

10 Phonological and morphological analysis by skilled readers of Serbo-Croatian

Laurie B. Feldman

Abstract

Special properties of Serbo-Croatian were exploited in order to investigate the ability of skilled readers to appreciate the phonological and morphological components of Serbo-Croatian words. In the first experiment bivalent letter strings (namely, those composed exclusively of characters that appear in both the Roman and Cyrillic alphabet) were presented in a lexical decision task. Latencies were prolonged relative to unique alphabet transcriptions of those same words and the outcome was interpreted as a phonological effect. In the second experiment, decision latencies to phonologically bivalent root morphemes with alphabetically bivalent and unique inflectional affixes were compared. Results of ambiguity obtained only when the affix as well as the base morpheme were bivalent suggesting that recognition of the base morpheme alone does not govern decision latencies. In the third experiment, phonologically bivalent words and their unique alphabet controls were presented in the context of semantic associates which were printed in an alphabet that was either congruent or incongruent with the target. For bivalent words, semantic association and alphabetic congruency facilitated in an additive fashion but for unique alphabet controls semantic association but not alphabetic congruency was significant. Results indicate distinct loci for associative and alphabetic effects. Finally, in the fourth experiment, facilitation due to repetition of a base morpheme was observed but repetition of the same orthographic pattern in unrelated words produced no facilitation. Collectively, the studies provide evidence for phonological and morphological analysis by skilled readers of Serbo-Croatian.

Serbo-Croatian, the major language of Yugoslavia, possesses two distinctive properties that have been exploited as tools in the study of reading.

Table 1. Types of letter strings and their lexical status.

Composition of letter string	Phonemic interpretation		Meaning
<i>Ambiguous and common</i> ¹			
BETAP	Roman	/betap/	meaningless
	Cyrillic	/vetar/	wind
POP	Roman	/pop/	priest
	Cyrillic	/ror/	meaningless
POTOP	Roman	/potop/	flood
	Cyrillic	/rotor/	motor
PAJOC	Roman	/pajotc/	meaningless
	Cyrillic	/rajos/	meaningless
<i>Common</i>			
MAMA	Roman	/mama/	mother
	Cyrillic	/mama/	mother
TAKA	Roman	/taka/	meaningless
	Cyrillic	/taka/	meaningless
<i>Unique and common</i> ²			
VETAR	Roman	/vetar/	wind
	Cyrillic	impossible	
ПОП	Roman	impossible	
	Cyrillic	/pop/	priest
ROTOR	Roman	/rotor/	motor
	Cyrillic	impossible	
ПОТОП	Roman	impossible	
	Cyrillic	/potop/	flood
RAJOS	Roman	/rajos/	meaningless
	Cyrillic	impossible	
ПАЈО П	Roman	impossible	
	Cyrillic	/pajots/	meaningless

¹ Phonologically bivalent letter strings.

² Phonologically unequivocal controls.

Adapted with permission of the American Psychological Association from Feldman and Turvey, 1983.

word, problems of equating familiarity, richness of meaning, length and number of syllables are eliminated.³ To reiterate, the letter strings exemplified by BETAP and VETAR are the same word and, therefore, identical in all respects but one, namely, the number of phonological interpretations.

Phonological analysis in skilled readers

When bi-alphabetic adult readers of Serbo-Croatian performed a lexical decision task (i.e., Is this letter string a word by either a Roman or a

Cyrillic reading?). single letter strings composed of ambiguous and common characters (i.e., those letter strings that could be assigned both a Roman and a Cyrillic alphabet reading) typically incur longer latencies than the phonologically unequivocal alphabet transcription of the same word. This outcome has been reported both in a mixed alphabet context where the lexical interpretation of a letter string was sometimes in Roman and sometimes in Cyrillic (Feldman and Turvey, 1983; Lukatela *et al.*, 1980) and a pure alphabet context where the lexical interpretation was always in Roman (Feldman, 1983; Lukatela *et al.*, 1978). The effect of phonological ambiguity is significant both for bivalent words and pseudowords but it is more robust for words. In characterizing the effect of ambiguity in lexical decision, several outcomes prove essential. First, the effect of phonological ambiguity did not vary as a function of word familiarity. For each word, decision latency to its phonologically unequivocal form was used as an index of familiarity and was correlated with the difference in decision latency between the bivalent and unequivocal forms of the word. In lexical decision, that correlation approached zero (Feldman and Turvey, 1983).⁴ Second, words composed entirely of common letters (with no ambiguous or unique letters) such as MAMA were accepted as words no more slowly than letter strings that included common and unique letters. Likewise, pseudowords composed entirely of common letters such as TAKA were rejected as words no more slowly than letter strings that included common and unique letters. Note that the distinction between common and ambiguous letters derives from their phonology: each type of letter occurs in both alphabets but only the latter have two phonemic interpretations. The foregoing discrepancy of outcomes suggest that it is *phonological* bivalence rather than a visually based alphabetic bivalence that governs the slowing of decision latencies (see Lukatela *et al.*, 1978, 1980, for a complete discussion). Third, lexical decision latencies to letter strings composed entirely of ambiguous and common letters were always slowed whether both alphabet readings yielded a positive response such as 'POTOP' (Lukatela *et al.*, 1980) or a negative response such as 'PAJOC' (Feldman, 1981; Lukatela *et al.*, 1978, 1980) or the Cyrillic reading and the Roman reading yielded opposite responses such as 'BETAP' or 'POP' (Feldman and Turvey, 1983; Lukatela *et al.*, 1978, 1980). This outcome invalidates a decision stage account of the detriment due to bivalence that posits some type of post-lexical interference between conflicting lexical judgments. Moreover, insofar as lexical decision is alleged to be susceptible to decision-stage influences in a way that naming is not (Balota and Chumbley, 1984; Seidenberg, Waters, Sanders and Langer, 1984) it is noteworthy that the detriment due to bivalence is generally enhanced in naming relative to lexical decision. Finally, the difference in decision

latency between the bivalent and unequivocal forms of a word increased as the number of ambiguous (but not common) characters increased (Feldman and Turvey, 1983). It was eliminated, however, by the presence of a single unique letter (Feldman, Kostić, Lukatela and Turvey, 1983). These findings imply that a segmental phonology is assembled from an analysis of a letter string's component orthographic structure and that sometimes (multiple) phonological interpretations are generated. The foregoing results of lexical decision experiments with phonologically bivalent letter strings provide evidence that access to the lexicon in Serbo-Croatian necessarily involves an analysis that is (1) sensitive to phonology and component orthographic structure; (2) is not sensitive to the lexical status of the various alphabetic readings. These results have been interpreted as evidence for an assembled segmental phonology in Serbo-Croatian.

In an attempt to understand conditions under which phonological codes and lexical knowledge do interact in Serbo-Croatian, we have begun to explore associative priming of phonologically bivalent words (Feldman, Lukatela, Katz and Turvey, forthcoming). In this procedure, target words are sometimes presented in the context of another word that is associated with it and decision latencies to the target with and without its associate are compared. Phonologically bivalent words and the unequivocal alphabet transcription of those same words were presented as targets in a lexical decision task. Half of the bivalent targets were words by the Cyrillic reading and half were words by their Roman reading. On some proportion of trials, target words were presented in the context of another word that was associatively related to it and preceded it by 700 ms. Sometimes, the alphabet of the associate was congruent with the alphabet in which the target reading was a word. Sometimes the associate and the target reading were alphabetically incongruent. Results showed significant facilitation in the context of associates, evidence of lexical mediation. More interestingly, decision latencies for bivalent letter strings that are words by one of their alphabet readings were reduced less when those words are preceded by an associate printed in the other, incongruent alphabet than when the associate was printed in the same alphabet as the word reading of the target. This outcome suggests alphabetic congruency as a second source of facilitation. For example, bivalent BETAP which means WIND when read as Cyrillic was preceded by the word for STORM. Inspection of word means in Table 2 reveals that target decision latencies for BETAP type words were 64 ms faster when it was preceded by the Cyrillic form of the word for STORM than by the Roman form of the same word. By contrast, target decision latencies for the same words written in their phonologically unequivocal form were facilitated equally by the prior presentation of an associated word printed in either the congruent or incongruent alphabet.

For example, WIND written in Roman, namely VETAR, is phonologically unequivocal and decision latencies were not significantly different when the word for STORM appeared in its Cyrillic or Roman form. Likewise for pseudowords, alphabet congruency had no effect (see Table 2).

Table 2. Lexical Decision (ms) to bivalent words and their unequivocal controls in the context of alphabetically congruent and alphabetically incongruent associates.

	Bivalent (BETAP)	Unequivocal (VETAR)
Alphabet of associate:		
Congruent	709	672
Incongruent	775	685
(No associate)	845	765

From Feldman, Lukatela, Katz, and Turvey (in preparation).

In summary, lexical decision latencies for phonologically bivalent letter strings are reduced significantly more when preceded by associates that are alphabetically congruent with the word reading of the letter string, than by associates that are not congruent. By contrast, decision latencies for phonologically unequivocal letter strings are not influenced by the alphabet of the associate. Associative and alphabetic sources of facilitation can be identified. Whereas facilitation by association occurs for all the words and is assumed to be lexical in origin, facilitation by alphabet congruency of associate and target was important *only* for bivalent letter strings. The special dependency of alphabetic congruency on ambiguity suggests that alphabetic priming and phonological ambiguity have a common origin.

In summary, studies of phonological ambiguity indicate that skilled readers of Serbo-Croatian analyse words phonologically. In judging letter strings composed exclusively of ambiguous and common letters for a lexical decision, adult readers appear to assign a phonological interpretation (or several) to each character (Feldman and Turvey, 1983). At the same time, the alphabet in which a prior occurring associate is printed appears to bias the generation or the evaluation of various phonological interpretations of a bivalent letter string. An analogous effect is absent in phonologically unequivocal words and in all pseudowords.

Morphological analysis in skilled readers

The effect of phonological ambiguity has provided a means to evaluate the analytic skills of readers with respect to morphological components. As

noted above, the Serbo-Croatian language, in a manner that is characteristic of Slavic languages generally, makes extensive use of inflectional and derivational morphology. A noun can appear in any of seven cases in the singular and in the plural where the inflectional affix varies according to its gender, number, and case. For example, the words STAN and KORA which mean 'apartment' and 'crust' respectively in nominative case can be inflected into six other cases in the singular and in the plural and different inflectional affixes mark each case (with some redundancy of affixes). Similarly, derived forms for 'little apartment' or 'thin crust' can be generated by adding one of the diminutive affixes (namely, ČIĆ, ICA, ENCE, AK) to the base word to produce STANČIĆ and KORICA respectively. The prevalence of inflectional and derivational formations in Serbo-Croatian are evidence of its productiveness (see Table 3).

Table 3. Examples of morphologically-related words formed with the base morpheme 'PIS' meaning 'write'.

Example	Derivational prefix	Base morpheme	Derivational suffix	Inflectional suffix	Meaning
OPIS	O	PIS			description
OPISI	O	PIS		I	descriptions (nom. plural)
PIŠEM	PIŠ			EM	I write (1p. sing)
PIŠETE		PIŠ		ETE	you write (2p. plural)
PISAC		PIS	AC		writer
PISCIMA		PIS	C	IMA	writers (dat. plural)
PISMO		PIS	MO		letter
POPIS	PO	PIS			inventory
POTPIS	POT	PIS			signature
SPISAK	S	PIS	AK		list

† All words are in nominative singular unless otherwise noted.

One way in which sensitivity to morphological constituents is construed is in terms of a morphological parser that operates prior to lexical access such that affixes are stripped from a multimorphemic word and the base morpheme serves as the primary unit for lexical search (see Caramazza, Miceli, Silveri and Laudanna, 1985). Frequency of the base unit and the whole word as well as the difficulty of segmenting the appropriate base unit figure significantly in decision latency (Taft, 1979; Taft and Forster, 1975). In one experiment (Feldman, Kostić, Lukatela and Turvey, 1983) the effect of phonological ambiguity was exploited to assess whether the base morpheme or the whole word serves as the unit for lexical access of

inflected words in Serbo-Croatian. Words were presented in nominative and dative case for a lexical decision. Words were selected so that the nominative case and the base morpheme (i.e., nominative minus inflectional affix for most singular nouns) were phonologically bivalent in the Cyrillic alphabet and phonologically unequivocal in Roman. For example, the nominative case of the word meaning VEIN is composed entirely of ambiguous and common letters when printed in Cyrillic (i.e., BEHA) and is therefore phonologically bivalent. In Roman, by contrast, it comprises unique and common letters (i.e., VENA) and is, therefore, phonologically unequivocal. Importantly, in the dative case, neither alphabet rendition is bivalent because the inflectional affixes for words of its class are the phonemes /u/ and /i/ both of which are represented by a unique letter in each alphabet, although the base morpheme of the Cyrillic form (i.e., BEH) is still bivalent.

The major outcome of that experiment was a significant interaction of alphabet and case. The difference in latency between dative nouns presented in Cyrillic and Roman was -28 ms which was not significant whereas the difference between nominative nouns was 304 ms which was significant. In that dative nouns always included a unique letter, it appears that the effects of phonological bivalence do not occur if letter strings composed of ambiguous and common characters contain even one unique character. Importantly, in that experiment, the unique character always constituted an inflectional morpheme. Stated in terms of morphological units, the outcome of that experiment was that an inflectional affix composed of a unique character and appended to a bivalent base morpheme cancelled the detriment due to ambiguity. Evidently, the reader could use the alphabet designation of the inflectional affix to assign a reading to the base morpheme. In conclusion, bivalence defined on the word but not on the base morpheme alone slowed performance on a lexical decision task. This outcome indicates that lexical access of inflected nouns is not restricted to information in the base morpheme unit. Rather, it encompasses the *entire* word.

An alternative perspective on a reader's appreciation of morphology assumes that lexical entries are accessed from whole word units and that the principle of organization among lexical entries or the lexical representations themselves capture morphological structure. The final experiment (Feldman and Moskovljević, in press) exploits the complex derivational morphology of Serbo-Croatian to provide further evidence that whereas the morphological structure of words is accessible to the skilled reader, lexical entries are not accessed from base morphemes. The experiment incorporated a comparison of three types of nouns all in nominative case: First, base forms (e.g., STAN, KORA). Second, the diminutive form of

those same nouns which, as described above, is formed (productively) by adding one of the suffixes ČIĆ, ICA, ENCE, AK to the base form (e.g. STANČIĆ, KORICA) where choice of suffix is constrained by gender of the noun. And third, an unrelated monomorphemic word whose construction inappropriately suggests that it contains the same base form and a diminutive affix (e.g., STANICA, KORAK). The latter are referred to as pseudodiminutive nouns. The examples mean 'station' and 'step' respectively.

The experimental design was a variation on the primed lexical decision task borrowed from Stanners and his colleagues (Stanners, Neiser, Hernon and Hall, 1979) and known as repetition priming. In the present adaptation of the task, base forms appeared as target words and they were preceded 7 to 13 items earlier in the list by a prime which was either the identical word again in its base form, its diminutive or a pseudodiminutive form. Decision latency to the target as a function of which type of prime preceded it was examined. In addition, decision latencies to the first presentation of the word in its base, diminutive, and pseudodiminutive forms were compared. Results are summarized in Table 4.

Table 4. Lexical Decision (ms) to target words preceded by identity, diminutive, or pseudodiminutive primes.

Prime		Target		Type of prime
STAN	610	STAN	563	Identity
STANČIĆ	754	STAN	585	Diminutive
STANICA	718	STAN	609	Pseudodiminutive

From Feldman and Moskovljević (in press).

Decision latencies on primes were fastest for base forms, followed by pseudodiminutives and lastly, diminutives. The pattern corresponded with that predicted by frequency and provided no evidence that monomorphemic pseudodiminutive forms were slowed by an inappropriate parsing of morphemic structure. In addition, latencies for base and diminutive forms correlated significantly and neither correlated with pseudodiminutive forms. An examination of target latencies provided further evidence that pseudodiminutive words are not associated with an inappropriate base morpheme (and affix) whereas true morphological relationships are appreciated. Decision latencies to target words that were preceded by pseudodiminutive words were as slow as target words presented for the first time. In contrast, both base word and diminutive primes significantly reduced target decision latencies. In summary, results in the repetition priming variation of lexical decision showed significant facilitation for morphological relatives and no facilitation for unrelated pseudodiminutive words. In

light of the claim that semantic relatedness of prime to target does not facilitate target decision latencies at lags as long as those introduced in the present task (Dannenbring and Briand, 1982; Henderson, Wallis and Knight, 1984), the foregoing results are interpreted as morphological in nature. In conclusion, the present experiment showed that the skilled reader of Serbo-Croatian is sensitive to morphological structure as evidenced by the results in repetition priming but offered no evidence that morphological analysis entails decomposition to a base morpheme prior to lexical access.

In summary, an examination of results from lexical decision and naming tasks that take advantage of the bi-alphabetic condition in Serbo-Croatian provides evidence that skilled reading in Serbo-Croatian proceeds with reference to phonology. Specifically: (1) Skilled readers are slowed when a letter string is phonologically bivalent relative to when it is phonologically unequivocal. (2) The alphabet congruency of a prior-occurring associate can speed decision latencies for phonologically bivalent (but not unequivocal) words. Moreover, it appears that phonological bivalence is defined on the entire word, not in the base morpheme alone which suggests that (3) Skilled readers do not attempt lexical access from an isolated base morpheme. Concurrently, they consider its affix. Failure to find evidence that base morphemes are the units for lexical access should not be construed as a claim against morphological analysis by the reader, however. The results from repetition priming indicate that prior presentation of a morphological relative but not of a visually similar word facilitates decision latency to a target. The foregoing results support the claim that the skilled reader of Serbo-Croatian analyses words both phonologically and morphologically.

Acknowledgements

This research was supported by funds from the National Academy of Sciences and the Serbian Academy of Sciences to Laurie B. Feldman; by NICHD Grant HD-01994 to Haskins Laboratories and by NICHD Grant HD-08495 to the University of Belgrade.

Notes

¹ The introduction of two alphabets into Yugoslavia reflects the influence of the Orthodox Church in the Eastern regions and the Catholic Church in the Western regions. The Cyrillic script is probably an adaptation of the Greek uncial alphabet of the 9th century AD and the Roman script is a variation of the Latin alphabet which was also derived from the Greek, probably via Etruscan (Diringer, 1948). In both cases, the scripts had to be adjusted to represent sounds not

- present in the Greek language and several mechanisms have been identified. (1) Combining two or more characters to represent a single phoneme such as DZ and, arguably, LJ and NJ. (2) Adding a diacritical mark to an existing letter to form a new letter such as Ć, Č, Š. The creation of new letters by inclusion of a diacritic is particularly prevalent in the adaptation of the Roman script to languages whose repertoire of phonemes differs greatly from the Latin. Palatal-alveolar fricatives and affricates are represented in this fashion in many Slavic languages, including Serbo-Croatian (Wellisch, 1978). (3) Taking an existing symbol which was not used in the new language to represent a phoneme not present (or represented by multiple symbols) in the old language. For example, Roman C became /ts/ and Roman K remained /k/. (4) Borrowing characters from other scripts. Insofar as particular adaptations were made independently in each alphabet and the shape of some letters (e.g., D, S, R) were modified slightly in the transition to Latin (Diringer, 1948) the intersection of the two alphabet sets represents a complex of factors.
- ² One consequence of the consistent mapping of grapheme to phoneme is that many dialectal variations are represented in writing such that spelling as well as pronunciation can vary from region to region. For example, the word that means MILK is MLEJKO in the dialects near Belgrade and is MLIKO in dialects along the Dalmatian Coast. It is important to note that the orthography fully specifies segmental phonology but that accent (rising/falling; long/short) is not represented. While vowel accent may differentiate between two semantic interpretations of a written letter string, this distinction is often ignored especially in the dialects of the larger cities, however (Magner and Matejka, 1971).
- ³ By law, all elementary school students must demonstrate competence to read and write in both alphabets. With the exception of liturgical text, which is relatively uncommon, the choice of alphabet is not systematic across genres of printed material. Therefore, it can be assumed that the Roman and Cyrillic forms of a word are equally familiar to the skilled reader.
- ⁴ In naming, however, more familiar words showed smaller effects of phonological ambiguity (Feldman, 1981). Analogous to claims made from studies with English materials (Seidenberg, 1985; Stanovich and Bauer, 1978), those words that are recognized more slowly and are presumably less familiar are more susceptible to phonological effects in a naming task than are less-familiar words.

References

- Balota, D. A. and Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 340-57.
- Caramazza, A., Miceli, G., Silveri, C. and Laudanna, A. (1985). Reading mechanisms and the organisation of the lexicon: Evidence from acquired dyslexia. *Cognitive Neuropsychology*, 2, 81-114.
- Dannenbring, G. L. and Briand, K. (1982). Semantic priming and the word repetition effect in a lexical decision task. *Canadian Journal of Psychology*, 36, 435-44.
- Diringer, D. (1948). *The alphabet: A key to the history of mankind*. London: The Fleet Street Press.

- Feldman, L. B. (1981). Visual word recognition in Serbo-Croatian is necessarily phonological. *Haskins Laboratories Status Report on Speech Research*, SR-66, 167-202.
- Feldman, L. B. (1983). Bi-alphabetism and word recognition. In D. Rogers and J. A. Sloboda (Eds) *The acquisition of symbolic skills*. New York: Plenum.
- Feldman, L. B., Kostić, A., Lukatela, G. and Turvey, M. T. (1983). An evaluation of the "Basic Orthographic Syllabic Structure" in a phonologically shallow orthography. *Psychological Research*, 45, 55-72.
- Feldman, L. B., Lukatela, G., Katz, L. and Turvey, M. T. (in preparation). Alphabetic and associative priming of phonologically ambiguous words.
- Feldman, L. B. and Moskovljević, J. (in press). Repetition priming is not purely episodic in origin. *Journal of Experimental Psychology: Learning, Memory and Cognition*.
- Feldman, L. B. and Turvey, M. T. (1983). Visual word recognition in Serbo-Croatian is phonologically analytic. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 288-98.
- Henderson, L., Wallis J. and Knight, D. (1984). Morphemic structure and lexical access. In H. Bouma and D. Bouwhuis (Eds) *Attention and performance X*. London: Erlbaum.
- Lukatela, G., Popadić, D., Ognjenović, P. and Turvey, M. T. (1980). Lexical decision in a phonologically shallow orthography. *Memory & Cognition*, 8, 124-32.
- Lukatela, G., Savić, M., Gligorijević, B., Ognjenović, P. and Turvey, M. T. (1978). Bi-alphabetical lexical decision. *Language and Speech*, 21, 142-65.
- Magner, T. F. and Matejka, L. (1971). *Word accent in Serbo-Croatian*. University Park, PA: Pennsylvania State University Press.
- Seidenberg, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition*, 19, 1-30.
- Seidenberg, M. S., Waters, G. S., Sanders, M. and Langer, P. (1984). Pre- and post-lexical loci of contextual effects on word recognition. *Memory & Cognition*, 12, 315-28.
- Stanners, R. F., Neiser, J. J., Herson, W. P. and Hall, R. (1979). Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior*, 18, 399-412.
- Stanovich, K. E. and Bauer, D. W. (1978). Experiments on the spelling-to-sound regularity effects in word recognition. *Memory & Cognition*, 6, 410-15.
- Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory & Cognition*, 9, 263-72.
- Taft, M. and Forster, K. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14, 638-647.
- Wellisch, H. H. (1978). *The conversion of scripts — its nature, history, and utilization*. New York: Wiley.