

# Orthographic depth and the interaction of visual and auditory processing in word recognition

RAM FROST and LEONARD KATZ  
*Haskins Laboratories, New Haven, Connecticut  
and University of Connecticut, Storrs, Connecticut*

We studied the influence of word frequency and orthographic depth on the interaction of orthographic and phonetic information in word perception. Native speakers of English and Serbo-Croatian were presented with simultaneous printed and spoken verbal stimuli and had to decide whether they were equivalent. Decision reaction time was measured in three experimental conditions: Clear print and clear speech, degraded print and clear speech, and clear print and degraded speech. Within each language, the effects of visual and auditory degradation were measured, relative to the undegraded presentation. Both effects of degradation were much stronger in English than in Serbo-Croatian. Moreover, they were the same for high- and low-frequency words in both languages. These results can be accounted for by a parallel interactive processing model that assumes lateral connections between the orthographic and phonological systems at all of their levels. The structure of these lateral connections is independent of word frequency and is determined by the relationship between spelling and phonology in the language: simple isomorphic connections between graphemes and phonemes in Serbo-Croatian, but more complex, many-to-one, connections in English.

The aim of this study was to compare the relative importance of two variables that have been claimed as factors that determine whether phonology is generated pre- or postlexically in reading: orthographic depth and word frequency.

Orthographic depth concerns the directness and simplicity with which a writing system represents the phonology of a language.<sup>1</sup> Two contrasting views have been proposed on the importance of orthographic depth. Feldman and her colleagues (Feldman & Turvey, 1983; Katz & Feldman, 1983) have suggested that it is a major factor in determining whether or not a skilled reader generates phonological information prelexically. They demonstrated in several experiments that lexical access for printed words in the orthographically shallow Serbo-Croatian language is necessarily based on phonological information (Feldman & Turvey, 1983; Lukatela, Popadić, Ognjenović, & Turvey, 1980; see Feldman, 1987, for a summary).

In a recent study, Frost, Katz, and Bentin (1987) compared naming and lexical decision performance in the orthographically deep Hebrew (which omits letters representing vowels from the script), the less deep English

(which has several phonemes for many of its letters, and more than one letter for many of its phonemes), and the shallow Serbo-Croatian (which has a simple isomorphism between sets of letters and phonemes). Controlling for the phonemic composition of the stimuli across languages, Frost et al. found that it took less time to name words than to make lexical decisions in Serbo-Croatian, whereas the opposite was true in Hebrew. In English, performance was similar to Serbo-Croatian, but the difference between the two tasks was smaller. In a second experiment, Frost et al. (1987) demonstrated that in the naming task, greater semantic priming effects were obtained in Hebrew than in English, and in English than in Serbo-Croatian. These results were interpreted as evidence that phonological information, which is necessary for naming, is generated prior to lexical access in shallow orthographies, but not in deep orthographies. Similar conclusions were suggested by Katz and Feldman (1983). In this study, semantic priming was found to facilitate lexical decision, but not naming, in Serbo-Croatian, whereas semantic priming in English was effective in both tasks.

In contrast to this hypothesis, Seidenberg and his associates (Seidenberg, 1985; Seidenberg & Vidanović, 1985) have argued that the primary factor that determines whether or not phonology is generated prelexically is not orthographic depth, but word frequency. Their claim is that in any orthography, frequent words are more familiar as visual patterns, so that rapid visually based lexical access can occur before a phonological code is generated from the print. In this case, the phonological information is obtained postlexically from memory storage. On the other hand, low-frequency words and nonwords are less familiar, their visual lexical access is slow, and their pho-

We are indebted to Bruno Repp for his extensive help and useful criticism throughout this study, and to Laurie Feldman, Carol Fowler, and Guy Van Orden for insightful comments on an earlier draft. This work was supported in part by National Institute of Child Health and Human Development Grant HD-01994 to Haskins Laboratories. We gratefully acknowledge the generous help provided by Georgije Lukatela, Aleksandar Kostic, and all the other members of the psychology laboratory at the University of Belgrade. Correspondence concerning this article should be addressed to Ram Frost, Haskins Laboratories, 270 Crown Street, New Haven, CT 06511.

nology is assumed to be generated prelexically (Seidenberg, 1985; Seidenberg, Waters, & Barnes, 1984).

In support of this hypothesis, Seidenberg (1985) demonstrated that there were few differences between Chinese and English subjects in naming frequent words. This outcome was interpreted to mean that in both logographic and alphabetical orthographies, the phonology of frequent words was derived postlexically, after the word had been recognized visually. Moreover, in another study, Seidenberg and Vidanović (1985) found similar semantic priming effects in naming frequent words in English and Serbo-Croatian, suggesting again that the phonology of frequent words is derived postlexically, whatever the depth of the orthography. Recently, Carello, Lukatela, and Turvey (1988) have shown consistent associative priming effects for naming in Serbo-Croatian. Although Carello et al. did not manipulate word frequency in their study, their results question the inevitability of prelexical phonology in a shallow orthography.

In all of the experiments discussed above, lexical decision and naming tasks were employed. However, one can criticize these tasks as a methodology for studying the processing consequences of the relation between phonology and orthography. Because phonological variation is typically obtained through orthographic variation, one can never be certain which one controls the subject's responses. Moreover, doubts have been raised about the tasks' adequacy for measuring lexical as contrasted with prelexical involvement (see Balota & Chumbley, 1984, 1985).

In the present experiment, we endeavored to find a way of varying orthographic and phonological information independently. We attempted to do this by presenting subjects simultaneously with phonetic information in the form of speech and orthographic information in the form of print. Our assumption was that the phonetic input would be automatically recoded into phonological form so that, in effect, phonological information would be transmitted over the auditory channel. The simultaneous presentation in two modalities initiated independent activation in the visual (orthographic) and auditory (phonetic-phonological) systems. The subjects were required to determine whether the spoken and printed stimuli were the same or different; for example, whether both stimuli were "the same word or nonword." To do this, they had to recode the information presented in the print into phonology, so as to be able to compare it to the phonological information that was provided audibly. The matching task bears some resemblance to the homophony judgment task. It has been demonstrated that when subjects are required to determine whether two printed words match entirely in their phonology, they invoke whole-word phonology that is subsequently used for lexical comparison (Besner & Davelaar, 1982; Besner, Davies, & Daniels, 1981). Although the matching task in the present experiment could also be performed by means of recoding the spoken word into an orthographic code and comparing it to the visual stimulus, we find this less probable because the

transformation of speech into spelling is practiced much less than the transformation of spelling into speech.

Our aim was to compare the performance of native speakers of English and Serbo-Croatian on the matching task, using visual-auditory pairs of high-frequency words, low-frequency words, and nonwords. If the phonology of high-frequency words is generated postlexically in all orthographies, as suggested by Seidenberg, then performance in matching high-frequency words should be similar for English and Serbo-Croatian subjects. In both orthographies, a high-frequency printed word will be recognized visually, and its phonology will be recovered postlexically. The postlexical phonological information will then be compared to the phonology derived from the audibly presented stimulus to ascertain whether a match or mismatch has occurred. Differences between English and Serbo-Croatian would be expected only for low-frequency words and for nonwords, because their phonology is said to be generated prelexically, which is easier to do in a shallow orthography. If, however, differences between the deep and shallow orthographies do, in fact, occur for high-frequency as well as for low-frequency words, then orthographic depth would seem a more important determinant of performance than word frequency.

A direct test of these hypotheses is methodologically problematic, however. A cross-language difference in matching performance may be hard to detect under normal conditions, because reading and listening are such well-practiced skills for subjects in all orthographies. More importantly, even if such a difference were found, it cannot be interpreted unequivocally because of possible differences between English and Serbo-Croatian subjects in reading experience, motivation, and other factors related to stimulus selection. This is an intrinsic problem with any cross-linguistic study. In light of these considerations, we needed a methodology that would demonstrate a cross-language difference relative to each language's own normal baseline. Although, theoretically, differences from baseline can also be affected by differences in subject groups unrelated to the native language, this likelihood would at least be considerably reduced.

Our tentative solution to this methodological problem was to examine subjects' performances under normal and degraded conditions. We degraded either the print or the speech, thereby presenting impoverished sensory information to the visual or the auditory perceptual systems. We measured the effect of visual and auditory degradation in each language relative to performance under normal (i.e., nondegraded) conditions, so that each language would serve as its own baseline control.

This experimental manipulation provides a critical test for examining the role of orthographic depth and word frequency. If the phonology of high-frequency words is generated postlexically in both orthographies, then auditory degradation of frequent words should affect English and Serbo-Croatian subjects in the same way. Since there is no reason to assume that the speakers of the two lan-

guages differ in their speech-perception skills, the processing of phonological information from the speech will be slowed by degradation by the same amount in both languages. Also, the phonology of the printed word should be equally easy to retrieve postlexically in both languages from the clear print. Thus, no cross-language differences are expected in the process of matching the print with the degraded phonological input. By the same argument, cross-language differences are also not expected when the visual input is degraded. Degradation affects the early stage of letter-feature analysis (e.g., Becker & Killion, 1977). Therefore, a mild degradation of the orthographic structure should slow both the generation of any prelexical phonological code and the visual lexical access that provides postlexical phonology. Because the efficiency of perceiving the letters themselves should not, in itself, be different for the two orthographies (the same font of Roman letters was used), visual lexical access of high-frequency words should lag by the same amount of time in English as in Serbo-Croatian.

Note that by this account, a cross-language difference might still be expected for low-frequency words and nonwords. Because the phonology of these stimuli is generated prelexically, at the graphemic and phonemic levels, degradation might have greater effects in an orthography that represents the phonology in a complex manner, as does the English writing system. Thus, low-frequency words will not be particularly informative; the critical question is whether the two orthographies show the same effects of degradation for high-frequency words.

Our own views led us to predict that orthographic depth is an important factor in determining the processing relationship between the orthographic and the phonological systems, regardless of word frequency. Differences in depth of orthography can be captured in terms of the differences in the dynamic interaction of the phonological and orthographic systems that apply to all words, frequent as well as infrequent. We assume that such interaction can take place prior to lexical access, so that degraded information in one modality can be "filled in" by clear information from the other. In contrast to the reader of English, the Serbo-Croatian reader has an internal representational system that consists of less ambiguous (in fact, isomorphic) connections between graphemes and phonemes, which allow faster intermodal compensation of this kind. Given a visual stimulus whose graphemic identity is recognized slowly because of visual noise, visual identification can be enhanced by means of corresponding phonemes that have already been activated by the clearly spoken stimulus. Similarly, clearly printed stimuli should speed the identification of degraded spoken phonemes. Note that the efficiency of these compensatory processes does not depend primarily on word frequency, but on the relation of spelling to sound. For all words, frequent as well as infrequent, the process is based on a possible information exchange between the orthographic and the phonological systems at a prelexical level. We predicted, therefore, that the effects of both visual

and auditory degradation would be stronger in English than in Serbo-Croatian, and that this should be equally true for high- and for low-frequency words.

## METHOD

Because the research was conducted in different countries (Yugoslavia and the USA), special care was taken to standardize the procedures and the selection of stimuli. Thus, the same experimenter ran identical experimental conditions in both countries, using the same apparatus.

The stimuli in English and in Serbo-Croatian were all two-syllable nouns that had stop consonants as their first phonemes. The number of phonemes per word in both the English and the Serbo-Croatian stimulus sets ranged from four to six, and the means did not differ significantly. The distribution of phoneme types was similar in both languages. We selected only regular English words that had no homophones or homographs. Since a standard objective frequency count does not exist in Serbo-Croatian, subjective frequency was estimated in each language using the same procedure: A pool of 200 words was created in each language. Fifty undergraduate students rated the frequency of these 200 words on a 5-point scale ranging from least frequent (1) to most frequent (5). The estimated frequency was calculated by averaging the ratings of all 50 judges. From these words, we selected the stimuli for the experiment.

### Subjects

The subjects were undergraduate students who participated in the experiment as part of the requirements of an introductory course in psychology. Twenty-four subjects participated in each of three experimental conditions in each country. Thus, there were 72 students from the University of Belgrade and 72 students from the University of Connecticut; the subjects were, respectively, native speakers of Serbo-Croatian and English.

### Stimuli and Design

A total of 144 pairs of stimuli were presented simultaneously in the visual and the auditory modalities. Seventy-two pairs were the same—that is, the stimulus that was presented in print was also presented as speech, and 72 pairs were different. The 72 *same* pairs consisted of 24 pairs of high-frequency words, 24 pairs of low-frequency words, and 24 pairs of nonwords. The mean frequencies of the high-frequency words were 4.2 and 4.3, and the mean frequencies of the low-frequency words were 1.7 and 1.75, in Serbo-Croatian and English, respectively. For the purpose of comparison, the mean frequencies of the English words were 164 and 7.3 for the high- and low-frequency words, according to Francis and Kučera (1982). Nonwords were produced by replacing one letter of a legal word; they were orthotactically and phonotactically legal. The 72 *different* pairs of stimuli were three types: 24 word-word pairs (e.g., *prison-treason*), 24 word-nonword pairs (e.g., *captain-capin*), and 24 nonword-nonword pairs (e.g., *poner-pober*). The 24 nonword-nonword pairs were equally divided, so that in each modality 12 words and 12 nonwords were presented. Because it was difficult to equate the difference of the *different* pairs in the two languages, they were considered merely as fillers and introduced in order to achieve a probability of .5 for a same stimulus. The different types of pairs were randomly distributed across the experimental session.

### Stimulus Preparation

The visual stimuli were presented on a Macintosh computer screen. In both orthographies, the same Roman script (the bold face Geneva font) was used. The stimuli subtended a visual angle of approximately 2.5° on the average. The auditory stimuli were originally spoken by female native speakers of each language in an

acoustically shielded booth and recorded on an Otari MX5050 taperecorder. The stimuli were then digitized at a 20-kHz sampling rate. Each digitized stimulus was edited with a digital wave-form editor. A mark tone was inserted at the onset of the stimulus on a second audio channel. The purpose of this tone was to trigger the presentation of the printed word on the computer screen. The onset of the spoken word (defined as the stop-consonant release burst) was determined visually on an oscilloscope, and it was confirmed audibly through headphones. The digitized edited stimuli were recorded at 4-sec intervals on one track of an audio tape, while the other track received the mark tones.

Visual degradation was achieved by overlaying a pattern of random dots on each word; the whole word was thus embedded in a dotted background that obscured its graphemic structure. Auditory degradation was achieved by masking each spoken stimulus with signal-correlated noise—that is, white noise with the same amplitude envelope as that of the spoken stimulus (Schroeder, 1968). Speech and noise were added digitally in proportions of 0.3 and 0.7, corresponding to a constant signal-to-noise ratio of  $-7.5$  dB. The use of signal-correlated noise ensured that the characteristics of the auditory degradation were identical for the Serbo-Croatian and the English speech.

### Procedure and Apparatus

Each subject sat in a semidarkened room in the front of a Macintosh computer screen and wore headphones. The experimental task consisted of pressing a "yes" key if the visual and the auditory stimuli were identical, and pressing a "no" key if they were different. The dominant hand was always used for the "yes" (i.e., same) response.

The audio tape containing the verbal stimuli and the mark tones was played on a two-channel Sony tape recorder. The verbal stimuli were transmitted binaurally to the subject's headphones. The mark tones were transmitted through an interface to the Macintosh, where they triggered the visual presentation and the computer's clock for reaction time (RT) measurements.

The experimental session began with 36 practice pairs. After the practice trials, all 144 test trials were presented to the subject in one block, preceded by a dummy trial.

## RESULTS

In each of the three experimental conditions (normal presentation, visual degradation, auditory degradation), the subjects' means and standard deviations of RTs for correct responses were calculated for each of the stimulus groups. Within each subject-condition combination, outliers that were greater than two standard deviations from the mean were eliminated. Outliers and missing responses typically accounted for less than 5% of all

responses—that is, about 1 out of 24 stimulus pairs per condition.

The results for the *different* trials fillers are inherently less interesting for a comparison between languages, because it is not possible to equate two languages on the quality or quantity of the difference between the spoken and printed stimuli, and because the word-frequency effect cannot be assessed. We will focus, therefore, on the *same* trials.

Table 1 presents mean same RTs for the normal, visually degraded, and auditorily degraded presentations for each of the three stimulus conditions (high-frequency, low-frequency, nonwords). Inspection of Table 1 suggests that, for both languages, on each level of degradation, stimulus frequency affected the subjects' reaction times; RTs decreased with stimulus frequency.

The effects of visual and auditory degradation are reflected in increased RTs for both Serbo-Croatian and English. Also, in both languages, the effect of auditory degradation was greater than the effect of visual degradation. Because the noise levels for each modality were set independently and arbitrarily, this difference is without interest.

Our main prediction was concerned with the interaction of language degradation. It is evident that the effects of both visual and auditory degradation were much stronger in English than in Serbo-Croatian. This outcome occurred in both high- and low-frequency words.

In order to assess the relative effect of degradation in high- and low-frequency words, we performed, separately for words and nonwords, analyses of variance across subjects ( $F_1$ ) and across stimuli ( $F_2$ ). Both analyses contained the factors language (English and Serbo-Croatian) and degradation (no degradation, visual degradation, and auditory degradation). The interaction of language and degradation was significant for words [ $F_1(2,138) = 3.04$ ,  $MS_e = 25,577$ ,  $p < .05$ ;  $F_2(2,184) = 65.62$ ,  $MS_e = 1,834$ ,  $p < .001$ ], and also for nonwords [ $F_1(2,138) = 6.60$ ,  $MS_e = 22,587$ ,  $p < .002$ ;  $F_2(2,92) = 31.08$ ,  $MS_e = 2,762$ ,  $p < .001$ ]. Thus, in all the stimulus conditions there were greater visual or auditory degradation effects in English than in Serbo-Croatian.

For words, the interaction of degradation and word frequency was not significant, nor was the three-way interaction of language  $\times$  degradation  $\times$  frequency ( $F = 3.6$

Table 1  
Reaction Times (in msec) and SEMs of the English and Serbo-Croatian Subjects  
in Normal and Degraded Presentations

	Serbo-Croatian						English					
	High-Frequency Words		Low-Frequency Words		Nonwords		High-Frequency Words		Low-Frequency Words		Nonwords	
	RT	SEM	RT	SEM	RT	SEM	RT	SEM	RT	SEM	RT	SEM
Normal presentation	671	18.8	711	21.0	784	21.6	600	16.6	675	23.1	782	33.0
Visual degradation	691	16.2	731	16.6	807	22.4	660	27.4	764	18.3	862	39.3
Auditory degradation	733	18.7	764	18.3	849	18.7	768	30.0	848	29.2	1062	41.0
Visual degradation effect	20		20		23		60		89		80	
Auditory degradation effect	54		53		65		168		173		280	

Table 2  
Percent of Errors (PE) and SEMs of the English and Serbo-Croatian Subjects  
in Normal and Degraded Presentations

	Serbo-Croatian						English					
	High Frequency		Low Frequency		Nonwords		High Frequency		Low Frequency		Nonwords	
	PE	SEM	PE	SEM	PE	SEM	PE	SEM	PE	SEM	PE	SEM
Normal presentation	5.3	0.69	6.7	0.70	8.2	0.92	6.4	1.00	6.0	0.49	12.6	1.37
Visual degradation	5.2	0.67	7.9	0.97	13.3	1.28	8.3	0.90	9.2	0.87	16.3	1.62
Auditory degradation	6.4	0.66	5.9	0.43	16.4	1.81	10.2	1.17	9.7	0.96	28.8	1.73
Visual degradation effect	-0.1		1.2		5.1		1.9		3.2		3.7	
Auditory degradation effect	1.1		-0.8		8.2		3.8		3.7		16.2	

in the stimuli analysis, but  $F < 1.00$  in the subjects analysis). Thus, within each language the disruptive effect of degradation—whether visual or auditory—was similar for high- and low-frequency words, although in English, a slightly greater visual degradation effect occurred for the low-frequency words.

An interaction of language  $\times$  word frequency was found [ $F_1(1,138) = 39.0$ ,  $MS_e = 1,197$ ,  $p < .001$ ;  $F_2(1,92) = 6.8$ ,  $MS_e = 15,101$ ,  $p < .01$ ]. This is due to a stronger frequency effect in English than in Serbo-Croatian. Although we had used identical procedures for estimating frequency, it might have been the case that the frequency difference between high- and low-frequency words was greater in English. Note that even with this difference, the three-way interaction of language  $\times$  frequency  $\times$  degradation was not significant.

We analyzed the error percentages (Table 2) in separate analyses of variance. The results are similar to the RT analyses. Degradation yielded more errors in English than in Serbo-Croatian. This interaction was significant in the subjects analysis of words [ $F_1(2,138) = 4.71$ ,  $MS_e = 4.6$ ,  $p < .01$ ], but not in the stimulus analysis. It is clear from the results that for both languages, degradation did not cause more errors in low- than in high-frequency words. The interaction of degradation  $\times$  frequency was nonsignificant, as was the three-way interaction of language  $\times$  degradation  $\times$  frequency. These results accord with the results obtained for RTs.

The analysis of nonwords revealed a higher percentage of errors for nonwords than for words in both English and Serbo-Croatian. However, the effects of degradation were again more pronounced in English than in Serbo-Croatian. The interaction of language  $\times$  degradation was significant in the subject analysis [ $F_1(2,138) = 5.7$ ,  $MS_e = 17.4$ ,  $p < .004$ ], but not in the stimulus analysis.

## DISCUSSION

The aim of this study was to determine whether orthographic depth or word frequency affects the prelexical generation of phonology from print. We chose a task in which print had to be matched to speech in order to ensure that phonological information was involved in the process. We hindered orthographic or phonological pro-

cessing by degrading the stimulus at the graphemic or phonemic level, and examined whether this manipulation interacted with word frequency, orthographic depth, or both.

The results indicated that degradation of either speech or print had a more deleterious effect on matching performance in the orthographically deep English than in the shallower Serbo-Croatian. In addition, although there was a strong effect of word frequency in the matching task, it did not interact with degradation. For Serbo-Croatian, visual degradation had a stable effect, causing an increase in RTs of about 20 msec for both high- and low-frequency words. For English, the increase due to visual degradation was three to four times larger, and, once more, it was independent of word frequency. The interlanguage differences that were found for visual degradation were replicated for auditory degradation: the degradation effects in English were three to four times larger than in Serbo-Croatian. Most importantly, no interaction with word frequency was found.

Collectively, these results question the assumption of the inevitability of orthographic access for printed high-frequency words, and its application to all orthographies whatever their orthographic depth. Seidenberg (1985) has presented a detailed parallel interactive model in which the time course of phonological and orthographic activation is the key factor in determining which of the two kinds of information is responsible for lexical access. The time-course activation model specifies that recognition of printed words is initiated with the extraction of visual information from the input. Only after the orthographic units are recognized do they activate their phonological representations. In this way, phonological access will usually lag behind visual access. According to this model, high-frequency words will always access the lexicon orthographically because orthographic information develops faster—for the skilled reader—than phonological information. Furthermore, the model explicitly predicts that orthographic information will dominate no matter what the depth of the orthography.

If postlexical phonology were obligatory for high-frequency words, both English and Serbo-Croatian should have shown the same effects of degradation for high-frequency words. When the speech is degraded, development of phonological information from the speech should

be equally slowed in the two languages, but the process of orthographically based lexical access (equivalent for the two languages) should be as rapid as normal. Therefore, no interlanguage differences would be expected in the matching task. Similarly, when print is degraded, letter identification that is the basis for either phonological recoding or orthographic lexical access should be slowed uniformly for both English and Serbo-Croatian. As noted above, for high-frequency words, we would expect the effects of degradation to be equivalent across orthographies. Our results suggest otherwise. Moreover, if Serbo-Croatian and English differed only in the processing of low-frequency words, where phonology must be derived prelexically, an interaction of orthography  $\times$  word frequency  $\times$  degradation should have been obtained. No such effects were significant in our study. Instead, the effect of orthographic depth appears to persist, even for high-frequency words. This result is consistent with other studies which demonstrate that a prelexical phonological code is used for lexical access irrespective of word frequency (Van Orden, 1987; Van Orden, Johnston, & Hale, 1988). Van Orden (1987) had demonstrated that the likelihood of mistaking the stimulus-like *rows* for the flower *rose* in a categorization task is not affected by the frequency of *rows*. If high-frequency homophones of this kind had been accessed visually, then such phonologically based errors would have been expected to decrease as a function of stimulus frequency. Similar results have been obtained by Van Orden and his colleagues, in which possible confounds such as category typicality or production frequency were eliminated (Van Orden et al., 1988).

An important finding is the similar pattern of results obtained independently in the visual and auditory modalities. Although it is commonly assumed that visual degradation slows the early processes of letter identification, thereby slowing generation of both pre- and postlexical access to phonology, it might be argued that it slows, mainly, late orthographic activation. According to this argument, visual degradation induces prelexical processing for both high- and low-frequency words in both orthographies, resulting in a nonsignificant interaction of degradation  $\times$  language  $\times$  frequency. This argument, however, would still not account for the results obtained independently in the auditory modality.

The results of the present experiment can be understood in terms of structural differences in the relationship between orthographic and phonological information, rather than in terms of the temporal course of activation. The class of interactive cascade network models proposed by Rumelhart, McClelland, and their associates (McClelland & Elman, 1986; McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982; Seidenberg & McClelland, 1987) provides a heuristic framework for discussing the interrelations of orthography and phonology in the mental lexicon. Interactive models can account for processing that is based on orthographic or phonological information, or on a combination of the two. An orthographic network for perceiving print can be modeled in terms of

a hierarchy that consists of letter features, letters, and whole words. A parallel network of phonological knowledge would consist of phonetic feature nodes that are associated with phoneme nodes, which, in turn, are associated with whole phonological words. Hence, the mental lexicon can be regarded as composed of two interactive systems, orthographic and phonological, which are interconnected at all levels of the two systems.

The concept of orthographic depth has direct bearing on the question of the relation between the phonological and orthographic systems. Within the interactive model, the way in which connections are made between the two systems should be constrained by the depth of the orthography that is being modeled. In a shallow orthography, connections are isomorphic: a graphemic node can be connected to only one phonemic node, and vice versa. Also, because words are spelled uniquely, each word node in the orthographic system must be connected to only one word node in the phonological system. In contrast, in a deeper orthography, a graphemic node may be connected to several phonemic alternatives, a phonemic cluster may be connected to several orthographic clusters, and finally, a word in the phonological system may be connected to more than one word in the orthographic system, as in the case of homophony (e.g., SAIL/SALE) or, alternatively, the case of homographic words with two pronunciations (e.g., WIND, READ, BOW, etc.). A diagram of the two intersystems is given in Figure 1 for a word that exists in both the English and the Serbo-Croatian languages. The Serbo-Croatian word, KLOZET, is composed of unique letter-sound correspondences, whereas the corresponding English word, CLOSET, is composed of graphemes, most of which have more than one possible phonological representation, and phonemes, most of which have more than one orthographic representation.

Given this model, how will the orthographic and the phonological systems be affected by degradation? The ability of interactive systems to compensate for partial loss of information within a single modality has been described for visual perception (McClelland & Rumelhart, 1981) and for speech perception (McClelland & Elman, 1986). This ability derives from both top-down and lateral processes within each system. However, in addition to these within-system processes, activation in one system can also be affected by activation in the other. Thus, compensatory activation below the whole word level can play a role when the information in one domain is impoverished, in the present experiment because of stimulus degradation. Because the effectiveness of compensatory activation depends on the complexity of the subword connections between the two systems, interlanguage differences in performance under degraded conditions arise. This claim follows from an extension of the interactive model that rationalizes the relationship between the orthographic and phonological systems in terms of lateral connections between the two networks at the sublexical levels, which apply to all words independently of word frequency. We propose that the structure of these lateral

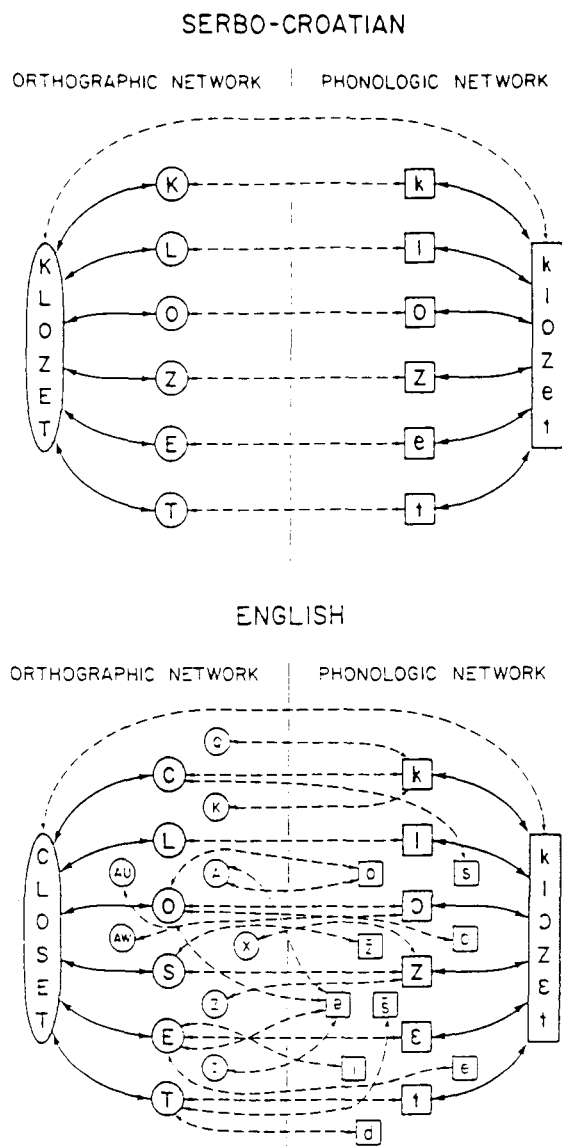


Figure 1. Interaction between orthographic and phonological networks in English and Serbo-Croatian.

connections is determined by the form of the relationship between spelling and phonology in each language.<sup>2</sup>

The poor matching performance in English relative to Serbo-Croatian was revealed by either visual or auditory degradation. This suggests that both systems, the orthographic and the phonological, could be activated one from the other; that is, activation between systems is not directionally constrained. This proposal is supported by findings that demonstrate not only phonological effects in visual word perception (Spoehr, 1978; Spoehr & Smith, 1975), but also orthographic effects in auditory word perception (Tanenhaus, Flanigan, & Seidenberg, 1980).

The results revealed for nonwords provide additional insights. For these stimuli, both our and Seidenberg's models would make similar predictions: since nonwords lack a top-level representation, their phonology can

only be retrieved prelexically. Hence, the depth of the English orthography should cause a greater effect of degradation in English than in Serbo-Croatian. However, compare the effect of degradation obtained for nonwords to the effects obtained for high-frequency words in the two languages. Here, our and Seidenberg's models make clearly contrasting predictions: if the phonology of high-frequency words is derived postlexically in deep as well as in shallow orthographies, then in Serbo-Croatian, just as in English, nonwords should be affected by degradation more than high-frequency words are, but our results suggest otherwise. In fact, one striking result in our study is that in Serbo-Croatian, the effects of visual and auditory degradation were similar for high-frequency words and for nonwords. In contrast, greater effects of degradation for nonwords were found in English. The significance of this difference was assessed in separate analyses of the RT data. The three-way interaction of language  $\times$  degradation  $\times$  stimulus type was indeed significant [ $F_1(2,138) = 6.41, p < .002$ ;  $F_2(2,184) = 14.3, p < .001$ ]. As hypothesized, a comparison in Serbo-Croatian revealed that the effects of auditory degradation for high-frequency words and nonwords did not differ significantly [ $F_2(1,184) < 1.0$ ], whereas in English they did [ $F_2(1,184) = 10.41, p < .001$ ].

In the present experiment, we used a cross-language design in order to study word recognition in different orthographies. Cross-language designs allow a direct comparison of native speakers' performances under identical experimental conditions. Hence, they can provide evidence concerning the effects of the orthography's characteristics on the process of word recognition. However, cross-language designs are not without potential pitfalls: language differences may be confounded with nonlinguistic factors. For example, differences in the subjects' populations (e.g., motivation, education, etc.) might interact with the experimental manipulation. The interpretation of the results thus hinges on whether they are likely to be free of such confounding.

A second problem involves potential confounds that are themselves linguistic factors. In the present study, the putative structural difference in the relationship between orthography and phonology is not the only variable that can potentially account for the differential effect of degradation. For example, there is a purely orthographic difference between the two languages (e.g., they have different numbers of graphemes). This might have caused the differential effect of print degradation. Also, there is a purely phonological difference (e.g., they have different numbers of phonemes). This might have caused the differential effect of auditory degradation. Although we cannot rule out such alternative explanations, we believe that the explanation based on structural differences is more plausible.

First, consider the possibility that the greater susceptibility to visual and auditory degradation in English arises from its greater orthographic complexity. We believe that this is unlikely. The English orthography in itself is not



more complex than the Serbo-Croatian. In fact, the Serbo-Croatian orthography has more graphemes than the English, and some of these graphemes consist of identical letters that differ only by additional signs that are printed above them (e.g., c, ć, č, s, š). Furthermore, the Serbo-Croatian reader learns two orthographic systems, Cyrillic and Roman, which share letters with identical or different phonologies. Therefore, it is clear that English subjects cannot simply be more susceptible to a visual degradation of print.

Consider a possible explanation of the differential degradation effect that is based on the greater number of phonemes in English. English has, indeed, more phonemes than Serbo-Croatian, but the possibility that English speakers are therefore more susceptible to auditory degradation can be rejected by means of a linguistic argument. Although languages do differ in phonetic structure, the overall amount of phonetic information (phonemes, stress patterns, amplitude envelope, etc.) conveyed by the native speaker to the native listener is similar across languages. Moreover, the native listener is presumably tuned to his or her own language's phonetic characteristics. Therefore, it is improbable that native speakers of different languages differ in the efficiency of perceiving their own spoken languages whether clear or degraded. Degraded speech is not an artificial experimental condition; rather, it is a real condition of everyday speech perception.<sup>3</sup> In addition, in order to avoid different effects of noise on speech in the two languages, we used signal-correlated noise instead of white noise. The signal-correlated noise is correlated with the amplitude envelope of each word, resulting in a similar degradation not only in each language, but also in each stimulus. Therefore, the noise we used disrupted the speech in itself in a similar way in both English and Serbo-Croatian.

Thus, we believe that the separate differences for degraded print and speech that we found between the two languages in the present study can more likely be attributed to differences in their efficiency in restoring degraded print from speech and vice-versa, differences that derive from their orthographic depth.

# REFERENCES

- BALOTA, D. A., & 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 & Performance*, **10**, 340-357.
- BALOTA, D. A., & CHUMBLEY, J. I. (1985). The locus of word frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory & Language*, **24**, 84-106.
- BECKER, C. A., & KILLION, T. H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **3**, 389-401.
- BESNER, D., & DAVELAAR, E. (1982). Basic processes in reading: Two phonological codes. *Canadian Journal of Psychology*, **36**, 701-711.
- BESNER, D., DAVIES, J., & DANIELS, S. (1981). Reading for meaning: The effects of concurrent articulation. *Quarterly Journal of Experimental Psychology*, **33A**, 415-437.
- CARELLO, C., LUKATELA, G., & TURVEY, M. T. (1988). Rapid naming is affected by association but not by syntax. *Memory & Cognition*, **16**, 187-195.
- FELDMAN, L. B. (1987). Phonological and morphological analysis for skilled readers in Serbo-Croatian. In A. Allport, D. MacKay, W. Prinz, & E. Sheerer (Eds.), *Language perception and production: Relation among listening, speaking, reading, and writing*. London: Academic Press, 1987.
- FELDMAN, L. B., & TURVEY, M. T. (1983). Word recognition in Serbo-Croatian is phonologically analytic. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 288-298.
- FRANCIS, W. N., & KUČERA, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston, Houghton Mifflin.
- FROST, R., KATZ, L., & BENTIN, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 104-114.
- KATZ, L., & FELDMAN, L. B. (1983). Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **9**, 157-166.
- LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D. P., & STUDDERT-KENNEDY, M. (1967). Perception of the speech code. *Psychological Review*, **74**, 431-461.
- LIBERMAN, A. M., & MATTINGLY, I. G. (1985). The motor theory of speech perception reconsidered. *Cognition*, **21**, 1-37.
- LUKATELA, G., POPADIĆ, D., OGNJENOVIC, P., & TURVEY, M. T. (1980). Lexical decision in a phonologically shallow orthography. *Memory & Cognition*, **8**, 124-132.
- MCCLELLAND, J. L., & ELMAN, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, **18**, 1-86.
- MCCLELLAND, J. L., & RUMELHART, D. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, **88**, 375-407.
- RUMELHART, D., & MCCLELLAND, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, **89**, 60-94.
- SCHROEDER, M. R. (1968). Reference signal for signal quality studies. *Journal of the Acoustical Society of America*, **43**, 1735-1736.
- SEIDENBERG, M. S. (1985). The time course of phonologic activation in two writing systems. *Cognition*, **19**, 1-30.
- SEIDENBERG, M. S., & MCCLELLAND, J. L. (1987, November). *A distributed, developmental model of visual word recognition and naming*. Paper presented at the Twenty-Eighth Annual Meeting of the Psychonomic Society, Seattle, WA.
- SEIDENBERG, M. S., & VIDANOVIĆ, S. (1985, November). *Word recognition in Serbo-Croatian and English: Do they differ?* Paper presented at the Twenty-Fifth Annual Meeting of the Psychonomic Society, Boston, MA.
- SEIDENBERG, M. S., WATERS, G. S., & BARNES, M. A. (1984). When does irregular spelling of pronunciation influence word recognition? *Journal of Verbal Learning & Verbal Behavior*, **23**, 383-404.
- SPOEHR, K. T. (1978). Phonologic encoding in visual word recognition. *Journal of Verbal Learning & Verbal Behavior*, **17**, 127-141.
- SPOEHR, K. T., & SMITH, E. E. (1975). The role of orthographic and phonotactic rules in perceiving letter patterns. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 21-34.
- TANENHAUS, M. K., FLANIGAN, H. P., & SEIDENBERG, M. S. (1980). Orthographic and phonological activation in auditory and visual word recognition. *Memory & Cognition*, **8**, 513-520.
- VAN ORDEN, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition*, **15**, 181-198.
- VAN ORDEN, G. C., JOHNSTON, J. C., & HALE, B. L. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory & Cognition*, **14**, 371-386.

# NOTES

1. The term *orthographic depth* has been used with a variety of related but different meanings. In this paper it denotes the degree of correspon-



dence between single graphemes and phonemes in a language's spelling system.

2. Although whole-word nodes do not contribute to the effects obtained in the present study, we include them in our model for the sake of completeness.

3. One would not like to assume that in the evolution of different languages, the native speakers and listeners of some languages cannot understand each other as well as the native speakers and listeners of others.

The transformation of speech (clear or degraded) into phonological structure appears to be an automatic process that occurs below the level of awareness (see Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Liberman & Mattingly, 1985). This modular process is assumed to be similar across languages.

(Manuscript received January 22, 1988;  
revision accepted for publication October 17, 1988.)