

Duration of fricative /ʃ/: analysis on isolated word corpus

Slobodan T. Jovičić, Zorka Kašić, and Silvana Punišić

Abstract — This paper presents the results of investigation in duration of fricative /ʃ/ on isolated word corpus. It was analyzed the initial, medial and final position of /ʃ/ in the words, and effects of word duration, syllable accent and coarticulation. Results showed the least duration /ʃ/ in medial position, largest and most stable in final position, and a larger one in shorter words and syllables with short accent. Coarticulation effects are different and significant depending on articulation distance of /ʃ/ and following voices.

Key words — Duration, fricatives, fricative /ʃ/, isolated words.

I. INTRODUCTION

PHONEME duration is an inherent characteristic of every language's phonetic system. It is a significant characteristic of every phoneme in realisation of continual verbal communication, not only in the linguistic, but also paralinguistic and extralinguistic aspect. Research on this phoneme dimension in the linguistic domain has lasted some time [1], [2] and, with development of contemporary speech technologies, has gained significance in the domains such as speech and speaker recognition, speech synthesis and natural language processing [3]. On the other hand, duration as a trait arouses interest in analysis of deviation and pathology in phoneme articulation [4].

Phoneme duration in spoken language is dependent on numerous factors such as position of phoneme in a word, the phonetic context, influence of accented syllable in a word, whether it is positioned before or after a pause, etc. [5]. Complexity of this issue lead to extensive research specific to every language, such as American English, Dutch, Spanish, German, French [6], etc., as well as comparative analysis for these languages [7]. Phoneme research on the linguistic level [8], [9] is also present for the Serbian language on the technological level, in optimisation of a TTS (text-to-speech) synthesis of the Serbian language [10], and research on distortions in articulation of certain phonemes [11].

This work was supported by Ministry of Science of the Republic Serbia under Grant number 148028G.

S. T. Jovičić, School of Electrical Engineering, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia (telefon: +381-64-5377689; e-mail: jovicic@etf.bg.ac.yu).

Z. Kašić, Faculty for Special Education and Rehabilitation, Visokog Stevana 2, 11000 Belgrade, Serbia (e-mail: zor.ka@sezampro.yu).

S. Punišić, Institute for experimental phonetics and speech pathology, Gospodar Jovanova 35, 11000 Belgrade, Srbija (e-mail: punisic@infosky.net).

This study is focused on analysis of fricative /ʃ/ duration, in isolated words in Serbian. Fricative /ʃ/ has been selected as a typical representative of unvoiced fricatives, being the most intense, of most greatly defined articulation and easily segmented in the context of other phonemes. Analysis was carried out for the position of /ʃ/ in a word and influence of three dimensions on /ʃ/ duration: word length, accented syllable where /ʃ/ is located and coarticulation of /ʃ/ with closest phonemes.

II. EXPERIMENT ORGANISATION

A. Creation of the speech corpus

The speech corpus for the study's needs consists of 85 two-syllable isolated words, 3 three-syllable words and 2 one-syllable words. The speech corpus is presented in the appendix. All the words in the speech base are accented with four Serbian accents: ˘ - short-falling, ˘˘ - short-rising, ˘˘˘ - long-falling and ˘˘˘˘ - long-rising.

The speech corpus words have been chosen with the fricative /ʃ/ being in all three positions: initial, medial and final. For the initial position, there are 25 words with /ʃ/ five times in front of all five vowels of the Serbian /a, e, i, o, u/. In the third position, there are consonants (C) from all phonetic groups (plosives, fricatives, affricates and sonants) behind the vowel (V), but only those which can form a combination /ʃVC.. / which can be found in the Serbian language. The next 25 words with the initial /ʃ/ have one of the plosives or sonants in the second position.

In the medial position, /ʃ/ has been placed in 20 words between the same vowels, different vowels or consonants and combinations of vowels and consonants. In the final position, /ʃ/ has been placed in the next 20 words, always following the vowels, which is the most frequent form in the Serbian language. This selection of words was used for analysis of coarticulatory effects between /ʃ/ and other voices.

B. Recording of the speech corpus

There were 12 speakers with standard Serbian pronunciation, producing words for the speech corpus (4 men and 8 women). Their ages were from 24 to 60. Pronunciation of all speakers was correct without distortions or pathological deviations.

The speech corpus has been recorded on a standard PC with 16 bits resolution and sampling frequency of 22050 Hz. The recording was carried out in a soundproof room with microphone type AKG C568EB. The speakers

pronounced the words with a neutral tone, without emotions, using their own tempo and intensity with a pause of 3 to 5 seconds.

C. The segmentation procedures

The segmentation of fricative /f/ for the whole corpus (1080 words) was carried out auditorily and visually by a phonetician with the help of Praat software [12]. The greatest problem in segmentation was establishing the final boundary of /f/ at the end of words, as well as prolonged duration of friction in occlusions of unvoiced plosives behind /f/. The main criterion of cessation of final /f/ articulation was ending of its spectral structure consistency.

III. EXPERIMENT RESULTS

A. Position in the word

Mean duration of fricative /f/ in the initial, medial and final position in isolated words of the recorded speech corpus is presented in table 1. Average values have been given for every position by stimulus and by speaker. The table also shows standard deviations for each three positions: the standard deviation for all measured results SD_{all} , the standard deviation between average values of stimulus $SD_{stimuli}$ as well as the standard deviation between speakers $SD_{speaker}$.

TABLE 1. DURATION OF FRICATIVE /f/ IN (ms).

	Initial	Medial	Final
Mean	188,24	177,01	272,64 (266,17)
SD_{all}	41,72	34,2	46,43 (40,82)
$SD_{stimuli}$	23,43	19,96	22,82 (11,71)
$SD_{speaker}$	26,64	24,1	33,43 (32,30)
$SD_{all}/Mean$	0,2216	0,1932	0,1703 (0,1534)

Results show that duration of fricative /f/ is the shortest for the medial position, to some extent longer for the initial position and the longest for the final position. The difference between the initial and medial position is not statistically significant and is generally caused by stressed pronunciation of the first sound in a word pronounced

isolated.

When duration of fricatives /f/ in all three positions are compared based on coefficient of variation, defined as a ratio between the standard deviation and mean ($SD_{all}/Mean$), it can be seen that fricative /f/ in the initial position is liable to greatest changes.

Results show that duration of fricative /f/ in the final position is the greatest and most stable. This result can be explained by the fact that in the final position, /f/ only has vowels which precede it, which is a characteristic of the Serbian language. In addition, for the final position, results have shown that duration of the final /f/ in one-syllable words /tu_f/ and /ku_f/ are much greater compared to all other two-syllable and three-syllable words and is about 330 ms (this result will be discussed later). Excluding these results from the analysis, the values are lower for all results in table 1 for the final position. They are lower especially for standard deviation between the stimuli $SD_{stimuli}$ and coefficient of variation. These results show a great level of autonomy in articulation of the final /f/.

B. /f/ in initial position

Figure 1 shows the distribution of average durations of all 50 initial fricatives /f/ presented by a growing trend. The distribution is nonlinear. Three segments of the distribution can be interpolated by linear regressions with significantly different slope coefficients. This fact points to different effects of coarticulation of voices following the initial /f/.

In the first word group, where the mean duration of /f/ is the shortest, there are 9 words with plosives /p, t, k/ in the second position following /f/ (combination /fP./, P - plosive) not appearing in other words in the second and third group. This is, therefore, a highly compact group. The second group is also compact (having the regression line with the smallest slope) having words with a vowel or sonant in the second position. The third group consists of four two-syllable words, which is to be discussed later.

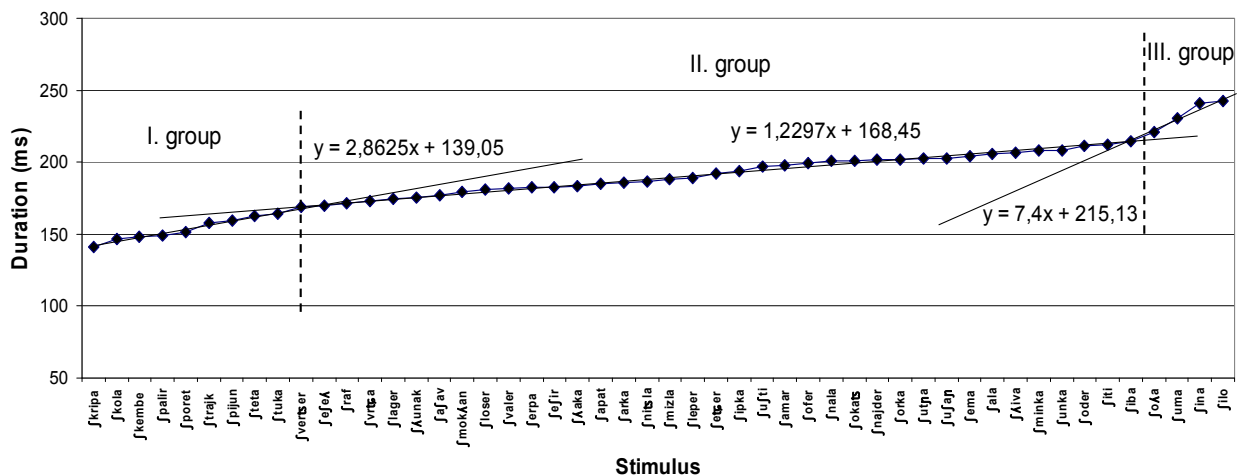


Fig. 1. Mean duration of /f/ in initial position for all 50 words in the speech base.

C. /ʃ/ in medial position

Figure 2 shows the distribution of mean duration of all 20 medial fricatives /ʃ/ presented by a growing trend. Like with the initial position, the distribution is nonlinear. The shortest duration for /ʃ/ is shown in words where it is followed by a plosive (or affricate), while duration is the longest in words where it is followed by a sonant. The greatest duration of /ʃ/ is seen in the two-syllable word /arʃin/.

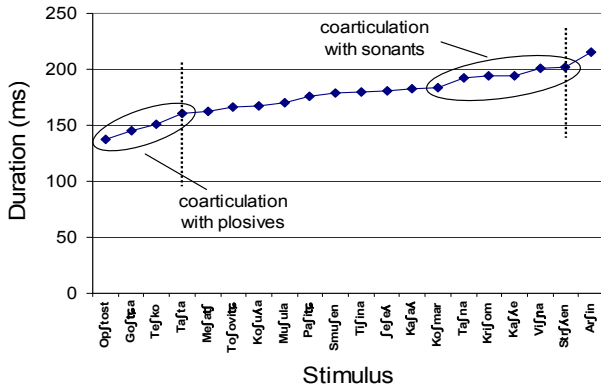


Fig. 2. Mean duration of /ʃ/ in medial position for all 20 words in the speech base.

D. /ʃ/ in final position

Distribution of mean durations of all 20 final fricatives /ʃ/ presented by a growing trend is shown in graph 3. The distribution is highly uniform with a minimal slope, which is confirmed by the result in table 1 where we concluded great stability in articulation of the final /ʃ/. The only deviation from this distribution is shown by monosyllable words /tuʃ/ and /kuʃ/.

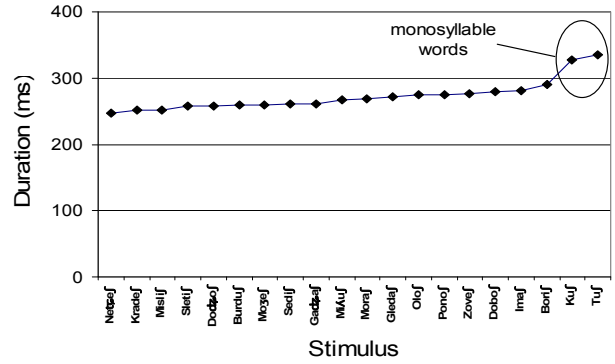


Fig. 3. Mean duration of /ʃ/ in final position for all 20 words in speech base.

IV. DISCUSSION

Results of fricative /ʃ/ duration, when being in the initial position preceding a vowel (combination /ʃV..), are shown in table 2. The table shows duration of fricative /ʃ/ sorted by a growing trend from 1 to 5. We can note grouping of fricatives /ʃ/ in front of vowel /a/ and /e/, /o/ and /u/, and /i/ especially. The result for /i/ can be explained by usage of words with four voices (ʃiti, ʃiba, ʃina), not seen in other vowels, and will be explained in discussion of the example presented in table 3.

Table 3 includes words with the initial /ʃ/ in front of the vowel /u/ with different number of phonemes, 5 or 4. Duration of these words was measured, showing that words including 5 phonemes have duration over 800 ms with the initial /ʃ/ lasting somewhat more than 200 ms. Also, all four words following the initial /ʃ/ include the vowel /u/ with the long accents (two long-rising and two long-falling). The fifth word /ʃuma/ is phonetically shorter

TABLE 2. DURATION OF FRICATIVE /ʃ/ IN FRONT OF VOWELS IN (ms).

	/ʃ/ in front of /a/	/ʃ/ in front of /e/	/ʃ/ in front of /i/	/ʃ/ in front of /o/	/ʃ/ in front of /u/
1	ʃaʃav	ʃeʃeA	ʃiʃpka	ʃoʃer	ʃuʃti
2	ʃapat	ʃerpa	ʃiti	ʃokats	ʃutpa
3	ʃarka	ʃeʃir	ʃiba	ʃorka	ʃuʃaʃ
4	ʃamar	ʃetæer	ʃina	ʃoder	ʃunka
5	ʃala	ʃema	ʃilo	ʃola	ʃuma
Mean	190,56	186,3	220,68	206,65	208,62

TABLE 3. /ʃ/ IN FRONT OF VOWELS /u/.

No	Word	Number of phonemes	Initial /ʃ/	Word duration
1	ʃuʃti	5	199,17	827,58
2	ʃutpa	5	202,50	819,83
3	ʃuʃaʃ	5	202,67	838,75
4	ʃunka	5	208,50	810,25
5	ʃuma	4	230,25	684,83

and includes 4 phonemes lasting under 700 milliseconds, the vowel /u/ has a short-falling accent and duration of the initial /ʃ/ is about 230 ms.

These results show that shortening of words by number of phonemes, is followed by growth of duration of fricative /ʃ/ with initial position in an isolated word. This

explains the previous result for vowel /i/.

Table 4 shows 5 words with the same content, consisting of 4 phonemes (words from the fifth row in table 2 with the greatest duration of /ʃ/). The difference between words is that the words /ʃilo/ and /ʃuma/ have short-falling accented vowels after /ʃ/, while other words have vowels with long-falling accents following /ʃ/. It can be pointed out that words /ʃilo/ and /ʃuma/ with short-falling accented vowels are of shorter duration (under 700 ms), that vowels with the short-falling accent following /ʃ/ are significantly shorter (about 146 ms) compared to vowels with a long-rising accent (over 230 ms) and that duration of initial /ʃ/ is significantly longer than in other words with long-rising accented

vowels.

TABLE 4. DURATION /f/ & ACCENTS.

No	Word	Number of phonemes	Word duration	Duration of initial /f/	Duration of accented vowel
1	šála	4	725,25	205,42	262,17
2	šéma	4	729,92	204,17	232,00
3	šilo	4	688,50	242,16	146,83
4	šóla	4	762,00	221,08	254,17
5	šuma	4	684,83	230,25	146,67

These results show an effect of accented vowels behind the initial /f/ on its duration. Namely, appearance of the vowel with a short-falling accent in the second position in an isolated word prolongs the duration of the initial /f/.

Result in table 3 concerning the duration of /f/ which is significantly greater compared to other examples (in the word /šuma/) can be interpreted as an integral effect of word shortening and the influence of short accent on duration increase for /f/.

The coarticulation effect can most easily be noted in figure 1 in group I words, where the initial /f/ is found in front of unvoiced plosives /p, t, k/; the same phenomenon can be found in medial /f/ (figure 2). We derived the following hypothesis.

In articulation of the Serbian fricative /f/, the frontal part of the tongue is raised towards the alveoli and is slightly thrust backwards. From this position, the articulatory organs need to reach (acquire “articulatory distance”) the new position for generating an unvoiced plosive /k/ in the word /škripa/ (the first stimulus in figure 1), for example, where the rear part of the tongue rises towards the soft palate, creating a barrier, while the tip of the tongue is lowered to the mouth cavity bottom. This transformation of articulator positions has an effect on duration of fricative /f/: more complex transformation (greater articulatory distance) is followed by shorter duration of fricative /f/. This hypothesis is confirmed by an example of the word /òpftost/ (first stimulus in figure 2) where /f/ is located between two unvoiced plosives and where the medial /f/ has clearly the shortest duration.

V. CONCLUSION

Analysis of duration of the fricative /f/ in isolated words of the Serbian language shows the shortest duration in the medial position and the longest and most stable in the final position. Also, the duration of fricative /f/ is greater for shorter words by number of phonemes and for syllables in words with a shorter accent. Coarticulation effects are different and greatly dependent on articulatory distance /f/ and voices which follow. The integral effect of /f/ position in words, word length, accented syllable and coarticulation, models the inherent duration of /f/. Influence of sentence prosody needs to be added to these effects, which is to be the subject of future research.

APPENDIX

In Serbian, there are 25 consonants grouped in five phonetic classes according to manner of articulation: plosives (unvoiced /p, t, k/ and voiced /b, d, g/), fricatives (unvoiced /f, s, š(š), h/ and voiced /z, ž(ž)/), affricates (unvoiced /ts(c), tɕ(ć), tʃ(č)/ and voiced /dʒ(dž), dʒ(dž)/), nasals (/m, n, ŋ(nj)/) and semivowels (vibrant /r/, laterals /l, ʎ(lj)/, approximants /v, j/) (all consonants are with IPA notation and with Serbian notation in parenthesis [13]).

šèter	šlâger	švîšnja
šèfir	šlôser	tâšna
šéma	šlâiva	kôšmar
šérpa	šlâka	tâfta
šèfeš	šminka	têško
šála	šmizla	kâšle
šamar	šnitsla	âršîn
šapat	šnála	štršlen
šârka	švâler	òpftost
šâfav	švîtea	gòštea
šiba	šraf	imaš
šiti	škembe	nêšes
šipka	škripa	sêdiš
šina	šlêper	òloš
šilo	šmòkšan	bûrduš
šóla	šlúnak	môraš
šofer	šnajder	zòveš
šoder	špâlir	slêtiš
šórka	štrâjk	pònoš
šòkats	švèšer	kûš
šúfti	tîšina	gîšedâš
šuma	kâšal	krâdeš
šunka	šèšes	mîšliš
šútna	mûšula	dòdžoš
šúšan	tòšoviteš	tûš
šporet	kòšula	gâdžâš
šúka	pâšiteš	mòžes
šteta	krîšom	bòriš
špijun	smûšen	dòboš
škòla	mêšatš	mîluš

LITERATURA

- [1] I. Lehiste, *Suprasegmentals*. MIT Press, Princeton, 1970.
- [2] R.L. Whitehead, N. Schiavetti, B.H. Whitehead, D.E. Metz, Temporal characteristics of speech in simultaneous communication. *J. Speech Hear. Res.*, 38(5), 1995, pp. 1014-1024.
- [3] X. Huang, A. Acero, H.W. Hon, *Spoken Language Processing: A Guide to Theory, Algorithm, and System Development*. Prentice Hall, 2001.
- [4] J.M. Christensen, B. Weinberg, Fricative duration in esophageal speech. *J. Commun. Desord.*, 14(2), 1981, pp. 678-689.
- [5] I. Lehiste, P. Ivić, *Word and Sentence Prosody in Serbo-croatian*. MIT, 1986.
- [6] E. Mendoza, G. Carballo, A. Cruz, M.D. Fresneda, J. Munoz, V. Marrero, Temporal variability in speech segments of Spanish: context and speaker related differences. *Speech Communication* 40, 2003, pp. 431-447.
- [7] P. Delattre, *Comparing the Phonetic Features of English, German, Spanish and French*. Julius Groos Verlag, Heidelberg, 1965.
- [8] Lj. Mihajlović, Duration of voiced and unvoiced phonemes in speech. *Glasnik SANU*, V(1), 1953. (in Serbian)
- [9] Z. Kašić, S. Peter, Z. Urošević, M. Filipović, Variability of phonemes duration in words. Proc. of XXXI Yugoslav Conference ETAN, Bled, 1987, pp. 229-233. (in Serbian)
- [10] S. Sovilj-Nikic, M. Secujski, V. Delic, I. Sovilj-Nikic, Analysis of different factors influencing vowel duration in Serbian language. Proc. of *Symp. on Social Communication*, Santiago de Cuba, Cuba 2007.
- [11] S. Jovičić, Z. Kašić, S. Punišić, Perceptual identification of distortion in duration of fricative /f/, prepared for publication.
- [12] Praat, <http://www.fon.hum.uva.nl/praat/>
- [13] IPA, http://en.wikipedia.org/wiki/Serbian_language