JOURNAL OF MEMORY AND LANGUAGE 32, 766–780 (1993)

Syllable Frequency and Visual Word Recognition in Spanish

MANUEL CARREIRAS, CARLOS J. ALVAREZ, AND MANUEL DE VEGA

Departamento de Psicología Cognitiva, University of La Laguna, 38201 Tenerife, Spain

This paper examines the role of the syllable as a sublexical representational unit during visual word recognition in a shallow language: Spanish. Five experiments were carried out to test the effect of the frequency of the syllables on naming latencies and lexical decision times. The orthographic redundancy hypothesis (Seidenberg, 1987, 1989) claims that the effects of syllabic structure are merely illusory and that they can be explained as effects deriving from the frequency of cooccurrence of letter patterns. The reliable effects of syllabic frequency that we found cannot be accounted by the frequency of cooccurrence of letter patterns. The implications of such findings for the dual route theory and for the PDP model proposed by Seidenberg and McClelland are discussed.

In the past two decades there has been a growth of interest in visual word recognition. Nowadays, visual word recognition is a field of research well developed both empirically and theoretically (Besner & Humphreys, 1990; Posner & Carr, 1992). Several models have been proposed to explain, among other things, how to access the meaning of a word or how to build a lexical representation of a word from a string of letters. Despite the vast amount of work done in the field, a number of basic questions remain unresolved. One current debate focuses on whether there are two routes to a lexical entry or just a single activation system.

Dual route theories claim that readers contact lexical entries either by "assembling" phonology from grapheme–phoneme correspondences and then looking up the phonological output or by directly contacting a lexical entry based on the orthographic input (Coltheart, 1978; Besner, 1990; Paap & Noel, 1991). One route is phonological recoding, and the other route is direct access. The faster of the two routes determines access in any given case (Besner, 1990). PDP models contrast with Dual Route Theories in their mechanism, because they postulate parallel activation along networks of distributed sublexical information. Single route PDP models propose that word-identification is an emergent property of parallel activation networks (Seidenberg & McClelland, 1989; Van Orden, Pennington, & Stone, 1990). Words are merely "identified" as an automatic consequence of patterns of phonological and graphic input features.

Another important issue is whether or not some sublexical units—sometimes termed "access units"—are implicated in the process of lexical access. There have been several proposals that readers parse or decompose complex words into sublexical units, prior to comprehending word meanings. Some of the sublexical units proposed, among others, are morphemes, syllables, and the Basic Orthographic Syllable Structure (BOSS) (Taft, 1989).

The role of the syllable as a unit in speech

This research was supported by Grant DGICYT PB88/0425 from the Dirección General de Investigación Científica y Técnica and Grant 91/103 from the Gobierno de Canarias. We are grateful to Tom Carr, Morton A. Gernsbacher, Bruce McCandlis, Matthews Traxler, Mike Posner, Tony Sanford, Edward J. Shoben, and Victor Villa-Good for their helpful criticisms on earlier drafts of this manuscript. Some parts of the manuscript were written while the first author was visiting the Department of Psychology at the University of Oregon. Reprint requests should be addressed to Manuel Carreiras, Departamento de Psicología Cognitiva, Universidad de La Laguna, 38201 Tenerife, Spain.

0749-596X/93 $5.00

segmentation has received a considerable amount of attention. Its relevance for speech perception seems to be demonstrated in several languages, such as French (Mehler, Dommergues, Frauenfelder, & Seguí, 1981; Cutler, Mehler, Norris, & Seguí, 1986), Portuguese (Morais, Content, Cary, Mehler, & Seguí, 1989), and Spanish and Catalan (Sebastián-Gallés, Dupoux, Seguí, & Mehler, 1992), among others. The data are less conclusive for visual word recognition, although the results of some studies suggest that this is the case. Printzmetal, Treiman, and Rho (1986) reported five experiments using feature integration errors to analyze effects of syllabic structure. They presented words printed in two different colors tachistoscopically; the subjects had to report the color of a target letter which was contained in each word. The main finding was that subjects usually reported that the target letter was the color of the syllable in which it was embedded, even when the color of the target letter was different. Their conclusion was that syllables are perceptual units in reading, that is, orthographically defined syllables are identified at a perceptual level of processing.

However, other studies about the role of syllabic structure on reading have yielded inconsistent results (Seidenberg, 1987, 1989). According to Seidenberg, the basis for these inconsistencies is that, either these studies were based on incorrect definitions of the syllable, or that the view that words are parsed into syllables is simply wrong. He explains these findings by proposing that these sublexical units are only the by-products of a parallel activation process. In fact, this second hypothesis is held by some of the current connectionist models of visual word recognition (Seidenberg and McClelland, 1989). The only sublexical structures relevant to recognition are those defined by orthographic redundancy, which reflects facts about the distribution of letter patterns in the lexicon and tend to correspond, although not necessarily, to units such as syllables and morphemes (Seidenberg, 1989). In other words, the orthographic redundancy hypothesis assumes that the effects of syllabic structure are merely illusory and that they can be properly understood as effects deriving from the frequency of cooccurrence of letter patterns. There are different statistical measures which could tap the redundancy. One of these measures is positional bigram frequency: the frequency of a pair of letters in a certain position in a word. According to Adams (1981) and Seidenberg (1989), the frequencies of bigrams within syllables are higher than the frequencies between syllables. Letters within syllables tend to cooccur in the written language more often than letters that mark syllable boundaries. Consequently, syllable boundaries are typically marked by a pattern of bigram frequencies that can be referred to as a "trough." On the basis of this coincidence of bigram troughs with syllabic boundaries, it can be argued that effects resulting from the presence of bigram troughs may be erroneously attributed to syllabic factors. Therefore, from this point of view, it follows that syllables are only emergent effects rather than obligatory access units. Nevertheless, the orthographic redundancy hypothesis has been challenged recently by Rapp (1992), who showed reliable effects of syllabic and morphological structure independent of the presence or absence of bigrams troughs.

On the other hand, the arguments against the role of syllables as a sublexical processing unit in reading and/or the use of a visual direct route might not apply to all languages. Some of the inconsistencies among experiments concerning a syllabic effect could be due to the fact that breaks between syllables are not always clear in English. In other words, syllables are not clearly defined in English. In addition, English has a deep orthography; there is not a close correspondence between graphemes and phonemes. Because English syllables are not always clearly defined and because English has a "deep" orthography, readers may not use the syllable as an access unit economically. Likewise, readers may not economically access words most of the time via the phonological route. Nevertheless, there are shallow languages, such as Spanish, that have a closer relationship between letters and sounds, and clearly defined syllable boundaries. Shallow orthographies provide a direct or "shallow" mapping between spellings and pronunciations. The result is that shallow orthographies, such as Spanish, have more invariance in their grapheme–phoneme mappings than do deep orthographies, such as English. Readers of shallow languages may take advantage of these features to follow a phonological route and use the syllable as an access unit during word recognition. In general, in shallow orthographies phonology is activated directly from print (Frost, Katz, & Bentin, 1987). In fact, Spanish readers and speakers can very easily and very accurately segment words into syllables and correctly spell a word that they have not seen before.

The purpose of this research was to test whether Spanish-skilled readers use syllabic information during visual word recognition. To do so, we carried out five experiments designed to test the effects of the frequency of the syllables embedded in words, by using lexical decision and naming tasks. Words were constructed by manipulating the value of positional frequency—high and low—of syllables embedded in words which had high and low values of lexical frequency. The value positional frequency of syllables is the number of times that a syllable appears in a particular position in a word (first, second, final, etc.), whereas the total frequency is the number of times that a syllable appears in a word, independent of its position within the words. Some previous studies have shown that the positional frequency of syllables, but not the total frequency of syllables, influenced reading times for words (de Vega, Carreiras, Gutierrez, & Alonso, 1990) and lexical decision times (de Vega & Carreiras, 1989). Since only the positional syllable frequency was manipulated in these experiments, we will refer to it from now just as syllable frequency.

The effects of syllable frequency and lexical frequency were evaluated with a lexical decision task in Experiment 1 and with a naming task in Experiment 2. In the three last experiments new stimuli were used. For words, the bigram frequency within syllables was matched to the bigram frequency between syllables. This control was included to examine whether the effects of syllable frequency were due to the frequency of letter patterns as predicted by the orthographic redundancy hypothesis (Seidenberg, 1987, 1989). In addition, for nonwords, the syllable frequency and the bigram frequency were orthogonally manipulated. A lexical decision task was used in Experiment 3, whereas a naming task was used in Experiments 4 and 5. In Experiment 4 words and nonwords were presented for naming in two different sets, whereas in Experiment 5 both words and nonwords were presented in the same set.

Experiment 1

Our first question was whether the frequency of syllables embedded in words plays a role in visual word recognition and whether its effect is different from that of lexical frequency. To answer this question, subjects performed a lexical decision task on high- versus low-frequency words that contained high- versus low-frequency syllables. If subjects use syllables to visually recognize words, reaction times should show an effect of syllable frequency.

Method

Subjects. The subjects were 20 undergraduate students from the University of La Laguna who received course credit for their participation.

Design and materials. One hundred and forty-four words were constructed by combining two factors in a 2 × 2 repeated-measures design: high versus low frequency of the syllables embedded in words and high versus low printed frequency of the words.

Syllables were selected according to their frequency in the Dictionary of Frequency of Syllables in Spanish (Alvarez, Carreiras, & de Vega, 1992a). Syllables were considered high frequency only when they had a minimum mean frequency of occurrence of 125. Syllables were considered low frequency when they had a maximum mean frequency of 60. The Dictionary of Frequency of Syllables in Spanish was built from a sample of 25,000 Spanish words included in small paragraphs extracted from natural texts (modern books, magazines, and newspapers). The positional frequency of each syllable was computed as the number of times that the syllable appeared in that word position (first, second, final, etc.). On the other hand, words were selected according to their printed frequency in the Julliand and Chang-Rodriguez (1964) Dictionary of Frequency. In this experiment words were considered as high frequency when they had a minimum value of occurrence of 60 and low frequency when they had a maximum value of 35. Words with very extreme values of high frequency were avoided.

The number of syllables of words was balanced and therefore each cell had the same number of words with two and three syllables. In addition, 144 nonwords were constructed with the constraint that they matched the experimental words in length and number of syllables. All of them were constructed with only legal syllables in Spanish.

Procedure. Subjects were tested individually in a small quiet room. The subjects' task was to attend to a string of letters (words and nonwords) presented on a computer screen and make a lexical decision as quickly as possible, maintaining a reasonable level of accuracy. The experiment was controlled by an IBM compatible computer. Each trial started with the presentation of an asterisk in the center of the screen for 600 ms. Then, a string of letters was displayed until subjects responded by hitting one of two keys. They were instructed to press the L key when it was a word and the A key when it was a nonword. The L key was labeled "SI" (that means YES in Spanish) and the A key was labeled "NO." One second after the subject responded, a new trial started. Twenty practice trials were followed by two blocks of 144 experimental trials. Items were presented in random order. The computer recorded subjects' answers and reaction times.

Results

The mean reaction times for correct responses and the average error rates for words were calculated separately across subjects and items, and each data set was subjected to an analysis of variance (ANOVA). The mean reaction times and the percentage of error rates are shown in Table 1. The analysis of variance on reaction times showed that both main effects—lexical frequency and syllable frequency—were statistically reliable. Reaction times were faster for high frequency words than for low frequency words [$F1(1,19) = 8.68$, $p < .01$; $F2(1,140) = 10.95$, $p < .001$]. However the reverse effect was found for syllable frequency. Reaction times were faster for words with low frequency syllables than for words with high frequency syllables [$F1(1,19) = 4.79$, $p < .05$; $F2(1,140) = 7.91$, $p < .01$]. The interaction was not reliable. The same pattern of results was obtained for words with two and three syllables.

TABLE 1

MEAN REACTION TIMES AND PERCENTAGE OF ERRORS (IN PARENTHESES) AS A FUNCTION OF SYLLABLE FREQUENCY AND WORD FREQUENCY

	Word frequency	
Syllable frequency	High	Low
High	790 (13.5)	825 (12.1)
Low	734 (14.2)	783 (12.8)

No effects were reliable in the error rate analysis.

Discussion

As usual, words with high printed frequency were recognized faster than words with low printed frequency. In contrast, and more importantly, the effect of syllable frequency was also reliable, showing that words are recognized faster when they have low frequency syllables.

This is the opposite effect of that found for lexical frequency, and to that expected just for any frequency. Intuitively, frequency should help because the more times an event occurs the more accessible it should be for comprehension and production. So, it should follow that as frequency increases, speed in processing should also increase. However, our results go in the opposite direction intuitively expected, but they concur with results obtained by Broadbent and Gregory (1968) for bigrams in English and by de Vega and Carreiras (1989) and de Vega et al. (1990) for syllables in Spanish.

This result suggests that syllables are important processing units for Spanish, because they have a very strong effect in visual word recognition. In addition, these results are consistent with the idea of a mandatory phonological route for Spanish readers. It is the case that Spanish is a language which has a transparent syllabic structure and a very regular orthography: the pronunciation of a string of letters can always be derived from print. Further, it has been suggested that pronunciation might be directly derived from the printed word in other shallow languages (Feldman & Turvey, 1983; Lukatela & Turvey, 1990a,b).

The effect of syllable frequency we found could be explained from a PDP-type model in which high-frequency syllables fire a larger set of lexical candidates than low-frequency syllables. It is reasonable to assume that high-frequency syllables are shared by a larger number of words than low-frequency syllables. In fact, particular syllables owe their frequency to the number of words in which they are embedded. Assuming also that it takes more time to select a word from a large set than from a small one, then it follows that high-frequency syllables should produce slower processing times than low-frequency syllables in a lexical decision task.

Experiment 2

Experiment 1 showed an important effect concerning the frequency of syllables. Words are recognized faster when they have low-frequency syllables. But it could be argued that the effect was a by-product of a special strategy that subjects employed to solve the task. Because the lexical decision task has been subjected to several criticisms in terms of how good it is as a measure of lexical access (see Balota & Chumbley, 1984; Besner & McCann, 1987; Paap, McDonald, Schvaneveldt, & Noel, 1987), we decided to obtain converging evidence regarding the effect of syllable frequency in order to further establish its generality. A naming task was employed in Experiment 2 using the same stimuli as in Experiment 1.

Method

Subjects. The subjects were 18 undergraduate students from the same population as those tested in Experiment 1. They received course credit for their participation. None of them had participated in the previous experiment.

Design and materials. The design and materials were the same as those in the previous experiment, with the exception that nonwords were not used.

Procedure. The subjects' task was to attend to the words presented on a computer screen and pronounce each word aloud as quickly as possible, maintaining a reasonable level of accuracy. The experiment was controlled by an IBM-compatible computer. Each trial started by presenting an asterisk for 600 ms in the center of the screen. Then a word was displayed. Reaction times were recorded from word onset on the screen until the subjects started making an utterance that closed a voice-activated relay (voice key). There was a 1-s delay from stimulus offset until the beginning of the next trial. The subjects' responses were also tape recorded in order to exclude any errors from the analysis of reaction times. After the practice trials, subjects received the 144 experimental trials in two blocks. Within each block the order of presentation of the stimuli was randomized. They also were instructed to respond quickly but as accurately as possible.

Results

The mean naming latencies for correct responses were calculated separately across subjects and items, and each data set was subjected to an analysis of variance (ANOVA). The main effect of word frequency was reliable [$F1(1,17) = 16.21, p < .001$; $F2(1,140) = 6.44, p < .05$]. Naming latencies for low-frequency words were slower than for high-frequency words. The main effect of frequency of syllables was also reliable [$F1(1,17) = 8.79, p < .01$; $F2(1,140) = 5.03, p < .05$]. Again, words with low-frequency syllables were named faster than words with high-frequency syllables. The interaction between the two variables was not reliable. The same pattern of results was obtained for words with two and three syllables (Table 2).

Discussion

Both variables—printed lexical frequency and syllable frequency—showed a pattern similar to that in Experiment 1, in which a lexical decision task was used. The effect of syllable frequency was the opposite to that of word frequency and there was no interaction between them. That is, words with high-frequency syllables were named slower than those with low-frequency syllables. The fact that these results were also obtained with naming suggests that the effect is robust. But also, the fact that these results were obtained both for words with high printed frequency and for words with low printed frequency, also in a task that requires the use of phonological output information (naming), suggests that lexical access in Spanish might require readers to employ a phonological route. If Spanish readers were employing a direct access route for high frequent words, no syllabic effects should have been found for the high-frequency words.

In addition, the usual effect for the word printed frequency was found. That is, high-frequency words were named quicker than low-frequency words. However, the magnitude of the word printed frequency effect was less than the effect obtained in Experiment 1 with a lexical decision task. Similar weaker effects for word printed frequency when using a naming task rather than a lexical decision task are reported in the literature (e.g., Andrews, 1989; Balota & Chumbley, 1984; Frederiksen & Kroll, 1976; Paap et al., 1987).

TABLE 2
MEAN REACTION TIMES AS A FUNCTION OF SYLLABLE FREQUENCY AND WORD FREQUENCY

	Word frequency	
Syllable frequency	High	Low
High	721	732
Low	694	718

EXPERIMENT 3

The effect of syllable frequency obtained in Experiments 1 and 2 does not rule out the orthographic redundancy hypothesis. Bigram frequencies might account for these effects, if syllables exhibit a trough pattern. A trough pattern is exhibited if bigrams between syllables are of low frequency but within-syllables are of high frequency (Seidenberg, 1987, 1989). In order to determine whether apparent syllabic effects can, in fact, be attributed to the bigram organization of words, one would want to compare the behavior of words with bigram troughs at these sublexical boundaries with that of words in which troughs were absent at syllabic junctions. If the presence of a bigram trough serves to mark an orthographic boundary, then we should expect to find syllabic effects whenever, and only if, a bigram trough is present. Evidence indicating the salience of syllabic structure, independent of the presence of bigram troughs, would create problems for the orthographic redundancy hypothesis, in which the representation of sublexical entities, such as the syllable, plays no role.

The third experiment was carried out to test that possibility. Again, a lexical decision task was employed, but new stimuli that do not exhibit a trough pattern were constructed. The orthographic redundancy hypothesis was first tested by orthogonally manipulating printed word frequency and syllable frequency in words, while holding constant the bigram frequency within and between syllables. Second, by orthogonally manipulating the bigram and syllable frequency in nonwords. In that way, any possible effect of syllable frequency in the recognition of words could not be explained by the orthographic redundancy hypothesis, because the bigram frequency was kept constant for words with high- and low-frequency syllables. Also, the nonword manipulation allows us to evaluate the contribution, if any, of the bigram frequency.

Method

Subjects. Twenty-two undergraduate psychology students from the University of La Laguna participated in this experiment in fulfillment of a course requirement.

Design and materials. We constructed a new set of 60 bisyllabic words, four to five letters long, while adding a new constraint of keeping constant the bigram frequency between and within syllables. The two factors manipulated were again word frequency (high versus low) and syllable frequency (high versus low). The main difference between these new stimuli and those used in the first two experiments was that the mean of the bigram frequency within syllables was equivalent to the mean of the bigram frequency between syllables. Bigrams within syllables are pairs of contiguous letters comprised within a syllable, but a bigram between syllables is a pair of contiguous letters one of which belongs to a syllable and the other to another syllable. Their positional bigram frequency was obtained from the Alvarez, Carreiras, and de Vega (1992b) Bigram Frequency Dictionary in Spanish.

Another set of 60 nonwords matched in length and number of syllables with the words were also constructed. Only legal syllables in Spanish were used. Two factors were also manipulated in the nonword materials: the syllable frequency (high versus low) and the bigram frequency between syllables (high versus low). The frequency of bigrams within syllables was kept constant.

It should be noted that words and nonwords were selected so that the number of occlusive and fricative initial phonemes was similar across conditions.

Procedure. The procedure was similar to that of Experiment 1.

Results

Table 3 shows the mean of error rates in parenthesis and the mean of reaction times for correct responses. Separate analyses on the error rates and on the reaction times were performed for both words and nonwords.

The ANOVA on reaction times for words revealed a reliable main effect of the word frequency [$F1(1,21) = 14.17$, $p < .005$; $F2(1,56) = 5.56$, $p < .05$]. Reaction times were faster for high-frequency words than for low-frequency words. The main effect of syllable frequency was also reliable, but only when subjects were treated as a random factor [$F1(1,21) = 5.74$, $p < .05$; $F2(1,56) = 1.54$, $p < .1$]. Again, reaction times were faster for words with low-frequency syllables than for words with high-frequency syllables. The interaction between both factors was not reliable.

The ANOVA on reaction times for non-

TABLE 3
MEAN REACTION TIMES AND PERCENTAGE OF ERRORS (IN PARENTHESES) AS A FUNCTION OF SYLLABLE FREQUENCY AND WORD FREQUENCY FOR *WORDS* AND AS A FUNCTION OF SYLLABLE FREQUENCY AND BIGRAM FREQUENCY FOR *NONWORDS*

	Words		Nonwords	
	Word frequency		Bigram frequency	
Syllable frequency	High	Low	High	Low
High	618 (2.1)	658 (6.4)	771 (10)	762 (12.1)
Low	612 (5.0)	635 (5.7)	709 (5)	734 (10)

words showed that only the main effect of syllable frequency was reliable [$F1(1,21) = 18.39, p < .001$; $F2(1,56) = 5.13, p < .05$]. Reaction times were faster for nonwords with low-frequency syllables than for nonwords with high-frequency syllables. Neither the main effect of bigram frequency nor the interaction between both factors was reliable.

The ANOVAS on the error rates for words and nonwords showed that none of the factors yielded significant effects.

Discussion

Again, the effect of lexical frequency was reliable, as in the previous experiments. But more important, after controlling for bigram frequency, there was a reliable effect of syllable frequency in words, although only in the subjects analysis. In addition, the same pattern of results was obtained for nonwords, but the effect of syllable frequency was stronger than it was for words. The effect of bigram frequency between syllables, manipulated in the nonwords set, was not reliable.

The results are similar to those obtained in the previous experiments. Therefore, bigram frequency cannot by itself explain the syllable effect in nonwords, where there was no reliable effect of the frequency of bigrams. But it could contribute to the effect in words because the effect was attenuated by holding constant the bigram frequency.

It is important to mention that the effect of syllable frequency is similar for nonwords with low frequency bigrams (a trough pattern) than for nonwords with high frequency bigrams. The effect of syllable frequency is significant, independent of bigram frequency, suggesting again that the syllable is a very important sublexical unit for lexical access in Spanish. Similarly, as in the previous experiments there was no interaction between syllable frequency and lexical frequency in words suggesting again that lexical access in Spanish might require readers to employ a phonological route. If Spanish readers were employing a direct access route for high frequent words, syllabic effects should have been found only for the low frequency words.

EXPERIMENT 4

This experiment was carried out to replicate the syllable frequency results found in Experiment 3, but with a naming task. Another purpose of this experiment was to evaluate the effects of the syllable frequency in naming nonwords, because in this case, subjects could resolve the task without lexical access, by just following phonological rules. In fact, nonword naming provides one of the clearest cases of involvement of the phonological system with orthography. To pronounce a nonword, the reader must convert an orthographic input into a phonological output.

Method

Subjects. The subjects were 20 undergraduate students. They received course credit for participating.

Design and materials. The design and materials were similar to those of Experiment 3. The same 60 words and 60 nonwords were used.

Procedure. The procedure was similar to that of Experiment 2, but involved a naming task. The experimental stimuli were presented randomly within two different blocks. One block of stimuli contained only words and the other only nonwords. First the words were presented to the subjects and then the nonwords.

Results

Mean naming latencies are shown in Table 4. Separate analyses were performed for word and nonwords.

The ANOVA on naming latencies for words showed only a reliable main effect of syllable frequency, when subjects were treated as a random factor [$F1(1,19) = 9.63$, $p < .01$; $F2(1,56) = 2.88$, $p < .1$]. Again, naming latencies were faster for words with low-frequency syllables than for words with high-frequency syllables. But surprisingly, the main effect of word frequency was not reliable, although the means are in the same direction as in the previous experiments. Naming latencies were slightly faster for high-frequency words than for low-frequency words. The interaction between both variables was not reliable.

The ANOVAS on nonwords showed that the main effect of syllable frequency was reliable [$F1(1,19) = 64.61$, $p < .001$; $F2(1,56) = 8.66$, $p < .005$]. In contrast with the results obtained from words, naming latencies were faster for nonwords with high frequency syllables than for nonwords with low frequency syllables. The main effect of bigram frequency was not reliable. However, the interaction between both factors was reliable [$F1(1,19) = 9.68$, $p < .01$; $F2(1,56) = 4.70$, $p < .05$].

TABLE 4
MEAN REACTION TIMES AS A FUNCTION OF SYLLABLE FREQUENCY AND WORD FREQUENCY FOR *WORDS* AND AS A FUNCTION OF SYLLABLE FREQUENCY AND BIGRAM FREQUENCY FOR *NONWORDS*

	Words		Nonwords	
	Word frequency		Bigram frequency	
Syllable frequency	High	Low	High	Low
High	638	644	646	663
Low	622	625	688	676

Discussion

The pattern of data obtained in this experiment both for syllable frequency and lexical frequency in words is similar to that of the previous experiments. However, the lexical frequency effect was not statistically reliable. The absence of differences in lexical frequency seems a bit odd, especially when it was found in the previous experiments and also in some other studies (e.g., Andrews, 1989; McRae, Jared, & Seidenberg, 1990). Nevertheless, it should also be noted that word frequency commonly generates smaller effects in naming than in lexical decision (Paap et al., 1987; Andrews, 1989).

The main effect of syllable frequency was reliable in the subjects analysis even though words did not show the trough pattern. The fact that bigram frequency was controlled in this experiment could explain why a weaker effect of syllable frequency was obtained here as compared to Experiment 2. A naming task was employed in Experiments 2 and 4. The bigram frequency alone cannot explain the effect of syllable frequency in word recognition, but as it happened in Experiment 3, the control for bigram frequency decreases to some extent the syllable frequency effect. Furthermore, the results obtained with nonwords, where the effect of syllable frequency was clearly reliable even with stimuli with a big trough pattern, cannot be accounted for by the orthographic redundancy hypothesis. The interaction found was not the one that the orthographic redundancy hypothesis would predict. We found naming time differences between high- and low-frequency syllables to be bigger for high-frequency bigrams than for low-frequency bigrams. But, the orthographic redundancy hypothesis would predict the opposite result.

It is very important to notice that the effect of syllable frequency for nonwords was the opposite to that of words. It seems to us that lexicality affects the syllabic effect. When there is no access to a lexical entry, then the typical frequency effect is obtained. That is, more frequent syllables are read faster than less frequent syllables. However, when the task involves access to lexical entries, the reverse effect occurs. This suggests that the syllabic effect—slower naming and lexical decision times for high frequency syllables—occurs at the lexical level during the activation of lexical candidates.

Experiment 5

Experiment 4 showed a different pattern of the syllabic frequency effect for words and nonwords and no differences in the case of the lexical frequency. The purpose of this experiment is to replicate the effect found in Experiment 4 for nonwords. This effect is different to that found for nonwords when a lexical decision task was used (Experiment 1 and Experiment 3) and also to that found for words, both when a lexical decision task and a naming task were used. In addition, we wanted to see whether the syllable frequency effect was similar when phonological cues are stressed in the naming task. Words and nonwords were presented for naming in the same block in order to make the naming task phonologically more demanding than that of Experiment 4. When words and nonwords are presented in the same block, we assume that the most efficient strategy for subjects to cope with the task is to use only a phonological route.

Method

Subjects. Seventeen undergraduate psychology students participated in this experiment and they received course credit for participating.

Design and materials. The design and materials were similar to those employed in Experiments 3 and 4.

Procedure. The procedure was similar to that of Experiment 4, except that words and nonwords were presented randomized within the same block.

Results

Mean naming latencies are shown in Table 5. Separate analyses were performed for words and nonwords.

The ANOVA on naming latencies for words showed that neither the main effects nor the interaction was reliable.

The ANOVA on naming latencies for nonwords showed a reliable main effect of syllable frequency [$F1(1,16) = 14.56$, $p < .005$; $F2(1,56) = 6.704$, $p < .05$]. Naming latencies were faster for nonwords with high-frequency syllables than for nonwords with low-frequency syllables. The main effect of bigram frequency was not statistically reliable. The interaction between both variables was reliable in the subjects analysis only [$F1(1,16) = 5.09$, $p < .05$].

Discussion

As in Experiment 4, word frequency was not statistically reliable. The effect of syllable frequency when naming words was not reliable. The effect of syllable frequency for nonwords was highly reliable. Finally, the interaction between syllable

TABLE 5
MEAN REACTION TIMES AS A FUNCTION OF SYLLABLE FREQUENCY AND WORD FREQUENCY FOR *WORDS* AND AS A FUNCTION OF SYLLABLE FREQUENCY AND BIGRAM FREQUENCY FOR *NONWORDS*

	Words		Nonwords	
	Word frequency		Bigram frequency	
Syllable frequency	High	Low	High	Low
High	747	746	808	822
Low	747	751	854	843

frequency and bigram frequency showed the same pattern as that in Experiment 4. Subjects had shorter naming latencies for frequent syllables than for infrequent syllables; this difference was bigger for nonwords with high-frequency bigrams than for nonwords with low-frequency bigrams.

All these results suggest that the syllable is an important sublexical unit in visual word recognition for Spanish readers, and that the syllable frequency effect is different when readers are induced to access lexical entries than when they are able to solve the task by accessing only to the phonological representation of the syllables. When there is no access to lexical entries, as happens with nonwords, then the usual frequency effect occurs. It is easier to pronounce a unit which is more frequent in language.

General Discussion

Taken together, the results of the present experiments consistently indicate that syllable frequency influences visual word recognition in Spanish. Experiments 1 and 3 show that reaction times both for words and nonwords were slower when they had high-frequency syllables. Experiment 2, and to a lesser extent Experiment 4 (in this case the effect was only reliable in the subjects analysis) also showed this effect in naming words. In addition, the opposite effect of syllable frequency was found for the naming of nonwords in Experiments 4 and 5. Naming times for nonwords were slower when they had low-frequency syllables.

These results suggest that Spanish readers routinely use the syllable as an access unit during visual word recognition. The same effects were found both for high-frequency and low-frequency words as well as for nonwords. Some PDP models of lexical processing abandoned entirely the notion of lexical access and any type of intermediate representation corresponding to syllables or morphemes (McClelland & Rumelhart, 1981; Seidenberg & McClelland, 1989; Seidenberg, 1987, 1989). The orthographic redundancy hypothesis claims that the identification of words and any sublexical units (e.g., syllables) is an emergent property of the parallel activation process (Seidenberg, 1987). According to this hypothesis, any syllabic effect could be reduced to the statistical distribution or cooccurrence of letters, which could be captured through the bigram frequencies: syllables would be composed of letters which cooccur more often than those which mark the syllable boundaries. To test this assumption empirically we constructed stimuli that did not exhibit the trough pattern. The frequency of bigrams and the frequency of syllables were orthogonally manipulated for nonword stimuli. For words, the frequency of syllables was manipulated while the frequency of bigrams within and between syllables was held constant. The results showed that the effect of syllable frequency was still reliable and that the frequency of bigrams could not account for this effect. Therefore, it seems that the syllable is an important sublexical access unit for Spanish readers and perhaps also for readers of other shallow languages. As we pointed out, Spanish is a language with a shallow orthography and a clear syllabic structure, as opposed to English. The syllabic method is one of the most frequently employed methods for teaching children to read Spanish (Hernández & Jiménez, 1987). We are not claiming that this is a universal strategy. It is quite possible that readers in different languages use different strategies, depending on, among other things, the deep–shallow dimension.

Our results could be explained within the framework of a PDP model, which includes the syllable as a sublexical unit routinely used in visual word and nonword recognition or identification and naming (production). According to that model high-frequency syllables fire a larger set of lexical candidates than low-frequency syllables. It is reasonable to assume that high-frequency syllables are shared by a larger number of words than low-frequency syllables. In fact, this is likely the main reason why they became high-frequency syllables. Assuming also that it takes more time to select a word from a large set than from a small one, then it follows that high-frequency syllables should produce slower processing times than low-frequency syllables in a lexical decision task. The mechanism of mutual inhibition between activated word level nodes implemented in the interactive–activation model (McClelland & Rumelhart, 1981) provides an apparently accurate account of the processes whereby word candidates compete for identification.

However, it is important to notice that the effect of syllable frequency is different depending on the lexicality of the task. When the task does not require access to lexical entries as seems to happen in our naming tasks for nonwords, the frequency of the syllable speeds processing. In contrast, when the task requires access to lexical entries, the syllable frequency slows processing because its effect occurs at the lexical level during the activation of lexical candidates. This could explain why the same effect was obtained for words and nonwords in Experiment 3. Lexical decision seems to require lexical processing and subjects might use this strategy for all stimuli (words and nonwords), because they do not know if the next stimulus will be a word or a nonword. However, just a phonological strategy could be used to cope with the naming task. Subjects can easily name words and nonwords because, as we pointed out, there are transparent translation rules from orthography to phonology in Spanish. This explanation could account for the results obtained for nonwords in naming tasks. It should be stressed that the typical frequency effect—faster reaction times for higher frequency syllables—was obtained for nonwords with a naming task, but the reverse effect—faster reaction times for lower frequency syllables—was obtained for nonwords and words in lexical decision tasks.

On the other hand, this previous explanation could apply to the results obtained with words in the naming tasks. Even when naming words, subjects can follow the same rules without accessing word lexical entries. But this is not what seems to happen. If so, the same results should have been obtained both for words and nonwords in the naming task. The results suggest that when subjects see a word, even when they try to follow the orthography–phonology rules to name the word without consciously accessing word lexical entries, they can't prevent access to their lexical entries, as happens in Stroop tasks. This could explain why the effects of lexical frequency and syllable frequency sometimes prove to be statistically reliable (as in Experiment 2), but not at other times (as in Experiments 4 and 5). It seems that when both effects are reliable, the same pattern arises in the lexical decision task (Experiments 1 and 3) and in the naming task (Experiment 2). However, when the effect of lexical frequency is not reliable, the effect of syllable frequency is either reduced (Experiment 4) or is not significant (Experiment 5).

The syllable frequency effect may be related to some degree to the concept of neighborhood. The concept of neighborhood has been used to refer mainly to orthographic similarity among words. Coltheart, Davelaar, Jonasson, and Besner (1977) operationalized neighborhood as the number of words of the same length that can be generated by changing one letter of a target word, preserving letter positions. Nonetheless, it has also been used to refer to phonetic similarity among words (Goldinger, Luce, & Pisoni, 1989), defined in this case as a collection of words that are phonetically similar to a given stimulus word.

Neighborhood size, neighborhood frequency, and syllable frequency might be related. Because high-frequency syllables are shared by a larger number of words than are low-frequency syllables, it is likely to assume that neighborhood size will be larger for a word that contains high-frequency syllables. In addition, it is plausible to assume that it is more probable to find a high-frequency neighbor for a word within a large neighborhood. However, it is also the case that the number of a word's neighbors that can be created by changing a letter is more restricted than is the number of lexical candidates that could be produced by changing a syllable. For instance, the word VENA (VEIN) would have VELA (SAIL) as a neighbor, but other plausible lexical candidates generated by changing the second syllable could be VELO (VEIL) and VELLO (HAIR). Therefore, although they might be related, they are not entirely equivalent.

The manipulation of neighborhood size has not produced consistent results. For example, Coltheart et al. (1977) found no effects of neighborhood size on lexical decision times for words, but they did find an inhibitory effect for nonwords. Nonwords from large neighborhoods were slower than nonwords from small neighborhoods. Grainger, O'Regan, Jacobs, & Segui (1989) also found that lexical decision times and eye gaze durations on single words were not influenced by the number of orthographic neighbors. In contrast, Luce and Pisoni (1989) observed an inhibitory effect of neighborhood size; words from small neighborhoods were recognized more quickly than words from large neighborhoods.

In contrast to these effects, Andrews (1989, 1992) has reported facilitatory effects of neighborhood size in lexical decision for low-frequency words. Neighborhood size speeded naming times for both high- and low-frequency words, but produced null or inhibitory effects for nonwords in lexical decision. Neighborhood frequency, on the other hand, showed inhibitory effects in several experiments. Words with high-frequency neighbors were harder to recognize than words with low-frequency neighbors (Grainger, 1990; Grainger & Segui, 1990; Grainger et al., 1989; Luce & Pisoni, 1989).

The inhibitory effects of syllable frequency on naming and lexical decision times that we found for words and nonwords, thus appear to conflict with the facilitatory effects of neighborhood size for words reported by Andrews (1989, 1992), although our findings are consistent with the inhibitory effects of neighborhood size observed for words (Luce & Pisoni, 1989) and nonwords (Coltheart et al., 1977; Andrews, 1989). They are also consistent with the inhibitory effects found when the neighborhood frequency was manipulated and with the inhibitory effects observed for target words when preceded by an orthographic related prime (Colombo, 1986; Segui & Grainger, 1990; Grainger, 1990).

Finally, our results, suggest that the phonological route might be mandatory for visual word recognition in Spanish. The effects of syllable frequency did not interact with lexical frequency and they were similar for words and nonwords. Phonological activation could be a routine, and essentially, an automatic part of word identification in Spanish, as it could be in other shallow languages. Shallow orthographies such as Spanish and Serbo-Croatian (see Lukatela & Turvey, 1990a,b) might encourage more phonological processing than relatively deep orthographies such as English or a deeper-still orthography such as Hebrew (Frost et al., 1987). However, even with deep orthographies, such as English, there is strong evidence in support of phonological activation which has come from backward masking (Perfetti, Bell, & Delaney, 1988; Perfetti & Bell, 1991), priming studies (Perfetti & Bell, 1991), and from homophonic interference in category decisions (Van Orden, 1987; Van Orden, Johnston, & Hale, 1988).

We do not claim that our results overturn the conclusions drawn in favor of the dual route hypothesis or the orthographic redundancy hypothesis in reading English. However, we do think that our results dispute the universal application of both hypotheses, and keep open the case for a phonological activation. At the same time, our results are consistent with dual route theories if they assume that phonological activation always occurs, as proposed by Perfetti, Zhang, and Berent (in press). The visual route could be an optional characteristic of the model that changes depending on the deep–shallow dimension, among other things. Our data are also consistent with a single route PDP model if it includes syllables as sublexical units. Finally, our results also imply a psychological reality for syllables in Spanish and perhaps in other shallow languages.

References

Adams, M. (1981). What good is orthographic redundancy? In H. Singer and O. J. L. Tzeng (Eds.), *Perception of print.* Hillsdale, NJ: Erlbaum.

Alvarez, C. J., Carreiras, M., & De Vega, M. (1992a). Estudio estadístico de la ortografía castellana: (1) La frecuencia silábica. *Cognitiva,* **4,** 75–105.

Alvarez, C. J., Carreiras, M., & De Vega, M. (1992b). Estudio estadístico de la ortografía castellana: (2) La frecuencia de bigramas. *Cognitiva,* **4,** 107–125.

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory and Cognition,* **15,** 802–814.

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity of orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory and Cognition,* **18,** 234–254.

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 and Performance,* **3,** 340–357.

Besner, D. (1990). Does de reading system need a lexicon? In D. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading.* Hillsdale, NJ: Erlbaum.

Besner, D., & Humphreys, G. W. (1990). *Basic processes in reading: Visual word recognition.* Hillsdale, NJ: Erlbaum.

Besner, D., & McCann, R. S. (1987). Word frequency and pattern distortion in visual word identification and production: An examination of four class of models. In M. Coltheart (Ed.), *Attention and performance: Reading* (Vol. 7). Hillsdale, NJ: Erlbaum.

Broadbent, D. E., & Gregory, M. (1968). Visual perception of words differing in letter digram frequency. *Journal of Verbal Learning and Verbal Behavior,* 7, 569–571.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies in information processing.* New York: Academic Press.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and Performance* (Vol. 6). Hillsdale, NJ: Erlbaum.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language,* **25,** 385–400.

De Vega, M., & Carreiras, M. (1989). The role of graphemic frequency in visual word processing. Paper presented at the 3rd European Conference for Learning and Instruction, Madrid, Spain.

De Vega, M., Carreiras, M., Gutierrez, M., & Alonso, M. L. (1990). *Lectura y comprensión: Una perspectiva cognitiva.* Madrid: Alianza Editorial.

Feldman, L. B., & Turvey, M. T. (1983). Word recognition in Serbo-Croatian is phonologically analytic. *Journal of Experimental Psychology: Human Perception and Performance,* **9,** 288–298.

Frederiksen, J. R., & Kroll, J. F. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance,* **2,** 361–379.

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multidimensional comparison. *Journal of Experimental Psychology: Human Perception and Performance,* **13,** 104–115.

Goldinguer, S. D., Luce, P. A. & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language,* **28,** 501–518.

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language,* **29,** 228–244.

Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J. (1989). On the role of competing word units in visual word recognition. *Perception and Psychophisics,* **45,** 189–195.

Grainger, J., & Segui, J. (1990). Neighborhood frequency effects in visual word recognition: A comparison of lexical decision and masked identification latencies. *Perception and Psychophisics*, **47**, 191–198.

JULLIAND, A., & CHANG-RODRÍGUEZ, E. (1964). *Frequency dictionary of Spanish words*. La Haya: Mouton.

HERNÁNDEZ, P., & JIMENEZ, J. (1987). Influencia de los métodos de lectura en los hábitos de trabajo intelectual de los escolares. *Infancia y Aprendizaje*, **39**, 15–26.

LUCE, P. A., & PISONI, D. B. (1989). Neighborhoods of words in the mental lexicon (cited in Goldinguer, Luce, & Pisoni, 1989).

LUKATELA, G., & TURVEY, M. T. (1990a). Phonemic priming with words and pseudowords. *European Journal of Cognitive Psychology*, **2**, 325–343.

LUKATELA, G., & TURVEY, M. T. (1990b). Phonemic similarity effects and prelexical phonology. *Memory and Cognition*, **18**, 128–152.

MCCLELLAND, J. L., & RUMELHART, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic finding. *Psychological Review*, **88**, 375–407.

MCRAE, K., JARED, D., & SEIDENBERG, M. S. (1990). On the roles of frequency and lexical access in word naming. *Journal of Memory and Language*, **29**, 43–65.

MEHLER, J., DOMMERGUES, J., FRAUNFELDER, U., & SEGUI, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, **20**, 298–305.

MORAIS, J., CONTENT, A., CARY, L., MEHLER, J., & SEGUI, J. (1989). Syllabic segmentation and literacy. *Language and Cognitive Processes*, **4**, 57–67.

PAAP, K. R., MCDONALD, J. E., SCHVANEVELDT, R. W., & NOEL, R. W. (1987). Frequency and pronounceability in visually presented naming and lexical decision tasks. In M. Coltheart (Ed.), *Attention and Performance: Reading* (Vol. 7). Hillsdale, NJ: Erlbaum.

PAAP, K. R., & NOEL, R. W. (1991). Dual-route models of print and sound: Still a good horse race. *Psychological Research*, **53**, 13–24.

PERFETTI, C. A., & BELL, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and masked priming. *Journal of Memory and Language*, **30**, 473–485.

PERFETTI, C. A., BELL, L., & DELANEY, S. (1988). Automatic phonetic activation in silent word reading: Evidence from backward masking. *Journal of Memory and Language*, **27**, 59–70.

PERFETTI, C. A., ZHANG, S., & BERENT, I. (in press). Reading in English and Chinese: Evidence for a "Universal" phonological principle. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology and meaning*. Amsterdam/New York: Elsevier.

POSNER, M. I., & CARR, T. H. (1992). Lexical access and the brain: Anatomical constraints on cognitive models of word recognition. *American Journal of Psychology*, **105**, 1–26.

PRINZMETAL, W., TREIMAN, R., & RHO, S. H. (1986). How to see a reading unit. *Journal of Memory and Language*, **25**, 461–475.

RAPP, B. C. (1992). The nature of sublexical orthographic organization: The bigram trough hypothesis examined. *Journal of Memory and Language*, **31**, 33–53.

SEBASTIAN-GALLES, N., DUPOUX, E., SEGUI, J., & MEHLER, J. (1992). Contrasting syllabic effects in Catalan and Spanish. *Journal of Memory and Language*, **31**, 18–32.

SEGUI, J., & GRAINGER, J. (1990). Priming word recognition with orthographic neighbors: The effects of relative prime-target frequency. *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 65–76.

SEIDENBERG, M. S. (1987). Sublexical structures in visual word recognition: Access units or orthographic redundancy? In M. Coltheart (Ed.), *Attention and performance: The psychology of reading* (Vol. 7). Hillsdale, NJ: Erlbaum.

SEIDENBERG, M. S. (1989). Reading complex words. In G. Carlson & M. Tanenhaus (Eds.), *Linguistic structure in language processing*. Boston: Kluwer Academic Publs.

SEIDENBERG, M. S., & MCCLELLAND, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, **96**, 523–568.

TAFT, M. (1979). Lexical access via an orthographic code: The BOSS. *Journal of Verbal Learning and Verbal Behavior*, **18**, 21–39.

VAN ORDEN, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory and 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 and Cognition*, **14**, 371–385.

VAN ORDEN, G. C., PENNINGTON, B., & STONE, G. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review*, **97**, 488–522.

(Received June 6, 1992)

(Revision received February 1, 1993)