

Bojana M. Jakovljević (student doktorskih studija) UDC 811.163.41:811.111]’42  
Filozofski fakultet, Novi Sad originalni naučni rad  
medieval1985@yahoo.com

## AUDITORY PERCEPTION OF SERBIAN AND ENGLISH VOICELESS STOPS BY SERBIAN SPEAKERS AND INTERFERENCE<sup>1</sup>

*ABSTRACT:* The aim of the paper is twofold: firstly, to examine the relevance of CV formant transitions and bursts in the auditory perception of Serbian word-initial /p t k/ and English [ph th kh] by the speakers of Serbian; and secondly, to explore the relation between native language experience of listeners and their perceptual abilities in other languages. The subjects were ten native speakers of Serbian (five males and five females) and the corpus consisted of Serbian-English pairs of words, illustrating the aforementioned sets of stops before Serbian /i/ (long-rising accent) vs. English /i:/ and Serbian /ǔ/ (short-falling accent) vs. English /ɒ/. Despite the differences in the perceptual salience of the transitions and bursts between the languages in question, the results of the research point to the strong tendency of Serbian speakers to rely on the acoustic cues relevant in Serbian (L1) in the auditory perception of English (L2) voiceless stops.

*Key words:* interference, auditory perception, formant transitions, burst.

### 1. INTRODUCTION

The paper examines the acoustic cues relevant in the auditory perception of Serbian and English voiceless stops by the native speakers of Serbian and explores the presence of interference in the perception of the latter. Before presenting the results of the research directly, it is instructive to set a brief theoretical introduction in order to clarify the most relevant notions. Accordingly, the introductory part of the paper has two most significant goals: firstly, to provide a brief account of the concept of interference; and secondly, to discuss the acoustic cues relevant in the perception of stop consonants.

The notion of interference according to which learners rely on their L<sub>1</sub> experience when encountered with the system L<sub>2</sub> is one of the most influential and

---

<sup>1</sup> Slightly modified version of the paper was presented at the 3<sup>rd</sup> International Congress of Applied Linguistics, *Applied Linguistics Today: Between Theory and Practice*, held in Novi Sad, Serbia, 31 October – 1 November 2009. It was published in the form of an abstract and the original title of the paper was *L1-L2 Interference in Auditory Perception of Voiceless Stops in Serbian and English*.

longstanding concepts in the field of  $L_2$  learning. According to the Contrastive Analysis Hypothesis (CAH) where it was originally proposed, the system of  $L_2$  is filtered through the system of  $L_1$ , with  $L_1$  facilitating the learning of  $L_2$  when the target structures are similar and interfering with it when they are different or completely absent (Celce-Murcia, Brinton and Goodwin 1996: 19–20). Although the CAH has been challenged primarily because it cannot accurately predict the degree of difficulty (i. e. subtle differences between the structures of  $L_1$  and  $L_2$  are sometimes more difficult to acquire than the more prominent ones) or its directionality (i. e. whether it will be more difficult for the native speakers of one language to master a dissimilar or non-existent structure in another or vice versa), the concept of interference, nowadays more commonly referred to as transfer, is accepted as valid by most of the researchers in the field. Even more importantly, its mechanisms can be observed in the areas of both language production and language perception.

Shifting the focus to the auditory perception of stops, there are two most informative acoustic cues responsible for distinguishing between distinct places of articulation: the spectra of the burst, i. e. explosion, and formant transitions of the following vowel (Dorman et al. 1977, Hayward 2000, Johnson 2003). Equally importantly, the two acoustic cues tend to exhibit a complementary relation in a sense that the strength of one of them is complemented by the weakness of the other and vice versa (Dorman et al. 1977).

The concept of a burst refers to the abrupt release of pressure in the production of stop consonants, which is typically identified as a narrow and dark vertical spike on a spectrogram. Depending on the place of articulation, the bursts of stops differ from one another in the distribution of energy. Thus, the bursts of bilabials are the faintest with energy evenly distributed over a wide range of frequencies, the energy of alveolar bursts is generated at higher frequencies around 4000 Hz, whereas the concentration of energy of velar bursts is in the middle of the frequency range, roughly around 2500 Hz. This acoustic information can be obtained either from conventional spectrograms or spectral analysis of the burst.

Another acoustic cue responsible for the perception of stops is the shape and range of the following vowel's formant transitions, primarily the ones of  $F_2$  and  $F_3$ .<sup>2</sup> Formants represent resonant frequencies of the vocal tract, which are identified as dark horizontal bands on a spectrogram. In CV sequences, the shapes of formants before reaching the steady state are influenced by the place of articulation of the preceding consonant, which corresponds to the concept of a formant transition. In other words, the stops with different places of articulation influence the shape of the following vowel's formant transitions in different ways, which is reflected in their perception. Thus, for bilabial stops, both formant transitions tend to rise due to the

<sup>2</sup> Due to the fact that the  $F_1$  transition simply indicates the presence of complete closure and thus rises after every stop regardless of the place of articulation, it is considered as irrelevant (Hayward 2000: 184, Ladefoged 2001: 49).

fact that their locus frequencies are rather low. The  $F_3$  transition under the influence of alveolars is level, whereas  $F_2$  rises if the following vowel is front or falls if it is back (Figure 1). Finally, the transitions under the influence of velars tend to originate closely together and then they move in opposite directions, with the  $F_2$  transition falling and the  $F_3$  transition rising. In languages with aspirated voiceless stops, such as English, the interval of aspiration, i. e. the additional energy of the breath accompanying the burst, overlaps with the following vowel's formant movements (Figure 1).

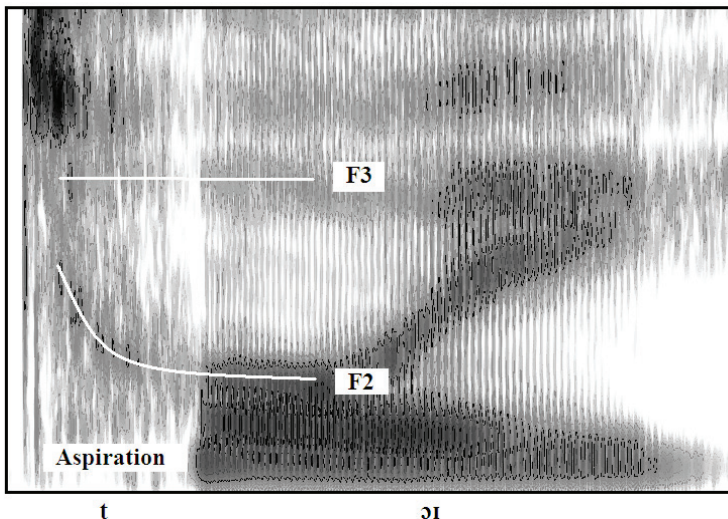


Figure 1: Spectrogram of the English word *toy* pronounced by the female speaker; the interval of aspiration and the transitions of  $F_2$  and  $F_3$  are highlighted.

It is often possible to identify a given stop solely by listening to the formant movements of the following vowel. This is the case in English whose speakers, as Hayward (2000: 184) points out, have no difficulty in distinguishing between the initial stops of e.g. *bay*, *day* and *gay* only on the basis of the CV formant transitions. The transitions are especially relevant in the perception of English voiceless stops. Due to the fact that they are characterized by considerable aspiration, the interval coinciding with the formant movements tends to be rather long and, therefore, rather informative. However, languages of the world do differ with respect to the degree of informativeness of the aforementioned acoustic cues. What is equally important, if the mechanisms of interference really work in the area of auditory perception, it is expected that the speakers who find bursts more informative in their native language will rely on the same acoustic cue in the perception of the corresponding stops in other languages.

## 2. GOALS AND METHODOLOGY

As already mentioned, the paper has two most significant goals: (a) to examine the informativeness of the burst and CV formant transitions in the auditory perception of Serbian and English voiceless stops by the native speakers of Serbian and (b) to explore the presence and nature of interference in the perception of the latter.

The subjects were 10 Serbian native speakers of the same age, 5 males and 5 females. All of them learned English as L<sub>2</sub> for 8 years, but have not received any training during the last 5 years. They were given two separate questionnaires, one for each language, and asked to listen to the isolated bursts and CV formant transitions in order to circle what they hear: /p/, /t/ or /k/.

For each language, there were 24 tokens, half of which were isolated bursts and another half the corresponding CV transitions, ordered randomly. The acoustic cues in question were extracted from Serbian-English pairs of words illustrating each voiceless stop before Serbian /i/ (long-rising accent) and /ǝ/ (short-falling accent) and English /i:/ and /ɒ/ (Table 1). The choice of one high front and one significantly lower back vowel in each language results from the fact that the vocalic environment can have considerable influence on the shapes of the formant transitions. The inter-linguistic choice of vowels results from both articulatory and acoustic similarities, i. e. their average formant frequencies in the languages in question. Lastly, both bursts and formant transitions were isolated in their entire duration. Thus, in the latter case, the starting point was the one where the formant transition begins, whereas the final point was the one where the steady state of a given vowel is reached.

Stop	Following vowel: Serbian /i/ - English /i:/	Following vowel: Serbian /ǝ/ - English /ɒ/
/p/	<i>Piza-Pisa, pila-peeler</i>	<i>pot-pot, pop-pop</i>
/t/	<i>Tina-teeny, Tića-teacher</i>	<i>top-top, tok-tock</i>
/k/	<i>Kina-keener, kipi (Present-3sg)-keeper</i>	<i>konjak-cognac, koš-cosh</i>

Table 1: The list of Serbian-English pairs of words from which the bursts and the corresponding CV formant transitions of word-initial voiceless stops were isolated

Eventually, there were 120 listenings of bursts and 120 listenings of the corresponding CV formant transitions in each language. The total of correct identifications was calculated by simple addition of the correct identifications made by all the subjects, with the relevant results expressed in percentage terms.

## 3. THE RESULTS OF THE RESEARCH

The analysis of the data shows that the native speakers of Serbian find bursts considerably more informative than the corresponding CV formant transitions in the perception of Serbian word-initial /p t k/ (Table 2). The number of correct identifications on the basis of the burst is 88 out of maximum 120, which is approximately 73%. As for the following vowel's formant transitions, the stops were correctly perceived in 47 out of 120 listenings, which makes up only 39% of the analyzed tokens.

	/p/	/t/	/k/	TOTAL
Burst	27	30	31	88 out of 120
CV transitions	23	16	8	47 out of 120

Table 2: The number of correct identifications of Serbian word-initial /p t k/ on the basis of the burst and the corresponding CV formant transitions; the maximum of correct identifications is 120 in each case.

Considering the place of articulation, the number of correct identifications on the basis of the burst increases from /p/ to /t/ and /k/, whereas in the case of the formant transitions, the ordering is reversed. This may point to the complementary relation between the two acoustic cues. However, what poses difficulties is the ratio between the numbers of correct identifications characterizing different places of articulation. While the bursts of /p t k/ appear to be more or less equally informative, the same observation does not apply to the formant transitions. The transitions under the influence of /k/ are significantly less informative than the ones under the influence of /t/ and especially /p/. Thus, it is possible to conclude that the true complementary relation between the relevance of Serbian bursts and the corresponding formant transitions does not really exist. In other words, the strength of one acoustic cue for different places of articulation is not complemented by the equal weakness of another (Chart 1). The same tendency is observed in Bakran's study of the perception of Croatian voiceless stops (1996: 89–90).

Another distinction between the two acoustic cues in Serbian is their sensitivity to the vocalic environment (Table 3). The number of correct identifications of word-initial /p t k/ on the basis of the burst is not significantly different in the contexts when the following vowel is /i/ and /ö/. Regardless of the place of articulation, the distinction does not surpass 13%, which points to rather low sensitivity of the bursts to the vocalic context. As for the perception on the basis of the formant transitions, the choice of the following vowel is a much more significant factor. This observation is particularly obvious in the case of /k/, whose identification in front of /i/ is approximately 67% more successful than in front of /ö/. Such behavior can be explained by the general tendency of velar stops to adjust their articulation to the following vowel. The opposite behavior

is exhibited by /t/ whose number of correct identifications in front of /i/ is 40% smaller in comparison to the context where the following vowel is /ö/. Lastly, the perception of /p/ proves to be least sensitive to the vocalic environment, which is probably caused by rather fixed place of articulation exhibited by bilabial stops in general.

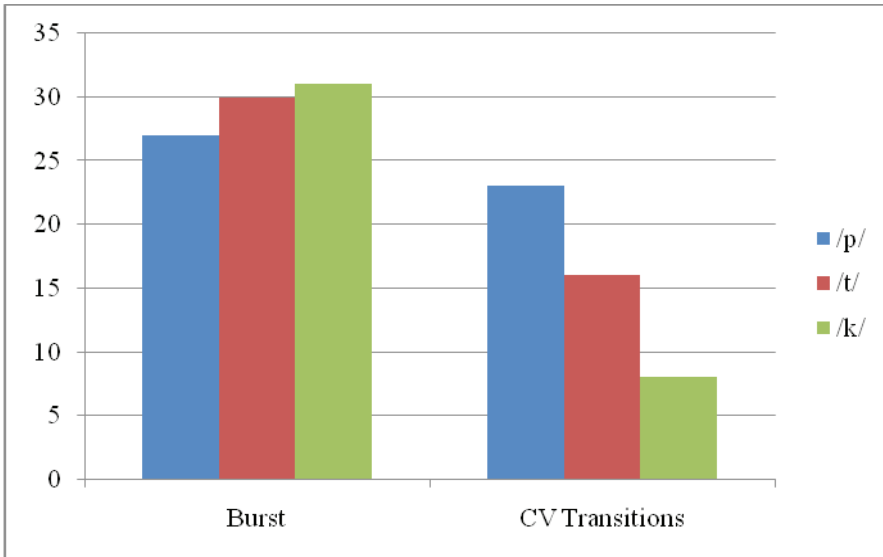


Chart 1: The relevance of the burst and corresponding CV formant transitions in the perception of different places of articulation of Serbian voiceless stops

Burst						CV transitions					
/p/		/t/		/k/		/p/		/t/		/k/	
/i/	/ö/	/i/	/ö/	/i/	/ö/	/i/	/ö/	/i/	/ö/	/i/	/ö/
14	13	14	16	16	15	11	12	6	10	6	2

Table 3: The number of correct identifications of Serbian word-initial /p t k/ on the basis of the burst and the corresponding CV formant transitions in different vocalic environments

Shifting the focus to the perception of English voiceless stops, the results of the research indicate that Serbian native speakers find the bursts of English word-initial /p t k/ more or less equally informative as the bursts of the corresponding Serbian counterparts, which points to the strong influence of their native language. Out of 120 listenings of bursts, there were 86 correct identifications, which makes up approximately 72% of the analyzed tokens (Table 4). The influence of Serbian is also observed in the case of CV formant transitions whose informativeness is significantly

lower in comparison to the one of bursts. According to the conducted measurements, the number of correct identifications on the bases of transitions is 32, which constitutes approximately 27% of the tokens. As in the case of Serbian voiceless stops, the number of correct identifications on the basis of the burst increases from /p/ to /t/ and /k/, but there are no significant differences between different places of articulation.

	/p/	/t/	/k/	TOTAL
Burst	27	29	30	86 out of 120
CV transitions	15	11	6	32 out of 120

Table 4: The number of correct identifications of English word-initial /p t k/ on the basis of the burst and the corresponding CV formant transitions; the maximum of correct identifications is 120 in each case.

Interestingly enough, Serbian speakers find the CV formant transitions characterizing English voiceless stops even less informative than the corresponding transitions characterizing their Serbian counterparts. The comparison of the results for Serbian and English shows that the formant transitions under the influence of English /p t k/ are approximately 12% less informative. What appears to be similar in the languages in question is the relevance of the formant movements with respect to the place of articulation. Thus, the transitions under the influence of /k/ are far less informative than the ones under the influence of /t/ and particularly /p/, which indicates that, as in Serbian, there is no complementary relation between the relevance of the burst and the corresponding formant transitions.

Another similarity is related to the sensitivity of the two acoustic cues to the vocalic environment. As in Serbian, the bursts of English voiceless stops are considerably less sensitive to the choice of the following vowel in comparison to the corresponding CV formant transitions (Table 5). Thus, the difference in the number of correct identifications in the contexts where the following vowel is /i:/ and /ɒ/ does not surpass 7%, whereas in the case of /k/, there are no differences whatsoever. As for the formant transitions, the identification of /k/ is considerably more successful if the following vowel is /i:/, the alveolar stop /t/ exhibits the opposite behavior, whereas the formant transitions under the influence of /p/ are, as it is expected, least sensitive to the vocalic context.

Burst						CV transitions					
/p/		/t/		/k/		/p/		/t/		/k/	
/i:/	/ɒ/	/i:/	/ɒ/	/i:/	/ɒ/	/i:/	/ɒ/	/i:/	/ɒ/	/i:/	/ɒ/
13	14	14	15	15	15	8	7	4	7	4	2

Table 5: The number of correct identifications of English word-initial /p t k/ on the basis of the burst and the corresponding CV formant transitions in different vocalic environments

## 4. CONCLUSIONS

Despite the fact that Serbian and English differ with respect to the relevance of the burst and the corresponding CV formant transitions in the perception of word-initial voiceless stops, Serbian native speakers show significantly heavier reliance on the burst spectra in both languages (Chart 2). Interestingly enough, although the transitions under the influence of English /p t k/ are longer due to aspiration and thus expected to be more relevant, Serbian speakers find them approximately 12% less informative than the corresponding transitions under the influence of their Serbian counterparts.

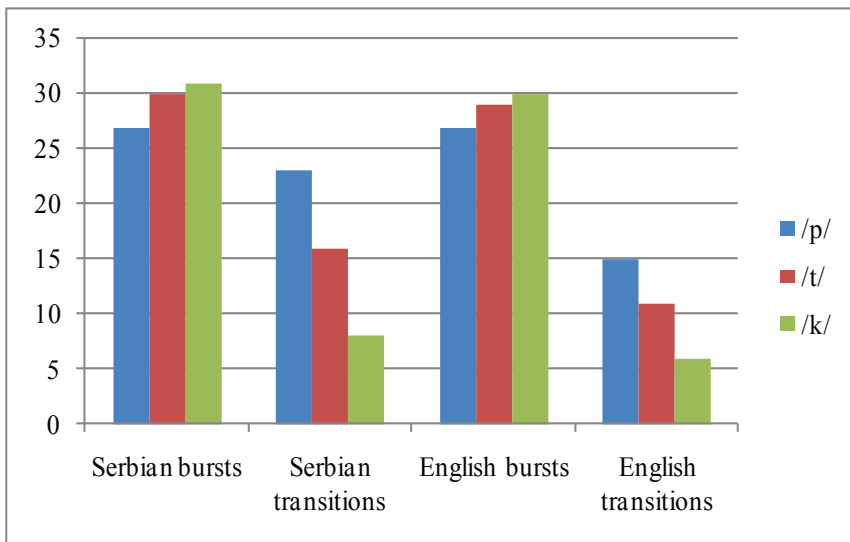


Chart 2: The relevance of the burst and corresponding CV formant transitions in the perception of Serbian and English word-initial stops by the native speakers of Serbian

Summarizing the obtained observations, it appears that the mechanism of interference works equally well in the areas of both language production and language perception. Thus, when perceiving the sounds of a foreign language, listeners tend to rely on the phonetic cues that are relevant in their native language. Other general tendencies observed in the perception of stops of both languages are the lack of complementary relation between the relevance of the burst and the corresponding formant transitions for different places of articulation as well as significantly higher sensitivity of the latter to the vocalic context, primarily in the case of alveolars and velars. The perception of word-initial /t/ is thus considerably more successful in front of a back vowel, whereas /k/ exhibits the opposite behavior in both languages.



REFERENCES

- Bakran, J. 1996. *Zvučna slika hrvatskoga govora*. Zagreb: Ibis grafika.
- Celce-Murcia, M., Brinton, D. and Goodwin, J. 1996. *Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Languages*. Cambridge: Cambridge University Press.
- Dorman, M. F., Studdert-Kennedy, M. and Raphael, L. J. 1977. Stop-consonant recognition: Release bursts and formant transitions as functionally equivalent, context-dependent cues. *Perception and Psychophysics* 22: 109–122.
- Hayward, K. 2000. *Experimental Phonetics: An Introduction*. London: Longman.
- Johnson, K. 2003. *Acoustic and Auditory Phonetics*. Oxford: Blackwell.
- Ladefoged, P. 2001. *Vowels and Consonants: An Introduction to the Sounds of Languages*. Oxford: Blackwell.

Bojana M. Jakovljević

PERCEPTION AUDITIVE DES OCCLUSIVES SOURDES EN SERBE ET EN ANGLAIS  
PAR LES LOCUTEURS NATIFS DE SERBE ET INTERFÉRENCE

*Résumé*

Le travail analyse les signes acoustiques importants pour la perception des occlusives sourdes initiales en serbe et en anglais par les locuteurs natifs de serbe et aborde l'influence de la langue serbe sur la perception des occlusives anglaises. Bien que les transitions formantiques CV jouent un rôle décisif dans la perception des occlusives anglaises, les résultats de la recherche montrent que les locuteurs serbes s'appuient sur l'explosion dans la perception des occlusives serbes et anglais, ce qui montre une interférence importante de leur langue maternelle. Les autres tendances sont: absence de la relation complémentaire entre la pertinence de l'explosion et les transitions formantiques correspondantes concernant les différents points d'articulation et la sensibilité plus importante des transitions formantiques au milieu vocalique, ce qui est surtout noté en cas des occlusives alvéolaires et vélares.

*Mots clés*: interférence, perception auditive, transitions formantiques, explosion.