A CORPUS-BASED APPROACH TO DIPHTHONG ANALYSIS OF STANDARD SLOVENIAN

Bojan Petek[1] *Rastislav Šuštaršič*[2]

University of Ljubljana
[1]*Faculty of Computer and Information Science, Tržaška 25* [2]*Faculty of Arts, Aškerčeva 2*
1000 Ljubljana, Slovenia

Email: bojan.petek@uni-lj.si rasto.sustarsic@uni-lj.si

ABSTRACT

This paper presents an inventory and relative frequency estimation of glides on the 527,190 word-form Standard Slovenian lexicon. Detailed acoustic-phonetic measurements for the first four most frequent glides /ai/, /au/, /ou/, and /ei/ in stressed syllables are given. Inspection of their formant trajectory plots enabeled measurements of the first four formants in the onset and offset steady-states. Normalized duration patterns for the onset steady-state, glide and offset steady-state are also given. Results represent a broader view to the recently published JIPA paper [4] and are an initial step towards the decision on the most appropriate allophonic symbols to be used in narrow transcription for the glides of Standard Slovenian.

1. INTRODUCTION

Slovenian is a Slavic language of 2 million population. Traditionally, the diphthongs in Slovenian are treated as sequences of vowels followed by phonetic realizations of the labiodental approximant /v/ or the alveolar lateral /l/ (e.g. [au], [eu]) and the semivowel /j/ (e.g. [ai], [ei]), occurring before a consonant or a word boundary. That is why they are generally referred to not as phonemic but *phonetic* diphthongs. The main arguments for such treatment seem to be the following specific features of Slovenian inflected forms of nouns, verbs, adjectives and pronouns.

Firstly, bilabial realization of the approximant /v/ results from the vocalization of /l/ and /v/ when these are followed by a pause or a consonant. Thus, for example, /v/ occurs in all forms of the noun 'glava' ('head') except the gen. dual and the gen. pl., which have a zero inflectional morpheme, with /v/ in final position ('glav' [glau]). Similarly, /l/ is realized as a bilabial vocalic element in the masculine sg. forms of the past participles (PP) of verbs such us 'videti' (to see), ie. in 'videl', PP, m. sg., where it is final, but retained as /l/ before the vowels occurring in all other inflections, as for example before /a/ in 'videla', PP, f. sg.

Second, the vocalic realization of the palatal semivowel largely parallels that of /l/ and /v/. As in a number of other languages, the /j/ in Slovenian words like 'jama' ('cave') is considered a consonant (sonorant, semivowel) rather than a vowel mostly on the basis of its phonemic alternation with other consonants in this position (cf. 'sama' ('alone'), 'rama' ('shoulder'), etc.). The treatment of its occurrence as a vocalic allophone in other positions is again based on morphology. Thus in the possessive adjective or pronoun 'moj' [moi] ('my, mine', nom. m. sg.), /j/ is supposedly realized as [i], and in 'moja' [mɔja] (nom. f. sg.) as a consonant [j].

In both the [au] and the [ai] series, the phonetic status of the glides is reinforced by rules of syllable division in Slovenian, according to which the two elements of a glide are separated before a vowel, with the second element shifted to the right syllable, thus becoming consonantal (e.g. in gla-va, vide-la, mo-ja; cf. syllabification in English: 'pow-er', 'play-er', not *po-wer, *pla-yer).

Although one can easily find minimal pairs to show that at least some of the glides can be distinctive (e.g. 'pajka'-'pika', 'spider', nom. dual, 'dot', nom. sg.), certain glides can only alternate with other glides (e.g. 'prišel'-'prišil' ('came'-'sewed'), both m. sg.). In the latter case, it could equally be argued that the distinctive segment is not the whole glide but only one or the other of its parts (e.g. /e/ in the example 'prišel' above, or the underlying /l/ in 'prišil' vs. 'prišit' ('sewn'), m. sg.). On the other hand, it is much easier to find examples where the vocalic realizations of /l/, /v/, and /j/ unambiguously perform a distinctive role. Thus, for example, the /j/ in, e.g. 'daj' ('give', imp.) can be replaced with a number of consonants to yield other forms of the verb ('dam', 'daš', 'dat') or other lexical items ('dac', 'dač', 'dah', 'dam', 'dan', 'dar').

With two exceptions, the first elements of the glides in Slovenian show no specific articulatory features that would justify their classification as autonomous sound units. The exceptions are the glides [ei] and [ou], in which the first elements have degrees of opening between the qualities of the monophthongs /e-ɛ/ and /o-ɔ/ respectively. Thus, for example, the [e] in 'Tejka' is considerably more open than the /e/ in 'teka' (noun 'run', gen. sg.) but closer than /ɛ/ in 'teče' (verb 'run', 3rd p. sg., pres.). Similarly, the [o], e.g. in 'rovka' ('shrew') is more open than the one in 'Roka' (name, gen. sg.), and closer than in 'roka' ('hand'). It should be noted, however, that the articulation of [e] in the first glide is the same before a consonant/pause and before a vowel ('medial' degree of opening) but the glide can only function as a distinctive segment before a consonant or a pause. The latter, of course, also holds for [ou], which is articulated before a vowel either as /ɔ:v/ (e.g. in 'lova' ('hunt'), noun, gen. sg.) or /ɔ:l/ (e.g. in 'stola' ('chair'), noun, gen. sg.).

The paper is organized as follows. Section 2 presents an inventory of glides that can be found in Standard Slovenian. Their relative frequencies in Slovenian language are estimated on a 527,190 word-form lexicon (Section 3). Finally, instrumental results of acoustic-phonetic analysis for the four most frequent ones are given in the Section 4.

2. INVENTORY OF GLIDES IN STANDARD SLOVENIAN

Slovenian language has 21 consonants and 8 vowels. In general the orthography represents the segmental pronunciation quite faithfully. Glides in Standard Slovenian occur when approximants /v/ and /j/ are preceded by a vowel and followed by a consonant or a word boundary. The labiodental /v/ becomes a rounded second element of a diphthong, i.e. [u], and the palatal /j/ becomes [i].

Table 1 presents a broader view to the recently published inventory of phonetic diphthongs in the Standard Slovenian [4]. Additional element not listed here is the [əu] (e.g. as found in the pronunciation of word 'rekel' [re:kəu] ('said')). Its realization in Standard Slovenian, however, can be found in the examples of /ɛu/, as the /ɛu/ in an unstressed syllable is weakened to an [əu] in the standard pronunciation.

GLIDE	STRESSED SYLLABLES	UNSTRESSED SYLLABLES
/ai/	pajka ('spider'), N, gen. sg.	likajte ('iron'), V, imp., 2pl
/ei/	Tejka ('Tea'), N	zlodej ('devil'), N
/ij/	pijte ('drink'), V, imp., 2pl	Belgijka ('Belgian'), N, f. sg.
/ɔi/	pojte ('sing'), V, imp., 2pl	dvoboj ('duel'), N
/oi/	spojka ('clip'), N	
/ui/	pujsa ('pig'), N, gen. sg.	
/au/	davki ('taxes'), N, 1pl	pikal ('bite'), V, PP, 1sg
/eu/	pevka ('singer'), N, f. sg.	
/ɛu/	bevska ('yelp'), V, 3sg	legel ('lay'), V, PP, m. sg.
/iu/	pivka ('drinker'), N, f. sg.	rabil ('use'), V, PP, m. sg.
/ou/	tolkel ('beat'), V, PP, m. sg.	makovka (a kind of pastry), N
/uw/	vsul ('pour'), V, PP, m. sg.	

Table 1: Inventory of glides in the Standard Slovenian. The symbols ɛ, ɔ denote the open /e/ and /o/, respectively. (V=verb, N=noun, gen=genitive, PP=past participle, imp=imperative, 1sg=1st person singular, 2pl=2nd person plural)

In this work, Table 1 served as a starting point for an instrumental phonetic analysis, i.e. the goal was to investigate whether or not a certain pronunciation conformed with the diphthong definition of a vocalic nucleus containing two target positions, as further discussed in the Section 4.

3. WORD-FORM LEXICAL ANALYSIS

Counts for estimation of the relative frequencies of glides in the Standard Slovenian pronunciation are summarized in Table 2.

GLIDE	COUNT
/ai/	23,229
/ei/	14,123
/ij/	4,736
/oi/, /ɔi/	2,454
/ui/	3,294
/au/	17,136
/eu/, /ɛu/	8,753
/iu/	8,436
/ou/	15,812
/uw/	26

Table 2: Counts of glides in the Standard Slovenian word-form lexicon of 527,190 entries.

These results were obtained on a word-form lexicon with 527,190 entries (containing 41,357 canonical forms). Only lexical spelling of the words was provided in this lexicon. Table 2 enables estimation of the relative frequencies of glides on the word-form lexicon rather than on a large text corpus. Therefore, only a rough estimate of the relative frequencies of glides in a written text can be made. Problems also occur with a naive interpretation of a rule predicting a glide from the orthography. For example, /ou/ could be predicted in both words 'molk' ('silence') and 'Molk' (family name) using a simple rule that predicts a glide on the basis of orthography alone. However, the standard pronunciation allows for /ou/ in the first case only. Despite some deviations from the basic general rules the results showed that the most frequent glides in Slovenian words are the /ai/, /au/, /ou/, and /ei/.

4. ACOUSTIC PHONETIC ANALYSIS

Instrumental acoustic analysis was completed on a database containing the tokens listed in Table 1.

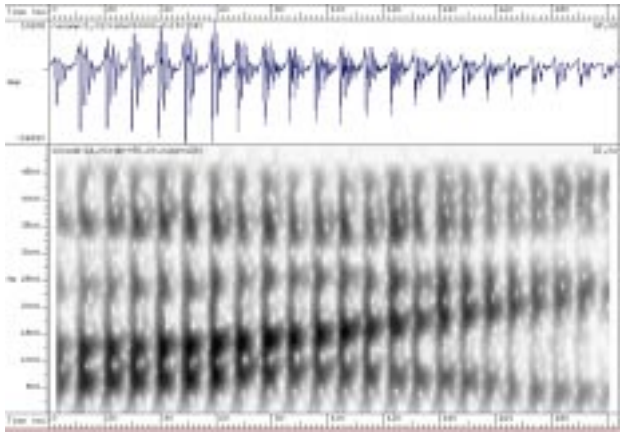

Figure 1: An example pronunciation of /ai/ in a stressed syllable, excised from the Slovenian word 'pajka' ('spider').

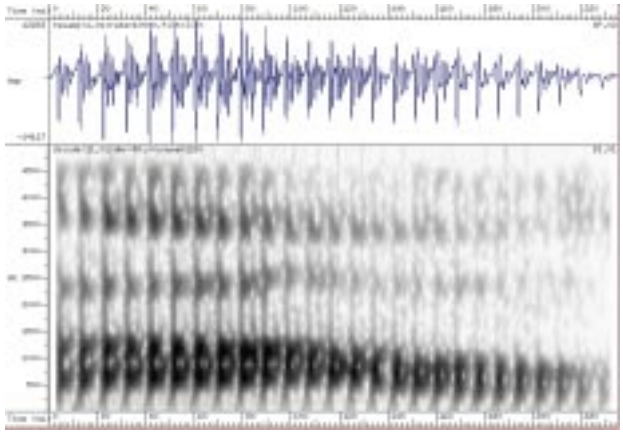

Figure 2: An example pronunciation of /au/ in a stressed syllable, excised from the Slovenian word 'davki' ('taxes').

Three tokens per candidate diphthong per speaker were digitized at a 16kHz sampling rate with falling, rising and level pitch using a close talking microphone. Tokens were embedded in a

sentence frame of the form 'Rekel sem...' ('I said...'). Therefore, each speaker entry contained 60 utterances with realizations of 20 glides in stressed and unstressed syllables within the carefully selected phonetic environments. As of the present this database includes three reference speakers of Standard Slovenian.

Spectral analysis was performed on manually segmented vocalic nuclei using the UCL's Speech Filing System Vs3.0 (SFS) [5]. All tokens were downsampled to a 10kHz sampling rate prior to the formant analysis. Formants were measured using a 12 pole LPC analysis with a 20ms window. Formant trajectories were checked for consistency and inspected visually using the SFS. Example realizations of the first four most frequent glides in Standard Slovenian are given in Figures 1 to 4.

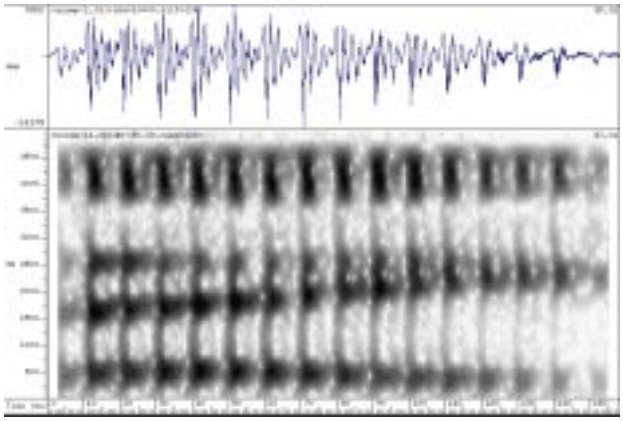

Figure 3: An example pronunciation of /ei/ in a stressed syllable, excised from the Slovenian word 'Tejka' ('Tea').

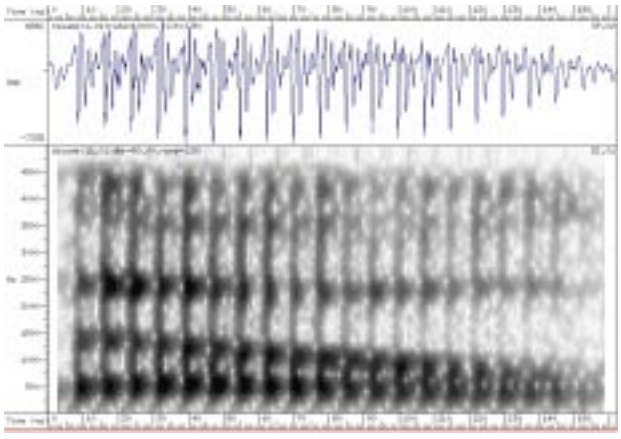

Figure 4: An example pronunciation of /ou/ in a stressed syllable, excised from the Slovenian word 'tolkel' ('beat').

4.1. Formant Analysis Results

The estimation of boundaries between the steady-state, glide, steady-state parts is in general a problematic task. Since constant formant values are rare the segmentation becomes ultimately arbitrary [2]. Tables 3 to 6 give detailed measurements of target formant values for /ai/, /au/, /ei/, /ou/ in stressed syllables.

SPKR	FMT	SS1			SS2		
		1	2	3	1	2	3
RS	F1	665	661	659	372	447	396
	F2	1,269	1,266	1,229	2,208	2,078	2,044
	F3	2,386	2,442	2,393	2,498	2,395	2,348
	F4	3,575	3,731	3,566	3,826	3,852	3,736
BP	F1	680	617	678	436	430	445
	F2	1,129	1,104	1,143	1,940	2,035	2,046
	F3	2,590	2,436	2,571	2,530	2,585	2,489
	F4	3,775	3,659	3,654	3,671	3,680	3,673
MS	F1	634	605	616	413	440	450
	F2	1,190	1,006	1,197	1,767	1,833	1,609
	F3	2,353	2,333	2,210	2,290	2,346	1,913
	F4	3,299	3,354	3,098	3,537	3,467	3,259

Table 3: Initial and final target formant positions for the Standard Slovenian [ai] in the stressed syllable positions (in Hz). Three measurements with falling (1), rising (2), and level pitch (3) were made to estimate the onset steady-state (SS1) and offset steady-state (SS2) formant values.

SPKR	FMT	SS1			SS2		
		1	2	3	1	2	3
RS	F1	668	677	692	470	470	430
	F2	1,188	1,189	1,185	771	752	717
	F3	2,347	2,357	2,364	2,422	2,268	2,231
	F4	3,512	3,655	3,591	3,540	3,604	3,631
BP	F1	719	659	646	447	440	431
	F2	1,160	1,088	1,093	814	709	706
	F3	2,629	2,577	2,532	2,333	2,386	2,358
	F4	3,747	3,727	3,659	3,345	3,571	3,351
MS	F1	663	605	574	446	452	463
	F2	1,155	1,063	1,090	862	839	879
	F3	2,354	2,269	2,174	2,012	1,976	1,866
	F4	3,252	3,293	3,205	3,424	3,277	3,355

Table 4: Initial and final target formant positions for the Standard Slovenian [au] in the stressed syllable positions (in Hz). Three measurements with falling (1), rising (2), and level pitch (3) were made to estimate the onset steady-state (SS1) and offset steady-state (SS2) formant values.

SPKR	FMT	SS1			SS2		
		1	2	3	1	2	3
RS	F1	485	497	500	380	345	364
	F2	1,642	1,628	1,669	2,146	2,275	2,128
	F3	2,539	2,545	2,523	2,427	2,486	2,378
	F4	3,857	3,992	3,857	3,807	3,982	3,817
BP	F1	507	499	485	416	369	407
	F2	1,571	1,579	1,539	2,019	1,937	1,970
	F3	2,575	2,545	2,556	2,391	2,294	2,237
	F4	3,633	3,587	3,608	3,507	3,452	3,504
MS	F1	457	449	462	326	351	325
	F2	1,525	1,593	1,501	1,867	1,897	1,824
	F3	2,353	2,466	2,331	2,228	2,256	2,152
	F4	3,314	3,364	3,230	3,257	3,395	3,213

Table 5: Initial and final target formant positions for the Standard Slovenian [ei] in the stressed syllable positions (in Hz). Three measurements with falling (1), rising (2), and level pitch (3) were made to estimate the onset steady-state (SS1) and offset steady-state (SS2) formant values.

4.2. Duration Analysis Results

Tables 7 and 8 give estimates of the normalized durations of the first four most frequent glides. Duration results are important to e.g., text-to-speech applications for Standard Slovenian.

SPKR	FMT	SS1			SS2		
		1	2	3	1	2	3
RS	F1	501	544	540	417	455	384
	F2	1,354	1,328	1,340	755	732	816
	F3	2,458	2,418	2,434	2,338	2,129	2,367
	F4	3,714	3,498	3,726	3,686	2,944	3,765
BP	F1	550	514	502	441	402	434
	F2	1,168	1,260	1,203	731	730	791
	F3	2,456	2,442	2,161	2,228	2,226	2,172
	F4	3,463	3,431	3,497	3,076	2,739	3,193
MS	F1	–	470	495	–	327	385
	F2	–	1,221	1,221	–	733	748
	F3	–	2,388	2,403	–	1,981	1,898
	F4	–	3,229	3,325	–	2,972	3,422

Table 6: Initial and final target formant positions for the Standard Slovenian [ou] in the stressed syllable positions (in Hz). Three measurements with falling (1), rising (2), and level pitch (3) were made to estimate the onset steady-state (SS1) and offset steady-state (SS2) formant values.

SPEAKER		DURATION [AI]			DURATION [AU]		
		1	2	3	1	2	3
RS	SS1	41	30	33	41	48	41
	GL	47	54	61	40	34	41
	SS2	12	16	6	19	18	18
BP	SS1	33	28	35	25	42	44
	GL	56	61	57	55	41	32
	SS2	11	11	8	20	17	24
MS	SS1	34	26	38	48	22	48
	GL	50	55	48	37	50	27
	SS2	16	19	14	15	28	25

Table 7: Normalized duration values for the Standard Slovenian [ai] and [au] in the stressed syllable positions (in %). Three measurements with falling (1), rising (2), and level pitch (3) were made to estimate the onset steady-state (SS1), glide (GL), and offset steady-state (SS2) normalized durations.

SPEAKER		DURATION [EI]			DURATION [OU]		
		1	2	3	1	2	3
RS	SS1	18	25	18	14	12	15
	GL	67	45	65	65	74	66
	SS2	15	30	17	21	14	19
BP	SS1	23	16	14	17	12	14
	GL	62	65	72	59	74	60
	SS2	15	19	14	24	14	26
MS	SS1	16	17	9	–	15	15
	GL	65	69	79	–	66	70
	SS2	19	13	12	–	19	15

Table 8: Normalized duration values for the Standard Slovenian [ei] and [ou] in the stressed syllable positions (in %). Three measurements with falling (1), rising (2), and level pitch (3) were made to estimate the onset steady-state (SS1), glide (GL), and offset steady-state (SS2) normalized durations.

5. CONCLUSION

This work is a step towards the decision on the most appropriate allophonic symbols to be used in narrow transcription and can be regarded as a preliminary step for further research in the fields of speech synthesis, automatic speech recognition of Standard Slovenian, and for the current project of contrastive phonetic analysis of English and Slovenian.

Presented are detailed acoustic-phonetic measurements for the first four most frequent glides /ai/, /au/, /ou/, and /ei/ in stressed syllables whose relative frequency estimates were obtained on the 527,190 word-form lexicon. Formant trajectory plots enabeled measurements of the first four formants in their onset and offset steady-states. Additionally, normalized duration patterns for the onset steady-state, glide and offset steady-state are also estimated.

Results presented here are preliminary. More material including more speakers and phonetic contexts has to be analysed before the final conclusions can be drawn. Nevertheless, the work presented here complements research that considers the Standard Slovenian language which has recently been published in the JIPA [4] and the ICSLP'96 papers [3].

ACKNOWLEDGEMENTS

Miro Romih (Amebis, d.o.o.) and Tomaž Erjavec (Jožef Stefan Institute) provided the word-form lexicons for research presented in Section 3 and the paper proposal, respectively. The authors are indebted to the members of the Laboratory of Artificial Perception, Faculty of Electrical Engineering, University of Ljubljana, Slovenia, for providing prerequisites for the acoustic-phonetic research.

6. REFERENCES

1. Lehiste, I. and Peterson G. E. (1961). Transitions, Glides, and Diphthongs. Journal of the Acoustical Society of America 33 (3): 268-277.
2. Peeters, W. J. M. and W. J. Barry (1989). Diphthong dynamics: production and perception in Southern British English. Proc. EUROSPEECH'89, 1: 55-58.
3. Petek, B., R. Šuštaršič, S. Komar (1996). An Acoustic Analysis of Contemporary Vowels of the Standard Slovenian Language. Proc. Int. Conf. on Spoken Language Processing ICSLP'96, 1: 133-136.
4. Šuštaršič, R., S. Komar, B. Petek (1995). Illustrations of the IPA: Slovene. Journal of the International Phonetic Association 25 (2): 86-90.
5. http://www.phon.ucl.ac.uk/resources/sfs.html