

SerBial: VLD RT Database of 2400 Serbian Words in Two Alphabets Reveals Interference of Codes

Dušica Filipović Đurđević¹, Milica Popović Stijačić², Sunčica Zdravković³, Maja Savić³, Ksenija Mišić⁴, Sara Božić¹

¹ Department of Psychology, Faculty of Philosophy, University of Belgrade, Serbia

² Department of Psychology, Faculty of Media and Communications, University Singidunum, Serbia

³ Department of Psychology, Faculty of Philosophy, University of Novi Sad, Serbia

⁴ Faculty of Philology, University of Belgrade, Serbia

Background: Speakers of Serbian are fluent in two alphabetic scripts –Cyrillic and Roman. Both alphabets code for the same phonology in a shallow orthographic mapping of one phoneme onto one grapheme [1]. Some of the graphemes are unique for one alphabet (e.g. Cyrillic *Г, Ф, Л*, and Roman *G, F, L*), some are shared (e.g. *A, E, M* are written and pronounced in the same way across alphabets), whereas some exist in both alphabets, but are pronounced differently (e.g. *H* would map onto /n/ in Cyrillic, but /h/ in Roman). Previous studies revealed interference from the non-target alphabet in visual word processing both in the single- and mixed- alphabet context, suggesting that two alphabets within a single language are simultaneously activated, similar to two languages in bilingual processing [2]. However, the interference was demonstrated with a specific type of items (those that could be fully mapped onto both alphabets, with a lexical reading in only one). Following the trend of collecting large datasets, we presented large number of Serbian words in two single-alphabet contexts. We operationalized the interference from the other alphabet as the difference between two summed diphone frequencies: one calculated by applying the phonological mapping within the target alphabet, and another calculated by applying the phonological mapping across the two alphabets (see Table 1 for illustration).

Method: We selected 2400 Serbian words and created the same number of pseudowords matched for length and summed bigram frequency. We presented each item in both Cyrillic and Roman alphabet, albeit in different sessions, counterbalancing the alphabet across items and participants. We created 32 lists of stimuli (16 in Roman and 16 in Cyrillic). Participants were presented with a visual lexical decision task to items from one list in Cyrillic and a different list in Roman, with at least two weeks between the two sessions and with counterbalanced order of alphabets across the two sessions. The ongoing data collection will continue during the spring, and we are presenting the results based on the data collected from 306 participants, most of which participated in both Roman and Cyrillic conditions (NCyr = 235; Nrom = 232; with ~15 participants per list).

Results: The results of the best-fitting linear mixed-effects regression model with maximum structure of random effects justified by the data are illustrated in Figure 1. Words presented in Roman alphabet were recognised faster than words presented in Cyrillic. Additionally, response latencies decreased during the experiment, increased with an increase in word length in phonemes, and decreased with an increase in (log) lemma frequency. Crucially, we observed slowing of the responses with an increase in the level of interference from the non-target alphabet, which was present for words from the lower and middle range of word length in phonemes (~4-8 phonemes in length), and absent for longer words (above ~8 phonemes in length).

Discussion: In addition to classical effects of word length and lemma frequency, we observed inhibitory effect of the interference from the non-target alphabet. As interference was operationalized as the summed frequency of bigrams that had phonological mapping in non-target alphabet, we interpret this finding as evidence of simultaneous activation of two alphabets, analogous to simultaneous activation of two languages in bilingual processing [2]. This is in accordance with previous research on Serbian bialphabetic processing [3–5] and the first demonstration of interference from non-target alphabet on sublexical level. In addition to this novel finding, our database will be useful in providing data on response latencies and response accuracy for a large number of items in Serbian, an underrepresented

language regarding the large available databases. Moreover, it will be the first large database resource entailing data on words presented in two alphabets within the same language, enabling further insights in the future.

Phonology in Roman mapping	Orthographic form (and bigrams)	Phonology in Cyrillic mapping
/lopta/ (a playing ball)	LOPTA	/
/lo/	LO	/
/op/	OP	/or/
/pt/	PT	/rt/
/ta/	TA	/ta/

Summed diphone frequencies within Roman alphabet: F/lo/ + F/op/ + F/pt/ + F/ta/
Summed diphone frequencies across alphabets: F/lo/ + F/op/ + F/pt/ + F/ta/ + F/or/ + F/rt/ + F/ta/
Interference from non-target alphabet: F/or/ + F/rt/

Table 1: Illustration of the operationalization of the interference from non-target alphabet.

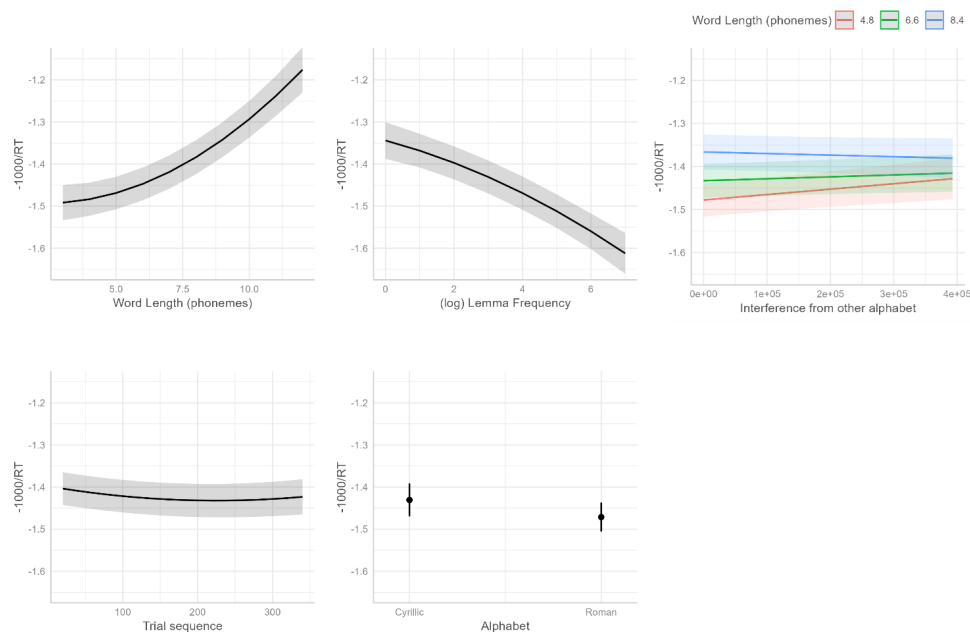


Figure 1: The effect of word length in phonemes, (log) lemma frequency, interference from non-target alphabet, trial sequence, and alphabet to response latencies in visual lexical decision task.

References

- [1] Helms-Park, R., Tucker, S.-K., & Dronjic, V. (2016). *From proto-writing to multimedia literacy: scripts and orthographies through the ages*. Routledge Publishers.
- [2] Heuven, W. V., Dijkstra, T., & Grainger, J. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, 39(3), 458–483. <https://doi.org/10.1006/jmla.1998.2584>
- [3] Feldman, L., & Turvey, M. (1983). Word recognition in Serbo-Croatian is phonologically analytic. *Journal of Experimental Psychology: Human Perception and Performance*, 9(2), 288–298. <https://doi.org/10.1037/0096-1523.9.2.288>
- [4] Đurđević, D. F., & Feldman, L. (2024). Do readers exert language control when switching alphabets within a language? *Journal of memory and language*, 139. <https://doi.org/10.1016/j.jml.2024.104546>
- [5] Đurđević, D. F., Milin, P., & Feldman, L. (2013). Bi-alphabetism: A window on phonological processing. *Psihologija*, 46(4), 419–436. <https://doi.org/10.2298/PSI1304421F>