

A Contrastive Acoustic-Phonetic Analysis of Slovenian and English Diphthongs

Robert Modic* and Bojan Petek

Interactive Systems Laboratory
Faculty of Natural Sciences and Engineering
University of Ljubljana
Snežniška 5, 1000 Ljubljana, Slovenia
robert.modic@guest.arnes.si bojan.petek@uni-lj.si

Abstract

This paper aims to narrow the gap between specific and general language theory by initiating contrastive acoustic-phonetic research of Slovenian and English diphthongs. Our ultimate goal is to investigate acoustic-phonetic similarities of the diphthongs across languages in the context of possible portability of resources between the languages. In general, the paper addresses the possibility of using language resources of well-resourced languages to efficiently bootstrap the human language technologies (HLT) of the under-resourced language. Therefore, as initial step we performed the contrastive analysis using English as an example "donor" language that is well researched with extensive language resources and Slovenian, the official and widely used language of Slovenia that is challenged by the significant lack of resources such as spoken, written or multimedia language corpora.

1. Introduction

The ultimate goal of this paper is to explore the possibility of using language resources of well-resourced languages to efficiently bootstrap the human language technologies (HLT) of an under-resourced language. For preliminary experiment we have chosen the English and Slovenian languages.

Slovenian is a Slavic language spoken by about 2 million people. It has more than 40 dialects divided into seven major dialect groups. Standard Slovenian (Toporišič, 2000) is official and widely used language in Slovenia yet in view of the mature HLT it is challenged by the lack of sufficient amount of the HLT resources.

On the other hand, English is typical example of the global language in view of the HLT since it is well researched and supported by extensive resources. It has many standard pronunciations: Received Pronunciation or RP English, general American, Scottish, and Australian (Gimson, 1994; Clark, 1995). It is one of the most prevalent languages of communication worldwide.

The paper presents a brief overview of the Standard Slovenian (SS) and the RP English (RPE) vowels since they represent an onset (SS1) and offset steady states (SS2) of the diphthongal vowel glides. Following this, an inventory of the SS and RPE diphthongs is presented. The candidate list of similar Slovenian and English diphthongs is then proposed and discussed. Next, instrumental analysis of live broadcast television speech is performed. The measurements obtained yielded that the SS1 and SS2 states of the SS and RPE cluster reasonably well. For observed deviations we tried to provide a linguistic explanation.

2. SS and RPE vowels

Vowel sounds described in this section are monophthongs (or pure vowels) since their quality remains relatively constant. The SS and the RPE have many common properties (Toporišič, 2000; Gimson, 1994). Vowels are in general voiced sounds produced without closure or narrowing typical for the consonantal

sounds. Table 1 gives an overview of the SS and RPE vowels as referenced in (Toporišič, 2000; Gimson, 1994), respectively. Their positions within Cardinal Vowel diagram (CVD) are shown in figure 1.

SS	i	e	ɛ	a	ə	ɒ	ʊ
RPE	i	e	ɛ	ɑ	ɒ	ɔ	u

Table 1: SS and RPE vowels

The SS includes vowel /ə/ which is different from the /ɑ/ found in the RPE. List of the presented RPE vowels comprises of primary Cardinal Vowels while their secondary realizations can also be obtained.

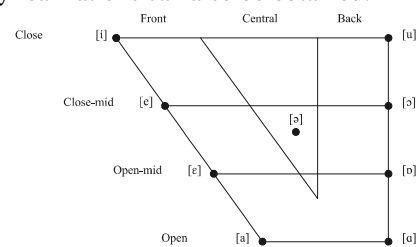


Figure 1: Positions of the SS and RPE vowels in the CVD

The vowel sounds and their allophonic realizations represent the SS1 and SS2 parts of the gliding vowels as described in the following section.

3. Diphthongs

In both SS and RPE often exist the gliding transition between the consecutive vowel sounds. These are so-called gliding vowels or diphthongs. Figure 2 shows examples of such diphthong realizations.

3.1. Diphthongs of SS

One of the first attempts to classify the SS diphthongs phonologically was made in (Petek and Šuštaršič, 1997; Šuštaršič, et.al., 1999). Research revealed that the diphthongs in SS are rather phonetic than phonemic since

* Young Investigator supported by the MŠZŠ of the Republic of Slovenia and Socrates/Erasmus exchange student under the multilateral agreement UL D-IV-1/99-JM/Kc.

they can be explained as a sequence of vowels followed by phonetic realization of labiodental approximant /v/ or the alveolar lateral /l/ and the semivowel /j/. Additionally, with two exceptions, they do not show specific articulatory features that would justify them to be autonomous sound units. Table 2 gives an inventory of the diphthongs in SS and their relative frequencies estimated using a word form lexicon.

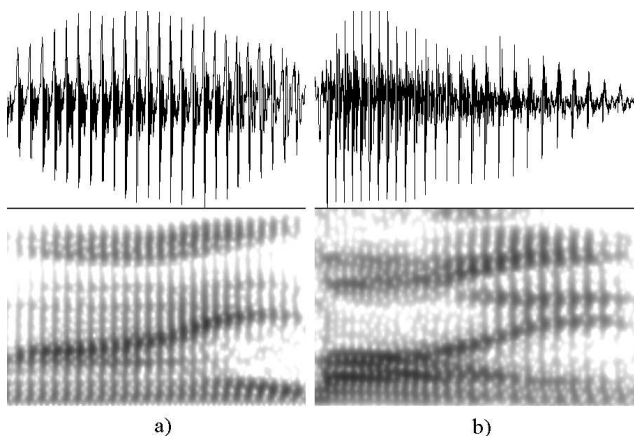


Figure 2: Waveform and spectrogram for the diphthong /ai/ in a) SS word 'Majda' and b) RPE word 'mind'

D	W	%
/ai/	pajka ('spider'), N, gen. sg.	4,41
/ei/	Tejka ('Tea'), N	2,68
/ij/	pijte ('drink'), V, imp., 2pl	0,90
/di/	pojte ('sing'), V, imp., 2pl	0,47
/ɔi/	spojka ('clip'), N	
/ui/	pujsa ('pig'), N, gen. sg.	0,62
/au/	davki ('taxes'), N, 1pl	3,25
/eu/	pevka ('singer'), N, f. sg.	1,66
/ɛu/	bevška ('yelp'), V, 3sg	
/iu/	pivka ('drinker'), N, f. sg.	1,60
/ɔu/	tolkel ('beat'), V, PP, m. sg.	3,00
/uw/	vsul ('pour'), V, PP, m. sg.	0,00

Table 2: The SS diphthongs (D), example realizations (W) and their relative frequencies (%) obtained on the word form lexicon of 527190 entries (Petek and Šuštaršič, 1997)

3.2. Diphthongs of RPE

The term diphthong in RPE refers to a glide between two vowels within one syllable. RPE diphthongs exist as autonomous sound units. Table 3 gives an overview of diphthongs in the RPE and their relative frequencies.

In terms of the Cardinal Vowels the RP English diphthongs can be generally phonetically described as:

CLOSING

glides to /ɪ/: /eɪ/, /aɪ/, /ɔɪ/
glides to /ʊ/: /əʊ/, /aʊ/

CENTERING

glides to open-mid /ə/: /ɪə/, /eə/, /ʊə/

The diphthongs in RPE are very susceptible to variation in different regional and social types of speech.

D	W	%
/eɪ/	pay	1.71
/aɪ/	pie	1.83
/ɔɪ/	coy	0.14
/əʊ/	low	1.51
/aʊ/	bough	0.61
/ɪə/	peer	0.21
/eə/	pair	0.34
/ʊə/	poor	0.06

Table 3: The RPE diphthongs (D), example realizations (W) and their relative frequencies (%) (Gimson, 1994)

3.3. The choice of diphthongs for contrastive analysis

We first observe that the SS diphthongs /ij/, /ui/, /eu/, /iu/, /uw/, /ɔu/, /ɔi/ do not appear in the RPE. The RPE centering diphthongs /ɪə/, /eə/, /ʊə/ are not present in the SS. Candidate list of similar diphthongs suitable for the contrastive analysis is presented in table 4.

SS	D	/ai/	/ei/	/ɔi/	/au/	/ɛu/
	U	[ai]	[ei]	[ɔi]	[au]	[əu]
RPE	D	/aɪ/	/eɪ/	/ɔɪ/	/aʊ/	/əʊ/
	U	[aɪ]	[eɪ]	[ɔɪ]	[aʊ]	[əʊ]

Table 4: The candidate list of SS and RPE diphthongs for contrastive analysis (D- broad transcription, U – the most common allophonic variant)

Realization of the SS element [əu] can be found in, e.g., pronunciation of word 'rekel' [re:kəʊ] ('said') where /ɛu/ is weakened in unstressed syllable to the [əu]. It should be noted that no precise definitions of the SS1 and SS2 parts exist for the diphthongs in SS. On the other hand, the RPE diphthongs and their allophonic variants are thoroughly researched and their SS1 and SS2 are well defined within the Cardinal Vowel diagram (Gimson, 1994; Clark, 1995).

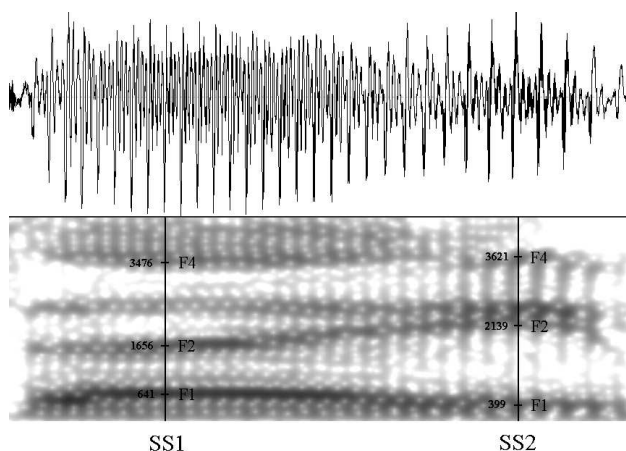


Figure 3: Example of formant analysis for the diphthong [eɪ] as pronounced in the English word 'day'

4. Instrumental setup

The term Received Pronunciation suggests that what is right or wrong is more a result of social judgement than an official decision. It has become increasingly accepted through the arrival of radio and television. The BBC recommended this form of pronunciation to its announcers, thus the RP English became recognized in public mind as the 'BBC English' (Gimson, 1994).

We therefore decided to record television broadcast speech for instrumental analyses. Informants selected were professional announcers of the SS main daily informative broadcast 'TV Dnevnik'. The RPE included the BBC News from the BBC television channel.

TV audio signal was captured by a SCART TV output on 16-bit PC sound card at a sample rate of 44 kHz and stored in a MS PCM .WAV file format. We used PRAAT software (Paul Boersma and David Weenink, University of Amsterdam) for the instrumental analyses.

Three male and three female speakers per language were analysed. Diphthongs were segmented and measured manually. The first (F1), second (F2) and fourth (F4) formant frequencies were measured for the steady states SS1 and SS2 of each diphthong candidate. Figure 3 shows an example of such formant measurement. We also calculated an average F4 value per speaker and used it as normalization of the vocal tract length before the comparison across speakers.

Contrastive analysis included words in the final position or words where sounds of interest resided within the voiceless plosives to diminish the effect of co-articulation. We strived to meet this criterion by judicious selection of words for analysis out of all recorded material. The words selected are listed in tables 5-8.

DIPHTHONG	Male1 SL	Male2 SL	Male3 SL
/ai/	veljajo ('valid'), V, gen. 3pl.	vsaj ('at least')	enajsti ('eleventh'), ord. num.
/ei/	meji ('border'), N, gen. sg.	meji ('border'), N, gen. sg.	poglejmo ('look'), V, 1pl
/di/	vojske ('army'), N, gen. sg.	obstoj ('existence'), N, nom. sg.	dvoboj ('duel'), N
/au/	javna ('public'), adj.	držav ('country'), N, acc. pl	igralke ('actress'), N, nom. pl
/əu/	ministrov ('minister'), N, acc. 3pl	milijonov ('million'), N, gen. pl.	svetovno ('world'), adj.

Table 5: Selection of words for the SS male speakers

DIPHTHONG	Female1 SL	Female2 SL	Female3 SL
/ai/	pogajanj ('negotiation'), N, acc. pl.	včeraj ('yesterday'), adv. of time	Kitajske ('China'), N
/ei/	meja ('border'), N	torej ('therefore')	meji ('border'), N, gen. sg.
/di/	zastoji ('standstill'), N, loc. sg.	svojo ('one's own'), poss. pron.	vojske ('army'), N, acc. sg
/au/	molčal ('silent'), V, 3sg	Vipavskem (countryside), adv. of place	reševalci ('rescuer'), N
/əu/	neplačnikov ('nonpayer'), N, acc. 3pl	domov ('home'), adv. of place	hotelov ('hotel'), N, gen. pl

Table 6: Selection of words for the SS female speakers.

DIPHTHONG	Male1 EN	Male2 EN	Male3 EN
/aɪ/	minds	rely	islands
/eɪ/	day	evaluation	Malesia
/ɔɪ/	oil	avoid	joined
/aʊ/	thousands	allow	cloud
/əʊ/	goat	growth	also

Table 7: Selection of words for the RPE male speakers

DIPHTHONG	Female1 EN	Female2 EN	Female3 EN
/aɪ/	spies	time	rise
/eɪ/	age	rates	place
/ɔɪ/	points	point	envoy
/aʊ/	south	now	down
/əʊ/	hope	closed	Moscow

Table 8: Selection of words for the RPE female speakers

5. Results

Normalized (by F4) average gender dependent SS1 and SS2 measurements of the RPE and SS diphthongs are plotted and analysed in two separate F1 vs. (F2-F1) diagrams and shown in figures 4-5.

For the proposed list of similar Slovenian and English diphthongs given above, measurements indicated that the SS1 and SS2 states of the SS and RPE cluster reasonably well. Some deviations can be found for which we try to provide a suitable linguistic explanation.

SS1 of RPE /aɪ/ is shifted backwards of /aʊ/ for both genders. This is contrary to the initial expectations. Yet considerable latitude is permitted between /a/ and /ɑ/ starting point of /aɪ/ and /aʊ/ therefore fronting or retracting of initial position is quite common. The RPE /eɪ/ has articulatory features of [ɛɪ] and for that kind of

articulation of /eɪ/ retracted initial position of /aɪ/ is likely to occur. Also in refined RPE the very back starting point of /aɪ/ is most common.

With RPE glide /əʊ/ the unrounded first element seems to be produced further forward whilst its final position is rather central and raised. We hypothesise that this is a reduced variant of /əʊ/ characteristic of the refined RPE that is also influenced by the co-articulation with adjacent velars.

Difficulty with having a limited quantity of speech is evident in the case of the SS glide /əʊ/ where its SS1 is shifted backwards in particular for the female speakers. In word 'domov' [domɔ:u] produced by 2nd Slovene female speaker the realization of the first element is not really [ə] but is closer to [ɒ]. This is consequently evident in the figure 5.

6. Conclusion

Work presented here is an extension to the effort of describing the SS diphthongs and to compare them to their RPE counterparts. The ultimate research goal is directed towards portability of HLT resources across languages. Contrastive acoustic-phonetic measurements confirmed that the quality of TV signal is sufficient to perform such analysis. One of the problems encountered was a very small amount of clean studio speech per broadcast record (e.g., max. 5 min out of 30 min of the recorded material). Consequently the analysis of radio broadcast signal could also be considered. Experimental acoustic-phonetic results support perceptually-based similarities between the SS and RPE diphthongs and should become non-preliminary when more data is included in the final analysis.

7. References

- Clark, John; Colin, Yallop, 1995. *An Introduction to Phonetics and Phonology*. Blackwell Publishers Ltd., 2nd edition.
- Gimson, Alfred Charles; Cruttenden, Alan, 1994. *Gimson's Pronunciation of English*. Arnold: London, 5th edition.
- Petek, Bojan; Šuštaršič, Rastislav, 1997. *A corpus-based approach to diphthong analysis of standard Slovenian*. Proc. Eurospeech '97, pp. 767-770.
- Šuštaršič, Rastislav; Komar Smiljana; Petek Bojan, 1999. Illustrations of the IPA: Slovene. *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge University Press, pp. 135-139. (ISBN: 0521652367).
- Toporišič, Jože; 2000. *Slovenska slovnica*. Založba obzorja, 4th edition. Maribor.

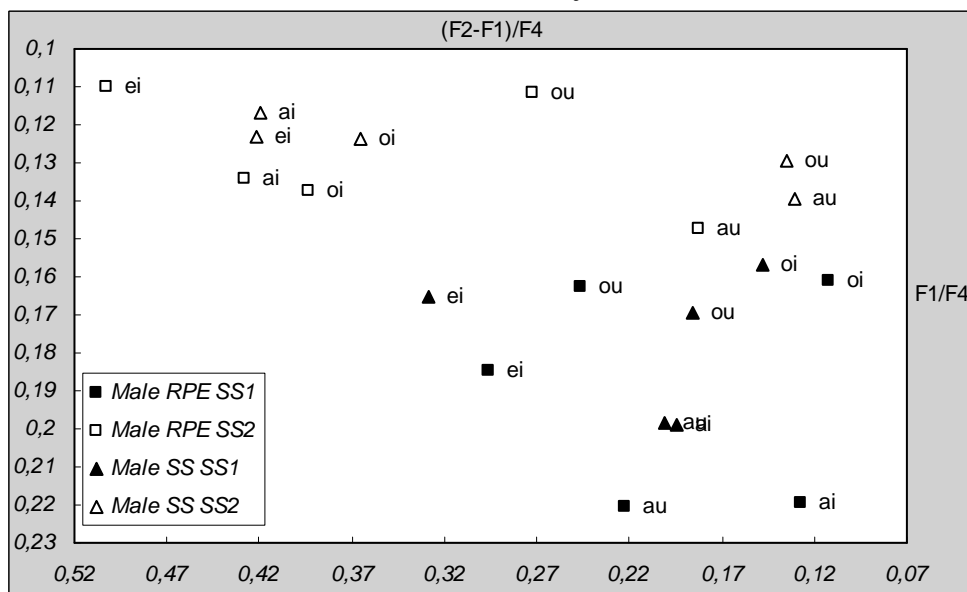


Figure 4: SS1 and SS2 positions for the male speakers of SS and RPE¹

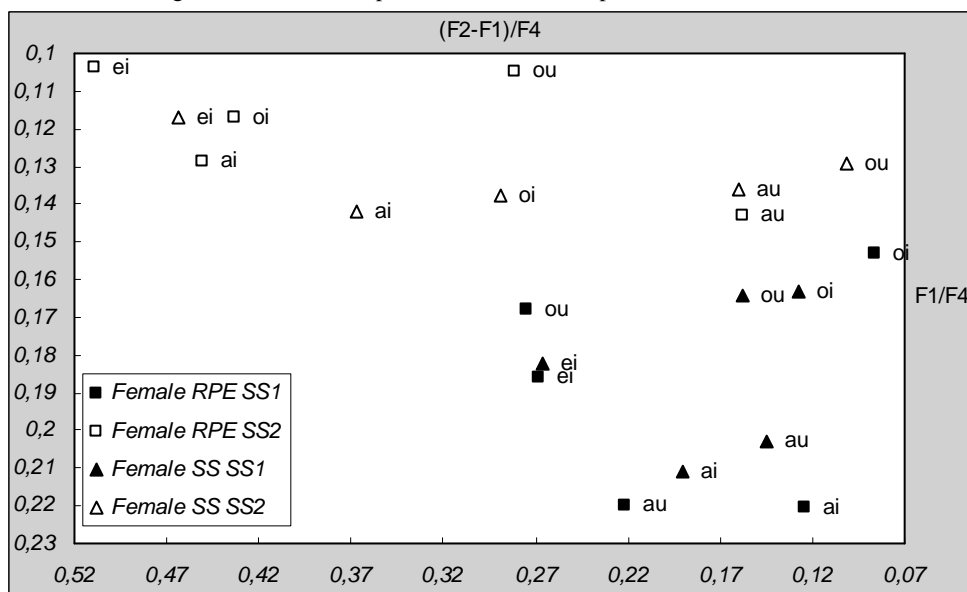


Figure 5: SS1 and SS2 positions for the female speakers of SS and RPE.

¹ 'ou' denotes 'əu'; respective centralized RPE vowel sounds 'o', 'u' are denoted as 'o', 'u'.