxed open sound and fairly close to the Slovene [au̯].[6] Both are common in word-final position in English (e.g. in *now*) and in Slovene (e.g. in *dal* 'gave' and *rokav* 'sleeve').

As for the place of articulation of /t/ and /d/, we propose that very few, if any, Slovene students of English make the distinction between the dental (StS) and alveolar (GB) articulations (see 2.2.1), except perhaps in the word-initial position where GB /t/ is aspirated (not a phonologically distinctive feature). Thus, a single set of recordings represented both /t/ and /d/. It should also be noted here that the focus of the experiment was on /θ/ and /ð/, the pair of phonemes from the third category of difficulty (Šuštaršič 2005, 24).

Table 5. List of English and Slovene monosyllabic words with tested consonantal phonemes.

GB unvoiced consonant	/haʊ/	/θaʊ/	
StS unvoiced consonant	/xaʊ/	/taʊ/	/faʊ/

[6] StS does not have diphthongs but categorises such combinations as two phonemes (vowel + /j/ or [u̯] – a bilabial allophone of the approximant /ʋ/). Nevertheless, /aʊ/ can safely be used as a substitute for [au̯] (Šuštaršič 2005, 10), and vice versa.

GB voiced consonant	/ðaʊ/	/ɹaʊ/	
StS voiced consonant	/daʊ/	/ɾaʊ/	

3.4 Procedure

The participants listened to pairs of monosyllabic recordings which differed in one phoneme (hereinafter IDs) and identified them as the same or different. The AX method was employed for the experiment; the program ZEP (Veenker 2013) was used for the testing procedure.

The testing was divided into two phases: the preliminary phase with feedback, followed by the main testing phase without feedback. The preliminary phase was comprised of four IDs, and the main phase of 60 IDs. All IDs were played in a pseudo-randomised order. The interstimulus interval for all trials was 500 ms. After hearing the recordings, the participants had to respond in an interval of 2000 ms. The responses were identified as "correct" (1) or "incorrect" (0). If they failed to answer before the next ID was played, the response was identified as "unanswered" (–1). A counter tracking the progress of the experiment was placed at the bottom of the screen.

The experiment was conducted in May 2017 over a period of 14 days. Each participant was tested individually in a controlled environment. First, the participants were asked to fill in a short, anonymous questionnaire. Before the preliminary phase, they were given instructions in Slovene. The instructions were also given in written form before the beginning of the experiment.

The participants were then given a set of headphones and a computer mouse, and asked to complete the preliminary phase with the experiment supervisor in the room to ensure that all the conditions for successful execution of the experiment were met. The preliminary phase consisted of four IDs with feedback. After the participant completed the four sets, the supervisor answered any questions that arose and checked that all instructions were clear and the volume was comfortable. Then the supervisor left the room and the participants listened to 60 IDs without feedback.

3.5 Data Processing

The results for each participant were automatically saved as a .cvs file at the end of each trial. The answers were identified as "correct" (1), "incorrect" (0), or "unanswered" (–1); any unanswered item (due to time-out) was considered to have been answered incorrectly with a response time (hereinafter RT) of 2000 ms.

The results were categorised by group (focus group, positive CG, negative CG), by tested parameters (accuracy and response time), and by phonological categories (GB vs. StS/GB contrasted vowels, place of articulation, voicing, control).

The statistical analysis was conducted using the various functions of Microsoft Excel 2016. We calculated the total number of correct responses by all participants per ID (in numbers and percentages) to determine the easiest and most difficult phonemes. Then we applied the same method to the RT data (the average RT by all participants per ID, and the total average RT of all IDs). Next, we took the percentage of accurate discriminations by all participants per ID, and divided it by the average RT by all participants per ID. Moving average trendlines were included in all charts to investigate if the patterns align. Based on the RT trendline, four patterns were discerned with regard to perceived and actual difficulty of phonemic contrasts (also see Table 6 below):

- Pattern I: high accuracy and short RT;
- Pattern II: high accuracy and long RT;
- Pattern III: low accuracy and long RT;
- Pattern IV: low accuracy and short RT.

High/low accuracy was defined as the number of correct discriminations above/below average, respectively. Short/long RT was defined as the average RT per ID below/above the total average RT, respectively. All IDs were categorised in the correlation patterns according to these criteria.

Also, the average RT and the total number and percentage of accuracy per phonological category were calculated. To compare the results and prove the validity of both methods, phonological categories and correlation patterns were cross-referenced with the accuracy identified for the observed phonological categories.

Finally, three supercategories were established and studied: vowels (V), consonants (C) and control (CC). The average RT and the percentage of accuracy per supercategory were calculated, and the supercategories were cross-referenced with correlation patterns (expressed in percentages).

The data in the following sections is presented in the form of charts; for detailed results, see Appendices I and II.

4 Results

4.1 Focus Group

Figures 3, 4 and 5 show the percentage of correct discriminations, the average RT, and the relationship between the two.

The percentage of accuracy is presented in Figure 3.[7] Out of 44 differing IDs, nine were answered correctly 100% of the time, which accounts for 20%. IDs 3 and 7

[7] The data per ID are available in Appendix I. This includes the total sum and the percentage of correct discriminations per ID, as well as the average RT per ID. Furthermore, it shows the total average of correct discriminations, the total percentage of accuracy, and the total average RT.

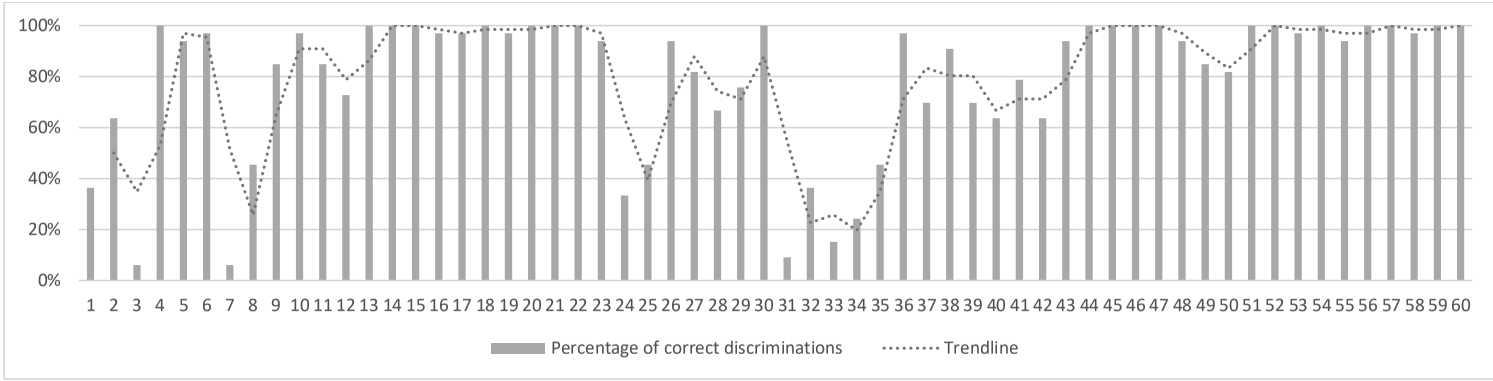


Figure 3. Percentage of correct discriminations (accuracy).

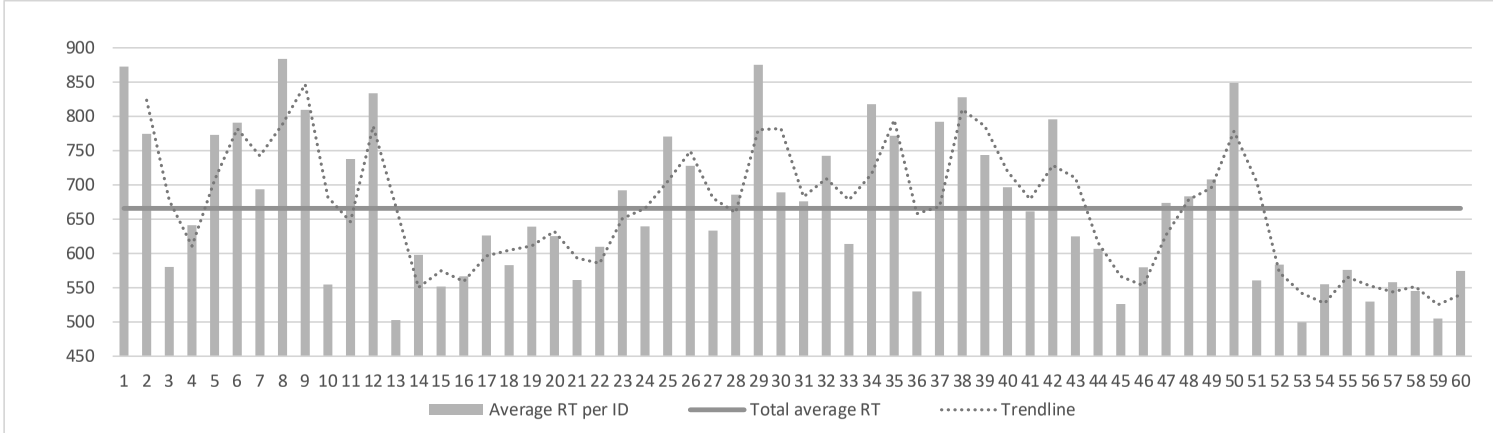


Figure 4. Average RT of all participants.

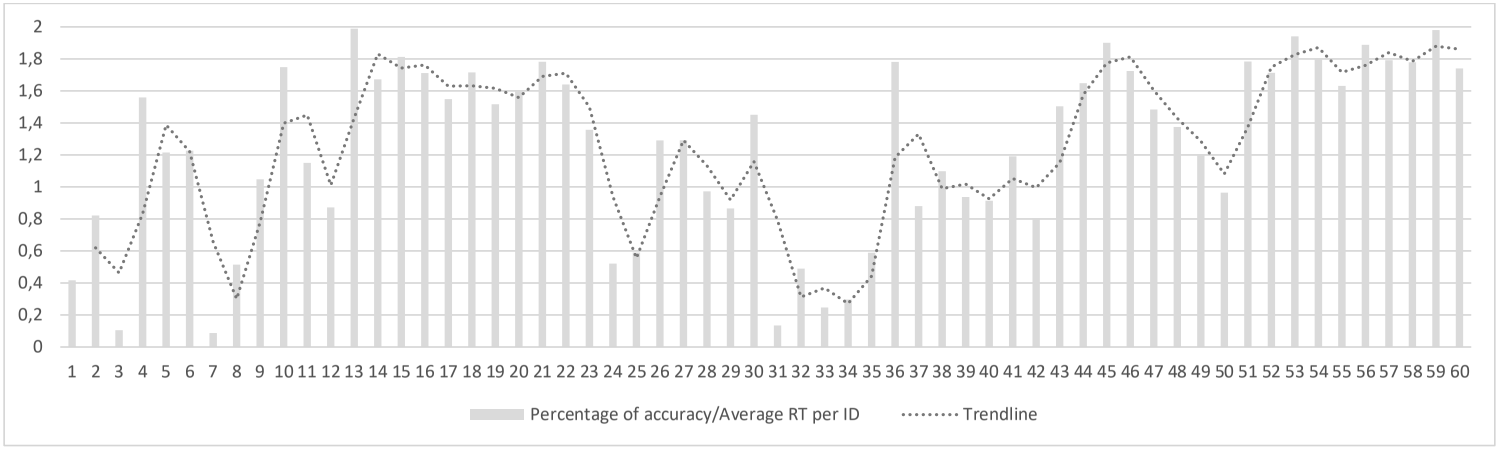


Figure 5. Percentage of accuracy/Average RT per ID.

were only answered correctly by two out of 33 participants, which is merely 6% – both entail the same two phonemes, /θ/ and /f/, only played in a different order. The accuracy for the 16 non-differing IDs ranged from 94 to 100%, averaging at 99%.

The average RT by all participants per ID is presented in Figure 4. The horizontal line signifies the total average RT by all participants for all IDs. For differing IDs, the minimal average RT per ID was 526.15 ms, and the maximum 883.85 ms. The average RT of non-differing IDs ranged from 499.67 to 626.27 ms, averaging at 557.24 ms. The total average RT was 665.74 ms.

Figure 5 presents the percentage of accurate discriminations per ID, divided by the average RT per ID. After taking into account the time component, the IDs grouped closely in Figure 3 are now farther apart. In differing IDs, the highest success coefficient was achieved for IDs 45, 21 and 36. The lowest success coefficient was achieved for IDs 7, 3 and 31.

Figures 3 and 5 show a great similarity with regard to trendline shapes.

Each ID was categorised into one of the four correlation patterns. The most numerous group of IDs follows Pattern I, and the least numerous falls into Pattern IV, see Table 6.

Table 6. Classification of IDs within correlation patterns.

Pattern	ID	Sum
I	4, 10, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 27, 36, 43, 44, 45, 46, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60	28
II	5, 6, 9, 11, 23, 26, 30, 38, 47, 48, 49, 50	12
III	1, 2, 7, 8, 12, 25, 28, 29, 31, 32, 34, 35, 37, 39, 40, 42	16
IV	3, 24, 33, 41	4

In the next step the phonological categories were considered. Table 7 links the IDs to their phonological categories. The number (ACC. [#]) and the percentage (ACC. [%]) of correct discriminations and the average RT per phonological category are presented.

Table 7. Accuracy and RT per phonological category.

CAT.	LABEL	ID	ACC. [#]	ACC. [%]	RT [ms]
1	DRESS VS. TRAP	19–22	32.75	99.24	608.82
2	DRESS VS. SQUARE	23–26	22.00	66.67	707.49
3	TRAP VS. SQUARE	27–30	26.75	81.06	720.86
4	DRESS VS. DEBLO	31–34	7.00	21.21	712.48
5	TRAP VS. DEBLO	35–38	25.00	75.76	734.05
6	SQUARE VS. DEBLO	39–42	22.75	68.94	724.27
7	DRESS VS. LES	43–46	32.50	98.48	584.35

8	KIT vs. LES	47–50	29.75	90.15	728.42
9	Place of articulation	1–10	20.80	63.03	737.40
10	Voicing	11–12	26.00	78.79	785.73
11	Control	13–18, 51–60	32.64	98.89	557.24

The highest accuracy rate (99.24%) was achieved in Category 1, while Category 4 reached a notably low accuracy of 21.21%. Excluding the control category, the best average RT was recorded for Category 7; the worst was recorded for Category 10.

To evaluate the relationship between accuracy and RT, we again cross-referenced the phonological categories with the correlation patterns. This is presented in Table 8.

Table 8. Cross-reference of phonological categories and correlation patterns.

CAT.	LABEL	Pattern I	Pattern II	Pattern III	Pattern IV
1	DRESS vs. TRAP	19, 20, 21, 22			
2	DRESS vs. SQUARE		23, 26	25	24
3	TRAP vs. SQUARE	27	30	28, 29	
4	DRESS vs. DEBLO			31, 32, 34	33
5	TRAP vs. DEBLO	36	38	35, 37	
6	SQUARE vs. DEBLO			39, 40, 42	41
7	DRESS vs. LES	43, 44, 45, 46			
8	KIT vs. LES		47, 48, 49, 50		
9	Place of articulation	4, 10	5, 6, 9	1, 2, 7, 8	3
10	Voicing		11	12	
11	Control	13–18, 51–60			

Categories 1, 7, 8 and 11 fall into a single pattern, while others appear in several patterns.

Finally, we separated the vocalic and consonantal contrasts into phonological supercategories: vowels, consonants, and the control category. Table 9 presents the accuracy and RT per supercategory.

Table 9. Accuracy and RT per phonological supercategory.

SUPERCATEGORY	$\%_{COR}$	RT [ms]
Vowels	75.19	690.09
Consonants	70.91	761.57
Control	98.89	557.24

Of the two tested supercategories, vowels have a higher accuracy rate and a lower average RT than consonants. The control category achieved a nearly perfect accuracy rate and an average RT well below the total average RT.

To assess the perceived difficulty, we again examined correlation patterns. Table 10 presents the distribution of supercategories across correlation patterns in percentages.

Table 10. Percentage of phonological supercategories in correlation patterns.

CODE	LABEL	Pattern I	Pattern II	Pattern III	Pattern IV
V	Vowels	31%	25%	34%	9%
C	Consonants	17%	33%	42%	8%
CC	Control	100%	–	–	–

The supercategories of vowels and consonants distributed across all four patterns. For a visual representation, we created a distribution chart for the patterns associated with vowels and consonants, see Figure 6 below.

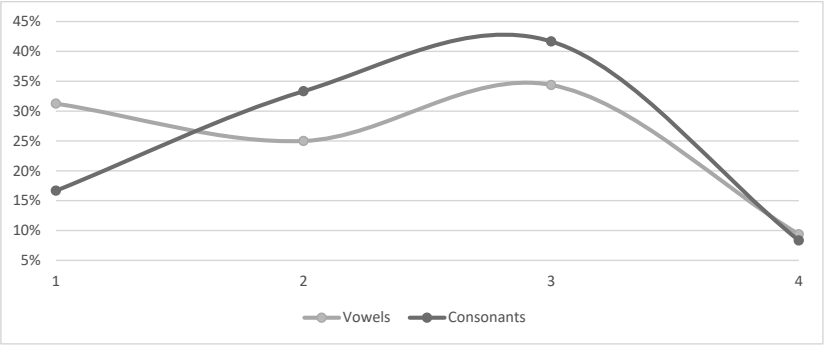


Figure 6. Distribution of supercategories across correlation patterns.

The vocalic line starts high in Pattern I, gently declines toward Pattern II, inclines toward Pattern III and steeply drops toward Pattern IV. On the other hand, the consonantal line starts low in Pattern I, then inclines all the way to Pattern III, and steeply drops toward Pattern IV.

4.2 Positive Control Group

Regarding accuracy, the results of the positive CG range from 60% to 100% per ID, with the average of 96%. The shortest RT was 402.60 ms, the longest 1292.20 ms, and the total average 665.04 ms. The worst performing phonological category was Category 4 (dress vs. deblo) with an 80% success rate.

In the supercategories, vowels achieved the average accuracy of 96% with the average RT of 690.62 ms. The accuracy rate for consonants was also 96%, with the average RT

of 718.68 ms. The control category achieved 98% with the average RT of 575.90 ms. Regarding patterns, both vowels and consonants peaked in Pattern II, with vowels starting higher in Pattern I. The control category started very high in Pattern I and steeply dropped toward Pattern III.

4.3 Negative Control Group

Regarding accuracy, the results of the negative CG range from 0% to 100% per ID. IDs 7 and 31 achieved 0%, with below average RT times. The shortest RT was 195.80 ms, the longest 959.20 ms, and the total average 596.49 ms. The worst performing phonological category was Category 4 (DRESS vs. DEBLO) with a 20% success rate. Other categories with the success rate below 65% were Categories 6 (SQUARE vs. DEBLO), 8 (KIT vs. LES), 2 (DRESS vs. SQUARE), and 9 (place of articulation). Category 7 (DRESS vs. LES) outperformed the control category (96%) and achieved a perfect score. The pattern analysis showed the worst two categories were 4 and 6.

In the supercategories, vowels achieved an average accuracy of 63% with the average RT of 631.29 ms. The accuracy rate for consonants was 71%, with the average RT of 637.63 ms. The control category achieved 95.67% with the average RT of 490.44 ms. Regarding patterns, vowels began low in Pattern I and peaked in Pattern III. Consonants began higher in Pattern I, peaked in Pattern II and then dropped steeply. ID 13 from the control category fell into Pattern IV.

5 Discussion

The findings on accuracy rates and the response times provide useful information on the perception of phoneme contrasts. We believe that the data – even in their raw form – could be used to plan and design classroom activities targeting the most problematic sounds. But before doing so, it should be considered which is more desirable, a shorter RT with a lower percentage of correct discriminations, or a longer RT with a higher percentage of correct discriminations. Or, in more general terms, which has greater weight, the accuracy of a participant's perception or their RT?

The above issue was addressed by examining the correlations between accuracy and RT. The data on accuracy and RT were combined into a single number (see Figure 5), a correlation coefficient, which indicates that accuracy should be treated as more important than RT: the accuracy rates represent the actual difficulty of the tested phonemic contrasts, while the RT data reveal the difficulty of the contrasts as perceived by the participants. The two types of results were combined into four patterns, which are listed below from most to least desirable.

Pattern I – high accuracy and short RT: The pattern marks a non-problematic contrast. The participants had little or no trouble recognising the potential difference between

the two recordings. We assume that they successfully created a separate phonological category for the non-native phoneme (in line with Flege 1995).

Pattern II – high accuracy and long RT: The pattern points to some difficulty in discrimination. More time was needed to process the difference, but the participants mostly identified it correctly. This pattern could be linked to their training in phonetics, which tends to emphasise problematic FL phoneme contrasts. For the IDs in this group, it can be assumed that the participants were in the process of creating a phonological category.

Pattern III – low accuracy and long RT: The pattern is characteristic of difficult contrasts. The participants were not confident in their answers and made incorrect decisions. We assume that a separate phonological category was not created for the non-native phoneme; however, some awareness of a difference between the recorded stimuli existed.

Pattern IV – low accuracy and short RT: The pattern includes the most challenging pairs of stimuli. The short RT reveals that the participants were confident in their decision (fast responses), but this decision was ultimately incorrect. A separate phonological category for the non-native phoneme was not created; the participants were not aware of the misperception.

5.1 Phonological Categories: Vowels

The above patterns become a valuable source of information when cross-referenced with the studied phonological categories (cf. Tables 7 and 8). Categories 1–8 focused on the perception of GB/StS vowels. The following categories were addressed.

DRESS vs. TRAP (Category 1) – Pattern I: The contrast was perceived accurately and with a fast RT. A factor contributing to the successful discrimination may be explicit instruction (in the sense of Saito 2012) combined with the number of years spent in the Department of English. The two phonemes have been previously discussed in both perception and production studies (Šuštaršič 2005; Stopar 2015; 2019; Komar 2017). Stopar's (2015; 2019) results indicate comparatively low perception rates (70% correct for DRESS and 75% correct for TRAP at the end of Year 1), while Komar's (2017) production study yielded a perfect score for DRESS and a poor result for TRAP (49% correct for reading from orthography). It should be noted, though, that these studies focus on first-year students, while our participants were at the same department two to three years longer.

DRESS vs. SQUARE (Category 2) – Patterns II, III, IV: The contrast is distributed among three patterns indicating various issues with either accuracy, RT, or both. We can conclude that both the quality and quantity of the GB vowels DRESS /e/ and SQUARE /ɛ:/ mitigate the perception; it is especially noteworthy that the results for this pair are better than those for the pair GB SQUARE /ɛ:/ and StS DEBLO /ɛ/ (see Category 6 below).

TRAP vs. SQUARE (Category 3) – Patterns I, II, III: This is an entirely non-native contrast with the main difference in both length and degree of openness. As mentioned by Šuštaršič (2005), a phoneme which differs from a familiar phoneme in two aspects is much more difficult to acquire than a phoneme which only differs in one aspect. While the participants performed well with the stimuli with an initial approximant /l/ (Patterns I and II), the stimuli with the initial plosive /p/ were more challenging (Pattern III).

DRESS vs. DEBLO (Category 4) – Patterns III, IV: The vowels in GB DRESS and StS DEBLO are close on the vowel charts, although /e/ is closer to C2 than /ɛ/. As mentioned above, Šuštaršič also found this contrast to be a problem for Slovene speakers of English, whose pronunciation of the GB /e/ tends to be too open. We can conclude that the participants do not create a separate phonological category in this case; they use the StS /ɛ/ instead of the GB /e/.

GB DRESS is located below C2, StS DEBLO /ɛ/ above C3, and GB TRAP below C3. Considering the three, we can establish that the contrast /e/–/æ/ (see Category 1) was successfully discriminated; the contrast /ɛ/–/æ/ (see Category 5) was noticeably less successful; and the contrast /e/–/ɛ/ (Category 4) was least successful. The StS vowel /ɛ/ which lies between the two GB vowels is therefore confused both with the more closed GB /e/ and with the more open GB /æ/. It can be concluded that the StS /ɛ/ acts as a perceptual magnet (Kuhl 1991), replacing GB /e/ and GB /æ/; such a result indicates that separate phonological categories for the non-native vowels were not created successfully.

TRAP vs. DEBLO (Category 5) – Patterns I, II, III: StS DEBLO contains a more open vowel than GB DRESS (see Category 1 above), so it may be easier to confuse it with that of GB TRAP. While the DRESS–TRAP contrast receives ample attention in phonetics courses, perhaps special attention should also be paid to the native StS vowel /ɛ/ in contrast to GB /æ/.

SQUARE vs. DEBLO (Category 6) – Patterns III, IV: The contrast between GB SQUARE and GB DEBLO is highly problematic. In comparison to StS /ɛ/, the GB /ɛ:/ is slightly lower on the vowel chart, more fronted, and longer, as it originates from a diphthong. It should be noted that the participants were taught the phonological system with the SQUARE vowel pronounced as the diphthong /eə/ and were not trained to recognise the contrast, which may be a contributing factor explaining the poor results. Since many dictionaries, textbooks and other study materials still treat the SQUARE-vowel as a diphthong, such misperceptions are likely to remain an issue.

DRESS vs. LES (Category 7) – Pattern I: The contrast was perceived accurately and with a fast RT. The successful discrimination is likely be due to (positive) native language transfer: the pair is like the native contrast between StS DEBLO vs. StS LES.

KIT vs. LES (Category 8) – Pattern II: The contrast falls into Pattern II – the participants considered the potential difference between two stimuli for a longer time, but ultimately made the correct decision. It has been reported in the literature that this contrast is not very obvious – Šuštaršič (2005, 12) describes GB /ɪ/ as moving in the direction of StS /e/ – which can be distracting for the participants. The formation of a separate phonological category for GB /ɪ/ is thus somewhat hindered by StS /e/.

5.2 Phonological Categories: Consonants

The categories targeting the place of articulation (Category 9) and voicing (Category 10) are distributed among all four patterns.

Place of Articulation (Category 9) – Patterns I, II, III, IV: The contrast /t/–/θ/ (and /θ/–/t/) falls into Pattern I and caused no problems. Pattern II comprises contrasts /d/–/ð/ (and /ð/–/d/) and /ɾ/–/ɹ/, which required longer RTs but were generally discriminated correctly. In combination with plosives, the non-native dental consonants are easily discriminated.

The majority of the remaining contrasts (/ɹ/–/ɾ/, /f/–/θ/, and /h/–/x/), however, matched Pattern III. According to the information and advice provided in Collins, Šuštaršič and Komar (2014, 27), "[p]roblems with /h/ are very common and persistent with Slovene learners of English and careful training is required". As for contrasts /ɾ/–/ɹ/ and /f/–/θ/ (and /θ/–/f/), it was observed that the contrasts with the native stimulus in the first position were more successfully perceived than those with the non-native stimulus in the first position. The nature of the problem may be psychological and is outside the scope of our study.

The contrast /θ/–/f/ was initially not included in the experiment, as it had not been detected as a problem in the pronunciation of advanced students; however, it is known that it does appear in the speech of untrained Slovene speakers of English. The phenomenon of switching /θ/ for /f/ is common in Cockney (Cruttenden 2014, 90) and other regional dialects, pointing to a great resemblance between the phonemes that even native speakers detect and rely on. The poor results for this pair (Pattern IV) may be caused by the noise component in both phonemes (see Cruttenden, 2014). This component is not present in /t/, which was successfully contrasted with /θ/. The results of the control category of IDs also show that the difference between the recordings (or its absence) was clear: the participants were able to correctly determine that the repeated recordings for /f/ and /θ/ were identical. Nevertheless, due to the emphasis on the pronunciation of /θ/ in phonetics (and general FL) courses, it is unlikely that the observed perception problem would permeate into production in advanced students of English.

Voicing (Category 10) – Patterns II, III: This category includes two contrasts, /θ/–/ð/ and /t/–/d/. The non-native contrast /θ/–/ð/ was discriminated accurately, but after a longer consideration, thereby falling into Pattern II. In contrast, the native contrast /t/–/d/ falls into Pattern III: the results show long RTs and inaccurate discrimination. This is unexpected because the same contrast also appeared in the preliminary phase of the experiment, where it was always discriminated correctly. Any phonological factors can thus be excluded, as well as any potential issues with the average duration of the experiment (about 3.5 minutes). The finding should be examined further in subsequent studies.

5.3 Vowels and Consonants Compared

The examination of the supercategories shows that the control category had the shortest average RT and the highest accuracy rate, while the results for the supercategories of vowels and consonants confirm our hypothesis: vowels had a higher accuracy rate and a shorter average RT.

The four patterns were also used to examine the performance of the supercategories in more detail (see Table 10 and Figure 6). The consonant supercategory exhibits a poor result in the most desirable Pattern I, whereas the same pattern can be identified for almost a third of the stimuli in the vowel category. The same can be observed for Patterns II and III – the total number of consonants following these patterns is greater than that for vowels. Pattern IV is relatively rare for both groups.

The results of the control groups are as expected. The positive CG performed very well, with negligible differences in success rates of vowel or consonant discrimination. The few incorrect responses were spread across different IDs, which indicates that it was possible to discriminate all IDs correctly.

The results of the negative CG show that the average RTs per ID were much shorter than in the focus group. This may be attributed to psychological factors, e.g., the subjects being unburdened by their results and performance. The negative CG had many problems with vocalic contrasts, achieving an average success rate of only 63%. The most problematic phonological categories contrasted /ɛ/ with /e/ or /ɛ:/. Additionally, the responses in these categories were quick, showing that the subjects believed their responses to be accurate. Another notable problem was the contrast /e/–/ɪ/. Categories including /æ/ were less problematic; the subjects may have interpreted it as /a/. Their performance was better with consonants, and near perfect with control IDs. As we predicted, the vocalic contrasts are more difficult for inexperienced speakers. The analyses of supercategories and patterns show that these subjects were generally better and faster at discriminating consonants than vowels, which suggests that their initial ability to discriminate vowels is worse than that of consonants. If this is compared to the results of the focus group, we can conclude that the same ability progresses more

rapidly and reaches a higher level in advanced students. This signifies that students tend to prioritise vowels in the acquisition of non-native phonemes.

Although the general results confirm our hypothesis on the prioritisation of vowels, the strongest perceptual magnet pull was observed in StS /ɛ/ in relation to GB /e, ɛ:, æ/. This is unsurprising as several researchers have identified these phonemes as problematic (see Šuštaršič 2005; Stopar 2015). This can be interpreted as a consequence of the similarity between the four phonemes. Most likely, many students did not create separate phonological categories for the different similar phonemes; rather, they extended their original native category to encompass the similar foreign sounds. The participants who achieved higher discrimination rates in these categories, on the other hand, can be assumed to have successfully established a new category for each of these foreign phonemes.

Although the correlation between perception and production is not perfect, we can assume that the phonemes which were discriminated correctly (Patterns I and II) most likely do not cause problems in production. The incorrectly discriminated phonemes (Patterns III and IV), however, are likely to be problematic in production, as has been suggested and established in some of the literature in this field (see Stopar 2015, 87; Komar 2017).

Even with the strong influence of the StS perceptual magnet /ɛ/ lowering the general success rate for vowels, the results show that students acquire the ability to discriminate vowels at a better rate: the vowels have a significant advantage of five percentage points over consonants. The results of the negative and positive control groups support this finding. Our hypothesis that students prioritise vowels over consonants is therefore confirmed.

6 Conclusion

The purpose of this study was to assess the state of foreign phoneme acquisition in native Slovene students of English at the end of their undergraduate studies and to determine whether vowels are prioritised over consonants in the process of foreign language learning. These aims were approached experimentally.

The obtained results show that while the students of English who participated in the experiment have reached a high level of acquisition, there are still some common problem areas. The perceptual magnet effect is most evident in StS /ɛ/ in relation to GB /e, ɛ:, æ/, with several other phonemes also showing signs of this phenomenon. In answer to our research question, we have confirmed that, on average, students do perform significantly better in perceiving a difference between similar vocalic phonemes, even though the vocalic supercategory included a strong perceptual magnet effect. This can be interpreted as students prioritising vowels over consonants in the process of FL acquisition.

It should be noted that the study has some limitations. While the size of the focus group is comparable to widely recognised studies, it was difficult to find suitable participants for the control groups. Even with this size limitation, the results show internal consistency and are reliable. We also believe that the number of contrasts in some categories should be increased in any future studies, especially in the category targeting voicing. The experiment could be expanded to involve focus groups of participants at different stages of the language learning process (beginner, intermediate, advanced), and to also consider the subjects' aptitude for languages or psychological factors affecting perception.

References

Abramson, Arthur S., and Leigh Lisker. 1970. "Discriminability along the Voicing Continuum: Cross Language Tests." In *Proceedings of the VIth International Congress of Phonetic Sciences, Prague,* 569–73. Prague: Academia.

Best, Catherine T. 1995. "A Direct Realist View of Cross-Language Speech Perception." In *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues,* edited by Winifred Strange, 171–203. Baltimore: York Press.

Bohn, Ocke-Schwen. 1995. "Cross-Language Speech Perception in Adults: First Language Transfer Doesn't Tell It All." In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by Winifred Strange, 278–300. Timonium, MD: York Press.

Carlson, Rolf, Gunnar Fant, and Björn Gränstrom. 1970. "Some Studies Concerning Perception of Isolated Vowels." *Speech Transmission Laboratory Quarterly Progress and Status Report* 11 (2–3): 19–35.

—. 1975. "Two-Formant Models, Pitch and Vowel Perception." *Auditory Analysis and Perception of Speech*: 55–82. https://doi.org/10.1016/B978-0-12-248550-3.50008-8.

Collins, Beverly, Rastislav Šuštaršič, and Smiljana Komar. 2014. *Present-Day English Pronunciation: A Guide for Slovene Students. Revised and Enlarged 2nd Edition.* Ljubljana: Znanstvena založba Filozofske fakultete, Oddelek za anglistiko in amerikanistiko.

Cruttenden, Alan. 2014. *Gimson's Pronunciation of English.* 8th Edition. Abingdon: Routledge.

Ellis, Rod. 2015. *Understanding Second Language Acquisition.* Oxford: Oxford University Press.

Escudero, Paola. 2002. "The Perception of English Vowel Contrasts: Acoustic Cue Reliance in the Development of New Contrasts." In *New Sounds 2000: Proceedings of the Fourth International Symposium on the Acquisition of Second-Language Speech, Klagenfurt*, edited by Allan James and Jonathan Leather, 122–31. Klagenfurt: University of Klagenfurt.

—. 2009. "Linguistic Perception of 'Similar' L2 Sounds." In *Phonology in Perception*, edited by Paul Boersma and Silke Hamann, 152–90. Berlin: Mouton de Gruyter.

Flege, James E. 1988. "The Production and Perception of Foreign Language Speech Sounds." In *Human Communication and Its Disorders, A Review*, edited by Harris Winitz, 224–401. Norwood, NJ: Ablex.

—. 1992a. "Speech Learning in a Second Language." In *Phonological Development: Models, Research, and Application*, edited by Charles A. Ferguson, Lise Menn, and Carol Stoel-Gammon, 565–604. Timonium, MD: York Press,

—. 1992b. "The Intelligibility of English Vowels Spoken by British and Dutch Talkers." In *Intelligibility in Speech Disorders: Theory, Measurement, and Management*, edited by Raymond D. Kent, 157–232. Amsterdam: John Benjamins.

—. 1995. "Second Language Speech Learning: Theory, Findings and Problems." In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by Winifred Strange, 233–277. Timonium, MD: York Press.

Jošt, Saša. 2018. "Perception of Foreign Phonemes: Assessment and Comparison of Slovene Students of English and Russian." MA thesis, University of Ljubljana.

Jurančič Petek, Klementina. 2014. "The 'Magnet Effect' – A Powerful Source of L1 Dialect Interference in the Pronunciation." *ELOPE* 11 (1): 45–64. https://doi.org/10.4312/elope.11.1.45-64.

Kewley-Port, Diane, T. Zachary Burkle, and Jae Hee Lee. 2007. "Contribution of Consonant versus Vowel Information to Sentence Intelligibility for Young Normal-Hearing and Elderly Hearing-Impaired Listeners." *Journal of the Acoustical Society of America* 122 (4): 2365–75. https://doi.org/10.1121/1.2773986.

Komar, Smiljana. 2017. "The Relationship between the Perception and Production of Four General British Vowels by Slovene University Students of English." *Linguistica* 57 (1): 161–70. https://doi.org/10.4312/linguistica.57.1.161-170.

Kuhl, Patricia K. 1991. "Human Adults and Human Infants Show a »Perceptual Magnet Effect« for Prototypes of Speech Categories, Monkeys Do Not." *Perception & Psychophysics* 50 (2): 93–107. https://doi.org/10.3758/BF03212211.

Lindblom, Björn E., and Michael Studdert-Kennedy. 1967. "On the Role of Formant Transitions in Vowel Recognition." *Journal of the Acoustical Society of America* 42: 830–43.

Rochet, B. L. 1995. "Perception and Production of Second-Language Speech Sounds by Adults." In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by Winifred Strange, 379–410. Baltimore: York Press.

Saito, Kazuya. 2012. "Effects of Instruction on L2 Pronunciation Development: A Synthesis of 15 Quasi-Experimental Intervention Studies." *TESOL Quarterly* 46: 842–54. https://doi.org/10.1002/tesq.67.

Saito, Kazuya, and Roy Lyster. 2012."Effects of Form-Focused Instruction and Corrective Feedback on L2 Pronunciation Development of /r/ by Japanese Learners of English." *Language Learning* 62: 595–633. https://doi.org/10.1111/j.1467-9922.2011.00639.x.

Stevens, Kenneth N., Alvin Liberman, Sven E. G. Öhman, and Michael Studdert-Kennedy. 1969. "Crosslanguage Study of Vowel Perception." *Language and Speech* 12 (1): 1–23. https://doi.org/10.1177/002383096901200101.

Stopar, Andrej. 2015. "Perception of Four General British Vowels by Slovenian University Students of English as a Foreign Language." In *PTLC2015: Proceedings of the Phonetics Teaching and Learning Conference, London, 5–7 August 2015*, 87–90. London: Phonetics Teaching and Learning Conference.

—. 2017. "Perception of the GB Vowels [...] by Slovenian Speakers of English. In *PTLC2017: Proceedings of the Phonetics Teaching and Learning Conference UCL, London, 9–11 August 2017,* 104–9. London: Phonetics Teaching and Learning Conference.

—. 2019 [forthcoming]. "Perception of General British Monophthongs: The Case of Slovene Students of English as a Foreign Language." *Jezikoslovlje*.

Šuštaršič, Rastislav. 2005. *English–Slovene Contrastive Phonetic and Phonemic Analysis and Its Application in Teaching English Phonetics and Phonology*. Ljubljana: Znanstveni inštitut Filozofske fakultete.

Šuštaršič, Rastislav, Smiljana Komar, and Bojan Petek. 1999. "Slovene." In *Handbook of the International Phonetic Association, A Guide to the Use of the International Phonetic Alphabet*, 135–139. Cambridge: Cambridge University Press.

Tatham, Mark, and Katherine Morton. 2011. *A Guide to Speech Production and Perception.* Edinburgh: Edinburgh University Press.

Tivadar, Hotimir. 2004. "Fonetično-fonološke lastnosti samoglasnikov v sodobnem knjižnem jeziku." *Slavistična revija* 52 (1), 31–48.

Toporišič, Jože. 2000. *Slovenska slovnica.* Maribor: Obzorja.

Veenker, Theo J. G. 2013. "The Zep Experiment Control Application (Version 1.6.3)." Utrecht: Utrecht Institute of Linguistics OTS, Utrecht University.

Watkins, Kate E., Antonio P. Strafella, and Tomas Paus. 2003. "Seeing and Hearing Speech Excites the Motor System Involved in Speech Production." *Neuropsychologia* 41 (8): 989–94.

Wells, John C. 1982. *Accents of English.* Cambridge: Cambridge University Press.

—. 2008. *Longman Pronunciation Dictionary.* 3rd Edition. Harlow: Pearson Education Limited.

Appendix I: List of IDs

ID	Phon. cat. A	Phon. cat. X	Phonemic transcrip. A	Phonemic transcrip. X
1	S7	S9	/haʊ/	/xaʊ/
2	S8	S10	/ɹaʊ/	/ɾaʊ/
3	S7	S9	/θaʊ/	/faʊ/
4	S7	S9	/θaʊ/	/ʈaʊ/
5	S8	S10	/ðaʊ/	/daʊ/
6	S10	S8	/daʊ/	/ðaʊ/
7	S9	S7	/faʊ/	/θaʊ/
8	S9	S7	/xaʊ/	/haʊ/
9	S10	S8	/ɾaʊ/	/ɹaʊ/
10	S9	S7	/ʈaʊ/	/θaʊ/
11	S8	S7	/ðaʊ/	/θaʊ/
12	S9	S10	/ʈaʊ/	/daʊ/
13	S9	S9	/faʊ/	/faʊ/
14	S10	S10	/ɾaʊ/	/ɾaʊ/
15	S9	S9	/xaʊ/	/xaʊ/
16	S7	S7	/θaʊ/	/θaʊ/
17	S8	S8	/ðaʊ/	/ðaʊ/
18	S7	S7	/haʊ/	/haʊ/
19	S1	S2	/pet/	/pæt/
20	S1	S2	/fet/	/fæt/
21	S2	S1	/pæt/	/pet/
22	S2	S1	/pæt/	/fet/
23	S1	S3	/det/	/dɛ:t/
24	S1	S3	/let/	/lɛ:t/
25	S3	S1	/lɛ:t/	/let/
26	S3	S1	/fɛ:t/	/fet/
27	S2	S3	/læt/	/lɛ:t/
28	S2	S3	/pæt/	/pɛ:t/
29	S3	S2	/pɛ:t/	/pæt/
30	S3	S2	/lɛ:t/	/læt/
31	S1	S6	/let/	/lɛt/
32	S1	S6	/bet/	/bɛt/
33	S6	S1	/lɛt/	/let/
34	S6	S1	/bɛt/	/bet/
35	S2	S6	/bæt/	/bɛt/
36	S2	S6	/dæt/	/dɛt/
37	S6	S2	/bɛt/	/bæt/
38	S6	S2	/fɛt/	/fæt/
39	S3	S6	/pɛ:t/	/pɛt/
40	S3	S6	/fɛ:t/	/fɛt/
41	S6	S3	/pɛt/	/pɛ:t/
42	S6	S3	/fɛt/	/fɛ:t/
43	S1	S5	/fet/	/fét/
44	S1	S5	/pet/	/pét/
45	S5	S1	/fét/	/fet/

46	S5	S1	/pét/	/pet/
47	S4	S5	/dɪt/	/dét/
48	S4	S5	/fɪt/	/fét/
49	S5	S4	/dét/	/dɪt/
50	S5	S4	/fét/	/fɪt/
51	S1	S1	/fet/	/fet/
52	S5	S5	/fét/	/fét/
53	S6	S6	/bɛt/	/bɛt/
54	S2	S2	/bæt/	/bæt/
55	S3	S3	/fɛ:t/	/fɛ:t/
56	S4	S4	/fɪt/	/fɪt/
57	S6	S6	/lɛt/	/lɛt/
58	S3	S3	/lɛ:t/	/lɛ:t/
59	S1	S1	/let/	/let/
60	S2	S2	/læt/	/læt/

Appendix II: Accuracy and RT per ID

ID	N_{COR}	$\%_{COR}$	RT per ID [ms]
1	12	36%	872.52
2	21	64%	774.45
3	2	6%	580.21
4	33	100%	641.21
5	31	94%	772.97
6	32	97%	790.85
7	2	6%	693.67
8	15	45%	883.85
9	28	85%	809.58
10	32	97%	554.73
11	28	85%	737.85
12	24	73%	833.61
13	33	100%	502.70
14	33	100%	597.94
15	33	100%	551.61
16	32	97%	566.67
17	32	97%	626.27
18	33	100%	582.91
19	32	97%	639.18
20	33	100%	625.15
21	33	100%	561.24
22	33	100%	609.70
23	31	94%	692.09
24	11	33%	639.42
25	15	45%	770.61
26	31	94%	727.85
27	27	82%	633.36
28	22	67%	685.73
29	25	76%	875.21

30	33	100%	689.12
31	3	9%	675.91
32	12	36%	742.45
33	5	15%	613.76
34	8	24%	817.82
35	15	45%	771.73
36	32	97%	544.45
37	23	70%	792.09
38	30	91%	827.94
39	23	70%	743.55
40	21	64%	696.52
41	26	79%	661.39
42	21	64%	795.61
43	31	94%	624.79
44	33	100%	606.64
45	33	100%	526.15
46	33	100%	579.82
47	33	100%	673.79
48	31	94%	683.21
49	28	85%	707.94
50	27	82%	848.73
51	33	100%	560.70
52	33	100%	583.76
53	32	97%	499.67
54	33	100%	555.09
55	31	94%	575.94
56	33	100%	529.73
57	33	100%	558.06
58	32	97%	545.24
59	33	100%	504.97
60	33	100%	574.58

Total average N_{COR}: 26.07 (of 33)

Total$\%_{COR}$: 79%

Total RT: 665.74 ms

ELOPE

Oleksandr Kapranov
Western Norway University of Applied Sciences, Norway

2018, Vol. 15 (2), 77–99(164)
revije.ff.uni-lj.si/elope
doi: 10.4312/elope.15.2.77–99
UDC: 811.111'243:81'355

Self-Assessment of the Sounds of the English Language that Pre-Service EFL Teachers Consider Problematic to Pronounce

ABSTRACT

The article presents and discusses a mixed-method study that aimed at establishing how pre-service teachers of English as a Foreign Language (EFL) self-assessed those sounds of the English language that would cause problems for their pronunciation in EFL. Fourteen pre-service EFL teachers on the intermediate level of EFL proficiency whose first language (L1) was Norwegian were recruited for the study. They were asked to write reflective essays concerning the sounds of the English language that they considered problematic to pronounce. The participants' essays were contrasted with the essays written by the control group that was comprised of 14 in-service EFL teachers whose L1 was Norwegian. The results of the analysis revealed that the participants identified several English sounds that they self-assessed as problematic to pronounce, e.g. /z/, /ð/, /θ/, and /ʌ/. The analysis of the controls' essays yielded similar results. These findings and their linguo-didactic implications are discussed in the article.

Keywords: English as a Foreign Language (EFL); pre-service EFL teachers; pronunciation; self-assessment; sounds of the English language

Glasovi angleščine, ki jih bodoči učitelji angleščine samoocenjujejo kot težje izgovorljive

POVZETEK

Članek predstavlja študijo, ki z mešano metodo ugotavlja, kako bodoči učitelji angleščine kot tujega jezika samoocenjujejo angleške glasove, ki se jim zdijo težje izgovorljivi. V raziskavi sodeluje 14 bodočih učiteljev angleščine na srednji stopnji znanja angleščine, katerih prvi jezik je norveščina. Udeležence smo prosili, naj zapišejo svoje misli o angleških glasovih, ki se jim zdijo težje izgovorljivi. Njihove zapise smo primerjali z zapisi kontrolne skupine, ki jo je sestavljalo 14 učiteljev angleščine kot tujega jezika, katerih prvi jezik je prav tako norveščina. Rezultati razčlembe so razkrili, da so udeleženci našteli več angleških glasov, ki so jih ocenili kot težavne, npr. /z/, /ð/, /θ/ in /ʌ/. Analiza kontrolne skupine je pokazala podoben rezultat. Prispevek razčleni te ugotovitve in njihov jezikovno-didaktični pomen.

Ključne besede: angleščina kot tuji jezik; bodoči učitelji angleščine kot tujega jezika; izgovarjava; samoocenjevanje; glasovi angleščine

1 Introduction

This article presents and discusses a mixed-method study of English sounds that pre-service teachers of English as a Foreign Language (EFL) whose first language (L1) is Norwegian consider problematic to pronounce. From a theoretical perspective, this study is embedded into the general theme of the present journal volume, namely the investigation of how the sounds of English are addressed in research, EFL learning and teaching, and self-assessment of pronunciation in EFL. The following two notions are central in the present study: i) self-assessment in the context of foreign language (FL) teaching and learning, and ii) the notion of pronunciation difficulties by EFL students. In the study, self-assessment is regarded as "one's own evaluation of one's performance or capabilities" (Dolosic 2018, 194). The notion of pronunciation difficulties by EFL students is operationalised in this study as a range of those English sounds and suprasegmental units that EFL learners consider challenging to pronounce.

Self-assessment is amply reported in scientific research in education (Boud 2013; Harris 1997; Oscarson 1989) and applied linguistics (Cieślicka and Rojczyk 2017; Dolosic 2018; Kapranov 2015; Lintunen 2013; Saito 2011; Szpyra-Kozłowska 2011; Szyszka 2011). Previous research indicates that self-assessment plays an important role in learning (Boud 2013; Liu and Brantmeier 2019). In this regard, Boud (2013) notes that students are able to self-assess the process and outcomes of their own learning. Self-assessment facilitates the students' awareness of and responsibility for their own learning (Boud 2013; Dlaska and Krekeler 2008; Liu and Brantmeier 2019). Boud (2013) posits that self-assessment involves the development of knowledge and a learner's awareness of the existing standards in the field of learning, as well as the capacity to reflect upon whether or not the learner meets these standards (Boud 2013, 12). In concert with Boud (2013), Lappin-Fortin and Rye (2014, 301) argue that self-assessment is a robust tool in FL teaching and learning, since it appears to increase the students' motivation and FL awareness (Lappin-Fortin and Rye 2014).

A similar approach to self-assessment is proposed by Oscarson (1989), who demonstrates that self-assessment is involved in learning and the students' awareness of learning goals (Oscarson 1989, 3–5). In unison with Oscarson (1989), Dlaska and Krekeler (2008) argue that self-assessment facilitates student-centred learning, provides insight into the learning process, stimulates pro-active learning, and supports students in identifying those sounds of the foreign language that they consider problematic to pronounce (Dlaska and Krekeler 2008, 508). These ideas map onto a definition of self-assessment as "the involvement of students in identifying standards and or criteria to apply to their work and making judgements about the extent to which they have met these criteria and standards" (Boud 2013, 11). In applied linguistics, self-assessment is defined as self-regulatory behaviour that is guided by the students' beliefs concerning how they can control their learning and increased awareness of the learning process of

their second language (SL) and/or FL (Sweet, Mack, and Olivero-Agney 2019, 177). It should be noted that the terms "self-assessment", "self-evaluation", and "self-ratings" are regarded as synonyms in the present article, in accordance with the prior studies conducted by Boud (2013), and Dlaska and Krekeler (2008).

Self-assessment is associated with a certain number of roles and functions (Boekaerts 1997). In didactics and pedagogy, in particular, self-assessment is regarded as a means of goal-setting, as a constituent part of a language diagnosis system, as well as a part of the learner's portfolio (Oscarson 2013, 2). Following Oscarson (2013), self-assessment in the present study is deemed to be a part of the learner's self-diagnosis system. However, it should be made explicit that self-assessment as a self-diagnosis system is subjective (Szpyra-Kozłowska and Stasiak 2010). As with any subjective judgement, self-assessment is thought to involve inaccuracies and learners' inflated ratings concerning their performance in an FL (Trofimovich et al. 2016).

Extending Oscarson's (2013) approach to self-assessment, it seems reasonable to suggest that self-assessment is characterised by a metacognitive function (Flavell 1979). According to Dettori and Lupi (2013), the metacognitive function of self-assessment is manifested by the learners' knowledge about knowledge, i.e. the learners are able to identify their knowledge gaps, and to distinguish factors and variables that affect the learning outcomes (Dettori and Lupi 2013). From the vantage point of metacognition, self-assessment is seen as "a key learning strategy for autonomous language learning, enabling students to monitor their progress and relate learning to individual needs" (Harris 1997, 12). In relation to EFL oral skills, the metacognitive function of self-assessment is exemplified by the learners' awareness of those individual sounds and/or suprasegmental units that cause difficulties in EFL pronunciation (Dettori and Lupi 2013).

As previously mentioned, another central notion in the present research involves those English sounds that intermediate EFL learners consider difficult to pronounce. Prior literature in applied linguistics and EFL studies indicates that there is a range of variables that are associated with the degree of difficulty that EFL learners might experience in their speech production in EFL (Huang and Radant 2009; Khamkhien 2010; Lintunen 2013; Ohata 2004; Saito 2014; Szpyra-Kozłowska 2011). Specifically, these variables involve the learner's L1, age of acquisition, EFL exposure, phonetic ability, attitude towards the acquisition of correct pronunciation, motivation, and the level of anxiety (Khamkhien 2010; Szyszka 2011). Presumably, these variables, especially the learner's L1, map onto the EFL learner's ability to pronounce the sounds of the English language correctly (Saito 2014). It is inferred from the current research literature that the learner's L1 is theorised to have a substantial impact upon potential difficulties EFL learners might encounter in terms of pronunciation of certain sounds of the English language (Ohata 2004; Saito 2014). In this regard, Huang and Radant

(2009) argue that "EFL language learners are likely to encounter difficulties when pronouncing sounds that do not exist in their first language" (Huang and Radant 2009, 116). Specifically, on the intermediate level of EFL proficiency phonetically difficult words and sounds pose significant challenges that involve "intelligibility, comprehensibility, foreign-accentedness and acceptability judgements" (Szpyra-Kozłowska 2011, 286). Previous research is indicative of the importance of "the proper understanding of the mechanisms that lie behind such serious errors which contribute to the phonetic difficulty of words" (Szpyra-Kozłowska 2011, 287) and individual English sounds that EFL learners tend to associate with difficulties.

Informed by the notions of self-assessment and the English sounds that are potentially difficult to EFL learners, the present study seeks to identify those English sounds that pre-service EFL teachers whose L1 is Norwegian (henceforth – 'participants') consider difficult to pronounce. The identification of the sounds is executed by means of the participants' self-assessment. Given that there is little research on the topic of self-assessment of the English sounds that pose difficulties to pre-service EFL teachers whose L1 is Norwegian (Hopland 2016), this study aims to establish a repertoire of those difficult sounds. Additionally, the study seeks to juxtapose the participants' repertoire of the difficult English sounds with those of the in-service EFL teachers. In particular, a group of in-service EFL teachers whose L1 is Norwegian has been asked to provide expert judgement concerning those English sounds that pose challenges to an average Norwegian L1 EFL learner on the intermediate level of EFL proficiency. The expert judgement approach follows Saito's (2011) methodology "to elicit experienced L2 teachers' opinions to determine learners' problematicity" (Saito 2011, 365). It is hypothesised in the present study that the juxtaposition of the participants' and the in-service EFL teachers' repertoires would be indicative of a range of the English sounds that both these groups subjectively evaluate as problematic and challenging to Norwegian L 1 intermediate EFL learners.

Further, this article is structured as follows. First, I will outline recent research publications that are associated with self-assessment of pronunciation difficulties in EFL by EFL learners from a variety of L1 backgrounds. Second, I will introduce the present study and discuss its findings. Third, the article will be concluded with linguo-didactic implications that would be relevant to the teaching and learning of EFL pronunciation to those EFL learners whose L1 is Norwegian.

2 Self-Assessment of Pronunciation Difficulties in EFL: Literature Review

There is a growing line of research that focuses upon the application of self-assessment to the identification of pronunciation difficulties and associated variables experienced by EFL students (Cieślicka and Rojczyk 2017; Kapranov 2015; Lintunen 2013;

Salimi, Kargar, and Zareian 2014; Szpyra-Kozłowska 2011; Szpyra-Kozłowska and Stasiak 2010; Szyszka 2011). These research studies involve EFL learners whose L1 is Farsi (Salimi, Kargar, and Zareian 2014), Finnish (Lintunen 2013), Korean (Kapranov 2015), and Polish (Cieślicka and Rojczyk 2017; Szpyra-Kozlowska 2011; Szyszka 2011). In particular, Salimi, Kargar and Zareian (2014) aim at establishing a set of difficult English sounds that have been identified by a group of Farsi L1 ELF learners. They argue that Farsi L1 EFL learners predominantly experience problems with the English diphthongs, e.g. /au/ and /əʊ/. Additionally, Farsi L1 EFL students consider problematic those consonants that are absent from their L1, e.g. /w/, /ð/, and /θ/ (Salimi, Kargar, and Zareian 2014).

Similarly to Salimi, Kargar and Zareian (2014), Lintunen (2013) investigates EFL learners' self-assessment of pronunciation problems in English. Lintunen (2013) seeks to compare subjective and objective modes of assessments of Finnish L1 advanced EFL learners' pronunciation in order to establish the connection between self-assessment and the explicit teaching of EFL phonetics. The results of the study by Lintunen (2013) reveal that Finnish L1 EFL learners are aware of their problems associated with EFL pronunciation, and their awareness has increased through teaching. Lintunen (2013) argues that "consonants caused most of the problems for the subjects. The most problematic feature was the phonemic opposition /v/–/w/. In addition, the sibilants (excluding /s/), affricates and dental fricatives were among the most difficult phonemes" (Lintunen 2013, 3–4).

Self-assessment is employed in Kapranov (2015) in order to investigate the evaluation of EFL speech fluency by Korean L1 advanced EFL learners, with the results showing that they self-assess this negatively (Kapranov 2015). The participants in the study (Kapranov 2015) indicate that there are several variables that impede their speech fluency in the English language, e.g. pronunciation, insufficient vocabulary, and limited exposure to EFL speaking contexts. However, the participants do not mention those particular English sounds that might appear problematic to Korean L1 EFL learners (Kapranov 2015).

Self-assessment in relation to pronunciation difficulties experienced by Polish L1 EFL learners is a central concept in the studies conducted by Cieślicka and Rojczyk (2017), Szpyra-Kozłowska (2011), Szpyra-Kozłowska and Stasiak (2010), and Szyszka (2011). In particular, Cieślicka and Rojczyk (2017) examine how Polish L1 EFL learners self-assess their own accent in English. Cieślicka and Rojczyk (2017) suggest that whereas the Polish L1 EFL learners' general self-assessment of their pronunciation and accent is stable, they do not associate their problems with pronunciation with particular English sounds. Similarly to Cieślicka and Rojczyk (2017), Szyszka (2011) has not found any specific English sounds that are seen as problematic by EFL learners. Instead, the participants in Szyszka (2011) indicate that

the major sources of pronunciation difficulties are associated with the suprasegmental units, consonants, and vowels.

Szpyra-Kozłowska (2011) focuses upon self-assessment of phonetically difficult words by Polish L1 intermediate EFL learners. In particular, the study argues that the contrast /z/–/s/ poses challenges to the Polish L1 EFL learners. In addition, Szpyra-Kozłowska (2011) demonstrates that the learners experience difficulties with high front vowels, i.e. /i:/ and /ɪ/, and several liquids in one word, e.g. "rural", "regularly", etc.

Szpyra-Kozłowska and Stasiak (2010) report a case of self-assessment of English pronunciation by Polish L1 EFL learners who have been exposed to a phonetic training course. Szpyra-Kozłowska and Stasiak (2010) indicate that the learners positively self-assess their correct articulation of various frequent words that they used to mispronounce. Furthermore, the learners' positive self-assessment involves attention to the relationship between English spelling and pronunciation.

As evident from the literature review, there is ample research concerning learners' self-assessment of English sounds and suprasegmentals that account for their difficulties with EFL pronunciation and speech fluency. However, little is known about the self-assessment of difficult-to-pronounce English sounds by pre-service EFL teachers whose L1 is Norwegian. In the following section of the article, I will present a mixed-method study that addresses and examines this under-researched area.

3 The Present Study: Its Context and Specific Research Aims

The present mixed-method study was contextualised within the course in English phonetics offered at a large university in Norway. The course was comprised of lectures and seminars that followed the topics described in the course book "English Phonetics for Teachers" by Nilsen and Rugesæter (2015). The topics that were addressed during the course involved such chapters in the course book as "Sound Foundation", "Consonants", "Vowels", and "The Varieties of Spoken English" (Nilsen and Rugesæter 2015).

Following Cieślicka and Rojczyk (2017), the present study involved an assumption that the participants' reflective essays on the topic "Sounds of the English Language that I Consider Problematic to Pronounce" would be indicative of their difficulties with English pronunciation. Since all participants reported that they assessed their EFL students' pronunciation during teaching practice sessions at school, it was hypothesised that the participants would self-assess their own problems with pronunciation in English. Specifically, it was hypothesised that by means of self-assessment the participants would identify a repertoire of English sounds that would be associated with a degree of difficulty. Concurrently with that assumption, however, it was theorised in the study that the participants' self-assessment could be subjective

and inflated, as indicated by Trofimovich et al. (2016). Following that contention, it was decided to contrast the participants' self-assessed repertoire of difficult-to-pronounce English sounds with that of the control group that consisted of in-service EFL teachers. Arguably, the comparison of those two repertoires would facilitate a deeper understanding of self-assessment in relation to those English sounds that were deemed difficult by Norwegian L1 intermediate EFL learners.

Based upon these assumptions, the following specific research aims were formulated:

i) to identify a repertoire of English sounds that were subjectively self-assessed as causing problems in the participants' speech production in EFL;
ii) to identify a repertoire of English sounds that the controls subjectively perceived as posing difficulties to a typical Norwegian L1 EFL learner at the intermediate level of EFL proficiency;
iii) to compare the aforementioned repertoires in order to establish which English sounds would be subjectively perceived as posing difficulties to an intermediate EFL learner whose L1 was Norwegian.

3.1 Participants

The participants in the study were 14 EFL university students (11 females and three males) who were enrolled in the teacher training programme at a large university in Norway. All participants indicated that Norwegian was their L1 and English was their FL. The participants were deemed to be at an intermediate level of EFL proficiency that was referred to as the English B1 and B2 levels in accordance with the common EU framework of proficiency in a foreign language (The Council of Europe 2011). The participants' mean age at the time of the experiment was 23.5 years. There were no bilinguals among the participants.

The control group consisted of 14 in-service EFL teachers (mean age = 44.5 years, mean duration of in-service teaching experience = 11.5 years) who were matched in terms of gender with the group of participants, i.e. 11 females and three males. The controls reported that Norwegian was their L1 and English was their FL. Analogous to the group of participants, there were neither bilinguals nor native speakers of English among the controls.

All participants and their respective controls signed a consent form that allowed the author of the article to analyse their written data for scientific purposes. The participants' and controls' identities were coded to ensure confidentiality. The following codes were used in the study to refer to participants: P and the numbers from 1 to 14, e.g. P1, P2, … P14. The same coding procedure was applied to the control group. The controls were coded as C and the numbers from 1 to 14, e.g. C1, C2, … C14.

3.2 The Corpus

The corpus consisted of the participants' and controls' reflective essays. The participants' essays were on the topic "Sounds of the English Language that I Consider Problematic to Pronounce", whereas the controls were requested to write their reflections on the topic "Sounds of the English Language that Norwegian L1 Intermediate EFL Learners Consider Problematic to Pronounce". The application of computer program The Statistical Package for the Social Sciences (SPSS 2016) to the corpus yielded the descriptive statistics summarised in Table 1 below:

Table 1. The descriptive statistics of the corpus.

N	Measure	Participants	Controls
1	Total Number of Words	4299	5586
2	M Words	307	399
3	STD	48	167
4	Minimum	197	228
5	Maximum	392	923

M = mean; STD = standard deviation.

3.3 Procedure and Method

The procedure in the study involved the following steps. First, the participants were asked to write reflective essays of approximately 300 words on the topic "Sounds of the English Language that I Consider Problematic to Pronounce". The controls were instructed to write reflective essays of 300 words on the topic "Sounds of the English Language that Norwegian L1 Intermediate EFL Learners Consider Problematic to Pronounce". It should be emphasised that the controls did not reflect on their own problems with the English sounds. Instead, they were specifically instructed to write their reflections on the possible range of English sounds that they thought would pose problems for an intermediate EFL learner whose L1 was Norwegian. The participants and their controls were given one week to write the essays. The participants and controls delivered their reflective essays to the author of this article via e-mail.

In addition to the reflective essay, the participants were asked to transcribe two short texts in IPA, one text a month prior to the writing of the essay (see Text 1 below) and another text (see Text 2 in this subsection) one month after the essay. The participants were expected to use the so-called broad IPA transcription that presupposed that allophonic nuances (for instance, the dark /l/) could be omitted and/or ignored. The participants were given one week to transcribe each text, thus making it two weeks in total for the IPA transcription task. The participants were allowed to use

pronunciation dictionaries, study aids, and the Internet in this task. The final task the participants were asked to execute was the sit-in exam in English linguistics that involved functional grammar and phonetics. At the exam, the participants were instructed to transcribe a short text in English (see Text 3 below) in IPA without any study aids, e.g. dictionaries, the course book, and online resources. The exam's duration was approximately 6 hours.

> Text 1. In 1904 an earthquake of magnitude 5.4 on the Richter scale shook Oslo, with an epicenter in the "Oslo Graben" which runs under the Norwegian capital. There are now signs that indicate that we can expect a major future earthquake in Oslo (IMDB 2018a).
>
> Text 2. From the outer reaches of space to the small-town streets of suburbia, the hunt comes home. Now, the universe's most lethal hunters are stronger, smarter and deadlier than ever before, having genetically upgraded themselves with DNA from other species. When a young boy accidentally triggers their return to Earth, only a ragtag crew of ex-soldiers and a disgruntled science teacher can prevent the end of the human race (IMDB 2018b).
>
> Text 3. Once upon a time, there was a famous musician. He met Anna, a struggling artist, who gave up on her dream to make it big as a singer. However, the famous musician decided to make her known all over the world. He wrote songs for her and went on tour with her. Soon, Anna's career took off and she became famous in North America (IMDB 2018c).

Texts 1–3 were chosen for transcription in IPA based upon the following criteria: i) suitability. In particular, the texts were deemed suitable taking into consideration that they contained all English consonants and vowels; ii) an understandable topic that the participants could relate to. As pointed by Thomas (2014), EFL materials should be authentic, relevant, and understandable. All three texts chosen for transcription in IPA were film plot summaries taken from the Internet Movie Database (IMDB) site. Following Thomas (2014), film summaries were thought to be understandable and relevant to the participants.

The participants' reflective essays were manually examined by the author of this article for the explicit presence of those sounds that the participants self-assessed as posing problems and difficulties as far as their pronunciation in the English language was concerned. Similarly, the controls' reflective essays on the topic "Sounds of the English Language that Norwegian L1 Intermediate EFL Learners Consider Problematic to Pronounce" were manually investigated for the presence of the English sounds that the controls deemed problematic for an intermediate EFL learner whose L1 was Norwegian. The participants' IPA transcriptions were

examined by the author of the article and by a lecturer in English, who served on the examiner's board at a large university in Norway. The lecturer confirmed the results of the error analysis.

3.4 Results

The results of the mixed methods analysis yielded descriptive statistics that were summarised in Tables 2–7. In particular, Table 2 involved a summary of the participants' self-assessed problematic sounds and the objective error analysis of the IPA assignments by the course teacher.

Table 2. The participants' self-assessment of problematic sounds and the objective error analysis of the IPA assignments by the course teacher.

N	**P**	**Self-Assessed Problematic Sounds**	**Errors in IPA Assignment 1**	**Errors in IPA Assignment 2**	**Errors in IPA Assignment 3**
1	P 1	/ð/ /θ/ /w/	-	-	-
2	P 2	/ð/ /θ/ /z/	/ð/ instead of /θ/	-	/s/ instead of /z/ /a:/ instead of /æ/ /ə/ instead of /e/
3	P 3	/z/ /ə/ /tʃ/	/ʌ/ instead of /ə/ /æ/ instead of /ə/	/s/ instead of /z/ /ʌ/ instead of /ə/ /ɒ/ instead of /ə/	/a:/ instead of /ə/ /s/ instead of /z/ /ʌ/ instead of /ə/ /ð/ instead of /θ/
4	P 4	/ʌ/ /z/ /s/ /ð/ /θ/ /ə/	/ʌ/ instead of /ə/	-	/s/ instead of /z/
5	P 5	/z/ /tʃ/ /ʃ/ð/ /θ/	/θ/ instead of /ð/	-	/ʌ/ instead of /ə/ /u:/ instead of /ə/ /ə/ instead of /e/
6	P 6	/θ/ /z/	-	-	/u:/ instead of /ə/
7	P 7	/θ/ /z/ /w/	/θ/ instead of /ð/ /ʌ/ instead of /ə/	/θ/ instead of /ð/ /ɒ/ instead of /ə/	/z/ instead of /s/ /s/ instead of /z/
8	P 8	/θ/ /z/	-	-	/s/ instead of /z/ /ʌ/ instead of /ə/ /ə/ instead of /e/
9	P 9	/z/	/ʌ/ instead of /ə/	-	/v/ instead of /w/ /z/ instead of /s/ /ʌ/ instead of /ə/ /s/ instead of /z/ /ʊ/ instead of /ə/
10	P 10	/z/ /s/	/s/ instead of /z/ /ɒ/ instead of /ə/	/ɒ/ instead of /ə/	/z/ instead of /s/ /a:/ instead of /ə/
11	P 11	/z/ /ə/	/ɒ/ instead of /ə/ /ð/ instead of /θ/	/s/ instead of /z/ /ɒ/ instead of /ə/	/z/ instead of /s/ /ʌ/ instead of /ə/

12	P 12	/θ/ /w/ /z/ /s/	-	/z/ instead of /s/	/z/ instead of /s/ /ʌ/ instead of /ə/ /ə/ instead of /e/
13	P 13	/θ/ /z/ /s/	/ɒ/ instead of /ə/	-	/ʌ/ instead of /ə/ /s/ instead of /z/ /z/ instead of /s/
14	P 14	ð/ /θ/ /s/ /z/ /v/ /w/	/ɒ/ instead of /ə/	/ʌ/ instead of /ə/	/v/ instead of /w/ /ʌ/ instead of /ə/ /a:/ instead of /æ/ /a:/ instead of /ə/

Explanation of the abbreviations: P = participant.

In addition to the difficult English sounds that were self-assessed by the participants, Table 2 was comprised of the participants' errors in three IPA transcription tasks. Those tasks are referred to as Assignment 1, Assignment 2, and Assignment 3 in Table 2. It should be reiterated that the execution of Assignments 1–2 involved two weeks of preparation at home with the use of all available study aids, whilst Assignment 3 was an unprepared part of the sit-in exam without any access to course books, study materials, or the Internet.

In contrast to the group of participants, the controls were instructed to write their reflective essays upon those English sounds that would typically cause problems for a Norwegian EFL learner at the intermediate level of EFL proficiency, i.e. the controls did not reflect upon their own problems with the difficult-to-pronounce English sounds. The results of the corpus of the reflective essays written by the controls are presented in Table 3.

Table 3. The controls' assessment of the sounds of the English language that pose problems for Norwegian L1 intermediate EFL learners.

N	Controls	English Sounds that Cause Problems for Norwegian L1 Intermediate EFL Learners
1	C 1	/ð/ /θ/ /w/
2	C 2	/ð/ /θ/
3	C 3	/ð/ /θ/ /z/
4	C 4	/z/ /ʧ/ /ə/
5	C 5	/z/ /s/ /ð/ /θ/
6	C 6	/z/ / θ / /ʧ/
7	C 7	/θ/ /z/ /w/
8	C 8	/θ/ /z/ /w/ /v/
9	C 9	/θ/ /z/ /w/
10	C 10	/z/ /s/ /ð/ /θ/

11	C 11	/z/ /ə//ð/ /θ/
12	C 12	/z/ /s/ /w/
13	C 13	/ð/ /θ/ /s/ /z/ /v/ /w/
14	C 14	/ð/ /θ/ /s/ /z/

The mean number of sounds that caused problems for Norwegian L1 intermediate EFL learners is summarised in Table 4 below.

Table 4. Sounds that caused problems for Norwegian L1 intermediate EFL learners (by group).

N	Measure	Participants	Controls
1	M sounds that cause problems for intermediate EFL Learners whose L1 is Norwegian (per group)	3.2	3.5
2	STD	1.5	1

M = mean; STD = standard deviation.

The percentages of those English sounds that were given in Tables 2–3 was calculated in SPSS (2016), and the results are presented in Table 5.

Table 5. The Percentages for the difficult sounds that the participants and controls consider problematic for Norwegian L1 intermediate EFL learners.

N	Difficult Sound	Participants	Controls
1	/ʌ/	7%	-
2	/ə/	21%	14%
3	/s/	36%	36%
4	/ʃ/	7%	-
5	/tʃ/	14%	14%
6	/ð/	50%	57%
7	/θ/	71%	86%
8	/v/	7%	14%
9	/w/	29%	43%
10	/z/	93%	86%

Tables 6–7 outlined the participants' mean number of errors (Table 6) and the percentage of errors per group (Table 7) in the IPA transcription tasks.

Table 6. Errors in IPA transcription tasks performed by the participants

N	Measure	Errors in IPA Assignment 1	Errors in IPA Assignment 2	Errors in IPA Assignment 3
1	M errors per group in the task	1	0.7	2.6
2	STD	0.8	1	1.3

M = mean; STD = standard deviation

Table 7. The percentage of errors per group of participants in the IPA transcription tasks.

N	Error	Percentage of Errors in IPA Assignment 1	Percentage of Errors in IPA Assignment 2	Percentage of Errors in IPA Assignment 3
1	/ʌ/ instead of /ə/	29%	14%	50%
2	/a:/ instead of /æ/	-	-	14%
3	/a:/ instead of /ə/	-	-	21%
4	/æ/ instead of /ə/	7%	-	-
5	/ə/ instead of /e/	-	-	29%
6	/ɒ/ instead of /ə/	29%	29%	-
7	/s/ instead of /z/	7%	14%	50%
8	/ð/ instead of /θ/	14%	-	7%
9	/θ/ instead of /ð/	14%	7%	-
10	/u:/ instead of /ə/	-	-	14%
11	/ʊ/ instead of /ə/	-	-	7%
12	/v/ instead of /w/	-	-	14%
13	/z/ instead of /s/	-	7%	43%

Table 8 and Table 9 below illustrate the participants' and controls' explanations and comments in the reflective essays that concerned the causes of the difficulties associated with the problematic English sounds.

Table 8. The participants' comments and explanations concerning the causes of difficulties associated with the English sounds.

N	Participants	The Participants' Comments and Explanations
1	P1	"There is no similar English sound in Norwegian."
2	P2	-
3	P3	"Differences between the native language and English…"
4	P4	"The difficult English sound is never used in Norwegian."
5	P5	"The sounds that are difficult for Norwegian learners of English as a foreign language are typically sounds that are not used in the Norwegian language."
6	P6	-
7	P7	"…the main reason for this is that we do not have the same or similar sounds in the Norwegian language. The /z/ sound can be a good example of this. This is a sound we do not use in Norwegian."
8	P8	-
9	P9	"The first sound in English that I struggle with is the /z/ sound. It is a sound we don't really use in Norwegian."
10	P10	"I think it's sometimes difficult to pronounce words that start with /v/ or /w/ because they sound like the same."

11	P11	"One of the sounds most Norwegians, myself included, struggle with is the sound /z/. This is due to the fact that /z/ doesn't appear in the Norwegian language."
12	P12	"English has different pronunciation and sounds which we cannot find in the Norwegian language. Thus, there are some sounds I find more difficult than others."
13	P13	"Most of the sounds that can challenge a Norwegian language learner of English are sounds that don´t exist in the Norwegian language."
14	P14	"The one that I struggle the most with is /q/. I think I find this sound really difficult to pronounce is probably because it is not used in Norwegian."

It should be noted that whereas Table 8 presented the participants' comments concerning their own problems and difficulties with the sounds of the English language and possible causes of these, the controls' comments summarised in Table 9 concerned intermediate EFL learners (in other words, in Table 9 the controls did not comment on the variables that caused problems with pronunciation to themselves).

Table 9. The controls' comments and explanations concerning the causes of difficulties associated with English sounds.

N	Participants	The Controls' Comments and Explanations
1	C1	"…my students tend *not* to round their lips when pronouncing an English /w/ as in *why, were, when.* Instead they use the Norwegian /v/. Again, it is said to be a Norwegian problem."
2	C2	"Norwegians often seem to use Norwegian articulation when they speak English."
3	C3	"One of the sounds my students find most problematic is /z/. I think it is a difficult sound to pronounce. We do not have that sound in the Norwegian language."
4	C4	-
5	C5	-
6	C6	"A common mistake is to substitute the sound /ð/ with /d/. A reason that many Norwegians find this sound difficult to pronounce could be that the sound /ð/ is a dental fricative, and there are no dental fricatives in the Norwegian language…"
7	C7	"The English dental fricative sounds /θ/ and /ð/ are sounds we don`t find in the Norwegian language, so the students find them hard to pronounce…"
8	C8	-
9	C9	"Many Norwegian learners find it difficult to pronounce dental fricatives. In the Norwegian language, there are no dental fricatives when it comes to consonants."
10	C10	-

11	C11	"I think that eloquent examples of difficult sounds come from words which are written in a completely different way than they are pronounced."
12	C12	-
13	C13	"…my students tend to mix /v/ and /w/. The reason for this is that in Norwegian we pronounce words spelt with 'w' in the same way as those spelt with 'v'."
14	C14	"…[the sound] my students find most problematic is /z/. I think it is a difficult sound to pronounce. We do not have that sound in the Norwegian language."

3.5 Discussion

As previously mentioned, it has been assumed in the study that the participants would self-assess their pronunciation difficulties associated with the sounds of the English language. The assumption involves the fact that the participants are pre-service EFL teachers who have experienced several sessions of teaching practice at school. Consequently, it has been hypothesised that the participants would use their assessment skills they might possess as pre-service EFL teachers in order to self-assess their own problems with pronunciation in English. The results of the analysis of the participants' reflective essays are indicative of the repertoire of English sounds that the participants associate with a degree of difficulty. As evident from Table 2 and Table 5, the participants' repertoire of self-assessed difficult sounds in English consists of predominantly consonants (e.g., /z/, /θ/, /ð/, /s/, /w/, /tʃ/, /ʃ/, and /v/) and monophthongs (e.g., /ə/, /ʌ/).

As seen in the data summarised in Table 5, the participants' most problematic sounds in the English language, e.g. /z/ (93%), /θ/ (71%), and /ð/ (50%), are absent from their L1, Norwegian. In order to illustrate these findings, let us consider the following excerpt written by one of the participants.

> (1) The sounds that are difficult for Norwegian learners of English as a foreign language are typically sounds that are not used in the Norwegian language. Personally I have struggled with following phonemes: /θ/, /z/, /w/ and /v/ … /θ/ as in *three* is a difficult sound because it does not exist in the Norwegian language. We have similar sounds, /t/ and /f/, and it is easy to replace the /θ/ with one of these sounds. /z/ as in *quiz* is a difficult sound for the same reason as the /θ/: It simply does not exist in the Norwegian language. Norwegian students (including me) usually replace it with an /s/. What I find interesting is that the /z/ sound is usually not difficult to produce, but it is easily forgotten. Another difficult sound found in the word *quiz* is the /w/. Since /w/ is not found in the Norwegian language either, it is usually replaced with a /v/. As a result we end up with *quiz* being pronounced as [kvis]. (Participant P 5, female)

Excerpt (1) and the findings in Table 5 support the literature (Ohata 2004; Saito 2014) that emphasises the role of the learners' L1 in their ability to pronounce English sounds of the English language correctly (Saito 2014). The analysis of the reflective essays written by the controls provides further support for the role of the Norwegian language in the learners' difficulties with the English sounds. Specifically, the controls and participants assess as difficult those sounds that are absent from the phonological system of the Norwegian language (e.g., /z/, /θ/, /ð/, /w/, and /ə/), as well as the contrasts /s/–/z/ and /v/–/w/. These findings appear to be in concert with the prior research by Lintunen (2013), who finds that Finnish L1 EFL learners' difficulties with the English sounds involve the phonemic opposition /v/–/w/ (Lintunen 2013, 3). Whereas Finish is not an Indo-European language, both Finnish and Norwegian lack a bilabial sound that is analogous to the sound /w/ in the English language. It is inferred from Lintunen (2013) that the difficulties with the contrasts such as /v/–/w/ are associated with the typological distance between Finnish and English.

Similar to the observations found in Lintunen (2013), five out of 14 controls (36%) indicate that the main cause of the Norwegian L1 EFL learners' difficulties is associated with those English sounds that have no equivalents in the Norwegian language. It is evident from Table 8 and Figure 1 below that the majority of the participants (71%) share the same assumption concerning the cause of their difficulties with certain English sounds.

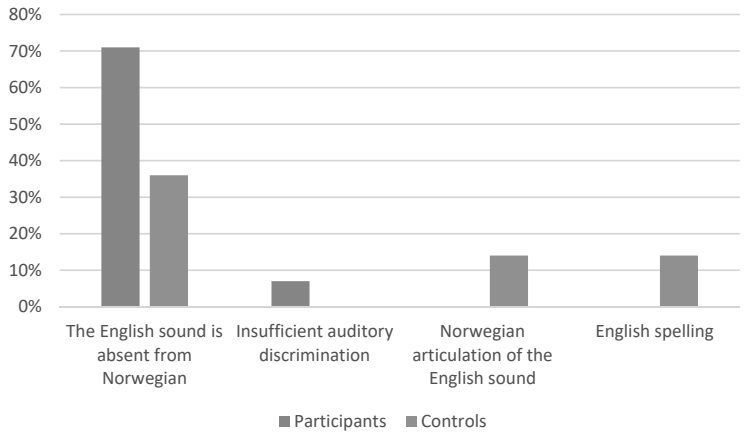


Figure 1. The causes of difficulties of problematic sounds according to the participants and controls.

Arguably, the findings shown in Figure 1 support previous studies that suggest that EFL learners "encounter difficulties when pronouncing sounds that do not exist in their first language" (Huang and Radant 2009, 116). These findings are further illustrated by excerpt 2, where a control indicates the following:

> (2) As an English teacher, I notice some difficulties that many Norwegian students, and adults, have. A person with a typical bad "Norwegian-English" or a student who starts to learn English often has difficulties with the sound /ð/. A common mistake is to substitute the sound /ð/ with /d/. A reason that many Norwegians find this sound difficult to pronounce could be that the sound /ð/ is a dental fricative, and there are no dental fricatives in the Norwegian language... (Control C 6, female)

Whereas the participants (71%) and controls (36%) attribute the causes of difficult English sounds to the fact that these do not exist in the Norwegian language, the data analysis indicates that 14% of the controls appear to associate the causes of difficulties with the English spelling conventions. In addition, 14% of the controls suggest that Norwegian L1 learners of English use typical Norwegian articulation to produce English sounds. In contrast to the control group, the participants do not refer to articulation and spelling as the source of the difficulties associated with the English sounds. Notably, the participants do not make explicit comments concerning their problems with the transcription of the English texts in IPA (see Table 2 and Table 7). Presumably, the English spelling could have triggered errors in IPA transcriptions (see Table 2). However, it does not follow from the data that the participants seem to be aware of the English spelling as a variable involved in their difficulties with certain English sounds. Obviously, the correlation between the spelling and pronunciation difficulties merits further attention. However, it is beyond the scope of the present study to offer scientific generalisations concerning this issue, since it has not been referred to by the participants. Whereas the participants do not identify the English spelling as cause of their difficulties, 7% of them point to insufficient auditory discrimination as a cause of difficulties associated with the certain English sounds. In particular, one participant writes in her reflective essay that "I think it's sometimes difficult to pronounce words that start with /v/ or /w/ because they sound the same" (Participant P 10, female).

As previously mentioned, the controls provided their expert judgements in order to determine intermediate EFL learners' problematicity in the sense postulated by Saito (2014; 2011). Notably, it is evident from Tables 2–7 that the participants and their respective controls share a common view concerning a range of English sounds that are deemed to be problematic. The participants' self-assessment and the controls' expert judgements are illustrated by Figure 2 below.

As seen in Figure 2, the English consonant sounds /z/, /θ/, /ð/, and /w/ are perceived as the most problematic by the participants and controls. The high occurrence of these sounds is explicable by the absence of these or analogous sounds in the Norwegian language. These findings lend support to the previous research literature that emphasises the connection between the problematicity of an English sound for EFL learners and the learners' L1 (Lintunen 2013; Szpyra-Kozłowska 2011).

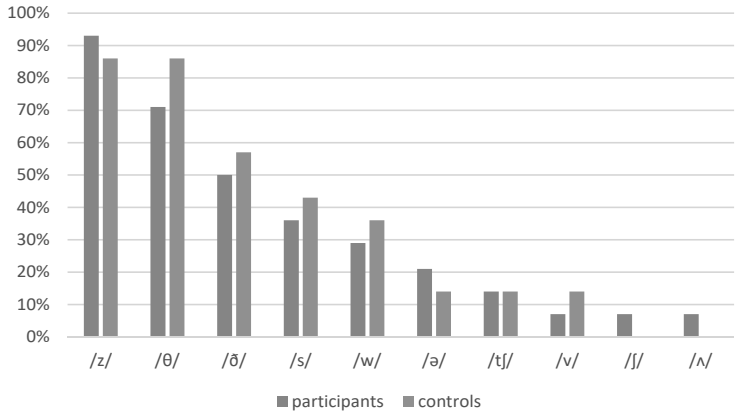


Figure 2. Difficult sounds in English according to the participants and the controls.

It is evident from Figure 2 that the participants self-assess the voiced fricative consonant /z/ as one of the most difficult English sounds. This finding supports previous research by Lintunen (2013) and Szpyra-Kozłowska (2011), who find that the contrast /z/–/s/ poses challenges to Finnish L1 and Polish L1 EFL learners, respectively. Moreover, the present findings provide indirect support to previous research that involves heritage speakers of Norwegian who reside in the USA (Haugen 1969; Moen 1988). Specifically, Moen (1988) reports a considerable number of errors related to the English fricative consonant /z/. Moen (1988) indicates that 46% of the first and second generation Norwegian Americans in the study substitute /z/ for /s/ in their oral communication in English.

The participants' subjective assessments of the voiced fricative consonant /z/ as a difficult sound are reflected in the objectively rated IPA transcription tasks, where the participants make a substantial number of mistakes that involve /z/. This finding is further exemplified by Figure 3 that illustrates the participants' self-assessment of the problematicity associated with the English consonant sound /z/, the controls' expert judgements concerning this sound, and the objective error analysis by the course teacher in the IPA task 3 (it should be remembered that this IPA task was executed by the participants without preparation and without any study aids).

As evident from Figure 3, 93% of the participants assess the English consonant sound /z/ as difficult. The errors that involve the incorrect use of /z/ by the participants account for 50% (/s/ instead of /z/) and 43 % (/z/ instead of /s/) in the IPA Task 3. These findings appear to be in unison with the research study conducted by Haugen (1969), who posits that "The most persistent difficulty of Norwegian Americans is the inability to pronounce a proper *z*, especially at the end of words" (Haugen 1969, 48).

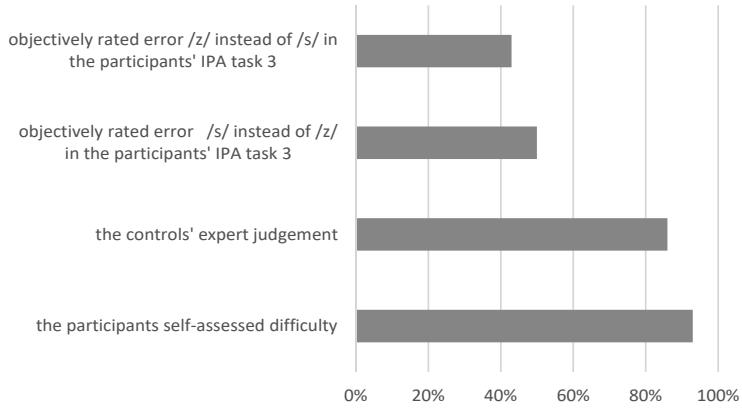


Figure 3. The English consonant sound /z/ in the participants' self-assessment, the controls' expert judgements and the objectively rated errors in IPA Task 3.

The participants' problems with the English consonant sound /z/ are evident in the context of cumulative errors in all IPA tasks, as seen in Figure 4 below. It follows from Figure 4 that in addition to the errors that involve /z/ and the /s/–/z/ contrast, the participants experience difficulties with the English short monophthong /ʌ/, especially in the IPA Task 3, and with the English neutral vowel /ə/. Whilst the short monophthong /ʌ/ is absent from the Norwegian phonological system, the participants' self-assessment of /ʌ/ as problematic is not frequent (just 7% of the participants and none of the controls).

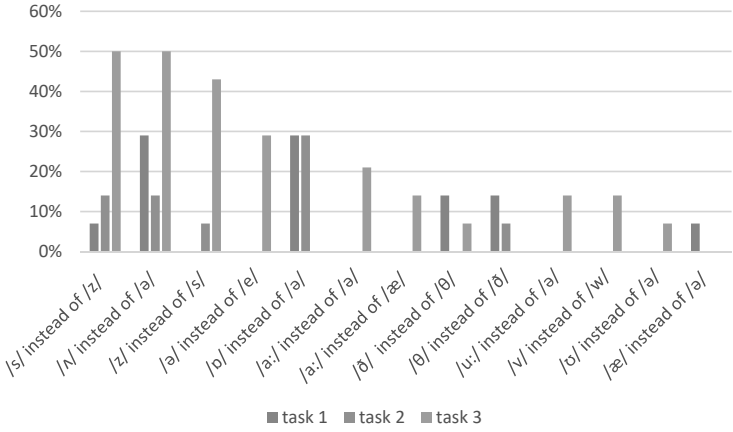


Figure 4. The percentage of errors in the IPA tasks per group of participants.

The participants' self-assessed difficulty with the English vowel /ə/ appears more frequent (21% of the participants). Even though the participants' self-assessment of the difficulties related to /ə/ seems less frequent in comparison with their self-assessment of other English sounds, especially /z/, it is, nevertheless, observed in

Figure 4 that the majority of their mistakes are associated with the incorrect use of /ə/. These findings are illustrated by Figure 5, which is based upon the participants' self-assessment of the neutral vowel sound /ə/, the controls' expert judgements, and the objective error analysis by the course teacher in the IPA Task 3.

Figure 5. The English neutral vowel sound /ə/ in the participants' self-assessment, the controls' expert judgements and the objectively rated errors in the IPA Task 3.

It is evident from Figure 5 that the participants make mistakes in the IPA Task 3 that involve several instances of incorrect representation of the English neutral vowel /ə/. Arguably, these mistakes are reflective of the participants' problems with this sound that is absent from Norwegian. Notably, previous research (Cieślicka and Rojczyk 2017; Kapranov 2015; Lintunen 2013; Salimi, Kargar, and Zareian 2014; Szpyra-Kozłowska 2011; Szpyra-Kozłowska and Stasiak 2010) does not report EFL learners' problems concerning /ə/.

Other English sounds that are frequently evaluated by the participants and their controls as difficult are the interdental fricative consonants /ð/ and /θ/ (see Table 5), which are not represented in the phonological system of the Norwegian language. This finding is in unison with the study by Salimi, Kargar and Zareian (2014), who also report EFL learners' self-assessed difficulties with /ð/ and /θ/ due to their absence in the learners' L1. Additionally, this finding lends indirect support to Szpyra-Kozłowska (2011), who reports that Polish L1 intermediate EFL learners assess the combination of "/θ/ + a consonant" as challenging to pronounce. However, it is evident from the results in the present study that the participants' subjective evaluation of the problematicity of the sounds /ð/ and /θ/ does not map onto multiple errors in IPA tasks. This finding is exemplified by Figure 6.

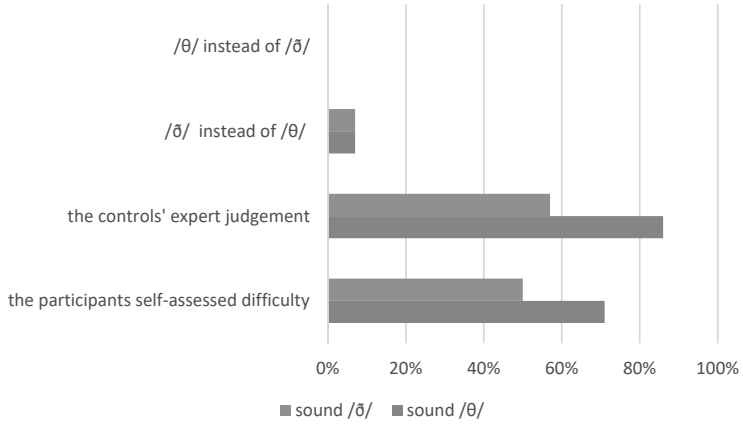


Figure 6. The English inter-dental consonant sounds /ð/ and /θ/ in the participants' self-assessment, the controls' expert judgements and the objectively rated errors in the IPA Task 3.

It is observed in Figure 6 that /ð/ and /θ/ do not seem to be associated with substantial errors in the IPA Task 3. However, it should be noted that the interdental fricative consonants /ð/ and /θ/ have traditionally been regarded as problematic for Norwegian L1 speakers of English (Haugen 1969). Judging from the present data, the same argument can be applied to the approximant /w/, which is assessed as difficult by 29% of participants and 43% of controls. Similarly to /ð/ and /θ/, the approximant /w/ is not involved in numerous mistakes in the IPA tasks (14% errors in the IPA Task 3 and no errors in other tasks).

4 Conclusions and Linguo-Didactic Implications

The study established that the application of self-assessment to the identification of difficult English sounds by intermediate EFL learners whose L1 was Norwegian resulted in a repertoire of English sounds that was similar to that of the control group. Their shared repertoire of the problematic English sounds consisted of predominantly consonants /z/, /θ/, /ð/, /s/, /w/, /tʃ/, /v/, and one neutral vowel /ə/. Concurrently with those findings, it was found that whilst the participants and the controls assessed the sounds /ð/ and /θ/ as highly problematic (e.g. 50% and 71% of the participants), the problematicity of those sounds did not map onto multiple errors in the IPA transcription tasks. In contrast, whereas the neutral vowel /ə/ was not assessed by all the participants as difficult, it caused a significant number of errors in the IPA tasks.

Since the study did not involve a substantial number of participants (N of participants = 14 and N of controls = 14, thus making it 28 in total), the results of the study should be treated with caution. Apart from the limited number of participants, another shortcoming in the study involved the lack of focus on the difficult-to-pronounce suprasegmental units. Arguably, an investigation of the Norwegian EFL

learners' difficulties with individual English sounds and suprasegmentals would be desirable. Nevertheless, the present study has several linguo-didactic implications that would be relevant to those pre-service EFL teachers whose L1 is Norwegian. These linguo-didactic implications are as followed. First, pre-service EFL teachers whose L1 is Norwegian should be provided with possibilities to use self-assessment as a means of identifying their problems with pronunciation in English, especially by means of reflection upon difficult-to-pronounce English sounds. Second, Norwegian L1 pre-service EFL teachers should pay specific attention to the English /s/–/z/ contrast, and, in particular, to the voiced fricative consonant /z/. Third, Norwegian L1 pre-service EFL teachers should be made aware of the distinction between the short monophthong /ʌ/ and the neutral vowel sound /ə/ in English.

Acknowledgements

The author of this article wants to acknowledge those pre- and in-service EFL teachers who took part in the study. Their participation is highly appreciated.

References

Boekaerts, Monique. 1997. "Self-Regulated Learning: A New Concept Embraced by Researchers, Policy Makers, Educators, Teachers, and Students." *Learning and Instruction* 7 (2): 161–86. https://doi.org/10.1016/S0959-4752(96)00015-1.

Boud, David. 2013. *Enhancing Learning through Self-Assessment*. London/NY: Routledge.

Cieślicka, Ewa Grażyna, and Arkadiusz Rojczyk. 2017. "Self-Reported vs. Self-Rated Pronunciation in a Non-Native Language." *Theory and Practice of Second Language Acquisition* 3 (2): 69–86.

Dettori, Giuliana, and Valentina Lupi. 2013. "Self-Observation and Shared Reflection to Improve Pronunciation in L2." In *International Handbook of Metacognition and Learning Technologies*, edited by Roger Azevedo and Vincent Aleven, 615–25. Berlin: Springer.

Dlaska, Andrea, and Christian Krekeler. 2008. "Self-Assessment of Pronunciation." *System* 36 (4): 506–16. https://doi.org/10.1016/j.system.2008.03.003.

Dolosic, Haley. 2018. "An Examination of Self-Assessment and Interconnected Facets of Second Language Reading." *Reading in a Foreign Language* 30 (2): 189–208.

Flavell, John H. 1979. "Metacognition and Cognitive Monitoring: A New Area of Cognitive
Development Inquiry." *American Psychologist* 34: 906–11.

IMDB. 2018a. "Skjelvet." Accessed 1 December 2018. https://www.imdb.com/title/tt6523720/.

—. 2018b. "The Predator." Accessed 1 December 2018. https://www.imdb.com/title/tt3829266/plotsummary.

—. 2018c. "A Star is Born." Accessed 1 December 2018. https://www.imdb.com/title/tt1517451/?ref_=nv_sr_1.

Harris, Michael. 1997. "Self-Assessment of Language Learning in Formal Settings." *ELT Journal* 51 (1): 12–20. https://doi.org/10.1093/elt/51.1.12.

Haugen, Einar Ingvald. 1969. *The Norwegian Language in America. A Study in Bilingual Behavior.* Bloomington: Indiana University Press.

Hopland, Amalie A. 2016. "Spoken English in the Classroom – A Study of Attitudes and Experiences of Spoken Varieties of English in English Teaching in Norway." MA thesis, University of Bergen.

Huang, Hui-Ling, and James Radant. 2009. "Chinese Phonotactic Patterns and the Pronunciation Difficulties of Mandarin-Speaking EFL Learners." *The Asian EFL Journal Quarterly* 11 (4): 115–34.

Kapranov, Oleksandr. 2015. "Self-Evaluation of Speech Fluency in English as a Second Language by Korean Exchange Students Studying in Sweden." In *Language, Thought and Education: Exploring Networks,* edited by Leszek Szymanski and Marek Kuczynski*,* 61–77. Zielona Gora: University of Zielona Gora.

Khamkhien, Attapol. 2010. "Thai Learners' English Pronunciation Competence: Lesson Learned from Word Stress Assignment." *Journal of Language Teaching & Research* 1 (6): 757–64. https://doi.org/10.4304/jltr.1.6.757-76.

Lappin-Fortin, Kerry, and B. J. Rye. 2014. "The Use of Pre-/Posttest and Self-Assessment Tools in a French Pronunciation Course." *Foreign Language Annals* 47 (2): 300–320. https://doi.org/10.1111/flan.12083.

Lintunen, Pekka. 2013. "The Effect of Phonetic Knowledge on Evaluated Pronunciation Problems." In *Proceedings of the Phonetics Teaching and Learning Conference,* edited by John A. Maidment, 55–58. London: UCL.

Liu, Huan, and Cindy Brantmeier. 2019. "'I Know English': Self-Assessment of Foreign Language Reading and Writing Abilities among Young Chinese Learners of English." *System* 80: 60–72. https://doi.org/10.1016/j.system.2018.10.013.

Moen, Per. 1988. "The English Pronunciation of Norwegian-Americans in Four Midwestern States." *American Studies in Scandinavia* 20 (2): 105–21.

Nilsen, Thor Sigurd, and Kåre N. Rugesæter. 2015. *English Phonetics for Teachers*. Oslo: Fagbokforlaget.

Ohata, Kota. 2004. "Phonological Differences between Japanese and English: Several Potentially Problematic Areas of Pronunciation for Japanese ESL/EFL Learners." *The Asian EFL Journal* 6.

Oscarson, Mats. 2013. "The Challenge of Student Self-Assessment in Language Education." *Voices in Asia Journal* 1 (1): 1–14.

—. 1989. "Self-Assessment of Language Proficiency: Rationale and Applications." *Language Testing* 6 (1): 1–13. https://doi.org/10.1177%2F026553228900600103.

Saito, Kazuya. 2014. "Experienced Teachers' Perspectives on Priorities for Improved Intelligible Pronunciation: The Case of Japanese Learners of English." *International Journal of Applied Linguistics* 24 (2): 250–77. https://doi.org/10.1111/ijal.12026.

—. 2011. "Identifying Problematic Segmental Features to Acquire Comprehensible Pronunciation in EFL Settings: The Case of Japanese Learners of English." *RELC Journal* 42 (3): 363–78. https://doi.org/10.1177%2F0033688211420275.

Salimi, Marjan, Ali A. Kargar, and Arash Zareian. 2014. "Self-Assessment of Pronunciation and Its Impact on EFL Learners' Pronunciation Ability." *International Journal of Language Learning and Applied Linguistics World* 6 (4): 605–22.

SPSS. 2016. *IBM SPSS Statistics for Windows, Version 24.0*. Armonk, NY: IBM Corp.

Sweet, Gabriela, Sara Mack, and Anna Olivero-Agney. 2019. "Where Am I? Where Am I Going, and How Do I Get There?: Increasing Learner Agency Through Large-Scale Self-Assessment in Language Learning." In *Foreign Language Proficiency in Higher Education*, edited by Paula Winke and Susan M. Gass, 175–95. Switzerland: Springer.

Szpyra-Kozłowska, Jolanta. 2011. "Phonetically Difficult Words in Intermediate Learners' English." In *Speaking and Instructed Foreign Language Acquisition*, edited by Mirosław Pawlak, Ewa Waniek-Klimczak, and Jan Majer, 286–99. Bristol: Multilingual Matters.

Szpyra-Kozłowska, Jolanta, and Sławomir Stasiak. 2010. "From Focus on Sounds to Focus on Words in English Pronunciation Instruction." *Research in Language* 8: 1–12. https://doi.org/10.2478/v10015-010-0012-7.

Szyszka, Magdalena. 2011. "Foreign Language Anxiety and Self-Perceived English Pronunciation Competence." *Studies in Second Language Learning and Teaching* 2 (1): 283–300.

The Council of Europe. 2011. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR).* Accessed 1 January 2019. https://www.coe.int/en/web/common-european-framework-reference-languages.

Thomas, Catherine. 2014. "Meeting EFL Learners Halfway by Using Locally Relevant Authentic Materials." *English Teaching Forum* 52 (3): 14–23.

Trofimovich, Pavel, Talia Isaacs, Sara Kennedy, and Kazuya Saito. 2016. "Flawed Self-Assessment: Investigating Self-And Other-Perception of Second Language Speech." *Bilingualism: Language and Cognition* 19 (1): 122–40.

ELOPE

Smiljana Komar
University of Ljubljana, Slovenia

2019, Vol. 16 (1), 101–122(164)
revije.ff.uni-lj.si/elope
doi: 10.4312/elope.16.1.101–122
UDC: 811.111'355:37.09.3

The Challenges, Methods and Results of Teaching GB Pronunciation to Slovene EFL Students

ABSTRACT

The paper presents and discusses the results of a study whose main purpose was to test the oral production of General British (GB) sounds in connected speech by Slovene BA students of English. Previous studies in contrastive English-Slovene pronunciation were mainly concerned with the perception and production of individual sounds. Our study, on the other hand, focused on the production of GB sounds in connected speech. We were interested in the state of affairs of English pronunciation before and after a 60-hour intensive and systematic theoretical and practical instruction of English pronunciation. The results confirmed out initial two hypotheses that the influences of L1 phonological and phonetic system, orthography and General American pronunciation were stronger before the instruction, and that the phonemic transcription has a very positive influence on the acquisition of foreign sounds in EFL students.

Key words: foreign language learning; English-Slovene pronunciation error analysis; teaching English phonetics and phonology; phonemic transcription

Izzivi, metode in rezultati poučevanja standardne britanske izgovarjave pri slovenskih študentih angleščine kot tujega jezika

POVZETEK

Članek prikazuje in obravnava rezultate raziskave, katere namen je bil preveriti izgovarjavo standardnih britanskih glasov v strnjenem govoru pri slovenskih študentih angleščine na prvi bolonjski stopnji študija. Predhodne kontrastivne angleško-slovenske raziskave izgovarjave so se osredotočale predvsem na percepcijo in produkcijo posameznih glasov. Z našo raziskavo pa smo želeli ugotoviti, kakšna je izgovarjava standardnih britanskih glasov v strnjenem govoru. Zanimalo nas je stanje izgovarjave pred in po šestdeseturnem intenzivnem in sistematičnem pouku teoretičnih in praktičnih vsebin angleške izgovarjave. Rezultati so potrdili naši dve izhodiščni hipotezi: prvič, vplivi glasoslovnih značilnosti maternega jezika, pisave ter ameriške angleščine so večji pred začetkom pouka, in drugič, fonemska transkripcija ima pozitiven vpliv na usvajanje tujih glasov pri študentih angleščine kot tujega jezika.

Ključne besede: učenje tujih jezikov; angleško-slovenska analiza napak izgovarjave; poučevanje angleške fonetike in fonologije; fonemska transkripcija

1 Introduction

In the era when English has become the international language of mass communication or the so-called *Lingua Franca*, the teaching of the British standard pronunciation to EFL students presents many challenges, from linguistic to motivational and pedagogical.

This paper focuses mainly on linguistic challenges although the motivational and pedagogical issues will also be addressed in the context of the English pronunciation taught at the English Department at the Faculty of Arts, University of Ljubljana. The course on English Phonetics, Phonology and Pronunciation is scheduled in the first year of BA studies. The pronunciation model and hence the target pronunciation for L2 Slovene students has always been the General British (GB) pronunciation, formerly known as Received Pronunciation (RP). Most Slovene students of English, whose general knowledge of English is at the CEFR level B2, speak a hybrid between American and British varieties with features arising from their L1 (Šuštaršič 2005, 70). Bridging the gap between the prescribed pronunciation model and the students' pronunciation presents the first major challenge in the teaching of English pronunciation to Slovene students.

The second challenge concerns the awareness and acquisition of all the features of the selected pronunciation model. To meet this requirement, the practical teaching of English pronunciation to Slovene students of English consists of several strategies which include (i) different ear-training exercises whose purpose is to improve the learners' perception of GB phonemes, (ii) phonemic (and to some extent also phonetic) transcription of written texts, (iii) reading phonemically transcribed texts, and (iv) reading regular texts. All these strategies are supported by the theory of GB phonology and phonetics. In recent years, several perception and production studies have been conducted (Šuštaršič 2005; Stopar 2015, 2017, 2019; Komar 2017) and their findings suggest different degrees of overlap between the perception, oral production, spelling and phonemic transcription of sounds.

In this paper we present the results of a study the purpose of which was to validate the findings of previous studies, as well as to detect other reasons for typical pronunciations errors made by Slovene EFL students. Different from previous studies, which focused on the perception and production of individual sounds in the context of one word, the present study focused on the pronunciation of sounds in connected speech, more precisely, in the linguistic context of one sentence.

The paper presents an overview of the findings in perception, production and acquisition of foreign sounds (Section 2), comparison of General British and standard Slovene phonological and phonetic systems (Section 3), predicted errors made by

Slovene EFL speakers (Section 4), hypotheses and methodology of the study (Section 5), results and their discussion (Section 6 and Section 7), and Conclusion (Section 8).

2 Research on Perception, Production and Acquisition of Foreign Sounds

Several factors govern the acquisition of foreign sounds. The ability to discriminate between L1 and L2 sounds is among the most important ones. It is often hindered by the influence of the phonemic and phonetic system of L1, since the judgements on similarity or difference between two phonemes are made against the background knowledge and phonemic awareness of the mother tongue. Flege (1995) and Escudero (2002) both believe that the majority of production errors are based on a false perception of L2 phonemes. Similarly, L2 speakers can create new L2 phonemes solely by relying on some phonetic or phonemic category of their L1 (Flege 1995; Best 1995). Escudero (2002) also suggests that foreign learners can learn to perceive foreign contrasts. Stopar (2015, 2017, 2019), who has analysed the pre-training and post-training perception of GB monophthongs by Slovene EFL students, confirms that perception of foreign sounds can be successfully acquired by systematic perception training even in those foreign vowels which share the same vowel space with L1 vowels.

According to Baker and Trofimovich (2006) and Bion et al. (2006), who have looked into the relationship between the perception and production of L2 vowels, successful perception results in successful production of sounds. Other studies (see Lord 2005; Lipinska 2013) suggest that explicit instruction in English phonetics and phonology improves learners' production of sounds.

There has been very little research in second language acquisition on the influence of L1 orthography on the perception and production of L2 phonemes. Bassetti (2008) argues that orthographic input of L1 interacts with L2 acoustic input, leading to nonnative pronunciations which cannot be attributed to the influence of L1 phonological and phonetic system. Research has shown that learners' mental representations of L2 phonemes created under the influence of orthographic input may affect their perception in such a way that they hear non-existing phonemes or phonetic features (see Matthews and Brown 2004). These mental representations are so strongly imprinted in the brain that they affect the production and perception also in the absence of the orthographic input.

Practising phonemic transcription of a foreign language whose grapheme-phoneme correspondence is opaquer (e.g. English) is particularly beneficial for those EFL speakers whose L1 exhibits a close grapheme-phoneme correspondence (e.g. Slovene). Phonemic transcription not only reflects pronunciation errors resulting

from orthographic input, but also develops L2 phonetic awareness in L1 learners. Studies by Wells (1996), Šuštaršič (2005), Lintunen (2005) and Komar (2017) all confirm the positive effect of phonemic transcription on L2 pronunciation of a foreign language with an opaque-grapheme correspondence. The correlation between the phonemic transcription and the pronunciation of consonants is stronger than that between the phonemic transcription and vowels.

3 General British (GB) vs. Standard Slovene (StS) Phonological System

3.1 Vowels

Studies in contrastive English-Slovene analysis of the vocalic system have shown that the two languages differ not only in the number of vocalic phonemes but also in their quality and quantity (see Šuštaršič, Komar and Petek 1995; Šuštaršič 2005; Stopar 2015). Stopar (2015: 87) presents the comparison between StS and GB monophthongs as shown in Figure 1.

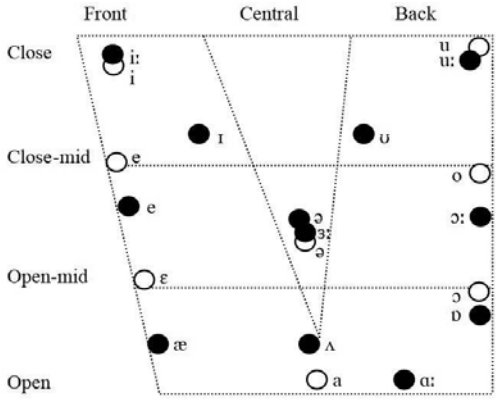


Figure 1. Vowel chart with StS (white dots) and GB (black dots) monophthongs.

Several studies (Šuštaršič 2005; Collins, Šuštaršič, and Komar 2014; Stopar 2015) in contrastive analysis of StS and GB vowel systems have shown that the most difficult GB vowel contrast for Slovene speakers of English to master is the contrast between the front close-mid to open-mid /e/ – or the dress-vowel – and the front just above open /æ/ – or the trap vowel – (transcribed in Cruttenden (2014) as /a/)[1]. The contrast is even more challenging since there is a similar vowel contrast in StS between /e/ (front close-mid) and /ɛ/ (front open-mid). Contrary to the expectation that Slovene EFL speakers would simply replace the two GB vowels with the StS

[1] dress and trap are two of 27 keywords introduced by Wells (1982) to represent a large number of words with the same vowel. Throughout the paper these keywords are used to refer to GB vowels.

equivalents, this is rarely the case[2]. It seems that the StS /e/ vowel is perceived by Slovene EFL speakers as too close to be used instead of the GB DRESS-vowel. Instead, both GB vowels are perceived, and as a consequence also pronounced, as the open-mid front vowel /ɛ/. This often results in homophonic realizations of minimal pairs, such as for example *bet/bat /ˈbɛt/* or *set/sat /*sɛt/*.

Slovene EFL learners may also find the central to back, open to mid-open vowel space problematic since the GB STRUT-vowel /ʌ/ and the GB PALM-vowel /ɑː/ may both be pronounced with the Slovene central open /a/ vowel. In addition, the GB STRUT-vowel is, under the influence of the letter <o>, often mispronounced as the GB LOT vowel /ɒ/.

There are two GB vowels in the back, open-mid to close-mid vowel space – GB LOT /ɒ/ and GB THOUGHT /ɔː/ – which have two very closely corresponding StS vowels: /ɔ/ and /o/, respectively. Although the situation mirrors the front vowel opposition between /e/ and /æ/, neutralization between /ɒ/ and /ɔː/ does not occur. These two GB back vowels are not problematic either in terms of production or perception (see Stopar 2017). Under the influence of General American (GA), however, the GB THOUGHT vowel is often pronounced as the GA /ɑ/, whereas the GB PALM-vowel is replaced by the GB TRAP-vowel in the words belonging to the BATH lexical set.

As presented in Figure 1, there are no corresponding StS vowels to GB KIT /ɪ/ and GB FOOT /ʊ/ vowels. These two vowels are often replaced by the nearest StS vowels, that is /i/ and /u/ which are both too close, as well as too front and back, respectively.

Maintaining the length of the GB long monophthongs before voiced consonants or silence is also a challenge for Slovene EFL learners since length is not a distinctive phonological feature in StS. In addition, all StS obstruents are voiceless and fortis before silence or another voiceless consonant. Hence, Slovene EFL learners often pronounce long GB monophthongs before silence or voiced consonants as too short, thus not differentiating between pairs of English words, such as *bead/beat, hard/heart, cord/caught, lose/loose*.

There are no phonemic diphthongs in StS. The nearest equivalents to GB closing diphthongs in StS are clusters of vowels followed by /j/ or [w][3]: [ej, oj, ɔj, aj, uj; ew, ɛw, aw, ɔw]. The StS diphthongs [aj, ej, ɔj, aw] are equivalent to GB diphthongs /aɪ, eɪ, ɔɪ, aʊ/, whereas the StS diphthong [ɔw] differs from the GB diphthong /əʊ/ in the first element which is in StS back and open-mid. There are no equivalents whatsoever to GB centring diphthongs. These represent a problem for Slovene EFL learners as

[2] Sometimes this can happen with Slovene speakers whose native accents do not recognize the two vowels, but only one.

[3] [w] is a bilabial realization of the labio-dental approximant /ʋ/.

they tend to pronounce them by StS monophthongs /i, ɛ, u/ followed by /r/ when there is the letter <r> in the spelling, or without the alveolar approximant when there is no letter <r> in the spelling.

3.2 Consonants

The phonological consonantal systems of the two languages are presented in Table 1 below.

Table 1. StS (grey) and GB (black) consonants.

	Bilabial	Labio–dental	Dental	Alveolar	Post–alveolar	Palato–alveolar	Palatal	Velar	Glottal
Plosive	p **p** b **b**		t d	**t d**				k **k** g **g**	
Affricate				ts		tʃ **tʃ** dʒ **dʒ**			
Fricative		f **f** v	**θ ð**	s **s** z **z**		ʃ **ʃ** ʒ **ʒ**		x	**h**
Nasal	m **m**		n **n**					**ŋ**	
Tap				ɾ					
Approximant	**w**	ʋ			**r**		j **j**		
Lateral approximant				l **l**					

Systemic differences between StS and GB consonants include the following:

- lack of consonants,
- additional consonants,
- different places of articulation,
- different manners of articulation.

There are three GB consonants which are non-existent in the StS. These are the voiced and voiceless dental fricatives /ð/ and /θ/, and the labio-velar (also referred to as bilabial) approximant /w/. There is also no voiced labio-dental fricative /v/ in StS. The GB velar nasal /ŋ/ occurs in StS only as an allophone of the dental nasal /n/ when followed by a velar consonant. As a result, Slovene EFL speakers tend to pronounce the GB velar nasal followed by a velar plosive also in distributions where the velar plosive is mute.

In addition to the voiced and voiceless palato-alveolar affricates /dʒ/ and /ʧ/, there is a voiceless alveolar affricate /ts/in StS.

StS consonants with different places of articulation than their near equivalents in GB are the voiced and voiceless plosives /d/ and /t/, and the nasal /n/ which all have dental place of articulation instead of alveolar. The nearest equivalent to the GB glottal fricative /h/ is the StS velar fricative /x/.

Consonants with different manners of articulation in StS concern the voiced alveolar tap /ɾ/, and the voiced labio-dental approximant /ʋ/ which Slovene EFL speakers often mispronounce as the GB labio-velar approximant /w/, as in the word *vowel* /**wawəl*/, for example.

3.3 Allophonic differences

With respect to allophonic differences concerning consonants, StS differs from GB primarily in:

- the lack of aspiration of voiceless plosives /p, t, k/ in stressed, syllable-initial and pre-vocalic positions;
- the lack of glottal reinforcement of voiceless plosives /p, t, k/ and voiceless affricate /ʧ/;
- the lack of dark [ɬ];
- nasals and approximants are not devoiced by the preceding voiceless and fortis obstruent;
- voiced obstruents in word-final position or followed by a voiceless obstruent are voiceless and fortis;
- voiceless obstruents followed by a voiced obstruent are voiced and lenis;

4 Predicted Errors Made by Slovene EFL Speakers

On the basis of the above-described phonemic and allophonic differences between StS and GB, we can predict the following pronunciation errors.

4.1 Vowels

- GB long vowels will be pronounced too short before lenis consonants,
- GB short vowels /ɪ, ʊ/ will be replaced by StS equivalents /i, u/,
- GB TRAP vowel /æ/ will be replaced by StS front open-mid /ɛ/,
- GB STRUT vowel /ʌ/ will replaced by /ɜ:/ or /ɒ/ - the latter mainly under the influence of the letter <o> in the spelling,
- GB centring diphthongs /ɪə, eə, ʊə/ will be replaced by StS monophthongs /i, ɛ, u/ followed by /ɾ/ when there is the letter <r> in the spelling.

4.2 Consonants

- GB dental fricatives /θ, ð/ will be pronounced as StS dental plosives /t, d/,
- GB glottal fricative /h/ will be pronounced as StS velar fricative /x/,

- GB voiced labio-dental fricative /v/ will be pronounced as StS labio-dental approximant /ʋ/,
- GB velar nasal /ŋ/ will always be followed by /k/ or /g/,
- GB voiced post-alveolar approximant /r/ will be pronounced in all positions.

4.3 Allophonic features:

- Lack of aspiration,
- Lack of glottal reinforcement,
- Fortition of voiced obstruents before voiceless obstruents and in word-final positions,
- Lenition of voiceless obstruents before voiced obstruents,
- Dark [ɬ] pronounced as clear.

In addition, we expect the influence of General American English, particularly on the BATH–TRAP and THOUGHT–LOT vowel oppositions.

5 The Study

The present study investigates the influence of English pronunciation teaching on the production of sounds in connected speech by Slovene students of English as a foreign language. Its purpose is to examine the correlation between the pronunciation of a model sentence before and after a 30-hour course in English phonetics and phonology. The aim of the study is also to verify the findings of previous research in the perception of individual GB vowels (Stopar 2015, 2017, 2019), the relationship between the perception and production of GB vowels (Komar 2017), and the interference of L1 phonological and phonetic features on L2 pronunciation (Šuštaršič 2005).

In addition, the study was designed to test two hypotheses:

H1: The interference of L1 phonemic and phonetic system, orthography and General American English will be present in the reading of the sentence from orthography.

H2: Systematic training of phonemic transcription has a positive influence on the production of sounds in connected speech.

5.1 Method and Participants

The study consisted of two production tests. In the first test, which took place at the beginning of the course in English phonetics and phonology (October 2018), the participants were given one orthographically transcribed English sentence. They were required to read and record it. The second test took place at the end of the course (January 2019). The course consisted of 30 hours of lectures and 30 hours of practical classes during which the students were trained in perception and production

of General British phonemes, as well as in phonemic and phonetic transcription of individual words and sentences. They also became familiar with the basic rules of tonality, i.e. the division of utterances into intonation phrases. In the second test, the participants were given the same sentence which they had to transcribe phonemically, divide it into intonation phrases, read it and record it.

The sentence, which the students were required to read and record, contained all GB vowels and consonants (Collins, Šuštaršič, and Komar 2014, 142).

The sentence:

> *When the lighthouse keeper's lovely young daughter Thelma makes crab and lobster rolls, tourists come from far and near to enjoy fresh air, good food and searching for treasure on the beach.*

The phonemic transcription:

> */ ˈwen ðə ˈlaɪthaʊs kiːpəz ˈlʌvlɪ ˈjʌŋ ˈdɔːtə ˈθelmə | ˈmeɪks ˈkræb ən ˈlɒbstə ˈrəʊlz | ˈtʊərɪsts[4] ˈkʌm frəm ˈfɑːr ən ˈnɪə | tʊ ɪnˈʤɔɪ ˈfreʃ ˈeə[5] | ˈgʊd ˈfuːd | ən ˈsɜːʧɪŋ fə ˈtreʒər ɒn ðə ˈbiːʧ /*

All the participants in the study were first-year BA students of English at the Faculty of Arts, University of Ljubljana. In the first test, 120 students took part, whereas in the second test only 50 students participated. Participation in the study was on a voluntary basis. In order to monitor individual changes in pronunciation and to get more reliable results which would indicate possible improvement, we decided to include in the final analysis only those 50 students who took part in both tests.

For the evaluation of results, the auditory method was used. The evaluator had vast and long experience in teaching British English phonetics and phonology to Slovene speakers of English and carried out research in contrastive English-Slovene phonetics, phonology and prosody.

6 Results

For the purpose of comparison, we are going to present the results of both readings of the sentence simultaneously. The results are grouped according to the predicted errors.

Figure 1 shows the percentage of correct pronunciation of the sentence in the first and second reading, respectively.

---

[4] The alternative correct pronunciation of <tourists> is also /ˈtɔːrɪsts/.

[5] Although Cruttenden (2014) no longer recognizes the centring diphthong /eə/ and proposes a long, front, open-mid monophthong /ɛː/ instead, the Slovene students of English are still trained to perceive and produce the diphthong.

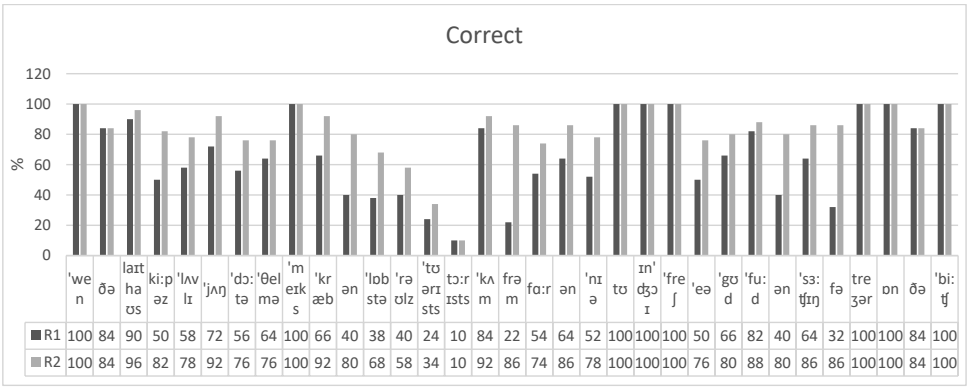


Figure 1. Overall pronunciation of the sentence – comparison of the first and second reading.

There were eight words which were correctly pronounced by all participants in both readings of the sentence. Five of them were lexical items (*makes, enjoy, fresh, treasure, beach*), whereas the remaining three were function words (*when, to, on*). In the majority of words, the participants made more than one pronunciation error.

Figure 2 shows the mean values of correct pronunciation of the whole sentence in the first and second reading. In the first reading, 66.24% of participants read the sentence correctly, while in the second reading the percentage was 82.48%. This represents an increase of 16.24%.

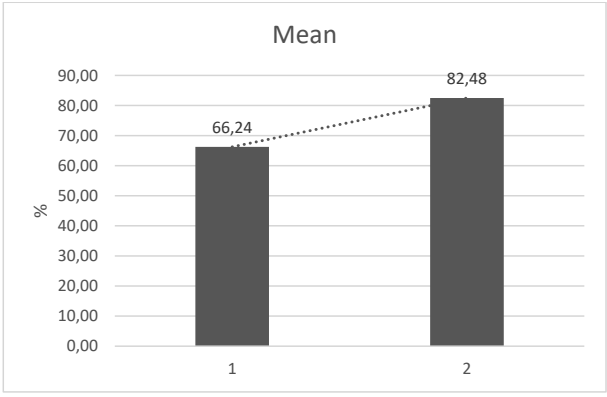


Figure 2. Mean values of correct pronunciation of the whole sentence in the first and second reading.

The pronunciation improvement of individual items in the second reading (see Figure 1 above) was best in the usage of weak forms of function words instead of their strong equivalents. The percentage of correct pronunciation in the second reading increased by 64% in *from*, 54% in *for,* and 22% in *and*. Among lexical items, the pronunciation most significantly improved in words *keepers* (by 32%), *lobster* (by 30%), *crab* and *air* (by 26%).

6.1 Errors in Vowels

The vowel quality of the word *good* and the vowel length in the word *food* were expected to be problematic for the participants. Table 2 shows that in the first reading 70% of participants pronounced the GB FOOT-vowel /ʊ/ correctly and only 30% replaced it with the StS vowel /u/. The percentage of correct pronunciation in the second reading increased to 82%.

The correct length of the GB GOOSE-vowel /u:/ in the first reading was achieved by 92% of the participants, and improved by 4% in the second reading.

TABLE 2. Vowel quality and length.

Reading 1			**Reading 2**	
Word	Correct %	/u/ %	Correct %	/u/ %
‘gʊd	70	30	82	18
‘fu:d	92	8	96	4
Mean	**81**		**89**	

Replacement of the GB TRAP /æ/ vowel by the StS vowel /ɛ/ was expected to occur in the word *crab.* Table 3 shows that in the first reading the percentage of correct pronunciation of the vowel in question was 66 %. In the second reading, the correct pronunciation increased by 26% to 92%.

TABLE 3. Replacement of the GB TRAP vowel with the Slovene vowel /ɛ/.

Reading 1			**Reading 2**	
Word	Correct %	/ɛ/ %	Correct %	/ɛ/ %
‘kræb	66	34	92	8

Mispronunciations of the GB STRUT /ʌ/ vowel were expected to occur in words *lovely*, *young* and *come.* Table 4 presents the percentage of correct pronunciations, as well as two mispronunciations of the vowel in question. The mean value of correct pronunciation of the GB STRUT-vowel was 78% in the first reading of the sentence, and 88.7% in the second reading. The worst result in the pronunciation of the vowel in question was in the word *lovely* where only 58% of participants pronounced the vowel correctly in the first reading. The most frequent mispronunciation of the vowel in this word was with the GB NURSE-vowel /ɜ:/ (40%), while only 1 student (2%) pronounced the word with the GB LOT-vowel. The correct pronunciation in the first reading of the GB STRUT-vowel in words y*oung* and *come* was 92% and 84%, respectively. The mean values show that 18.7% of participants mispronounced the GB STRUT-vowel with the GB NURSE-vowel /ɜ:/, and 3.3% with the GB LOT-vowel.

In the second reading, the mean value of correct pronunciation of the GB STRUT-vowel increased by 10.7% to 88.7%. The mean value of mispronunciation of the vowel by the GB STRUT vowel was 10%, and 1.3% by the GB LOT-vowel.

Table 4. Replacement of the GB STRUT vowel by /ɜ:/ or /ɒ/.

Reading 1				**Reading 2**		
Word	Correct %	/ɜ:/ %	/ɒ/ %	Correct %	/ɜ:/ %	/ɒ/ %
ˈlʌvlɪ	58	40	2	78	22	0
ˈjʌŋ	92	8	0	96	4	0
ˈkʌm	84	8	8	92	4	4
Mean	**78**	**18.7**	**3.3**	**88.7**	**10**	**1.3**

Mispronunciation of the GB centring diphthongs /ʊə/, /ɪə/, and /eə/ by StS monophthongs /u/, /i/, and /ɛ/ was expected to occur in words *tourists, near* and *air*. As presented in Table 5, the average correct pronunciation of the three centring diphthongs is 50.7% in the first reading. The weakest performance was the word *tourists* where 66 % of the participants replaced the diphthong /ʊə/ or the alternative long monophthong /ɔ:/ with the StS monophthong /u/. Mispronunciation of the word *air* with the StS monophthong /ɛ/ came second with only 50 % of correct pronunciations[6]. The word *near* came third with 68% of correct pronunciation, while 32% of the participants pronounced the diphthong in question as the combination of the StS monophthong /i/ and /r/.

In the second reading, the average correct pronunciation increased to 68%. The pronunciation of the word *tourists* remains to be most problematic since only 10% of the participants improved it whereas the percentage of correct pronunciation of the other two diphthongs in the words *near* and *air* raised by 14% and 28%, respectively.

Table 5. GB Centring diphthongs pronounced by StS monophthongs followed by /r/.

Reading 1						**Reading 2**				
Word	Correct %	/ur/ %	/ir/ %	/ɛr/ %	/ɛ:/ %	Correct %	/ur/ %	/ir/ %	/ɛr/ %	/ɛ:/ %
ˈtʊərɪsts ˈtɔ:rɪsts	34	66				44	56			
ˈnɪə	68		32			82		18		
ˈeə	50			40	10	78			20	2
Mean	**50.7**					**68**				

The results presented in Table 6 show that the pronunciation of the diphthong /əʊ/ in the word *rolls* was quite problematic. In the first reading, 68% of the participants pronounced the word correctly, whereas 14% mispronounced the diphthong either as the long GB monophthong /ɔ:/ or a diphthong with a rounded first element

[6] If we took into consideration the trend in the pronunciation of the GB diphthong /eə/ as a long front, open-mid monophthong /ɛ:/, then the percentage of correct pronunciations would increase to 60%.

instead of the central fist element. Pronunciation with the GB LOT-vowel occurred in 4% of the participants.

In the second reading, the correct pronunciation increased by 14% to 82%. Mispronunciations with either the long GB monophthong /ɔ:/ or a diphthong with a rounded first element instead of the central fist element prevail.

Table 6. The pronunciation of the GB diphthong /əʊ/.

Reading 1					**Reading 2**			
Word	Correct %	/ɔ:/ %	/ɒ/ %	/ou/ %	Correct %	/ɔ:/ %	/ɒ/ %	/ou/ %
‘rəʊlz	68	14	4	14	82	8	2	8

6.2 Errors in Consonants

The GB voiced /ð/ and voiceless /θ/ dental fricatives occurred in words *the* and *Thelma*. Table 7 shows that in the first reading 84% of participants correctly pronounced both definite articles, whereas the correct pronunciation of the voiceless dental fricative in the word *Thelma* was achieved by only 64% of participants. This means that 16% of the participants replaced the voiced dental fricative in the definite article by the StS voiced dental plosive /d/, while 36% of the participants pronounced the StS voiceless dental fricative in the word *Thelma* as the voiceless dental plosive /t/.

In the second reading, the percentage of correct pronunciation of the GB voiced dental fricative /ð/ in the definite article *the* remained unchanged. The correct pronunciation of the GB voiceless dental fricative /θ/ in the word *Thelma* increased by 12%, while the remaining 24% of the participants retained the StS voiceless dental plosive.

Table 7. Pronunciation of the GB dental fricatives.

Reading 1				**Reading 2**		
Word	Correct %	/t/ %	/d/ %	Correct %	/t/ %	/d/ %
ðə	84		16	84		16
‘θelmə	64	36		76	24	
Mean	**74**			**80**		

Table 8 shows that in the first reading the pronunciation of the GB voiceless glottal fricative /h/ in the word *lighthouse* was correct in 90% of participants. Only 10% replaced the sound with the StS voiceless velar fricative /x/.

In the second reading, the percentage of correct pronunciation increased by 6%, with only 4% of participants pronouncing the GB voiceless glottal fricative as the StS voiceless velar fricative.

Table 8. Pronunciation of the GB voiceless glottal fricative /h/.

Reading 1			**Reading 2**	
Word	Correct %	/x/ %	Correct %	/x/ %
ˈlaɪthaʊs	90	10	96	4

As shown in Table 9, the pronunciation of the GB voiced velar nasal /ŋ/ in words *young* and *searching* was in the first reading characterised by adding either the voiced velar plosive /g/ in the word *young* or the voiceless velar plosive /k/ in the word *searching*. In both words this mispronunciation occurred in 26% of participants. The remaining 74% of participants pronounced the words correctly.

In the second reading, the mean value of correct pronunciation increased by 20%. The pronunciation of the GB voiced velar nasal followed by the voiced velar plosive /g/ in the word *young* remained with 8% of participants, whereas the pronunciation of the GB voiced velar nasal followed by the voiceless velar plosive /k/ in the word *searching* was noticed with 4% or participants.

Table 9. Pronunciation of the voiced velar nasal /ŋ./

Reading 1				**Reading 2**		
Word	Correct %	/ŋg/ %	/ŋk/ %	Correct %	/ŋg/ %	/ŋk/ %
ˈjʌŋ	74	26		92	8	
ˈsɜːʧɪŋ	74		26	96		4
Mean	**74**			**94**		

Pronunciation of the GB voiced post-alveolar approximant /r/ resulted in two types of errors: pronunciation of the sound in positions before a consonant or silence, and the lack of its pronunciations in the position before a vowel, the so-called linking /r/. Tables 10 and 11 show the results of both mispronunciations of the sound in question.

The results of the first reading indicate that on average 71.25% of participants pronounced the GB voiced post-alveolar approximant /r/ in all distributions correctly. Its pronunciation in non-pre-vocalic positions was most frequent in words *keepers* (44%), *near* (48%), and *air* (40%), whereas in *daughter* (26%), *searching* (10%) and *for* (16%) the pronunciation of /r/ was less frequent.

The results of the second reading show that on average 86% of participants pronounced the GB voiced post-alveolar approximant /r/ in all distributions correctly. The pronunciation of the consonant in pre-vocalic positions in the word *keepers* decreased by 26%, in the word *near* by 30%, and in the word *air* by 20%. Results also indicate improved pronunciation in words *daughter* by 16%, *searching* by 2%, and *for* by 4%.

There were two instances of linking /r/ in the sentence: *far and* and *treasure on*. As presented in Table 11, results of the first reading show that the omission of linking /r/ in *far and* occurred with 46% of participants, while the pronunciation of the linking /r/ in *treasure on* occurred in all participants.

In the second reading, the pronunciation of the linking /r/ in *far and* increased by 20%, while its pronunciation in *treasure on* remained unchanged.

Table 10. Pronuniation of the GB voiced post-alveolar approximant /r/.

Reading 1			**Reading 2**	
Word	Correct %	/r/+C/0 %	Correct %	/r/+C/0 %
‘ki:pəz	56	44	82	18
‘dɔ:tə	74	26	90	10
‘nɪə	52	48	82	18
‘eə	60	40	80	20
‘sɜ:ʧɪŋ	90	10	92	8
fə	84	16	88	12
Mean	**69.3**		**85.7**	

Table 11. Pronunciation of linking /r/.

Reading 1			**Reading 2**	
Word	Correct %	No linking /r/	Correct %	No linking /r/
‘fɑ:r ən	54	46	74	26
‘treʒər ɒn	100		100	
Mean	**77**		**87**	

6.3 Errors in Allophonic Features

Among the predicted errors in allophonic features, only two were detected. The first one is lack of aspiration of voiceless plosives, the second one is the fortition of the voiced obstruents when they occur before voiceless and fortis obstruents or when they are in word-final positions. Interestingly, the dark allophone of /l/ in words *Thelma* and *rolls* was correctly pronounced by all participants. The lack of glottal reinforcement of voiceless plosives /t/ and /k/ in words *lighthouse, makes*, as well as the voiceless affricate /ʧ/ in the word *beach*, was not detected either.

Lack of aspiration was noticeable in the word *tourists*. The results in Table 12 show that in the first reading more than half (54%) of the participants pronounced the word-initial voiceless alveolar plosive /t/ without aspiration. In the second reading, only 10% of participants improved their pronunciation, while 44% of them continue to pronounce the initial voiceless alveolar plosive /t/ in the word *tourist* without aspiration.

Table 12. Lack of aspiration.

Reading 1			**Reading 2**	
Word	Correct %	No aspiration %	Correct %	No aspiration %
‘tʰʊərɪsts ‘tʰɔ:rɪsts	46	54	56	44

As mentioned in 3.3, fortition of voiced obstruents when they occur in a word-final position or before another voiceless obstruent is a very frequent pronunciation error made by Slovene speakers of English. There are five words in the sentence, which end in a voiced obstruent (*keepers, crab, rolls, good, food*) and one word (*lobster*) in which the first syllable ends in a voiced plosive, while the second syllable begins with a voiceless fricative. Results in Table 13 show that in the first reading the mean value of correct pronunciation was 65.2%. Least problematic were the words *good* and *food* where the percentage of correct pronunciation was 84%, whereas the word *lobster,* was mispronounced by 62% of participants.

In the second reading, the mean value of correct pronunciations increased by 18.4%. The pronunciation of the word *lobster* improved by 30%. Significant improvement occured with the words *keepers* (32%)*, crab* (26%) and, *rolls* (18%).

Table 13. Fortition of the voiced obstruents.

Reading 1						**Reading 2**				
Word	Correct %	s %	ps %	p %	t %	Correct %	s %	ps %	p %	t %
‘ki:pəz	50	50				82	18			
‘kræb	66			34		92			8	
‘lɒbstə	38		62			68		32		
‘rəʊlz	42	58				60	40			
‘gʊd	84				16	92				8
‘fu:d	84				16	92				8
Mean	**65.2**	**54**			**16**	**83.6**	**29**			**8**

6.4 Errors in Features of Connected Speech

There were two types of errors related to the connected speech: the absence of linking /r/ and the use of strong forms of grammatical items. The usage of the linking /r/ is presented in Table 11, whereas Table 14 shows the results of the usage of weak as opposed to strong forms.

In the first reading, the mean value of correct pronunciation of function words with their weak forms was only 39.4%. In the second reading, the percentage of correct pronunciation increased by 18.4%.

The conjunction *and* appears three times in the sentence and the percentage of its weak pronunciation varied according to its position in the sentence. The percentage of weak pronunciation was better in the closely bound phrase *far and near* (64% in the first reading) than in the other two distributions: *crab and lobster* and *good food and searching* (both 40% in the first reading). In the second reading, the pronunciation with the weak vowel /ə/ increased to 86 % in case of *far and near*, and to 80 % in the other two distributions of the conjunction *and*.

In the first reading, the prepositions *from* and *for* were correctly pronounced by 22% and 32%, respectively. In the second reading, their correct pronunciation increased by 64% in case of *from*, and 54% in case of *for*.

Table 14. Usage of weak forms.

Reading 1								**Reading 2**						
Word	Correct %	/æ/ %	/ɛnd/ %	/ɛnt/ %	/ɒ/ %	/ɔ:/ %	/ɔ:r/ %	Correct %	/æ/ %	/ɛnd/ %	/ɛnt/ %	/ɒ/ %	/ɔ:/ %	/ɔ:r/ %
and / ənd/, / ən/	40	0	40	20				80	0	14	6			
from / frəm/	22				78			86				14		
far **and** /ən/ near	64	0	36					86	0	14				
and / ənd/, / ən/	40	0	40	20				80	0	14	6			
for /fə/	32					52	16	86					12	2
Mean	**39.6**							**83.6**						

6.5 Influence of American English

The pronunciation of the word *daughter* reflected the influence of American English in three phonemes:

- replacement of the GB NORTH-vowel /ɔ:/ by the GB LOT-vowel /ɒ/;
- replacement of the GB voiceless alveolar plosive /t/ by the voiced alveolar tap /ɾ/;
- the pronunciation of the voiced post-alveolar approximant /r/ in the word-final position and before the initial voiceless dental fricative /θ/ of the following word *Thelma*.

As presented in Table 15, only 56% of participants pronounced the word correctly in the first reading, while 34% of them mispronounced the vowel, and 4% used

the voiced alveolar tap /ɽ/ instead of the voiceless alveolar plosive /t/. More than a quarter (26%) of participants pronounced the final voiced post-alveolar approximant although the next word began on a consonant.

In the second reading, the percentage of correct pronunciation increased by 20%. Mispronunciation of the first vowel remained with 16% of participants, whereas the pronunciation of the voiced alveolar tap occurred with only 2% of participants. The pronunciation of the final voiced post-alveolar approximant in the pre-consonantal position was retained by 10% of participants.

Table 15. Influence of American English.

Reading 1					**Reading 2**			
Word	Correct %	/ɒ/ %	/ɽ/ %	/r/ %	Correct %	/ɒ/ %	/ɽ/ %	/r/ %
ˈdɔːtə	56	34	4	26	76	16	2	10

7 Discussion

When interpreting the results, we have to take into consideration that the sample of the participants in both readings was only 50. It is reasonable to assume that the participants who decided to take part in the second test were better than their colleagues who failed to record the sentence at the end of the course. Thus the relatively high mean value of the correct pronunciation of the sentence in the first reading (66.34%) can be explained by good pronunciation of the selected participants before they started the course on English phonetics and phonology. The increase of correct pronunciation to 82.48% in the second reading can be explained by high motivation of the same participants to improve their pronunciation even more and to pass the exam. Nonetheless, the results of the first reading, when the participants read the sentence from the spelling, confirmed nearly all predicted mispronunciations arising from the phonetic and phonological differences between General British English and standard Slovene, as well as the influence of General American (see Table 15). In addition, the influence of the spelling on the pronunciation was particularly noticeable in the word *rolls* (68%), and to a much lower percentage in the words *lovely* (2%) and *come* (8%) in which the correct GB vowel were mispronounced by either /ɒ/ or **/ɔː/.** This additionally confirms the findings of previous studies (Komar 2017; Šuštaršič 2005) which suggested that the orthography had a misleading effect on the pronunciation. Results of the second reading indicate a significant improvement particularly in case of the word *rolls* when 82% of participants pronounced the correct diphthong. We believe that this improvement, as well as all other improvements of pronunciation in the second reading mentioned below, are the result of systematic training of perception and production of GB sounds, as well as phonemic transcription, and as such, confirm our second hypothesis.

There was only one instance of a long GB monophthong followed by a voiced consonant (*food*) which, as predicted, would be mispronounced by the participants as too short. However, this did not happen. The vowel was in the first reading correctly pronounced by 92% of participants. On the other hand, the GB FOOT-vowel, which occurred in the word *good*, was pronounced with the StS /u/ vowel by 30% of participants. In the second reading the correct pronunciation increased by 18% to 82%.

In line with the studies by Stopar (2015, 2019) and Komar (2017), which found that the GB TRAP vowel is the most difficult vowel for Slovene speakers to perceive and produce correctly, the results of the present study confirm these findings. The percentage of correct pronunciation of the word *crab* in the first reading was at 66%. Systematic training of pronunciation and phonemic transcription, as well as highly predictable spelling of the vowel by the letter <a>, resulted in an increase of correct pronunciation of the word in the second reading to 92%.

It was predicted that GB centring diphthongs /ɪə/, /eə/ and /ʊə/ would be pronounced by StS monophthongs /i/, /ɛ/ and /u/ followed by /r/. In the three words which had the three GB centring diphthongs (*tourists, near, air*), the mean value of correct pronunciation was 50.7% in the first reading, and raised to 68% in the second reading. The average improvement in the second reading was thus only 17.3%. Although there are no studies available which would look into the relationship between perception and production of these three diphthongs by Slovene speakers of English, we suggest two main reasons for these mispronunciations: first, the lack of the centring diphthongs or their near equivalents in L1, and second, the influence of the spelling, particularly the letter <r>, which was present in all three words.

Although the spelling of the two GB dental fricatives by the letters <th> is highly predictable, 36% of the participants pronounced the GB voiceless dental fricative /θ/ in the word *Thelma* as the StS voiceless dental plosive /t/ in the first reading. With a lot of pronunciation training, the pronunciation in the second reading improved by 12%. The GB voiced dental fricative /ð/ occurred only in the definite article *the* which was preceded by words ending on the voiced alveolar nasal /n/ (*when the, on the*). In Slovene the nasal has dental place of articulation rather than alveolar which resulted in the pronunciation of the GB voiced dental fricative /ð/ as the voiced dental plosive /d/ by 84% of participants in both readings.

Theoretical instructions, systematic training of perception and production, as well as phonemic transcription had a positive influence on the correction of the pronunciation of the GB voiced velar nasal /ŋ/ when it is not followed by the GB voiced velar plosive /g/. In the first reading, 74% of participants pronounced the GB velar nasal in the word *young* together with the voiced velar plosive /g/, and in the

word *searching* together with the voiceless velar plosive /k/. In the second reading, the mean value of the correct pronunciation reached 94%.

Another pronunciation feature, which was at the beginning of the course in English phonetics and phonology present on average in 69.3% of participants, was the pronunciation of the GB voiced post-alveolar approximant /r/ in word final or pre-consonantal positions. Reasons for that are to be found both in the spelling, as well as in the influence of General American pronunciation. Systematic training of production and phonemic transcription had a positive influence on the second reading when 85.7% of participants pronounced the words *keepers, daughter, lobster, near, air, searching,* and *for* without the post-vocalic /r/.

The two allophonic variations of GB obstruents, which are for Slovene speakers of English very difficult to master, are aspiration and devoicing of voiced obstruents in word-final position or before another voiceless consonant. Aspiration, which is not a feature of the standard Slovene language, proved to be particularly challenging. In the period between the two tests only 10% of participants acquired this important feature, while the percentage of correctly pronounced voiceless alveolar plosive in the word *tourists* reached 56% (see Table 11). The pronunciation of devoiced and lenis obstruents in the word final position or before another voiceless consonant on average improved by 18.4%, reaching 83.6% in the second reading.

There is no doubt that systematic training of phonemic transcription resulted in a significant increase (44% on average) in the usage of the weak forms of function words.

Another manifestation of the benefit of systematic pronunciation teaching is the awareness of the difference between standard British English and General American. The influence of the latter was most noticeable in the word *daughter* when only 56% of participants pronounced the word using standard British English pronunciation in the first reading. In the second reading, the percentage raised to 76% which means a 20% increase (see Table 15).

8 Conclusion

The main purpose of our study was to test the amount of improvement and development of General British pronunciation of Slovene students of English after a 60-hour intensive theoretical and practical course of English phonetics and phonology. With that in mind, the first-year students of the BA level were asked to read and record an English sentence, which contained all GB sounds, before and after the course. In the end, we selected 50 same students who took part in both tests.

Although the results of the first reading showed an overall high competence in GB pronunciation, they nonetheless confirmed our first hypothesis that L1 interference,

orthography and General American pronunciation will be more numerous in the first reading of the sentence. Pronunciation of individual sounds, which previous studies mentioned as most problematic, turned out to be problematic also in the first reading of the sentence. Overall results of the second reading improved which confirms our second hypothesis that systematic pronunciation training which includes theoretical knowledge, phonemic and phonetic transcription, as well as perception and production of individual sounds, has positive effects on the acquisition of foreign sounds. Our study also showed that some allophonic features (e.g. aspiration) are more difficult to acquire than others (e.g. devoicing and lenition of final voiced obstruents). The benefit of practising phonemic transcription and reading from it is certainly visible in the fact that the best improvement was achieved in the acquisition of weak forms of function words.

Although both our hypotheses were confirmed, two reservations have to be mentioned. First, the students whose readings were analysed probably do not represent the average competence of all our first-year students. It is safe to claim that our sample of participants consisted of highly motivated students whose initial English pronunciation competence was above average. Their results after the second reading indicate that they were also highly motivated to improve their English pronunciation.

Our second reservation concerns the sentence which the participants had to read. Although it contained all GB sounds, it did not contain certain problematic features for Slovene EFL learners, such as non-reduced long GB vowels. This is why we suggest that future research in the production of GB sounds should focus on the acquisition of foreign allophonic features in more defined contexts, ranging from individual syllables to phrases and sentences. In addition, future research of pronunciation should also move away from reading to spontaneous speech.

References

Baker, Wendy, Pavel Trofimovich. 2006. "Perceptual paths to accurate production of L2 vowels: the role of individual differences." *International Review of Applied Linguistics in Language Teaching* 44 (3): 231–50. http://dx.doi.org/10.1515/IRAL.2006.010.

Bassetti, Benedetta. 2008. "Orthographic input and second language phonology." In *Input Matters in SLA*, edited by Thorsten Piske and Martha Young-Scholten, 191–206. Clevedon, UK: Multilingual Matters.

Best, Catherine T. 1995. "A Direct Realist View of Cross-Language Speech Perception." In *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues,* edited by Winifred Strange, 171–203. Baltimore: York Press.

Bion, Ricardo A. H., Paola Escudero, Andréia S. Rauber, and Barbara O. Baptista. 2006. "Category Formation and the Role of Spectral Quality in the Perception and Production of English Front Vowels." *Proceedings of Interspeech 2006*, 1363–6. Pittsburgh: Interspeech.

Collins, Beverly, Rastislav Šuštaršič, and Smiljana Komar. 2014. *Present-Day English Pronunciation: A Guide for Slovene Students. Revised and Enlarged 2nd Edition.* Ljubljana: Znanstvena založba Filozofske fakultete, Oddelek za anglistiko in amerikanistiko.

Cruttenden, Alan. 2014. *Gimson's Pronunciation of English.* 8th Edition. Abingdon: Routledge.
Escudero, Paola. 2002. "The Perception of English Vowel Contrasts: Acoustic Cue Reliance in the Development of New Contrasts." In *New Sounds 2000: Proceedings of the Fourth International Symposium on the Acquisition of Second-Language Speech, Klagenfurt*, edited by Allan James and Jonathan Leather, 122–31. Klagenfurt: University of Klagenfurt.
Flege, James E. 1995. "Second Language Speech Learning: Theory, Findings and Problems." In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by Winifred Strange, 233–277. Timonium, MD: York Press.
Komar, Smiljana. 2017. "The Relationship between the Perception and Production of Four General British Vowels by Slovene University Students of English." *Linguistica* 57 (1): 161–70. https://doi.org/10.4312/linguistica.57.1.161-170.
Lipinska, Dorota. 2013. "Influence of Formal Instruction in English Phonetics and Phonology on Polish Learners' Production of English /TRAP/." In *Proceedings of the Phonetics Teaching and Learning Conference, UCL, London, 8–10 August 2013*, 59-62. London: Phonetics Teaching and Learning Conference.
Lintunen, Pekka. 2005. "Phonemic Transcription and Its Effect on Learning." In *Proceedings of the Phonetics Teaching and Learning Conference 2005,* edited by John A. Maidment, 1–4. London: University College London.
Lord, Gillian. 2005. "(How) Can We Teach Foreign Language Production? On the Effects of a Spanish Phonetics Course." *Hispania* 88 (3): 557–67. http://www.jstor.org/ stable/20063159.
Matthews, John, and Cynthia Brown. 2004. "When Intake Exceeds Input: Language Specific Perceptual Illusions Induced by L1 Prosodic Constraints." *International Journal of Bilingualism* 8 (1): 5–27. http://dx.doi.org/10.1177/1367006904 0080010201
Stopar, Andrej. 2015. "Perception of Four General British Vowels by Slovenian University Students of English as a Foreign Language." In *PTLC2015: Proceedings of the Phonetics Teaching and Learning Conference, London, 5–7 August 2015*, 87–90. London: Phonetics Teaching and Learning Conference.
—. 2017. "Perception of the GB Vowels [...] by Slovenian Speakers of English. In *PTLC2017: Proceedings of the Phonetics Teaching and Learning Conference UCL, London, 9–11 August 2017,* 104–9. London: Phonetics Teaching and Learning Conference.
—. 2019 [forthcoming]. "Perception of General British Monophthongs: The Case of Slovene Students of English as a Foreign Language." *Jezikoslovlje*.
Šuštaršič, Rastislav. 2005. *English–Slovene Contrastive Phonetic and Phonemic Analysis and Its Application in Teaching English Phonetics and Phonology*. Ljubljana: Znanstveni inštitut Filozofske fakultete.
Šuštaršič, Rastislav, Smiljana Komar, and Bojan Petek. 1999. "Slovene." In *Handbook of the International Phonetic Association, A Guide to the Use of the International Phonetic Alphabet*, 135–139. Cambridge: Cambridge University Press.
Wells, John C. 1996. "Why Phonetic Transcription Is Important." *Malsori (Phonetics), the Journal of the Phonetic Society of Korea* 31–32, 239–42.
—. 1982. *Accents of English.* Cambridge: Cambridge University Press.

Emilija Mustapić,
Frane Malenica
University of Zadar, Croatia

2019, Vol. 16 (1), 123–144(164)
revije.ff.uni-lj.si/elope
doi: 10.4312/elope.15.2.123–144
UDC: 81'221.24

The Signs of Silence – An Overview of Systems of Sign Languages and Co-Speech Gestures

ABSTRACT

The paper presents an overview of sign languages and co-speech gestures as two means of communication realised through the visuo-spatial modality. We look at previous research to examine the correlation between spoken and sign language phonology, but also provide an insight into the basic features of co-speech gestures. By analysing these features, we are able to see how these means of communication utilise phases of production (in the case of gestures) or parts of individual signs (in the case of sign languages) to convey or complement the meaning. Recent insights into sign languages as *bona fide* linguistic systems and co-speech gestures as a system which has no linguistic features but accompanies spoken language have shown that communication does not take place within just a single modality but is rather multimodal. By comparing gestures and sign languages to spoken languages, we are able to trace the transition from systems of communication involving simple form-meaning pairings to fully fledged morphological and syntactic complexities in spoken and sign languages, which gives us a new outlook on the emergence of linguistic phenomena.

Keywords: visuo-spatial modality; co-speech gestures; sign languages; spoken languages

Zvoki tišine – pregled sistemov znakovnih jezikov in obgovornih kretenj

POVZETEK

Prispevek predstavi pregled znakovnih jezikov in obgovornih kretenj kot sredstev komuniciranja, ki se udejanijo v vizualno-prostorski modalnosti. S pregledom nekaterih dosedanjih raziskav proučuje korelacijo med fonologijo govorjenih in znakovnih jezikov in ponudi vpogled v osnovne značilnosti obgovornih kretenj. Z analizo slednjih ugotavljamo, kako ta sredstva komunikacije uporabljajo faze produkcije (v primeru kretenj) ali dele posameznih znakov (v primeru znakovnih jezikov) pri izražanju ali dopolnjevanju pomena. Novejše raziskave znakovnih jezikov kot *bona fide* jezikovnih sistemov in obgovornih kretenj kot sistemov brez značilnosti jezika, ki govorjeni jezik le spremljajo, so pokazale, da komunikacija ne poteka le znotraj posamezne modalnosti, pač pa je multimodalna. S primerjavo kretenj, znakovnih jezikov in govorjenih jezikov lahko opazujemo prehod med sistemi komunikacije, ki vključujejo preproste pare oblika-pomen, in sistemi s polno razvitimi morfološkimi in skladenjskimi spleti tako v govorjenih kot v znakovnih jezikih, kar nam razkrije nov pogled na vznik jezikovnih pojavov.

Ključne besede: vizualno-prostorska modalnost; obgovorne kretnje; znakovni jeziki; govorjeni jeziki

1 Introduction

The role of language in society is one of the fundamental questions discussed by numerous linguists. In his *Course in General Linguistics*, one of the groundbreaking linguistic works of the 20th century, Ferdinand de Saussure describes language (*langue*) as a tool which enables humans to understand one another and speech (*parole*) as the individual use of language in communication. He further defines language as a system of signs acquired within the confines of society, on a par with the alphabet, Braille alphabet, military insignia, symbolic rituals, etc. A linguistic sign is a psychological entity embodied by the concept (the meaning of the element) and the vocal image (the psychological imprint of the sound, not the material sound itself). For a long time, one of the defining traits of language was the pairing of vocal images with mental concepts. However, William Dwight Whitney, one of de Saussure's predecessors, claimed that language is a social institution and that the choice of the vocal apparatus is a pure coincidence imposed by nature as mankind could have just as well chosen movement and visual imagery instead of vocal forms (De Saussure 1916 [2000]).

The existence and usage of sign language in everyday communication seems to confirm Whitney's viewpoint. Sign language primarily uses body movements as a linguistic means, unlike spoken language which generally transfers its message via the speech apparatus. Hearing is often emphasised as one of the most important senses, and if it becomes damaged other problems may arise and other forms of communication (such as manual communication) will obtain the primary role. However, just like hearing is not based solely on the acoustic sign, neither does sign language rely only on body movement, but also on the fact that people understand the content which is not spoken, i.e. realised by language (Jelaska 2004). The difference between spoken and sign languages thus might seem trivial when taking into consideration their common trait – transmitting a message and realising the chain of communication from its source to its goal. Perhaps even more importantly, spoken and sign languages seem to share quite few structural features – their utterances can be segmented into smaller individual signs (i.e. lexemes in spoken languages), which themselves can be further segmented into discrete units that bear no meaning on their own (i.e. phonemes in spoken languages). However, it should not be assumed that the distinction between spoken and sign languages represents a dichotomy with no middle ground in between. Co-speech gestures, which are the second main focus of our paper, can be viewed as an intermediary step between these two modes of communication – they share the same means of articulation with the sign languages but they co-occur with speech itself. As the discussion in the upcoming sections will show, both sign and spoken languages use gestures when conveying a message.[1]

[1] For more about use of gestures in sign languages, see Liddell and Metzger (1998), Liddell (2003), Goldin-Meadow and Brentari (2017).

In this paper, we are going to limit our description of gestures to spoken language only and focus primarily on co-speech gestures, which are accompanied by spoken utterances but do not constitute a coherent linguistic system. Still, it has been shown that some forms of manual communication developed within small communities can sometimes develop particular language-like properties. We believe these incremental developments of linguistic modules, which we describe in more detail in the following sections, might tell us something more about the nature of language itself.

Given that the traditional phonological approaches aim to describe the sounds of a particular language, the analysis of the two modes of communication presented in this paper stands in stark contrast with this tradition. By looking at the individual segments of these systems of communication and the meaning they convey, linguistic research dealing with sign languages and gestures reveals new evidence about the emergence of complexity in human systems of communication. Our goal in this paper is to provide a brief review of these lines of research focused on the two systems of signs (in the Saussurean sense) that involve no production of speech at all.[2] This goal is achieved through the following steps – in section 2, we provide an outline of sign languages and co-speech gestures as two modes of manual communication, their mutual differences and similarities. In section 3, we give an overview of attempts at developing a formal linguistic analysis of sign languages and provide an example of such an analysis in the form of Brentari's (1998; 2002) prosodic model. In section 4, we introduce the basic features of gestures and their interaction with spoken languages. In section 5, we take a more general look at sign languages and co-speech gestures and their correlation with spoken languages from the perspective of development of linguistic modules, and finally, we provide our concluding remarks in section 6.

2 Forms of Manual Communication

Both co-speech gestures and sign languages represent forms of manual communication, which is why the two are often conflated as a single, visual modality. Özyürek and Woll (in press) provide a detailed description of the visual modality by comparing sign languages and co-speech gestures in spoken languages. They claim that despite the close relationship that has been proven between gestures and language, gestural studies have been excluded from the majority of grammatical theories and descriptions since most of these theories take spoken and/or written language elements (such as words, phrases, or sentences) as their focal point. While there exist numerous aspects that distinguish them from spoken languages, gestures and sign languages are also mutually quite different and, in fact, represent the two end-points of Kendon's

[2] This claim and the title of the paper might seem a bit misleading to the reader as we state in §1 that gestures are obligatorily accompanied by speech. However, the main point here is that gestures (like sign languages) are not realised through speech.

continuum, shown in (1) (McNeill 1992, 37).[3] Before proceeding to the analysis of dimensions of Kendon's continuum, it is crucial that we define what is implied by each of these forms of manual communication, as the term gesture is sometimes used rather broadly for any kind of bodily movement to convey a message. Gestures or gesticulation refer to "motion that embodies a meaning relatable to the accompanying speech" (McNeill 2005, 5). Thus, the term 'gestures' in our paper denotes those cases of non-verbal communication which co-occur with speech and do not constitute a codified system of signs. Emblems are manual signs whose meaning is established by mutual agreement between members of a particular culture, such as the 'thumbs-up' sign, while pantomimes are sequences of gestures that express a particular story with no speech involved. While not completely codified as linguistic systems, emblems and pantomime are conventionalised to a certain extent as all participants in the conversational act need to be aware what each symbol means, and they tend to differ from culture to culture (Kita 2009; Özyürek 2012). Sign languages, according to McNeil (2005, 5) are systems "with their own linguistic structures, including grammatical patterns, stores of words, morphological patterns". Thus, they are fully codified systems with a coherent set of basic units and rules for combining those units.

(1) Gestures → Pantomimes → Emblems → Sign Languages

McNeill (2005) states that gestures and sign languages are located on the opposing ends of this continuum, given in (1), according to several criteria which we explain below: i) relationship to speech, ii) relationship to linguistic properties, iii) relationship to conventions, and iv) character of semiosis.[4] On this continuum, pantomimes and emblems occupy the middle ground and their relative location towards the either end of the spectrum, i.e. their characterisation as more gesture-like or more sign language-like, is somewhat variable along these dimensions (McNeill 1992; 2005).[5]

In terms of i), gestures represent one extreme of the spectrum as they are always necessarily accompanied by speech, while sign languages obligatorily involve no speech at all, which constitutes the other extreme (McNeill 1992). Between the two extremes lie emblems, such as the OK sign made by creating a circle with a thumb and the index finger, which can (but do not have to) be accompanied by speech, and pantomime, which obligatorily involves no speech at all. With respect to this dimension, emblems seem to be more gesture-like, while pantomimes are closer to sign languages. The second and third relevant dimensions along which

[3] Kendon's continuum was named after Adam Kendon, who presented one of the first contemporary classifications of gestures (McNeill 1992).

[4] Taking into consideration the number of dimensions, McNeill (2005, 6) believes this is better described as a set of several continua, rather than a single continuum. This distinction is of somewhat secondary importance for the discussion presented in this section.

[5] A similar discussion on the differences between gestures and sign languages can also be found in Özyürek (2012).

gestures and sign languages differ is their relationship to linguistic properties and their conventionalisation. Gestures have no conventionalised standards of form, they cannot be analysed and segmented into constituents, and cannot be combined into higher order structures to create more complex gestures. Contrary to gestures, sign languages are conventionalised systems with standards of well-formedness (cf. Section 3); they can be segmented into individual constituents (cf. Section 3), they have lists of basic units, equivalent to the lexicon in the traditional dictionary-and-grammar terminology, and these basic units can be combined into larger structures, i.e. syntax in traditional dictionary-and-grammar terminology. Emblems and pantomimes are again located between these two extremes but with inverse relative positions. Emblems have rules of well-formedness, which means certain rules for their creation have to be obeyed (e.g. a pointed middle finger conveys a specific message that a pointed index or ring finger do not), which makes them more akin to sign languages. However, they cannot be further embedded into larger hierarchical structures (McNeill 1992, 37–39). Pantomimes are in this respect also more gesture-like as there seem to be no conventionalised rules for them (McNeill 2005, 10).

Dimension iv) involves two dichotomies – global vs. segmented and synthetic vs. analytic. The term 'global' implies that the meaning of a particular gesture is not determined from individual movements, but from the gesture as a whole, while the term 'segmented' means that the meaning of the whole utterance is mapped out from the meaning of individual constituents (this issue is discussed in more detail in Section 4).[6] This is related to the second dichotomy, which represents another parallel with the spoken languages. Languages are divided into more synthetic or more analytic, depending on how much information they are able to 'pack' into words – morphologically more complex languages (like Croatian or Turkish) are closer to the synthetic end of the spectrum, and morphologically simpler languages (like English and Chinese) are analytic. Similarly, a sign (i.e. a bodily movement) is synthetic if it simultaneously expresses several meanings, while it is analytic if different meanings are represented by individual signs. Given this 2x2 array, each of the four types of manual communication represents a different combination of features – gestures are global and synthetic, pantomimes are global and analytic, emblems are segmented and synthetic and sign languages are segmented and analytic (McNeill 2005, 10–11). However, these categories of manual communication are motivated by more than just linguistic description. In fact, McNeill (1992) claims that damage to the cerebral area has a different effect on each of these categories – sign languages are affected similarly to spoken languages, and the condition can be considered equivalent to aphasia. Emblems and pantomimes are not affected differently by different kinds of

[6] This dichotomy is equivalent to the holistic vs. atomistic approaches in grammars of spoken languages. While McNeill (2005) takes the traditional standpoint that syntax of spoken languages is inherently atomistic, some approaches to grammar claim otherwise (cf. Goldberg 2006; Perek 2015).

aphasia, but their severity is correlated with the overall communicative impediment. Co-speech gestures, having a closer relationship with speech, are affected by these impediments in a similar way to the speech they accompany.

Although the positioning of gestures and sign languages on the opposing poles of Kendon's continuum might indicate the two are completely dissimilar, this is certainly not the case. Kendon (2004, 284) states that use of gestures in particular circumstances (e.g. without being accompanied by speech) may result in a somewhat conventionalised kinesic code, and further elaboration of this code may lead to the creation of sign languages. Among these, it is important to distinguish primary sign languages, which are developed within the community of deaf people, and alternate sign languages, which may be found in speaker-hearer communities. The different origins of kinesic codes have significant implications for their further development. Kendon cites the examples of research by Goldin Meadow and her colleagues (Feldman, Goldin-Meadow, and Gleitman 1979; Goldin-Meadow 1979, 1982, 1993; Goldin-Meadow and Mylander 1990) which show how deaf children who grew up with hearing parents and no training in sign language spontaneously develop a particular kinesic code with a coherent set of rules. A similar situation can be noted in the case of the Nicaraguan Sign Language which arose through interaction of the so-called 'home sign' systems. The home sign systems were kinesic codes that several communities of deaf people developed on their own but separately from one another. As they began to interact using their own systems, a new code emerged which would ultimately grow into a fully-fledged sign language – Idioma de Señas de Nicaragua (ISN) (Kendon 2004, 291). Liddell and Metzger (1998) provide evidence of another connection between gestures and sign languages. They argue that deictic gestures can be used in conjunction with ASL (American Sign Language) signs for indicating context-dependent referents, much like the gestures in the spoken languages. Their research was confirmed by a subsequent study of ASL, Australian Sign Language, Taiwan Sign Language and nonsigners' gestures (Schembri, Jones, and Burnham 2005), which showed that classifier constructions with verbs represent a blend of gestures and signs.

Codes similar to the home sign systems with varying degrees of complexity can also arise in speaking communities under special circumstances. According to McNeill (1992, 39), it is customary for Warlpiri women to relinquish speech altogether for a particular period of time when going into mourning and during other special occasions. During those periods, they use a system of signs McNeill calls the Warlpiri Sign Language (WLS). This code can also be used to accompany speech when speech is culturally permitted, or as an alternative means of communication. Kendon (2004) also mentions examples of codes that arise among the hearing in specialised professional circumstances – crane driver gestures, sawmill systems and monastic sign languages. The first represents a limited code with a restricted set of signs related to the

job in question, i.e. a limited system of form and meaning pairings which could hardly be called a proper linguistic system. The latter two, however, include more complex linguistic features – compound signs created through concatenation of simpler signs, e.g. a combination of GOD+UP+DAY signs for 'Easter' in the case of the monastic signs. However, their usage is still limited to specific domains, which means they are too narrow to be regarded as fully fledged languages (Kendon 2004, 291–98).

The common denominator in all of these examples is that there seems to be the incremental development of unsegmented and unconventionalised movements towards more coherent, segmentable and conventionalised codes and ultimately languages. This gradual development of linguistic features tells a more interesting story from a wider perspective – how linguistic features and complexities arise, to which we return in Section 5.

3 Sign Language Systems

Hearing impairments can be caused by various factors such as long-term exposure to noise with intensity over 85 dB, hereditary diseases, prenatal, perinatal, and postnatal diseases. Complete or partial deafness primarily inhibits communication with the hearing environment, which underlines the need for non-auditory channels of communications (Juriša 2012). Thus, three basic types of manual communication can be distinguished (Bradarić-Jončić 2000; Juriša 2012):

1) *Manual alphabet* – the three times slower manual form can be used to write about 60 words per minute (in comparison to vocal speech, with the rate of 180 words per minute). Distinction is made between a one-handed alphabet, or dactylology, and a two-handed alphabet, or chirology.

2) *Simultaneous sign-oral communication* (or manually coded vocal languages) – simultaneous usage of oral and sign communication in which syntactic structures of the spoken language follow lexical units of the sign languages. This is a visualised spoken national language accompanied by signs borrowed from the source sign language and the signs from the manual alphabet. These are exemplified by the Signed English Language, Signed Croatian Language, Signed Exact Italian, etc. This communication system is no longer regarded as appropriate for use as the primary communication system acquired by a deaf child, primarily due to exposure of the child to a mixture of two simultaneously transmitted incomplete language systems.

3) *Original sign language* – this is defined as a "[…] a standalone language system […] with its own rules of grammar, different from those in the hearing community." (Bradarić-Jončić 2000, 125). The examples of this are the American Sign Language (ASL), British Sign Language (BSL), Italian Sign Language (ISL), etc.

Up until the second half of the 20th century, the pervasive opinion was that sign language is holistic, that its complexity is based on speech conventions and that it has more resemblances to everyday gestures than to spoken languages. This perception was common not only among laymen but also among some of the renowned linguists of the period, like Leonard Bloomfield and Edward Sapir. However, the author of the first dictionary of sign language William Stokoe defines the notion of 'gesture' as a communicative movement and does not equate it to the sign as was the case with earlier approaches (Meier 2002; Sandler 2014; Fenlon, Cormier, and Brentari 2015). Subsequent research has shown that sign language has a phonological system very much like spoken language, while gestures represent a form of manual communication and not a linguistic system, and consequently, have no phonological module (McNeill 2005; 2006).

Stokoe's linguistic analysis from 1960 marked the beginning of gradual yet substantial changes in the way linguists perceived sign languages. Stokoe claimed that the signs of the ASL consist of a finite list of units with no meaning on their own, as is the case in the phonological domain of the spoken languages. Furthermore, sign languages also consist of two structural levels – the level of meaning and the level without meaning, and it is this duality that represents one of the fundamental features of the human language (Hockett 1960; Martinet 1960; Sandler 2014), and the second point of conjunction between spoken and sign languages. Stokoe's discovery also showed that comparing natural languages in two different physical modalities makes the theoretical questions about the universality and inherent features of linguistic structures more intriguing and interesting for further research (Meier 2002).

The discovery that both speech and sign take part in the realisation of language is a confirmation of previous ideas on the existence of multimodal linguistic means which changed the general understanding of what sign language really is. The analysis of different types of sign languages, such as American Sign Language or British Sign Language in countries in which English is the official language in use, is independent of the analysis of the English language in speech and writing.[7] This is because the process of learning the sign language necessarily involves acquisition of a communication system whose principles do not rely on the standard structure of grammar and vocabulary of the English language. This was one of the primary reasons why sign languages were not within the scope of too many linguistic analyses, as these were primarily focused on written and spoken varieties. However, there are artificial sign systems mentioned above, such as the Signing Exact English in the USA, or the Paget Gorman Sign System in the UK, which are closely related in terms of their grammar and vocabulary to the spoken English language (e.g. there are signs

[7] American Sign Language and British Sign Language are two unrelated sign languages, while the spoken varieties have vastly more overlapping features.

which represent affixes, they have the same order of sentential constituents, etc.). Other artificial sign systems, such as Cued Speech, are based on the phonology of the English language and contain signs which enable the deaf persons to recognise which phoneme was articulated in cases of insufficient visual cues. These systems were designed primarily for educational purposes so that the speaking-impaired children, who may or may not have hearing impairments, would be encouraged to communicate (Crystal 1995).

3.1 Types of Linguistic Modalities

Numerous empirical studies on the acquisition of the ASL and other sign languages had been conducted by the end of the 20th century, which ultimately led to the conclusion that the human language competence is plastic and that there exist at least two linguistic modalities or transmission channels through which language can be produced and perceived – the vocal-auditory modality in the spoken language, and the visual-gestural modality in the sign language (Meier 2002). Sign language differs from spoken language as it emerges in the visual-gestural modality, i.e. its phonological organisation is determined by the type of articulators used (Fenlon, Cormier, and Brentari 2015). According to Sandler (2012), the visually perceptible and moving parts of the body, such as hands, facial expressions, head and the upper body, represent the articulators of the sign language. Words are created, delimited and compared with one another through the use of these articulators. Sign language is formed through hand and body movements and facial expressions. Contrary to the belief that sign languages are the same as spoken languages, apart from being different in terms of the content of their features (Perlmutter 1992), Brentari (2002) believes that the type of modality used in communication affects the structure of the phonological system, as the auditory system has an advantage in horizontal processing, while the visual system has an advantage in vertical processing (it processes paradigmatic information at a faster rate and with more accuracy).[8] This is how the differences in frequency of occurrence of particular phenomena and differences in phonological features of sign and spoken languages arise.

3.2 Phonological System of Sign Languages

Considering how the acquisition and processing of the phonological features of a language are typically attributed to spoken languages and require mastering of their unique structure for producing and hearing distinctive phonological and prosodic features and ignoring non-distinctive features, the acquisition of differential and

[8] Horizontal processing refers to our ability to process temporally defined input into temporally defined events (e.g. sequencing of objects in a particular period of time, syntagmatic processing). Vertical processing refers to our ability to process different types of input displayed simultaneously, such as pattern recognition and paradigmatic processing (cf. Brentari 2002).

culturally determined intonations (Jelaska 2004, 11), the claim that sign languages have phonological systems was groundbreaking. The study of sign language phonology started in the 1960s, parallel to the first studies of sign languages in general (Fenlon, Cormier, and Brentari 2015). Sandler believes that the emergence of sign language phonology is in line with the principles of Blevins's (2004) theory based on the study and analysis of the phonological history of spoken languages, which claims that most of the synchronic phonological features of any language stem from the interaction of physical, cognitive and social elements of history and are not in and of themselves intrinsic. This supports the view that sign language phonology was created incrementally, developing fundamental phonological features through transfer and usage in everyday communication.

Stokoe made a huge step forward in understanding sign languages by producing the first dictionary of ASL. The explanations for key terms in ASL phonology have facilitated understanding and comparison of the phonological systems of spoken and sign languages. In (2), we have provided a part of his glossary with some of the basic yet crucial terms for understanding his view on sign languages (Stokoe Jr. 2005). Stokoe describes the minimal pairs which differ in terms of hand configuration (handshape), location (place of articulation) and type of movement, which was one of the first pieces of evidence for the existence of phonological system in ASL (Sandler 2012; 2014). In (3), we compare some of Stokoe's main terms in sign language with their equivalents in spoken languages. However, as Liddell (2003, 7) points out, Stokoe's terminology was not accepted by the wider linguistic community, which opted for the traditional linguistic terminology.

(2)

ASPECT – the structural distribution of activities of sign language (analogous to the segment) onto constituents for position, configuration and movement (analogous to vowels and consonants in the spoken language);

ALLOCHER – cheremes with identical realisation in language;

ASPECTUCAL CHEREME – tab, dez or sig (see below);

CHEREME – a distinctive unit which corresponds to the phoneme in the spoken language, signs are formed by combining cheremes;

CHEROLOGY – the structure and analysis of units in sign language;

DEZ – a configuration of hand(s) which makes a meaningful movement in a particular position;

SIG – a component of movement or aspect of activity of sign language; a specific movement of hand configuration (dez) in a particular position;

TAB – the place where the movement occurs (place of articulation), which designates

the aspect of the sign language activity; the place in which a configuration (dez) makes the movement (sig);

SIGN – the smallest unit of sign language which has a lexical meaning (analogous to word); one of the two types of morphemes used to construct an expression in sign language (the other is finger spelling).

(3)

SIGN LANGUAGE	SPOKEN LANGUAGE
cherology	phonology
chereme (later parameter)	phoneme
allocher	allophone
sign	word

3.3 Prosodic Model of Phonology of Sign Languages

In the post-1960 research, phonologists described the models of phonological systems of different sign languages, thereby confirming that units with no meaning really behave systematically and are best understood not in terms of phonemes (cheremes in Stokoe's terminology), but in terms of features which are assumed not to be innate (Sandler 2014). Unlike spoken languages, which stem from older languages or have centuries of history behind them, sign languages can emerge anew as a means of communication in a newly founded group of deaf persons (Sandler 2014). Sandler believes that the emergence of such sign languages provides the ability to study the emergence of phonology and other linguistic levels in real time.[9] Apart from this, the emergence of other novel approaches in the phonological theory by the end of the 20th century, such as autosegmental phonology (Goldsmith 1976), feature geometry (Clements 1985; McCarthy 1988; Clements and Hume 1995), and prosodic phonology (Nespor and Vogel 1986; Itô 1986) allowed for a more detailed comparison of the phonology of spoken and sign languages.

Several phonological models for the description of parameters of sign languages have been proposed, and the prosodic model of sign language phonology is regarded as one of the most influential (Brentari 1998).[10] The aim of this model is to integrate the

[9] Sandler (2014) spent ten years researching the Al-Sayyid Bedouin Sign Language (ABSL). She sees it as an extremely functional language with no prejudice against it as it is used by both deaf and hearing persons. Apart from the interesting sociolinguistic image of the Al-Sayyid tribe in south Israel, Sandler describes the emergence of the phonological system and, based on the evidence presented in her paper, claims that ABSL still does not have a fully developed phonological system.

[10] Before the emergence of the prosodic model of sign language phonology, the predominant models were the hold-movement model (Liddell Johnson 1989 in Fenlon, Cormier, and Brentari 2015) and the hand-tier model (Sandler 1989 in Fenlon, Cormier, and Brentari 2015).

systematicity of the paradigmatic and syntagmatic structures of the sign into a single model. The prosodic model distinguishes two types of features of the sign – inherent and prosodic. Inherent features are articulated simultaneously during the creation of a sign, while prosodic features are articulated consecutively (Brentari 1998; Šarac Kuhn, Alibašić Ciciliani, and Wilbur 2006). These features are determined once per lexeme (sign) and remain unchanged throughout the production of the sign. They have a more complex hierarchical structure, they occur simultaneously (Brentari 1998), they are susceptible to a smaller number of constraints, and they do not create time segments (Brentari 2002). In terms of articulation, inherent features include handshape (a primarily active articulator), place of articulation, and hand orientation (passive articulators). Brentari (1998; 2002) draws parallels between the articulators of sign and spoken languages, and claims that the vocal mechanism in speech has primary active articulators (tongue, lips and larynx) and passive articulators (teeth, palate, and the pharyngeal area). Lips and glottis can act as active or passive articulators, while others are constant – the tongue is always active and the palate is always passive in speech production. However, this is not the case in sign language, as every part of the body involved in the production of the sign can be active or passive. As an example, Brentari (2002) mentions the hand which is an active articulator in the lexeme THINK but a passive articulator in the lexeme TOUCH. In fact, as Sandler (2012) points out, the dual use of dominant and nondominant hand in articulation does not have a direct counterpart in the spoken language, which further demonstrates the specific role of articulators in the phonological structure of the sign language.

Hands as manual articulators are primarily active and generally regarded as the most frequent ones. Sometimes the sign uses the nonmanual articulators, such as head, face and/or body. The manual articulators branch out into dominant (H1) and nondominant (H2) hands. A one-handed sign only has the H1 features, which includes contrastive units, such as specific fingers that are moved, number of fingers moved and the form of the finger (straight, bent, curved). Fenlon, Cormier, and Brentari (2015) cite an example from British Sign Language (BSL) in which the signs GAY and UNSURE differ only in the number of selected fingers – GAY is signed with extended thumb and other fingers closed, while UNSURE is signed by having all five fingers extended. A two-handed sign will have both H1and H2 features (Fenlon, Cormier, and Brentari 2015). In terms of handshape, three groups of two-handed signs can be observed: i) same handshape and movement for both hands; ii) same handshape, only the dominant hand is moved; and iii) different handshapes, only the dominant hand is moved (Battison 1974; Šarac Kuhn, Alibašić Ciciliani, and Wilbur 2006).

The place of articulation in which inherent and prosodic features are realised is divided into three planes – horizontal (y-plane), which refers to upward and downward movement of the body; vertical or frontal (x-plane), which refers to forward and

backward movement of the body; and midsagittal (z-plane), which refers to leftward or rightward movement (Brentari 1998; Šarac Kuhn, Alibašić Ciciliani, and Wilbur 2006). The signs in the vertical plane are also specified for four main regions – head, arm, body and hand, and further specified for contrastive features (Fenlon, Cormier, and Brentari 2015), which we do not mention here for the sake of brevity.

Orientation in the Prosodic model represents the connection between the hand and the place of articulation and includes two types – orientation of the hand and orientation of the fingertips. Both types have six possible directions – up, down, left, right, front, and back (Šarac Kuhn, Alibašić Ciciliani, and Wilbur 2006). As an example of the relationship between orientation and place of articulation, Fenlon, Cormier, and Brentari (2015) mention the signs MUM and DANGER, which differ in terms of hand orientation (the fingers are oriented towards the place of articulation for the former, and away for the latter).

Prosodic features (PF) are the features of signs which are realised sequentially via dynamic features of the movement (Brentari 1998), in contrast to inherent features, which are realised simultaneously. Prosodic features are made up of types of movement (Brentari 1998; 2002), which include four main categories – straight, arc, circle and trilled, which is specified for every sign. The other PF values that the signs in sign language are marked for are path, setting orientation and apertures (Fenlon, Cormier, and Brentari 2015), but we do not discuss them here in full detail for the sake of brevity.

A combination of inherent and prosodic features determines some limitations of the sounds. There are two universal complementary conditions which define the complexity of the sign – the symmetry condition and the dominance condition as its corollary (Kyle and Woll 1985; Valli and Lucas 2000; Pribanić and Milković 2012). The first condition states that if both hands are active, they need to have the same location and the same type of movement, while the latter condition determines that if two hands have a different handshape, one has to be active and the other passive.

While not obvious at first glance, there are several aspects that the phonological systems of spoken and sign languages share. The main parallel lies in the structure that signalises paradigmatic contrast – in sign languages, that function is performed by the inherent features of handshape and place of articulation, while this is done by consonants in the spoken languages. A similar parallel can be drawn between movements (prosodic feature) in sign language and vowels in spoken language, both of which represent media for carrying signal over distance (Brentari 2002).[11] The inherent branch of the structure contains a bigger lexical contrast than prosodic

[11] Brentari (2002) also lists calculation of complexity and role of the root node as a point of conjunction between syntax and phonology.

features, much like consonants have a higher potential of lexical contrast than vowels in spoken languages. The movement (prosodic feature) represents a medium for signal, much like vowels function as the medium for the spoken language. The movement (prosodic feature) functions as the basis of the syllable in sign language. However, the main difference between these two phonologies is that consonants and vowels are realised simultaneously in sign language and sequentially in spoken language.

4 Gestures

When observing gestures from a scientific perspective, they are defined as body movements (mainly arm and hand movements, but also movements of some other body parts, to a lesser extent) which appear in communication and form part of the utterance. This definition excludes all non-verbal movements whose function is practical and non-communicational, such as fixing your hair, playing with jewellery, etc. (Gullberg 2009; Cooperrider and Goldin-Meadow 2017). Numerous scientists are of the opinion that gestures and language form a single, integrated system. This synchronised integration of the visuo-spatial modality depicted through gestures and the verbal modality in the form of spoken languages facilitates not only language production, but also enables better comprehension of the articulated message (McNeill 1992, 2005; Holler and Beattie 2003; Habets et al. 2010; Kelly, Özyürek, and Kelly 2010), which, consequently, makes language acquisition less problematic (Cooperrider and Goldin-Meadow 2017). Throughout history there have been many classifications of gestures which approached them from different standpoints (starting from Ancient Roman rhetorical studies to the more contemporary and scientifically-oriented studies from the 20th century onwards). The most cited is McNeill and Levy's classification (1982) which divides gestures into four categories or dimensions: a) *Iconic* – gestures that depict concrete objects and actions; b) *Metaphoric* – gestures which depict abstract concepts; c) *Deictic* – pointing gestures; and d) *Beats* – hand movements which accompany speech rhythm. One type of gestures rarely occurs as an isolated kinesic pattern, but more frequently in combination with other types (McNeill 1992, 2005).

4.1 Production of Gestures

The gesture movement hierarchy was first initiated by Kendon and modified by McNeill in the form of a diagram presented in Figure 1 (McNeill 1992). Arm use and body posture refer to the various patterns of arm usage and body position adopted by the speaker. Head movement usually starts from the centre of the gesture space. *Gesture-Unit* or *G-Unit* represents the period of time which starts when the limb begins to move and finishes when the limb gets to the resting position. According to Kendon, gesture production consists of three *phrases of gesticulation* or *G-Phrases* (1980, 212–15):

1) Preparation – the arm starts moving upwards from the resting place to a position in which the stroke phase of the gesture is about to start;
2) Stroke – upper arm goes inwards and outwards two times in order to move the hand into the centre of the gesture space; and
3) Retraction or recovery phase – the arm moves downwards to its resting position.

The stroke represents a basic part of gesture production, while the preparatory and recovery phases have proven to be optional. Hold (pre-stroke or post-stroke) represents any short-term pause in movement; pre-stroke hold usually occurs when the stroke is postponed, while the post-stroke hold appears at the end of the stroke, prior to retraction (McNeill, Levy, and Pedelty 1990, 209–11; McNeill 1992, 82–83, 2005: 29–36).

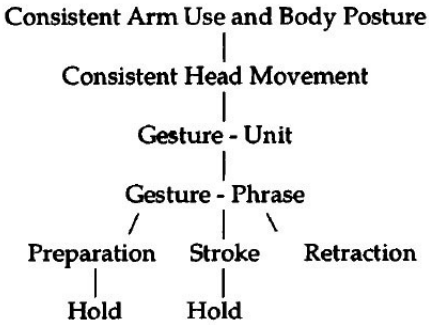


Figure 1. Gesture levels based on Kendon's kinesic hierarchy (McNeill, Levy, and Pedelty 1990, 209).

Sign languages also have a sequence of three elements (Hold – Movement – Hold) when making a sign, and they are phonetically realised[12] (Kita, van Gijn, van der Hulst 1998). Considering the production of gestures and sing languages, it can be concluded that the structure of movement (the onset of movement, stroke and the hold phase) is a common feature of both gestures and signs.

4.2 Gestures versus Linguistic Systems

Gestures and languages can both express a particular meaning, but how they convey it reflects a fundamental difference between them. When expressed in form of a language, the meaning of a single action or an event is divided into segments, i.e. hierarchically organised strings of words. A hierarchy based on segmentation and linearisation, a generally assumed common property of all linguistic systems (but see Section 5 for a discussion on this issue), stems from the premise that all languages (spoken or sign)

[12] See Section 2 for more details on the phonological structure of signs in sign languages.

are one-dimensional, i.e. they change in accordance with a single dimension of time, echoing the relationship between language units; phonemes, morphemes, words, phrases, sentences, and discourse. This restriction along with the multidimensionality of meanings is what forces languages to split the meaning into segments and combine them along a single timeline. Unlike the spoken and sign languages, gestures do not undergo segmentation and linearisation because they are multidimensional and can present complex meanings as wholes, which is what supports another important property of gestures – their non-combinatoric nature, the inability to form a more complex gesture out of two or more simpler gestures. As opposed to the sentences in which smaller units can form larger ones, gestural symbols are already complex enough and express a complete meaning with no need to combine with other gestures. Still, they tend to convey the meaning from different perspectives, with each perspective of meaning being complete and expressed on its own. The final dissimilarity between gestures and language systems worth mentioning is the duality of patterning. Words of language systems are usually organised in two potential patterns of contrast at once; phonological and semantic contrast. Phonological contrast implies that words differ from one another in terms of sounds (e.g. "dog" in contrast to "doll" or "dig"), whereas the semantic contrast indicates a distinction in meaning (e.g. "dog" in contrast to "cow" or "monkey"). Gestures do not have the duality of patterning; their kinesic form is not independent as the sounds are, and is dictated by the meaning of the gesture (McNeill 1992). The ability to express the meaning and, consequently, define its form is what makes the use of gestures more advantageous and 'less demanding' when compared to linguistic systems which have the separate structure of the form and meaning.

Despite all the differences, gestures and languages belong to the same system when considering a number of similarities which connect them: a) speech is always accompanied by gestures (co-speech gestures); b) semantic and pragmatic co-expressiveness marks the symbiosis of speech and gestures; c) they appear in synchrony; d) they develop together in early ages; and e) the neurological damage in aphasic patients affects both speech and gestures (McNeill 1992). Kita (2009) also claims that gestures and language are so correlated that no culture has been found that does not have co-speech gestures. McNeill's observations also speak in favour of the idea that gestures and speech form an "unbreakable bond" (2005, 24–29): a) speech-gesture synchrony is not interrupted by delayed auditory feedback[13] (DAF cannot break speech-gesture synchrony); b) gestures lessen stuttering (gesture stroke phase weakens the onset of stuttering); c) congenitally blind people gesture (i.e. lack of vision does not prevent the blind from gesturing); d) information exchange

[13] Delayed auditory feedback (DAF) includes hearing your own speech played back over the earphones after a short delay. This has a negative effect on speech fluency which tends to be interrupted and slowed down (McNeill 2005, 25).

(information transferred in a form of a gesture may be recalled in a form of speech and vice versa); and e) gesture and speech fluency is parallel, not reciprocal (i.e. when the speech fluency decreases so does the gesture fluency). As stated earlier in the paper, Liddell and Metzger (1998) and Liddell (2003) also provide evidence of gestures being used in conjunction with sign languages.

5 Sign Language and Gestures in a Wider Setting – Interaction with Other Linguistic Modules

The discussion presented in this paper shows some of the main features of co-speech gestures and sign languages. As can be seen from Section 3, the status of phonology of sign language as the study of its minimal units is beyond dispute. The presence of other linguistic aspects in sign language, such as syntax and morphology, is also rather uncontroversial (*inter alia*, Neidle 2000; Valli and Lucas 2000; Brentari 2012; Meir 2012; Steinbach 2012; Zwisterlood 2012; Mathur and Rathmann 2012; Cormier 2012; Neidle and Nash 2012; Quer 2012; Tang and Lau 2012). For instance, reduplication and conversion, two very well-established and in some languages very productive word-formation patterns, are well established in sign language as well (Valli and Lucas 2001; Goldin-Meadow and Brentari 2017; Tkachman and Meir 2018). Like spoken language, some varieties of sign language, such as American Sign Language and Italian Sign Language, also exhibit the possibility of embedding relative clauses into more complex syntactic structures (Goldin-Meadow and Brentari 2017, 5).

Despite these formal similarities, it is worth noting that sign and spoken language differ in certain aspects, besides the modality through which they are realised. According to Goldin-Meadow and Brentari (2017, 7), ASL is able to express polymorphic words, i.e. words containing more than one stem and/or affix, using a monosyllabic sign. In spoken languages like English, Hmong and Hopi, all three other possible combinations of syllable-morpheme correspondence (monosyllabic monomorphemic words, polysyllabic monomorphemic words, polysyllabic polymorphemic words) are attested, except for this, which makes this feature of ASL rather unique. Another morphological peculiarity of sign language is related to another frequently covered concept in spoken language – verb agreement. Like spoken language, sign language utilises particular units that mark the features of verb arguments in a particular setting. According to Goldin-Meadow and Brentari (2017), when the sign for ASK (bent index finger) is moved from the signer towards the interlocutor, it means *I ask you*; when the sign is moved from the interlocutor towards the signer, it means *You ask me*. However, this phenomenon differs from agreement in spoken language in several respects. The number of possible combinations of agreement features (e.g. number and person) is finite in spoken language – you can only get as many combinations as allowed by the grammar of a language in question. The number of possible locations

towards which the verb is directed (i.e. predicate arguments) in sign language is not finite. A different sign can be directed at any participant in the discourse. The form of the sign used also varies from referent to referent, which means that a different sign will be used for a tall person and a short person. This variability and lack of discreteness make this property of sign language very different from the categorical grammatical notions that agreement in spoken language entails (Goldin-Meadow and Brentari 2017). It also makes this aspect of sign language more akin to gestures than to grammar in spoken language.

An interesting view on the relevance of specific types of sign languages is provided by Jackendoff and Wittenberg in their paper on linear grammars (Jackendoff and Wittenberg 2016). Their hypothesis is that complex spoken languages have emerged gradually, through the evolution of linguistic systems which did not have the level of grammatical complexity of modern languages. These simple grammars, which they call *linear grammars*, involve simple pairings of form (sounds or signs) and meaning (concepts), and have very little morphology and syntax. For instance, Jackendoff (2009) argues that NN compounds in English might be regarded as vestiges of the what Bickerton (1990) calls a 'protolanguage' – a previous step in the evolution of language which involved simple form-meaning pairings that depended largely on pragmatics and had little or no morphological or syntactic complexities. On a similar note, Jackendoff and Wittenberg (2016) claim that some sign languages might be regarded as linear grammars. Home signs, the sign languages invented by deaf children with no exposure to actual signed languages, have the basic form-to-meaning mappings and involve very little morphology (Jackendoff and Wittenberg 2016).The village sign languages, like the ABSL (Sandler 2014) have similar pairings of signs and concepts, but seem to lack syntactic structure. The word order is typically agent first, action second, but the utterances involving two animate arguments of the verb are potentially ambiguous. Meir (2018) regards this lack of syntactic embedding in ABSL and ISL (Israeli Sign Language) as a strong argument against recursion as a crucial property of the language faculty (Hauser, Chomsky, and Fitch Hauser 2002). Speakers of ABSL use paratactic structures, i.e. sequences of two or more concatenated sentences with no formal embedding, to convey the same meaning as the syntactic structures in spoken language which include overt complementisers (such as *that*-clauses in English) (Meir 2018). The study on compounding in ISL and ABSL by Tkachman and Meir (2018) speaks in favour of this view that linguistic structure is an emergent phenomenon, and that it develops at different rates in different domains and languages.[14] Cases like this suggest that segmentability may not be an inherent feature of sign languages as such, but a by-product of their development.

[14] Similar claims can be found in research on grammaticalization (Bybee 2006; Traugott 2008).

6 Conclusion

Taking into consideration everything stated in the sections above, we believe adding co-speech gestures as the third point in the relationship between spoken and sign languages gives us a new outlook on the nature of complexity in systems of communication. While some phenomena in spoken language have their counterparts in sign language, some features of sign languages seem to be closely related to co-speech gestures (McNeill 2005; 2006). Gestures seem to have a complementary relationship with spoken languages, they are even used by blind persons who have no benefit whatsoever from the visuo-spatial modality (cf. Section 4), they seem to be affected by the same neural disfunctionalities, and yet there are systems of manual communication which have language-like features (cf. Section 2 and Kendon 2004). Furthermore, there are sign languages like ASL with a morphological and syntactic complexity similar to that of spoken languages, and there are spoken languages like Riau Indonesian with little to no morphological and syntactic complexity (Jackendoff and Wittenberg 2016). Cases such as these tell us that the differences between sign languages and gestures and sign languages and spoken languages are far from categorical.

Ultimately, this brings us back to the notion of linear grammars invoked by Jackendoff and Wittenberg (2016) and even further back to the Saussurean notion of the linguistic sign. While the Saussurean concept of language as a system of signs may not be complex enough to explain the phonological, morphological and syntactic intricacies of modern-day spoken languages, evidence provided by co-speech gestures and sign languages are a good indication that the form-meaning pairing may be the fundamental motivation behind human communication systems, regardless of the modality in which they take place.

References

Bickerton, Derek. 1990. *Language & Species*. Chicago: University of Chicago Press.

Blevins, Juliette. 2004. *Evolutionary Phonology*. Cambridge: Cambridge University Press.

Bradarić-Jončić, Sandra. 2000. "Manualna komunikacija osoba oštećena sluha." *Hrvatska revija za rehabilitacijska istraživanja* 36 (2): 123–36.

Brentari, Diane. 1998. *A Prosodic Model of Sign Language Phonology.* Cambridge, MA: MIT Press.

—. 2002. "Modality Differences in Sign Language Phonology and Morphophonemics." In *Modality and Structure in Signed and Spoken Languages,* edited by Richard P. Meier, 35–65. Cambridge: Cambridge University Press.

—. 2012. "Phonology." In *Sign Language: An International Handbook*, edited by Roland Pfau, Markus Steinbach and Bencie Woll, 21–54. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110261325.21.

Bybee, Joan L. 2006. "From Usage to Grammar: The Mind's Response to Repetition." *Language* 82 (4): 711–33. https://doi.org/10.1353/lan.2006.0186.

Clements, George N. 1985. "The Geometry of Phonological Features." *Phonology Yearbook* 2 (1): 225–52. https://doi.org/10.1017/S0952675700000440.

Clements, George N., and Elizabeth V. Hume. 1995. "The Internal Organization of Speech Sounds." In *The Handbook of Phonological Theory*, edited by John A. Goldsmith, 245–306. Cambridge, MA: Blackwell.

Cooperrider, Kensy, and Susan Goldin-Meadow. 2017. "Gesture, Language, and Cognition." In *The Cambridge Handbook of Cognitive Linguistics*, edited by Barbara Dancygier, 118–34. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316339732.

Cormier, Kearsy. 2012. "Pronouns." In *Sign Language: An International Handbook*, edited by Roland Pfau, Markus Steinbach and Bencie Woll, 227–43. https://doi.org/10.1515/9783110261325.227.

Crystal, David. 1995. *The Cambridge Encyclopedia of the English Language.* Cambridge: Cambridge University Press.

De Saussure, Ferdinand. 2000 [1916]. *Tečaj opće lingvistike*. Zagreb: ArTresor naklada, Institut za hrvatski jezik i jezikoslovlje.

Feldman, Heidi, Susan Goldin-Meadow, and Lila Gleitman. 1978. "Beyond Herodotus: The Creation of a Language by Linguistically Deprived Deaf Children." In Ac*tion, Symbol, and Gesture: The Emergence of Language*, edited by Andrew Lock, 351–414. New York: Academic Press.

Fenlon, Jordon, Kearsy Cormier, and Diane Brentari. 2017. "The Phonology of Sign Languages." In *The Routledge Handbook of Phonological Theory,* edited by S. J. Hannahs and Anna Bosch, 453–75. New York: Routledge.

Goldin-Meadow, Susan. 1979. "Structure in a Manual Communication System Developed Without a Conventional Language Model: Language Without a Helping Hand." In *Studies in Neurolinguistics, Volume 4, A Volume in Perspectives in Neurolinguistics and Psycholinguistics*, edited by Haiganoosh Whitaker and Harry A. Whitaker, 125–209. New York: Academic Press. https://doi.org/10.1016/B978-0-12-746304-9.50010-0.

—. 1982. "The Resilience of Recursion: A Study of a Communication System Developed Without a Conventional Language Model." In *Language Acquisition: The State of the Art,* edited by Eric Wanner and Lila R. Gleitman, 51–77. New York: Cambridge University Press.

—. 1993. "When Does Gesture Become Language? A Study of Gesture Used as a Primary Communication System by Deaf Children of Hearing Parents." In *Tools, Language and Cognition in Human Evolution*, edited by Kathleen R. Gibson and Tim Ingold, 63–85. New York: Cambridge University Press.

—. 2017. "What the Hands Can Tell Us about Language Emergence." *Psychonomic Bulletin & Review* 24 (1): 213–18. https://doi.org/10.3758/s13423-016-1074-x.

Goldin-Meadow, Susan, and Diane Brentari. 2017. "Gesture, Sign, and Language: The Coming of Age of Sign Language and Gesture Studies." *Behavioral and Brain Sciences* 40: e46. https://doi.org/10.1017/S0140525X15001247.

Goldin-Meadow, Susan, and Carolyn Mylander. 1990. "The Role of Parental Input in the Development of a Morphological System." *Journal of Child Language* 17 (3): 527–63. https://doi.org/10.1017/S0305000900010874.

Goldsmith, John. 1976. "Autosegmental Phonology." PhD diss., Massachusetts Institute of Technology.

Gullberg, Marianne. 2009. "Why Gestures are Relevant to the Bilingual Lexicon." In *The Bilingual Mental Lexicon – Interdisciplinary Approaches*, edited by Aneta Pavlenko, 161–84. Bristol: Multilingual Matters.

Hauser, Marc D., Noam Chomsky, and W. Tecumseh Fitch. 2002. "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?" *Science's Compass* 298 (5598), 1569–79. https://doi.org/10.1126/science.298.5598.1569.

Habets, Boukje, Sotaro Kita, Zeshu Shao, Aslı Özyürek, and Peter Hagoort. 2010. "The Role of Synchrony and Ambiguity in Speech-Gesture Integration during Comprehension." *Journal of Cognitive Neuroscience* 23 (8): 1845–54. https://doi.org/10.1162/jocn.2010.21462.

Hockett, Charles. 1960. "The Origin of Speech." *Scientific American* 203 (3): 88–97.

Holler, Judith, and Geoffrey Beattie. 2003. "How Iconic Gestures and Speech Interact in the Representation of Meaning: Are Both Aspects Really Integral to the Process?" *Semiotica* 146 (1/4): 81–116. https://doi.org/10.1515/semi.2003.083.

Itô, Junko. 1986. "Syllable Theory in Prosodic Phonology." PhD diss., University of Massachusetts Amherst.

Jackendoff, Ray. 2009. "Compounding in the Parallel Architecture and Conceptual Semantics." In *The Oxford Handbook of Compounding,* edited by Rochelle Lieber and Pavol Štekauer, 105–28. Oxford: Oxford University Press.

Jackendoff, Ray, and Wittenberg, Eva. 2017. "Linear Grammar as a Possible Stepping-Stone in the Evolution of Language." *Psychonomic Bulletin & Review* 24 (1): 219–24. https://doi.org/10.3758/s13423-016-1073-y.

Jelaska, Zrinka. 2004. *Fonološki opisi hrvatskoga jezika*. Zagreb: Hrvatska sveučilišna naklada.

Juriša, Mirjana. 2012. *Oštećenje sluha, gluhoća/nagluhost i savjeti za prevladavanje komunikacijskih barijera. Zbornik radova sa stručnih skupova*. Zagreb: Hrvatski savez gluhih i nagluhih.

Kelly, Spencer D., Aslı Özyürek, and Eric Maris. 2010. "Two Sides of the Same Coin: Speech and Gesture Mutually Interact to Enhance Comprehension." *Psychological Science* 21 (2): 260–67. https://doi.org/10.1177/0956797609357327.

Kendon, Adam. 1980. "Gesticulation and Speech: Two Aspects of the Process of Utterance." In *The Relation Between Verbal and Nonverbal Communication*, edited by M. R. Key, 207–27. The Hague: Mouton.

—. 2004. *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.

Kita, Sotaro. 2009. "Cross-Cultural Variation of Speech-Accompanying Gesture: A Review." *Language and Cognitive Processes* 24 (2): 145–67. https://doi.org/10.1080/01690960802586188.

Kita, Sotaro, Ingeborg van Gijn, and Harry van der Hulst. 1998. "Movement Phases in Signs and Co-speech Gestures, and Their Transcription by Human Coders." In *Gesture and Sign Language in Human-Computer Interaction, GW 1997, Lecture Notes in Computer Science, vol. 1371*, edited by I. Wachsmuth and M. Frölich, 23–35. https://doi.org/10.1007/BFb0052986.

Kyle, Jim G., and Bencie Woll. 1998 [1985]. *Sign Language: The Study of Deaf People and Their Language*. [1st paperback ed., reprinted]. Cambridge: Cambridge University Press.

Liddell, Scott K. 2003. *Grammar, Gesture, and Meaning in American Sign Language*. Cambridge: Cambridge University Press.

Liddell, Scott K., and Melanie Metzger. 1998. "Gesture in Sign Language Discourse." *Journal of Pragmatics* 30 (6): 657–97. https://doi.org/10.1016/S0378-2166(98)00061-7.

Martinet, André. 1960. *Éléments de Linguistique Générale.* Paris: Armand Colin.

Mathur, Gaurav, and Christian Rathmann. 2012. "Verb Agreement." In *Sign Language: An International Handbook*, edited by Roland Pfau, Markus Steinbach and Bencie Woll, 136–57. https://doi.org/10.1515/9783110261325.136

McCarthy, John. 1988. "Feature Geometry and Dependency: A Review." *Phonetica* 4: 84–105.

McNeill, David. 1992. *Hand and Mind: What Gestures Reveal about Thought.* Chicago: University of Chicago Press.

—. 2005. *Gesture and Thought*. Chicago: University of Chicago Press.

—. 2006. "Gesture and Communication." In *Encyclopedia of Language and Linguistics, 2nd Edition*, edited by Keith Brown, 58–66. Elsevier Science.

McNeill, David, and Elena Levy. 1982. "Conceptual Representations in Language Activity and Gesture." In *Speech, Place, and Action: Studies in Deixis and Related Topics*, edited by Robert J. Jarvella and Wolfgang Klein, 271–96. New York: John Wiley & Sons.

McNeill, David, Elena Levy, and L. L. *Pedelty.* 1990. "*Gestures* and Speech." In *Advances in Psychology 70: Cerebral Control of Speech and Limb Movements*, edited by Geoffrey R. Hammond, 203–56. Amsterdam: Elsevier/North Holland Publishers.

Meier, Richard P. 2002. "Why Different, Why the Same? Explaining Effects and Non-Effects of Modality upon Linguistic Structure in Sign and Speech." In *Modality and Structure in Signed and Spoken Languages*, edited by Richard P. Meier, Kearsy Cormier and David Quinto-Pozos, 1–27. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486777.001.

Meir, Irit. 2012. "Word Classes and Word Formation." In *Sign Language: An International Handbook*, edited by Roland Pfau, Markus Steinbach and Bencie Woll, 77–111. https://doi.org/10.1515/9783110261325.77.

—. 2018. "Topic-Openendedness: Why Recursion is Overrated." In *Proceedings of the 12th International Conference on the Evolution of Language (Evolang12)*, edited by Christine Cuskley, Molly Flaherty, Hannah Little, Luke McCrohon, Andrea Ravignani, and Tessa Verhoef, 295–305. Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika. https://doi.org/10.12775/3991-1.073.

Neidle, Carol, Judy Kegl, Dawn MacLaughlin, Benjamin Bahan, and Robert G. Lee, eds. 2000. *The Syntax of American Sign Language: Functional Categories and Hierarchical Structure*. Cambridge, MA: MIT Press.

Neidle, Carol, and Joan Nash. 2012. "The Noun Phrase." In *Sign Language: An International Handbook*, edited by Roland Pfau, Markus Steinbach and Bencie Woll, 265–91. https://doi.org/10.1515/9783110261325.265.

Nespor, Marina, and Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.

Özyürek, Aslı. 2012. 27. "Gesture." In *Sign Language: An International Handbook*, edited by Roland Pfau, Markus Steinbach and Bencie Woll, 626–46. https://doi.org/10.1515/9783110261325.626.

Özyürek, Aslı, and Bencie Woll. (in press). "Language in the Visual Modality: Co-Speech Gesture and Sign." In *Human Language: From Genes and Brain to Behavior*, edited by P. Hagoort, 67–83. Cambridge, MA: MIT Press.

Perlmutter, David. 1992. "Sonority and Syllable Structure in American Sign Language." *Linguistic Inquiry* 23 (3): 407–42. https://doi.org/10.1016/B978-0-12-193270-1.50016-9.

Pfau, Roland, Markus Steinbach, and Bencie Woll, eds. 2012. *Sign Language: An International Handbook*. Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110261325.

Pribanić, Ljubica, and Marina Milković. 2012. Što znamo nakon pet godina istraživanja gramatike hzj? *Zbornik radova sa stručnih skupova.* Zagreb: Hrvatski savez gluhih i nagluhih.

Quer, Josep. 2012. "Negation." In *Sign Language: An International Handbook*, edited by Roland Pfau, Markus Steinbach and Bencie Woll, 316–39. https://doi.org/10.1515/9783110261325.316.

Sandler, Wendy. 2012. "The Phonological Organization of Sign Languages." *Language and Linguistics Compass* 6 (3): 1–21. https://doi.org/10.1002/lnc3.326.

—. 2014. "The Emergence of the Phonetic and Phonological Features in Sign Language." *Nordlyd* 41 (1): 183–212.

Schembri, Adam, Caroline Jones, and Denis Burnham. 2005. "Comparing Action Gestures and Classifier Verbs of Motion: Evidence from Australian Sign Language, Taiwan Sign Language, and Nonsigners' Gestures Without Speech." *Journal of Deaf Studies and Deaf Education* 10 (3): 272–90. https://doi.org/10.1093/deafed/eni029.

Steinbach, Markus. 2012. "Plurality." In *Sign Language: An International Handbook*, edited by Roland Pfau, Markus Steinbach and Bencie Woll, 112–36. https://doi.org/10.1515/9783110261325.112.

Stokoe, William C., Jr. 2005. "Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf." *Journal of Deaf Studies and Deaf Education* 10 (1): 3–37. https://doi.org/10.1093/deafed/eni001.

Šarac Kuhn, Ninoslava, Tamara Alibašić Ciciliani, and Ronnie B. Wilbur. 2006. "Phonological Parametres in Croatian Sign Language." *Sign Language & Linguistics* 9 (1/2): 33–70. https://doi.org/10.1075/sll.9.1.05kuh.

Tang, Gladys, and Prudence Lau. 2012. "Coordination and Subordination." In *Sign Language: An International Handbook*, edited by Roland Pfau, Markus Steinbach and Bencie Woll, 340–65. https://doi.org/10.1515/9783110261325.340.

Tkachman, Oksana, and Irit Meir. 2018. "Novel Compounding and the Emergence of Structure in Two Young Sign Languages." *Glossa: A Journal of General Linguistics* 3 (1): 136. https://doi.org/10.5334/gjgl.632.

Traugott, Elizabeth Closs. 2008. "Grammaticalization, Constructions and the Incremental Development of Language: Suggestions from the Development of Degree Modifiers in English." In *Trends in Linguistics. Variation, Selection, Development. Probing the Evolutionary Model of Language* Change, edited by Regine Eckardt, Gergard Jäger and Tonjes Veenstra, 219–52. Berlin: Mouton De Gruyter.

Valli, Clayton, and Ceil Lucas. 2000. *Linguistics of American Sign Language: An Introduction*. 3rd edition. Washington, D.C.: Gallaudet University Press.

Zwitserlood, Inge. 2012. "Classifiers." In *Sign Language: An International Handbook*, edited by Roland Pfau, Markus Steinbach and Bencie Woll, 158–86. https://doi.org/10.1515/9783110261325.158.

ELOPE

Serkan Şen
Başkent University, Ankara, Turkey

2019, Vol. 16 (1), 145–157(164)
revije.ff.uni-lj.si/elope
doi: 10.4312/elope.16.1.145–157
UDC: 811.111'243'342.41:811.512.161

Helping Turkish EFL Learners with the Pronunciation of Four General British Vowels

ABSTRACT

The paper presents a contrastive analysis of Turkish and English with respect to their vowel systems and orthography. The author then focuses on effective ways to teach four General British (GB) vowels, /i:/, /ɪ/, /ʊ/, and /u:/, to twenty Turkish university students of English as a Foreign Language. The paper presents a practical approach to teaching the perception and production of the four vowels through a set of classroom activities. As practice shows, Turkish students experience difficulties in the recognition and pronunciation of the observed vowels, which can be mitigated using the described activities.

Keywords: English pronunciation; vowel system; orthography; Turkish; General British vowels

Kako pomagati turškim študentom angleščine pri izgovarjavi štirih britanskih samoglasnikov

POVZETEK

Prispevek predstavi kontrastivno analizo turščine in angleščine z vidika samoglasniških sistemov in ortografije. Avtor se nato osredini na učinkovite načine, s katerimi so štiri britanske samoglasnike, /i:/, /ɪ/, /ʊ/ in /u:/, približali dvajsetim turškim študentom angleščine kot tujega jezika. Prispevek predstavi praktičen pristop k poučevanju percepcije in produkcije opazovanih samoglasnikov s pomočjo več aktivnosti. Kot pokaže praksa, imajo turški študenti težave s prepoznavanjem in izgovarjavo opazovanih samoglasnikov, ki jih lahko izboljšamo z uporabo opisanih aktivnosti.

Ključne besede: angleška izgovarjava; samoglasniški sistem; ortografija; turščina; standardni angleški samoglasniki

1 Introduction and Literature Overview

1.1 Teaching Pronunciation

Jim Scrivener suggests that although grammar is taught through language skills such as reading, writing, and speaking, there is more to language teaching than this (2011, 24). He analyses a sentence such as 'Pass me the book' under five main categories:

Table 1. Analysis of a language sample.

Phonological	The sounds	/pa:s mi: ðə ˈbʊk/ The stress is probably on *book*, but also possible (with different meanings) on *pass* or *me*. The word *me* has a weak vowel sound.
Lexical	The meaning of the individual words or groups of words	pass=give; hand over; present me=reference to speaker the book=object made of paper, containing words and/or pictures and conveying information
Grammatical	How the words interact with each other within the sentence	Verb (imperative) + first person object pronoun + definite article + noun
Functional	The use to which the words are put in particular situations	A request or order
Discoursal	The way that communication makes sense beyond the individual phrase or sentence, analysing how the sentences relate (or do not relate) to each other	Although not a direct transparent answer to the request, we can still draw a meaning from this reply. The word *it*, referring to the book, helps us to make a connection to the request. Assuming that Mary has put it in her bag is intended as a genuine response to the request; it may suggest a reason why the book cannot be passed. In order to fully understand the meaning, we would need to know more about the situational context and more about the surrounding conversation.

(Scrivener 2011, 25)

As the table above suggests, language teaching starts with the study of sounds. Therefore, language teachers should start teaching language with phonology or, specifically, with pronunciation. To put it differently, pronunciation plays a crucial role in language teaching to convey the meaning besides grammar and lexis. This is mainly because successful communication requires good pronunciation. As Gerald Kelly states,

"a learner who consistently mispronounces a range of phonemes can be extremely difficult for a speaker from another language community to understand" (2000, 11). For example, when a learner says *chip* in a situation such as a market where they should have said *cheap*, the wrong articulation of a phoneme can lead to a misunderstanding.

Although crucial, there is a paradox in teaching pronunciation in the language classrooms. While teachers may not be interested in the subject, language learners are usually interested in improving their pronunciation skills because they are aware of the fact that pronunciation is important for them to communicate better. In order to provide the learners with pronunciation practice they need, teachers need to improve their teaching skills. As Kelly states "many experienced teachers would admit to a lack of knowledge of the theory of pronunciation and they may therefore feel the need to improve their practical skills in pronunciation teaching" (2000, 13). He further argues that "teachers of pronunciation need:

- a good grounding in theoretical knowledge
- practical classroom skills
- access to good ideas for classroom activities" (2000, 13).

Language teachers should have enough technical knowledge and there should be a shift from reactive to proactive teaching of pronunciation. Instead of only responding to errors that students make in the classroom, pronunciation instruction should be carefully planned just like grammar and lexis lessons. One way to achieve this is to focus on a specific area of pronunciation and make a careful lesson plan in which "a particular feature of pronunciation is isolated and practised for its own sake, forming the main focus of a lesson period" (Kelly 2000, 14).

Another reason why teachers do not integrate pronunciation practice into their classes is that they do not have enough knowledge about pronunciation. As Scrivener contends, "pronunciation can be an overlooked area of language teaching, partly because teachers themselves may feel more uncertain about it than about grammar or lexis, worried that they do not have enough technical knowledge to help students appropriately" (2011, 271). Besides, there is a misconception that in time learners can master the pronunciation by themselves and that pronunciation does not require the same attention as grammar or lexis. However, if there is a careful lesson planning, the results may be quite rewarding.

In countries like Turkey where English is learnt as a foreign language, teaching pronunciation is not easy. Listening and speaking skills are not included in the nation-wide foreign language exam, YDS. Hence, pronunciation is neglected in favour of grammar and vocabulary, and learners are not given the opportunity to test their proficiency in listening and speaking.

1.2 Turkish and General British Vowel Systems

As Cruttenden states, when compared to other languages, including Greek, Hindi, Japanese, and Spanish, all of which have just five vowels, English is more problematic because "[i]n whatever way the vowels of English are counted (i.e. even counting some or all of the diphthongs as sequences of short vowel plus semivowel (= consonant)), the English system is one of the less common and more complex types" (2014, 109). Not surprisingly, most foreign learners have difficulty in attaining the vowel system of English, especially with /i: i/, /e a ʌ/, /ɒ a: ɔ:/, /u: ʊ/. Turkish learners also have difficulty with the production of these vowels because of the difference between Turkish and English vowel systems.

In their analysis of the Turkish language system, Aslı Göksel and Celia Kerslake argue that "Turkish has a very symmetrical vocalic system, consisting of the eight vowels /a/, /e/, ('ı'), /i/, /o/, /œ/ ('ö'), /u/, /y/ ('ü') […]. These vowels differ from each other in terms of the height of the tongue, the roundedness of the lips and the frontness of the tongue", as can be seen in Table 2 below (2005, 9).

Table 2. The eight vowels of Standard Turkish.

	High		**Non-high (mid and low)**	
	Rounded	Unrounded	Rounded	Unrounded
Front	ü (y)	i	ö (œ)	e
Back	u	ı (ɯ)	o	a

(Göksel and Kerslake 2005, 9)

In contrast to Turkish, English has forty-four phonemes consisting of twenty vowels and twenty-four consonants, which are further divided into two categories as voiced and voiceless (Cruttenden 2014). Vowel sounds are all voiced, and they are either single vowels – twelve monophthongs – or a combination of two vowels – eight diphthongs. The twelve monophthongs are outlined in Table 3 below.

Table 3. Twelve[1] monophthongs of GB.

Tongue / Jaw	**Front**	**Centre**		**Back**
High / Close	i:	ɪ	ʊ	u:
Centre / Close-Mid and Open-Mid	e	ə	ɜ:	ɔ:

[1] The latest edition of *Gimson's Pronunciation of English* (Cruttenden 2014) adds /ɛ:/ to the list as the thirteenth GB monophthong and a substitute for the diphthong /eə/ (appearing in words such as *square*).

Low / Open	æ	ʌ	ɑ:	ɒ
Lips	Spread	Neutral		Rounded

In terms of their vowel systems, Turkish and English differ substantially. Twenty-nine phonemes in Turkish are represented by twenty-nine letters, while forty-four English phonemes are represented by only twenty-six letters. As Khalilzadeh states, "[e]very Turkish vowel has only one pronunciation, e.g. the vowel /u/ is pronounced with the features: (+high, + round, +back), none of the features changes because of the phonetic environment", while in English "the same vowel may sound differently in different words, e.g. the pronunciation of /u/ in the words: *bus, surrender, busy, occur, pure*" (2014, 12). Therefore, Turkish learners tend to pronounce English words as combinations of (Turkish) sounds represented by the letters in their spelling. From a related perspective, Kelly asserts that "the lack of a simple correspondence between the spelling system and the pronunciation system in English tends to cause problems for learners in that it can lead them to initially or repeatedly misspell words and mispronounce them" (2000, 125). When considering the relationship between English and Turkish, Rogerson-Revell explains that "although Turkish uses essentially the same alphabet as English, its orthographic system, which employs to a large extent one-to-one letter-sound correspondence, can cause interference with English pronunciation" (2011, 289).

This difficulty is greater for the learners of phonetic languages in which spelling and pronunciation are similar, such as Italian, Spanish and Turkish (see Marks and Bowen (2012, 33) for a list of some common sound-spelling relationships). Komar's study (2017) of Slovene also shows that orthography can influence the production of GB vowels by foreign language learners.

However, with regard to Turkish, not only the pronunciation but also the number of phonemes is different. Khalilzadeh (2014, 11) highlights this difference by saying that

> The main cause of pronunciation problems of Turks in English is the differences between the vowel systems of the two languages, both due to the difference in number and the way the vowels are pronounced. In Turkish, there is only one vowel type, i.e. monophthong, while there are three types of vowels in English: monophthong, diphthong and triphthong. The number of vowels (monophthongs) in Turkish is 8, while there are 12 vowels (monophthongs) in English. In English, there are long, mid and short vowels whereas all Turkish vowels are short.

Therefore, Turkish learners may have problems in pronouncing many words in English that include vowels which are non-existent in Turkish. The difference in vowel systems of Turkish and English leads to a problem in the pronunciation of four

monophthongs. Lewis states that although there are some exceptions, "[a]s a rule, Turkish vowels are short; i, for example, ordinarily has the sound heard in English *bit* (/i/), not that heard in *machine* (/i:/)" (1992, 5). Therefore, Turkish learners of English tend to pronounce these vowels with difficulty. For example, they may confuse *sheep* /ʃi:p/ and *ship* /ʃɪp/ or *pool* /pu:l/ and *pull* /pʊl/. Joanne Kenworthy similarly asserts that "the /ɪ/ – /i:/ contrast ('*pit – Pete*') is troublesome. [Turkish] learners will use /i:/ for both vowels […]. The /ʊ/ –/u:/ ('*full – fool*') contrast is [also] problematic [for Turkish learners of English because they] will tend to use a vowel which is close to /uː/ for both vowels" (1987, 158). However, in most cases, Turkish learners of English will pronounce long vowels /i:/ and /u:/ as /ɪ/ and /ʊ/. From a related perspective, Ian Thompson identifies two problematic vowel sounds of English for Turkish learners, i. e. /i:/ and /u:/ and makes a comparative analysis of these errors:

> /i:/ as in *key* is often pronounced like the diphthong /ɪə/, or in a closed syllable as /ɪ/ – the Italian error in reverse: /kɪə/ for *key*; *kip* for *keep*. The Turkish word *giy* contains a good approximation to English /i:/". [Moreover], /u:/ tends to become /ʊə/ when final and /ʊ/ in closed syllables: /dʊə/ for *do*; '*pullink*' for both *pooling* and *pulling*. Turkish speakers are able to pronounce the sound successfully after /j/, as in *few*. (2001, 215)

In the same vein, Pamela Rogerson-Revell states that "[learners have] difficulty distinguishing tense versus lax vowels (e.g. '*bit*' vs '*beat*') and [there is also] /u:/ and /ʊ/ confusion" (2011, 289). Such perception issues have been observed in other languages as well, and they point to the close relationship between perception and production of foreign language phonemes (e.g. Stopar 2019). Kelly highlights the pronunciation difficulties related to the above vowels which speakers of some languages, including Turkish, may face when speaking English. They are summarised in Table 4.

Table 4. Common pronunciation difficulties.

Potential Difficulty		**For speakers of:**												
Sound	Can be pronounced	A	C	F	G	Gk	Ind	It	J	P	R	Sc	Sp	Tu
/i:/	/ɪ/ /hɪt/ for heat		*			*				*	*			*
/ɪ/	/e/ /bed/ for bid	*												
	/i:/ /hi:t/ for hit		*	*		*		*	*	*		*	*	*
/ʊ/	/u:/ /pu:l/ for pull			*		*		*		*			*	*
/u:/	/ʊ/ /sʊt/ for suit		*							*				

Key to Languages: A=Arabic; C=Chinese; F=French; G=German; Gk=Greek; Ind=Indian; It=Italian; J=Japanese; P=Portuguese; R=Russian; Sc=Scandinavian; Sp=Spanish; Tu=Turkish (Kelly 2000, 145)

Turkish learners of English also have difficulties with the pronunciation of borrowed words. They tend to hear and pronounce the sounds of English according to the sounds of Turkish. Wells explains that "when we encounter a foreign language, our natural tendency is to hear it in terms of the sounds of our own language. We actually perceive it rather differently from the way native speakers do" (1999, 118). This is especially true for the pronunciation of borrowed or loanwords – words taken from one language and used in another. Turkish has borrowed many words from English and modified them, so they fit its own sound system. As a rule, Turkish vowels are short. The word *illegal* /ɪˈliːgəl/ is written as *illegal* but pronounced as /ɪllegʌl/ in Turkish. Similarly, English has borrowed the word *kebap* /kebʌp/ from Turkish. However, the word is written as *kebab* and pronounced as /kɪˈbæb/ in English. In this context, Beel and Felder contend that "the significant number of consonants and vowels which occur in English […] do not occur in Turkish. Consequently, as English loanwords are adopted into the Turkish vocabulary, phonological adjustments must be made in order to compensate for the additional English phone[me]s that are foreign to the native speaker of Turkish" (2013, n.p.).

This paper focuses on four GB vowels that were explored in an EFL classroom setting. They are as follows:

> Close Front Vowel /iː/: Cruttenden explains that in articulating the long GB vowel /iː/ "the front of the tongue is raised to a height slightly below and behind the front close position; the lips are spread; the tongue is tense, with the side rims making a firm contact with the upper molars". He further states that "[/iː/] is often noticeably diphthongised, especially in final positions" (2014, 111). The /iː/ vowel is found in many words such as *bead, cheese, keep, key, pea, people, piece, police, scene, sea, seat, quay*. Cruttenden argues that this vowel is not difficult for foreign learners, but he suggests that they should avoid exaggeration while producing the vowel as it may sound artificial, and try not to confuse it with /ɪ/.
>
> Close-mid Vowel /ɪ/: Cruttenden explains that the vowel /ɪ/ "is pronounced with a part of the tongue nearer to centre than to front raised just above the close-mid position; the lips are loosely spread; the tongue is lax (compared with the tension for /iː/), with the side rims making a light contact with the upper molars" (2014, 113). The short GB vowel /ɪ/ is used in many words such as *big, busy, hit, it, minute, mountain, rhythm, sausage, sick, sieve, wishes, women*. For foreign learners, Cruttenden suggests that they should be careful with the difference between /ɪ/ and /iː/.
>
> Close-mid Vowel /ʊ/: According to Cruttenden (2014, 130) the short GB vowel /ʊ/ "is pronounced with a part of the tongue nearer to centre than to back raised just above the close-mid position; it has, therefore, a symmetrical back relationship with the front vowel /ɪ/; the tongue is laxly held (compared with the tenser /uː/), no firm contact being made between the tongue and the upper molars". The most common examples for this vowel are *book, foot, full, good, look, pull, push, put, should, wolf,*

wo̲man, wo̲u̲ld. Just as the vowels / ɪ / and /i:/ tend to be confused by some learners, so can /ʊ/ and /u:/.

Close Back Vowel /u:/: In his description of the long GB vowel /u:/, Cruttenden explains that this vowel is "a close back vowel with varying degrees of centralisation and unrounding. Two types occur within GB: (i) a more centralised monophthongal vowel [ü:] or, with unrounding, [ɯ:]; and (ii) a short diphthong [ʊu] or, with unrounding, [ʉɯ] (these being particularly common in final position, e.g. in *do, shoe, who)*" (2014, 133). The most common examples are *bo̲o̲t, fo̲o̲d, fo̲o̲l, fru̲it, lo̲se, ne̲w, ru̲de, sho̲e, sho̲o̲t, so̲u̲p, tru̲e, who̲*. In his advice to foreign learners, Cruttenden suggests that "[t]he quality of this vowel should cause no difficulty to most learners, many of whom will have a close back rounded vowel in their own language" (2014, 134).

This paper aims to present how to teach the above GB vowels to Turkish learners by raising their awareness of the differences between Turkish and English in terms of their phonetic structures. To achieve this goal, a set of classroom activities was devised that aimed to improve the pronunciation of the four targeted GB vowels. The activities aim to address some specific problems that can be identified in Turkish language classrooms, including a lack of effective pronunciation activities, insufficient focus on the vowel systems of Turkish and English, and the influence of loanwords on foreign language pronunciation.

2 Classroom Activities

2.1 Students and Their Background

The activities described in the sections below were used in a class of twenty students at the Faculty of Health Sciences (Başkent University, Ankara, Turkey) with a similar background in foreign language instruction: in order to enrol in their BA programme, the students had to pass the national university entrance exam. Their ages ranged from 18 to 22; 17 were female and three were male. They were taking English classes four hours a week. They had not been familiarised with the international phonetic alphabet (IPA) and its symbols for GB phonemes before taking the course, which is representative of the inefficiency of pronunciation activities carried out in Turkish education system.

2.2 Procedures and Activities

The teacher relied on visual, kinaesthetic and auditory techniques that had the aim of raising the students' awareness of the importance of pronunciation and correcting their pronunciation of the targeted vowels.

During the recognition phase, the teacher first presented a phonemic chart. Starting from the layout of the chart, the teacher modelled the monophthongs focusing on four high vowels /i:/, /ɪ/, /ʊ/, /u:/, and asked students to listen to and distinguish

between the sounds. The teacher also relied on affective characteristics such as enthusiasm, encouragement and humour. Kelly's (2000, 38) techniques were used to make the teaching of vowel sounds more memorable and enjoyable for learners.

To address the issues related to the differentiation of long and short vowels, the teacher used the similarity between the /ʊ/ and 'gorilla sound' (U-U-U-U) and the /u:/ and the 'gossip sound' in Turkish (uuuUUUuuu); see Figure 1 below.

Sound	**Suggestion**
/i:/ – sheep	A 'smiling' sound. Smile widely, make and hold the sound. Demonstrate that it is a 'long' sound.
/ɪ/ – ship	Make the sound, make it obviously short. If necessary, contrast it with /i:/.
/ʊ/ – book	A short sound. Exaggerate the forward position of your lips. One way into this sound is to ask students what noise a gorilla makes.
/u:/ – boot	Make and hold the sound. Use a rising then falling intonation, as if you have heard something surprising, or some interesting gossip (uuUUuu). Demonstrate that it is a 'long' sound.

Figure 1. Sound and Suggestion activity (Kelly 2000, 38).

After the recognition phase, the teacher pointed to one of the four vowels and asked students to produce the sound in groups, pairs or individually. The teacher also asked individual students to do the same for their classmates.

At this stage rhyming sentences can also be used to make the sounds easier to remember. Some such examples can be found in Hancock's work (2003, 20, 22, 44):

/i:/: Steve keeps the cheese in the freezer

/ɪ/: Alex's lettuces tasted like cabbages

/ʊ/: That cook couldn't cook if he didn't look at a cook book

/u:/: Sue knew too few new tunes on the flute

While working with the described groups, the teacher also focused on the perception of the targeted sounds by using a minimal pair activity, which can be effective when practising new or problematic sounds. The teacher read one word from each pair and learners chose the word they thought they had heard. Some minimal pairs for the relevant vowels are presented in Figure 2 below.

Since simply saying the correct answers aloud might feel monotonous, the teacher gave the students small cards in two different colours, and they were asked to raise the one representing the uttered phoneme.

Vowels	Minimal pairs
/ɪ/ – /iː/	ship – sheep; bin – bean; chip – cheap; it – eat; sit – seat; live – leave; fill – feel; fit – feet; hit – heat
/ʊ/ – /uː/	look – Luke; pull – pool; full – fool; could – cooed; would – wooed; foot – food; soot – suit;

Figure 2. Minimal Pairs activity.

Similar to the use of cards to mark different sounds, a sound maze game (Figure 3) was also used to appeal to the different learning styles of the students. In this case, the activity focused on the /iː/ sound and involved group work. The teacher divided the class into groups; each group included both weak and strong learners, so that weaker students had the opportunity to get information from the stronger ones and improve their knowledge of pronunciation. The aim of the activity is to travel through the maze by identifying the correct vowel in the listed words.

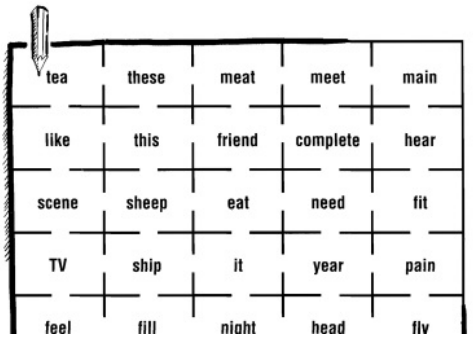


Figure 3. The Sound Maze activity (a fragment taken from Hancock's photocopiable (1995, 56)).

An activity similar to the one above can also be devised using short texts. In our case, the students in the participating class were presented with one of Hancock's task sheets (see Figure 4) and asked to do the task individually. The students then compared their answers with those of their partners, while the teacher gave them feedback.

Listen to your teacher. Which underlined words have the /uː/ sound?
I studied English at school in London last summer. I was there for two months: May and June. English is famous for bad food and weather, but I thought the food was good. The pub lunches were very nice. But it's true about the weather. Too much rain for me! (Hancock 2003, 45)

Figure 4. Find the /u:/ Words activity.

Finally, the teacher used drilling and repetition activities in the class in order to help students with the pronunciation of loanwords from English. According to Marks and Bowen (2012, 64), using loanwords or "bilingual minimal pairs" is a good way to discover the differences between L1 and L2. Accordingly, the teacher showed some bilingual English-Turkish pairs to students and pronounced them. Learners tried to identify whether the teacher was pronouncing the word in their mother tongue or in English by using the previously mentioned cards in two different colours, which brought along a dynamic variety. The pairs are presented in Figure 5 below.

English	**Turkish**
cheesecake /ˈtʃiːzkeɪk/	çizkek /ˈtʃɪzkek/
obese /əʊˈbiːs/	obez /ɒˈbez/
obesity /əʊˈbiːsəti/	obezite /ɒˈbezite/
deep /diːp/	dip /dɪp/
illegal /ɪˈliːgəl/	illegal /ˈillegʌl/
police /pəˈliːs/	polis /ˈpɒlɪs/
princess /prɪnˈses/	prenses /ˈpɪrenses/
rocket /ˈrɒkɪt/	roket /ˈrɒket/
prince/prɪns/	prens /ˈpɪrens/
unit /ˈjuːnɪt/	ünite /ˈynɪte/
kebab /kɪˈbæb/	kebap /keˈbʌp/
studio /ˈstjuːdiəʊ/	stüdyo /ˈsytydyɒ/
graffiti /ɡrəˈfiːti/	graffiti /ˈgraffiti/
shoot /ʃuːt/	şut /ʃʊt/
feedback /ˈfiːdbæk/	fidbek /ˈfɪdbek/

Figure 5. Bilingual Pairs activity.

3 Discussion and Conclusions

The activities and procedures presented in this work have the basic aim of raising students' awareness of pronunciation and improving both their perception and production of FL sounds. With this in mind, a set of pronunciation teaching activities was devised and used to focus on teaching four GB vowels – /iː/, /ɪ/, /ʊ/, /uː/ – to a group of Turkish university students of English.

The above vowels were identified as problematic based on a prior analysis and comparison of English and Turkish vowel systems. It was established that the sets of phonemes in the two languages differ in quality and quantity, and that the one-to-one letter-sound correspondence observed for Turkish may confuse Turkish learners of English as a foreign language. The theoretical assumptions were confirmed in practice by working on both the production and perception of the vowels.

Throughout the course, it was observed that Turkish university students were generally not aware of the difference between Turkish and English in terms of their phonetic structures. Although most of the students were familiar with learning vocabulary and grammar, they were not used to instruction in phonetics and phonology since pronunciation is often neglected in the classroom in favour of grammar- and vocabulary-centred activities.

In the course of completing the activities described in the previous section, it was confirmed that Turkish students tend to confuse the pairs of targeted vowels and frequently use /ɪ/ instead of /iː/, and /ʊ/ instead of /uː/. Evidently, one of the most important reasons for such confusion is the lack of /iː/ and /uː/ in the Turkish vowel system, but it should be noted again that the students had not had any targeted instruction in pronunciation before taking this course.

Considering the above factors, the teacher also focused on several affective strategies. It was concluded that using anecdotes, analogies from daily experiences, and videos helped the Turkish university students in becoming more involved in the learning process. Consequently, they became more aware of the characteristics of the GB vowel system, which also paved the way for their better pronunciation of English sounds.

This paper offers some practical solutions to specific problems related to pronunciation in the hope that they may be of help to other teachers of English as a foreign language. There is no need to shy away from teaching pronunciation. As Hancock points out:

> "teaching pronunciation can be interesting, playful and a real joy. You can use games, puzzles, rhymes and raps, drama and pair-works. It can be challenging, but it doesn't have to be frightening. You may find that pronunciation becomes the part of your lessons that the learners look forward to most." (2019, n.p).

The key is to plan carefully and have an enjoyable lesson for both teachers and students alike.

References

Beel, Rachel, and Jennifer Felder. 2013. "Phonological Adaptations of English Loanwords in Turkish." *Other Undergraduate Scholarship* 2. https://digitalcommons.liberty.edu/eml_undergrad_schol/2/.

Cruttenden, Alan. 2014. *Gimson's Pronunciation of English, Eighth Edition.* London and New York: Routledge.

Göksel, Aslı, and Celia Kerslake. 2005. *Turkish: A Comprehensive Grammar*. London and New York: Routledge.

Hancock, Mark. 1995. *Pronunciation Games.* Cambridge: Cambridge University Press.

___. 2019. "Top Ten Tips for Teaching Pronunciation." http://hancockmcdonald.com/ideas/top-ten-tips-teaching-pronunciation.

Kelly, Gerald. 2000. *How to Teach Pronunciation*. Pearson Education Limited.

Kenworthy, Joanne. 1987. *Teaching English Pronunciation*. London: Longman.

Khalilzadeh, Amir. 2014. "Phonetic and Non-phonetic Languages: A Contrastive Study of English and Turkish Phonology Focusing on the Orthography-induced Pronunciation Problems of Turkish Learners of English as a Foreign Language." *International Journal of Languages' Education and Teaching* 2 (1): 1–16.

Komar, Smiljana. 2017. "The Relationship between the Perception and Production of Four General British Vowels by Slovene University Students of English." *Linguistica* 57 (1): 161–70. https://doi.org/10.4312/linguistica.57.1.161-170.

Lewis, Geoffrey Lewis. 1992. *Teach Yourself Turkish*. Hodder & Stoughton Educational Division.

Marks, Jonathan, and Tim Bowen. 2012. *The Book of Pronunciation: Proposals for a Practical Pedagogy*. Delta Publishing.

Rogerson-Revell, Pamela. 2011. *English Phonology and Pronunciation Teaching*. Continuum International Publishing Group.

Scrivener, Jim. 2011. *Learning Teaching*. Macmillan Education.

Stopar, Andrej. 2019 [forthcoming]. "Perception of General British Monophthongs: The Case of Slovene Students of English as a Foreign Language." *Jezikoslovlje*.

Thompson, Ian. 2001. *Learner English: A Teacher's Guide to Interference and Other Problems*, edited by Michael Swan and Bernard Smith, 214–26. Cambridge: Cambridge University Press.

Wells, John C. 1999. "Overcoming Phonetic Interference." *Speech, Hearing and Language: Work in Progress* 11: 117–28. https://www.phon.ucl.ac.uk/home/shl11/pdf_files/wells(a).pdf

List of Contributors

LIST OF CONTRIBUTORS

Biljana Čubrović
University of Belgrade, Faculty of Philology
Studentski trg 3, Beograd 11000, Serbia
biljana.cubrovic@fil.bg.ac.rs

Ivo Fabijanić
University of Zadar, Department of English Studies
Obala Kralja Petra Krešimira IV. br. 2,
HR-23 000 Zadar, Croatia
ifabijan@unizd.hr

Nataša Hirci
University of Ljubljana, Faculty of Arts
Aškerčeva cesta 2, 1000 Ljubljana, Slovenia
natasa.hirci@ff.uni-lj.si

Saša Jošt
University of Ljubljana, Faculty of Arts
Aškerčeva cesta 2, 1000 Ljubljana, Slovenia
sasa.jost7632@gmail.com

Oleksandr Kapranov
Western Norway University of Applied Sciences
Høgskulen på Vestlandet, Postbox 7030,
5020 Bergen, Norway
oleksandr.kapranov@hvl.no

Smiljana Komar
University of Ljubljana, Faculty of Arts
Aškerčeva cesta 2, 1000 Ljubljana, Slovenia
smiljana.komar@ff.uni-lj.si

Frane Malenica
University of Zadar, Department of English Studies
Obala Kralja Petra Krešimira IV. br. 2,
HR-23 000 Zadar, Croatia
fmalenica@unizd.hr

Emilija Mustapić
University of Zadar, Department of English Studies
Obala Kralja Petra Krešimira IV. br. 2,
HR-23 000 Zadar, Croatia
emustapic@unizd.hr

Andrej Stopar
University of Ljubljana, Faculty of Arts
Aškerčeva cesta 2, 1000 Ljubljana, Slovenia
andrej.stopar@ff.uni-lj.si

Serkan Şen
Başkent University
Fatih Sultan Mahallesi Eskişehir Yolu 18. km
06790 Etimesgut/Ankara, Turkey
sens@baskent.edu.tr

GUIDELINES FOR CONTRIBUTORS

ELOPE English Language Overseas Perspectives and Enquiries
ELOPE publishes original research articles, studies and essays that address matters pertaining to the English language, literature, teaching and translation.

Submission of Manuscripts
Manuscripts should be submitted for blind review in electronic form using the Faculty of Arts (University of Ljubljana) OJS platform (https://revije.ff.uni-lj.si/elope/about/submissions). Only one contribution by the same author per volume will be considered. Each paper should be accompanied by abstracts in English and Slovene (between 150 and 180 words each) and keywords. Abstracts by non-native speakers of Slovene will be translated into Slovene by ELOPE. Please be sure to have a qualified native speaker proofread your English-language article. Suggested length of manuscripts is between 5,000 and 8,000 words.

Manuscript Style and Format
The manuscript should be in the following format:

- title in English,
- abstracts in English and Slovene,
- the text should be divided into introduction; body of the paper (possibly subdivided); and conclusion.

The text should be preferably written in Word format (OpenOffice and RTF files are also acceptable). Please observe the following:

- 12-point Times New Roman font size,
- 2.5 cm page margins on all sides,
- 1.5-point line spacing,
- left text alignment,
- footnotes should be brief (up to 300 words per page; 10-point Times New Roman font size).

For resolving practical style and formatting queries, please see the articles in the latest on-line issue (https://revije.ff.uni-lj.si/elope/issue/current) or contact the technical editor.

References
References should comply with *The Chicago Manual of Style* (17th edition, 2017) author-date system. A Quick Guide to CMS is available here: http://www.chicagomanualofstyle.org/tools_citationguide/citation-guide-2.html.

Final note
Please note that only manuscripts fully adhering to the ELOPE Guidelines for Contributors will be considered for publication.

ELOPE
English Language Overseas Perspectives and Enquiries

ELOPE Vol. 16, No. 1 (2019)

Editorial Policy

ELOPE: English Language Overseas Perspectives and Enquiries is a double-blind, peer-reviewed academic journal that publishes original research articles, studies and essays that address matters pertaining to the English language, literature, teaching and translation.

The journal promotes the discussion of linguistic and literary issues from theoretical and applied perspectives regardless of school of thought or methodology. Covering a wide range of issues and concerns, ELOPE aims to investigate and highlight the themes explored by contemporary scholars in the diverse fields of English studies.

ELOPE
English
Language
Overseas
Perspectives and
Enquiries

Published by
Ljubljana University Press, Faculty of Arts
Znanstvena založba Filozofske fakultete Univerze v Ljubljani
Aškerčeva 2, 1000 Ljubljana, Slovenia
For the Publisher: Roman Kuhar, the Dean of the Faculty of Arts

Issued by
Slovene Association for the Study of English
Slovensko društvo za angleške študije
Aškerčeva 2, 1000 Ljubljana, Slovenia

Department of English, Faculty of Arts, University of Ljubljana
Oddelek za anglistiko in amerikanistiko, Filozofska fakulteta, Univerza v Ljubljani
Aškerčeva 2, 1000 Ljubljana, Slovenia

The journal is published with support from the Slovenian Research Agency.

The publication is free of charge.

Universal Decimal Classification (UDC)
Kristina Pegan Vičič

Journal Design
Gašper Mrak

Cover
Marjan Pogačnik: *Zimsko cvetje,* 1994
7.6 x 10.0 cm; colour etching, deep relief
Owner: National Gallery, Ljubljana,
Photo: Bojan Salaj, National Gallery, Ljubljana

Printed by
Birografika Bori

Number of Copies
110

University of Ljubljana
FACULTY OF ARTS

https://doi.org/10.4312/elope.16.1
Online ISSN: 2386-0316
Print ISSN: 1581-8918