811.163.41'342.623
Оригинални научни рад

Наслеђе 26 • 2013 • 9-17

Биљана Чубровић[1]
Катедра за англистику
Филолошки факултет Универзитета у Београду

INITIAL AND MEDIAL VOICELESS STOPS IN SERBIAN: AN ACOUSTIC ANALYSIS

Stops make up a class of consonants that can be found in virtually all languages of the world. Their significant role in phonemic inventories is further supported by the fact that they are acquired at an early stage of language acquisition. This paper looks into the nature of the production of the voiceless stops /p t k/ in word-initial and word-medial positions in the Serbian language. A list of selected Serbian words/pseudowords is recorded and analyzed acoustically in terms of Voice Onset Time (VOT) and some conclusions reached as to the importance of VOT in Serbian, where it is not distinctive. Four different contexts are studied: word-initial accented, word-initial unaccented, word-medial accented and word-medial unaccented. The present data shows that Serbian is a short-lag language, as expected, but that VOT is not a function of stress, which would hardly be hypothesized. Furthermore, word-medial positions almost invariably show longer VOTs compared to the word-initial positions. The gradient scale of rising VOTs starts with the initial accented positions, to initial unaccented contexts, followed by medial accented and medial unaccented phonetic contexts.

Key words: stop consonants, voiceless, Serbian, VOT, aspiration, initial, medial, reverse pattern

1. Introduction

All languages contain a category of stops in their phonemic inventories, which makes a stop a typical, optimal or ideal representative of the consonantal class. Various parameters are implemented when describing stops in the world's languages: phonation type, airstream mechanisms, relative timing of the onset of voicing, and relative timing of velic closure. The relative timing of the onset of voicing is in the focus of this paper. Accordingly, stops can be classified as unaspirated, aspirated and pre-aspirated. We will concentrate on unaspirated and aspirated voiceless stops, even though aspiration is not phonological in Serbian. However, pre-aspirated stops form phonetic inventories of a very limited number of world's languages, and Serbian is not one of them. Preaspirated stops[2] are best described as "voiceless stops which are preceded by a period of glottal frication [hp ht hk], as in the Scottish Gaelic words

1 biljana.cubrovic@fil.bg.ac.rs
2 For a full account of the origin, evolution and phonological behaviour of preaspirated stops, see Clayton (2010).

bata [pahta] 'stick' or *put* [p^{h}uht] 'buoy,' or the Icelandic words *kappi* [k^{h}ahpi] 'contest' or *epli* [ɛhp̥i] 'apple'" (Clayton 2010: 2) and are typically found in word-medial position, even though word-final preaspirated stops also occur, but less frequently.

The parameter of Voice Onset Time (VOT), which is defined as the time interval between the stop release and the onset of vocal fold vibration for the following vowel (Lisker and Abramson 1964) has been a matter of debate in phonetic studies since it was first introduced in the 1950's in an attempt to deal with some heated issues in acoustically-based speech synthesis. Although the concept was originally designed for initial plosives, it was later implemented in other contexts, becoming the means of differentiating between voiced and voiceless stops in a large number of languages. Such a phonetic parameter as VOT was needed because acoustic measurements at the time were insufficient to account for the absence of vocal fold vibration in typically voiced consonants. In English, for instance, word-initial and word-final voiced stops may lack voicing, e.g. [b̥en], and this triggered the question about a phonetic feature or features that may play the role in distinguishing between the initial devoiced [b̥-] and initial voiceless [p-]. Acoustic studies also relied heavily on the state of affairs in English, which differentiates aspirated voiceless stops and unaspirated voiced ones. The answer lies in a timing cue that is referred to as VOT, and is expressed in milliseconds (ms). It measures the delay of the onset of voicing following a stop release burst. A longer delay is associated with voiceless stops, the stage at which the vocal folds remain open for a short period of time after the release of the stop. Johnson (2012: 101) draws some attention to VOT measurements, emphasizing that "many languages have a boundary between aspirated and unaspirated stops at about 30 ms VOT", the 30 ms VOT being the most common boundary between aspirated and unaspirated stops. This is where auditory phonetics may intervene and we will not further elaborate on this matter.

Generally speaking, some languages seem to make more use of VOT, where others do not. Starting from this assumption, Lisker and Abramson (1964) look into the VOT values for 11 different languages and set up a scale which has come to be used in later research on stop consonants. To sum their research up, Abramson (1977: 296) later pointed out that they adopted "the convention of assigning a timing value of zero to the moment of stop release, negative values to voicing lead, and positive values to voicing lag". What they found was "a tri-modal distribution of VOT values" (ibid: 296). The first category centers at -100 ms for a range of values by which voiced unaspirated stops are represented. The second mode amounts to +10ms and corresponds to what we know as voiceless unaspirated stops. The category of voiceless unaspirated stops seems to be present in the phonetic inventory of Serbian. Lastly, the third category of stops centers at +75ms and corresponds to voiceless aspirated stops. This goes in line with several studies analyzing the delay of voicing in bilingual speakers, but also providing data for languages that have unaspirated stops (French, Greek, Spanish), and compared them to the Eng-

lish long lag realizations. Olson (2013: 414) claims that Spanish-English bilinguals perceive aspiration differently in the two languages, and for his Spanish dominant group the mean value for word-initial /k/ is 35.1 ms in Spanish. This places Spanish in the group of short-lag languages. An examination of the French-English bilinguals and their VOT values for French reveal similar results in one of the experiments where word-initial position was studied (24 ms for /p/, 19 ms for /t/, and 28 ms for /k/) (Grosjean and Miller 1994: 203). Antoniou et al's (2010) study of VOT of the voiceless bilabial stop /p/ and voiceless coronal /t/ in word-initial and word-medial in two separate experiments, but with similar stimuli, reveal the results for Greek monolinguals. Word-initially, the mean of the VOT for the Greek /p/ is 14.8 ms, and 17 ms for the Greek /t/. In word-medial positions, the stressed /p/ has a short lag of 14.5 ms, and its unstressed counterpart is slightly longer (17 ms). In word-medial contexts, the mean value of VOT for stressed /t/ amounts to 16.3 ms, and it is slightly longer in unstressed positions, which does not go in line with the expected results for English VOT. It is widely known that the rise of VOT values is dependent on the presence of stress. However, this does not seem relevant for Greek stops.

Further research draws parallels with the results that Abramson and Lisker (1964) obtained in their cross-language study. Their research was only a trigger for numerous studies in which VOT was measured in normal adult and children's monolingual and bilingual speech, and also in a number of speech disorders both in adults and children. Kent & Read (1996: 108) assign similar values of VOT for stops, and points out that the delay in voiceless consonants is on the order of 25-100ms. Kent & Read (1996) truly emphasize that there can hardly be deduced a single value of VOT that could be used by all speakers or across phonetic contexts and this still holds true. Various aerodynamic factors are at play when it comes to the final output of VOT (the mobility of articulators, effects of the place of articulation of different stops, alongside with the temporal adjustment between the closure duration and VOT).

Claims are made that VOT also depends on the place of articulation. Whalen et al (2007: 342) report that VOT values of labial stops are consistently shorter than the values for lingual stops. However, it seems that some speaker variation exists, and some studies show that alveolar stops have a shorter VOT values than velars (Lisker & Abramson 1967; Zue 1976; Weismer 1979), whereas others find out that these values are the same (Docherty 1989; Cooper 1991). Such differences may be accounted for by the speech rate at which speakers enunciate the analyzed material. VOT, as a measurement of time (ms), is dependent on the temporal characteristics of speech. More often than not, the details of the recording procedures are not clearly stated in articles, so tempo as a factor in VOT values has to be considered. As briefly mentioned before, aerodynamics is another factor that explains the differences in VOT across various places of articulation: the lips are extremely movable due to their small mass, as compared to the soft palate which is not as agile. Velar contact is slow and the so-called "double burst" (even triple burst or multiple bursts) may ensue in the articulation of /k g/.

In Serbian there are three voiceless stops: bilabial /p/, dental /t/, and velar /k/. We will now look into the nature of these three stops in word-initial and word-medial positions with the aim of measuring VOT values in Serbian speech.

2. *Previous research on stops in Serbian*

Čubrović (2012) is a study that looks at VOT values in Serbian and Serbian English in word-initial stops. Four native speakers of Serbian, whose English is at the B2 level according to the Common European Framework of Reference for Languages, are recorded. VOT is measured in word-initial, accented positions in Serbian and Serbian English. The values for VOT in Serbian are given in Table 1, whereas the measurements for English as spoken by the same native speakers of Serbian are summarized in Table 2 below.

STOP	**Average (ms)**	**Range**
/p/	221	13:36
/t/	28	21:39
/k/	52	42:64

Table 1: VOT values for Serbian word-initial stops

STOP	**Average (ms)**	**Range**
/p/	31	19:46
/t/	49	34:65
/k/	63	54:76

Table 2: VOT values for Serbian English word-initial stops

The average VOT values are as predicted. Cumulative values of Serbian stops seem lower than those typical of Serbian English stops.

What is strikingly noticeable in Tables 1 and 2 is a systematically arranged increase in VOT values in Serbian English stops, starting from the bilabial /p/, to the dental/alveolar /t/, ending in the velar /k/. Similar results were obtained in an earlier mirror study conducted with proficient speakers of Serbian English. The average VOT values for this group of speakers for Serbian English were slightly higher (39 ms for /p/, 64 ms for /t/, and 71 ms for /k/) (Čubrović 2011: 15). This may lead to a conclusion that aspiration is a phonetic phenomenon that is acquired gradually. Non-native speakers of English are obviously

Наслеђе 26 • 2013 • 9-17

able to slowly stretch their VOT, until its numerical values reach the point where native speakers perceive the relevant consonants as voiceless aspirated.

3. Experiment and discussion

In the present study, we will look into the word-initial and word-medial stops in Serbian, with regard to aspiration. The informants are four native speakers of Serbian (average age, 20: 9). They were asked to read three trisyllables of Serbian, and wherever possible real words of Serbian were used (e.g. *tatama*). However, this was not possible in all contexts so we used pseudowords in such contexts. Stops were produced preceding and following a low vowel /a/ (low vowels are associated with lower velum positions than high vowels, which may influence the appearance of a nasal). Target syllables are embedded in carrier phrases and recorded at a sampling rate of 22,050 Hz. All stimuli are repeated five times and VOT values measured using Praat software 5.3.55.

Standard Serbian is described as a language with two types of pitch accents, 'falling' and 'rising', and even though these conventional labels often fail to reflect phonetic reality (Lehiste and Ivić 1986: 42), they are still widely used. Care was taken that the pitch accent in the corpus is unaltered in all trisyllables, and that the length of the vowel in any syllable is the same. All syllables analyzed in the present study have a short vowel. Four different positions in the trisyllables have been studies: initial accented (abbreviated to IA), initial unaccented (IU), medial accented (MA), and medial unaccented (MU).

The measurements are provided in Tables 3-6, for any chosen stop, bilabial/ dental/ velar respectively, depending on the stress and position in a particular word/string of syllables:

STOP	**Average (ms)**	**Range**
/p/	13	8:18
/t/	16	12:20
/k/	39	22:60

Table 3: VOT values for word-initial accented stops

It is worth noting that in the word-initial positions (either accented and unaccented), the bilabial stop has the shortest VOT, followed by the dental plosive. As expected, due to the aerodynamic conditions, the velar stop /k/ has the longest lag.

STOP	**Average (ms)**	**Range**
/p/	15	11:35
/t/	20	10:31
/k/	36	23:48

Table 4: VOT values for word-initial unaccented stops

The measurements for the word-medial positions as given in Tables 5-6. Serbian medial stops are characterized by slightly longer VOT values compared to word-initial positions. The only exception is the accented word-medial dental /t/ whose average values are equivalent to the word-initial accented position (20 ms). In the languages which differentiate between voiceless aspirated stops and voiced unaspirated stops, like English, initial accented positions are normally characterized by longer lags (Cooper 1991). According to the data, Serbian seems to show a reverse pattern, but the differences are very slight.

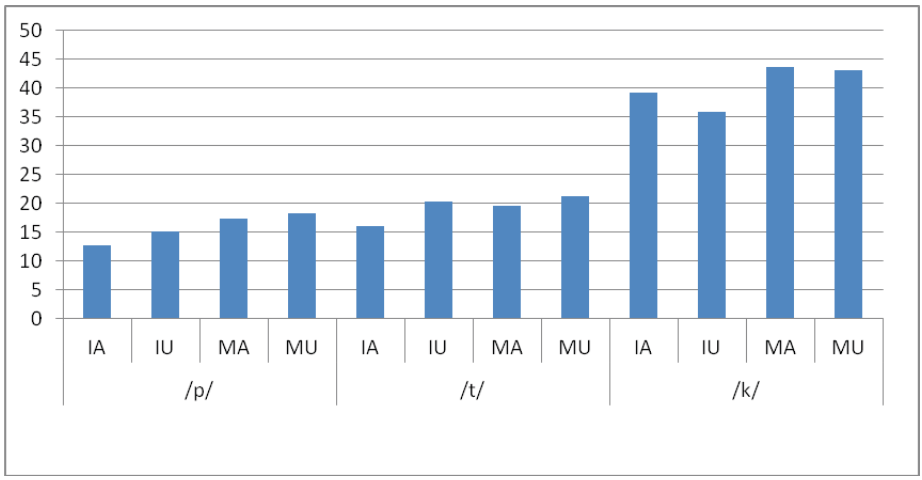
STOP	**Average (ms)**	**Range**
/p/	17	11:43
/t/	20	13:36
/k/	44	22:60

Table 5: VOT values for word-medial accented stops

STOP	**Average (ms)**	**Range**
/p/	18	11:31
/t/	21	12:32
/k/	43	28:59

Table 6: VOT values for word-medial unaccented stops

The VOT measurements in all four positions are very similar, to the effect that if shown on a gradient scale, the difference would most likely be lost. The lowest VOT values are characteristic of the bilabial stop /p/ for the four positions are 13 ms (IA), 15 ms (IU), 17 ms (MA), and 18 ms (MU) respectively. The dental /t/ shows the following values 16 ms (IA), 20 ms (IU), 20 ms (MA), 21 ms (MU). The findings for /p/ and /t/ run counter to what we would expect. The presence of accent does not seem to influence the rise in VOT values of Serbian /p t/. Moreover, medial stops, as opposed to initial stops, have higher VOT values. The results for the velar stop are as follows: /k/ 39 ms (IA), 36 ms (IU), 44 ms (MA), 43 ms (MU). In this case, accented syllables (both initial and medial) have a higher VOT value, and the lack of accent affects the decrease in VOT values. Graph 1 below presents all the numerical values for the Serbian VOTs in all four contexts relevant for the present study:



Graph 1: Cumulative VOT measurements (mean values in ms)

4. *Final remarks*

The results of the present study show that Serbian is a typical short-lag language, which is an expected finding. Long-lag VOT is not necessary in the Serbian realizations of /p t k/ due to the fact that it would not need to be implemented as a distinguishing feature between these three consonants, and their voiced counterparts /b d g/.

However, we assumed that the VOTs of the initial Serbian stops would have higher mean values compared to those in word-medial positions. The present findings are not compatible with this assumption in that the Serbian medial stops invariably show longer VOT values. Thirdly, stress was manipulated in the data to the effect of checking whether the patterns of VOT vary as a function of stress. Strikingly, Serbian accented syllables are characterized by shorter VOTs, unlike the English stressed syllables, for example. This con-

clusion goes in line with Antoniou et al's (2010) results for Greek. As a final note, the present study is only a small-scale pilot study, and for the claims to be confirmed we would need more informants and more data. Even though the number of informants is limited in the present study, it is worth pointing out that Serbian and Greek, show a similar phonetic tendency when it comes to the relationship between aspiration and distribution and the factor of stress.

References:

Abramson 1977: A. S. Abramson, Laryngeal Timing in Consonant Distinctions, *Phonetica*, 34, 295–303.

Antoniou et al 2010: M. Antoniou, C. T. Best, M. D. Tyler and C. Kroos, Language context elicits native-like stop voicing in early bilinguals' productions in both L1 and L2, *Journal of Phonetics*, 38, 640–653.

Clayton 2010: I. D. Clayton, *On the natural history of preaspirated stops*, Unpublished PhD dissertation, University of North Carolina at Chapel Hill.

Cooper 1991: A. M. Cooper, *An Articulatory Account of Aspiration in English,* Unpublished PhD dissertation, Yale University.

Čubrović 2011: B. Čubrović, Voice Onset Time in Serbian and Serbian English. Ljubljana: *English Language Overseas Perspectives and Enquiries*, VIII, Spring, 9–18.

Čubrović 2012: B. Čubrović, The Evolution of a Phonetic Phenomenon: The Case of Voice Onset Time in Serbian Intermediate EFL Learners, u: T. Paunović and B. Čubrović (eds.), *Exploring English Phonetics*, Newcastle upon Tyne: Cambridge Scholars Publishing, 47–55.

Docherty 1989: G. J. Docherty, *An experimental phonetic study of the timing of voicing in English obstruents*, Unpublished Ph. D. dissertation, University of Edinburgh.

Grosjean and Miller 1994: F. Grosjean and J. L. Miller, Going in and out of languages: an example of bilingual flexibility, *Psychological Science*, 5, 4, July 1994, 201–206.

Johnson ^{3}2012: K. Johnson, *Acoustic and Auditory Phonetics*, Malden, MA: Wiley Blackwell.

Kent and Read 1996: R. D. Kent and C. Read, *The Acoustic Analysis of Speech*, San Diego/London: Singular Publishing Group, Inc.

Lehiste and Ivić 1986: I. Lehiste and P. Ivić, *Word and Sentence Prosody in Serbocroatian*, The MIT Press, Cambridge, MA.

Lisker and Abramson 1964: L. Lisker and A. S. Abramson, A cross-language study of voicing in initial stops: acoustical measurements, *Word*, 20, 394–422.

Olson 2013: D. J. Olson, Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production, *Journal of Phonetics*, 41, 407–420.

Weismer 1979: G. Weismer, Sensitivity of voice onset measures to certain segmental features in speech production, *Journal of Phonetics*, 7, 194–204.

Whalen et al 2007: D. H. Whalen, A. G. Levitt & L. M. Goldstein. VOT in the babbling of French- and English-learning infants, *Journal of Phonetics,* 35, 341–352.

Zue 1976: V. W. Zue, *Acoustic characteristics of stop consonants: A controlled study*, *Technical Report 523*, Lexington, MA: Lincoln Laboratory/MIT.

Наслеђе 26 • 2013 • 9-17

Биљана Чубровић

АКУСТИЧКА АНАЛИЗА ИНИЦИЈАЛНИХ И МЕДИЈАЛНИХ ПЛОЗИВА У СРПСКОМ ЈЕЗИКУ

Резиме

Плозиви су консонанти који су заступљени у фонетском инвентару готово сваког нама познатог језика. Један од разлога њихове широке заступљености је вероватно и чињеница да деца ове консонанте усвајају на самом почетку језичког развоја. Овај чланак бави се безвучним плозивима /p t k/ у српском језику и то када се они налазе у иницијалној или медијалној позицији у речи, односно слогу. Одабрана класа речи, односно низова слогова предмет је анализе овог рада. Фонетска појава позната у англистичкој литератури као Voice Onset Time (VOT), тј. време наступа звучности, предмет је акустичких мерења у овом раду, јер није познато у којој мери говорници српског језика аспирују безвучне плозиве. Анализи су подвргнута четири фонетска контекста у којима се могу наћи плозиви српског језика: иницијални акцентовани слог, иницијални неакцентовани слог, медијални акцентовани и медијални неакцентовани слогови. Акустичка анализа одабраних слогова указује на то да време наступа звучности у српском језику није битан фонолошки фактор, будући да на њега не утичу ни позиција анализираног слога у речи ни место акцента. Највише просечне вредности времена наступа звучности измерене су у медијалној неакцентованој позицији, што је сасвим супротно очекиваним резултатима. Након тога, вредности су нешто ниже у медијалној акцентованој позицији. Иницијална акцентована позиција указује на најниже вредности, а благо су више оне које су измерене у иницијалним неакцентованим позицијама.

Кључне речи: плозиви, безвучни, српски језик, VOT, време наступа звучности, аспирација, иницијална позиција, медијална позиција, обрнути образац

Примљен у октобру 2013.
Ревидиран у новембру 2013.
Прихваћен у новембру 2013.