Sound symbolism

Edited by

LEANNE HINTON, JOHANNA NICHOLS,
AND JOHN J. OHALA
University of California at Berkeley

CAMBRIDGE
UNIVERSITY PRESS
1994

18

*Phonosyntactics**

JOAN A. SERENO

18.1. Introduction

Although the search for linguistic universals has been successful in discovering a wide variety of similarities among the languages of the world (Greenberg 1966, 1978), few generalities about the semantic properties of languages have been uncovered (Weinreich 1963). Some of the most striking examples of a structured semantic domain are to be found in the connection between sound and meaning. In principle, resemblances between meaning and sound should not exist. As Greenberg (1957) notes, the connection between sound and meaning is essentially arbitrary. A meaning can theoretically be represented by almost any set of sounds in a language.

Despite claims that this principle applies to all languages, the articles in this volume show that connections between the meanings and the sounds of language do exist. Two universal and well-documented cases of sound symbolism are relevant to the present study. First, Ultan (1978) found evidence for distance symbolism. The notion of distance symbolism must be considered within the broader framework of deixis. Deixis refers to those features of the language which reflect the spatio-temporal coordinates of the relative situation of the utterance. Distinctions are made between things that are near in space and/or time versus things that are far in space and/or time. In an analysis of 136 languages, Ultan found that 33.1% of the sample exhibited distance (spatial) symbolism in their demonstrative system. More importantly, the languages that overtly symbolized distance relationships predominantly used front or high front vowels to represent proximity to the speaker. Ultan also found universal correspondences for size symbolism in language. Some languages overtly mark words expressing diminution by changing the phonological features of the sound in the root. Ultan found that 27.3% of the 136 languages he sampled had diminutive marking. In almost 90% of these languages, the diminutive was symbolized by high front vowels. The widespread distribution of these consistently recurring patterns of sound symbolism suggests that the relation between sound and meaning in language is not completely arbitrary.

Taking these universal semantic–sound correspondences as a model, a different type of sound symbolism is suggested here. Correspondences between *sound* and the clearly delineated and linguistically salient categories of syntactic class are analyzed. Specifically, the syntactic classes of "noun and "verb" are compared in terms of the phonological classification of their vowels. The data to be presented in this chapter suggest that there exists a peripheral type of sound symbolism in which specified phonological features are associated with different syntactic categories.

A preliminary lexical database analysis was undertaken to carefully examine the sound–syntactic class relationship. The reference set of words was the list usually referred to as the Brown Corpus compiled by Francis and Kučera (1982). Francis and Kučera categorized over one million American English words, recording both lexical and grammatical information. The frequency of occurrence as well as the grammatical category of every word is listed.

Francis and Kučera rank-order the words in the Brown Corpus in terms of lemmata. They define the term "lemma" as a set of words belonging to the same major word class. For example, the base form of the word "comment" consists of two separate lemmata – a noun lemma with a frequency of 64 per million and a verb lemma with a frequency of 31 per million. A noun lemma includes singular or mass nouns, possessive singular nouns, plural nouns, possessive plural nouns, and singular proper nouns. A verb lemma includes verbs in the base form and inflected verbs, such as third-person singular forms, past tense forms, past participles, present participles, and gerunds. The modals ("can," "may," "shall," "will") and the non-standard verbs "be," "have," and "do" are categorized separately and are not included in the verb lemmata.

In the present study, the first 1000 noun and verb lemmata in the rank list (a descending order of lemmata by frequency) were first classified according to the phonological category of their stressed vowel (front vowel [i, ɪ, e, ɛ, æ] vs. back vowel [ɚ, ɑ, ʌ, ɔ, o, ʊ, u, aɪ, aʊ, ɔɪ].[1] This division into front and back vowels was motivated by previous universal sound-symbolism research in which front or high front vowels are implicated. In the present analysis, the base form of each word was used. Diphthongs were classified on the basis of the more prominent first vowel sound (Ladefoged 1993). A previous analysis (Sereno 1983), using the Lorge (1949) frequency word list, had shown that the exclusion of all possible controversial cases did not affect the categorization results. In that analysis, multisyllabic words, words containing diphthongs, and words that change the categorization of the stressed vowel (e.g. strong verbs with alternative forms, such as "see/saw") were excluded. This restricted sample of words nevertheless showed identical associations in the comparison of syntactic class membership and phonological categorization of the

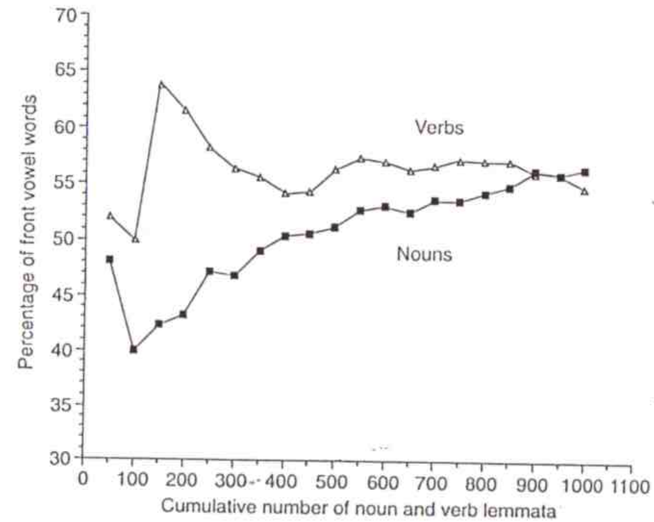

Figure 18.1. Percentage of nouns and verbs in English with front vowels as tabulated from the rank list of the Brown Corpus.

stressed vowel. The present analysis of the Brown Corpus used the broader sample of words.

The phonological analysis of nouns and verbs in the Brown Corpus revealed a systematic, skewed distribution. In general, nouns were more likely to have back vowels rather than front vowels, while verbs were more likely to have front vowels compared to back vowels. However, this distribution was frequency-dependent, occurring only in high-frequency words. The dichotomy between high-frequency nouns and verbs and their low-frequency counterparts with regard to the phonological category of their stressed vowel shows up clearly in figure 18.1. Figure 18.1 represents a cumulative listing (calculated every fifty lemmata) of the first thousand nouns and verbs categorized in terms of the front/back quality of the stressed vowel. A greater number of high-frequency nouns have back vowels, while high-frequency verbs have a greater number of front vowels. However, this pattern does not occur in lower-frequency words. As lower frequency words are included, an equal number of front and back vowels is found in nouns and verbs.

To quantify and evaluate this pattern, two sets of 200 stimuli (a high-frequency set and a low-frequency set) were analyzed in terms of vowel quality. For the high-frequency set, the first 200 nouns and verbs in the rank list representing words with frequencies greater than 250 per million were analyzed. In these high-frequency words, nouns were more likely to have back vowels rather than

Table 18.1 *Vowel categorization of the first 200 high-frequency nouns and verbs (frequency greater than 250 per million). Percentages represent the proportion of a single form class (i.e. noun or verb) having front or back vowels.*

Front vowels

	i	ɪ	e	ɛ	æ	Total
Nouns	9	9	12	15	10	55 (43%)
Verbs	17	8	7	7	6	45 (62%)

Back vowels

	ɚ	ɑ	ʌ	ɔ	o	ʊ	u	aɪ	aʊ	ɔɪ	Total
Nouns	8	12	10	3	4	10	5	13	4	3	72 (57%)
Verbs	3	3	3	3	1	5	3	7	0	0	28 (38%)

Table 18.2 *Vowel categorization of the first 200 low-frequency nouns and verbs (frequency less than 50 per million) in the rank list of the Brown Corpus. Percentages represent the proportion of a single form class having front or back vowels.*

Front vowels

	i	ɪ	e	ɛ	æ	Total
Nouns	4	20	25	17	9	75 (54%)
Verbs	3	5	8	7	6	30 (48%)

Back vowels

	ɚ	ɑ	ʌ	ɔ	o	ʊ	u	aɪ	aʊ	ɔɪ	Total
Nouns	9	12	6	0	8	1	11	11	3	2	63 (46%)
Verbs	3	5	4	2	5	1	2	7	2	1	32 (52%)

front vowels in their stressed syllable, while verbs were more likely to have front vowels than back vowels (see table 18.1). Only 43% of nouns have front vowels, while 57% have back vowels. This pattern is reversed for the verbs, of which 62% have front vowels while only 38% have back vowels. A Chi-Square test for the high-frequency words showed that there is indeed a significant relationship between the syntactic category of the word and the front/back quality of the stressed vowel of the word ($\chi^2 = 6.23$, $p < 0.025$). There is a significantly greater number of nouns with back vowels and verbs with front vowels.

A similar distribution, however, was not maintained for low-frequency nouns and verbs. For the low-frequency set, the first 200 nouns and verbs in the rank list having a frequency less than 50 per million were analyzed. In this set of words, the proportion of front and back vowels for nouns and verbs is virtually identical (see table 18.2): 54% of nouns have front vowels and 46% have back vowels, while 48% of verbs have front vowels and 52% have back vowels. A Chi-Square test for these low-frequency words showed that there was no significant relationship between the syntactic category of the word and the vowel quality of the stressed vowel ($\chi^2 = 0.61$, $p > 0.50$, n.s.).

To summarize, then: for high-frequency words, there is a significantly greater number of nouns with back vowels and verbs with front vowels, but this pattern does not hold for low-frequency nouns and verbs in English. The sound-symbolic correspondences found in studying linguistic universals suggest the possibility that languages of the world do not represent completely arbitrary symbolic systems. As shown above, an analysis of high-frequency words occurring in modern American English reveals a definite relationship between syntactic category and the front/back quality of the stressed vowel. To further investigate these correspondences, a psycholinguistic experiment was conducted to determine whether this relationship is effective in the processing of language.

In this experiment, a series of words was presented to subjects. Their task was to categorize the stimulus word as a noun or verb. Stimulus words were divided into separate groups: a high-frequency set and a low-frequency set. If it is the case that a systematic relationship obtains between syntactic class and phonological features of the language, it may be expected that nouns with back vowels and verbs with front vowels will be processed faster. In addition, a comparison of the high-frequency set to the low-frequency set might clarify whether such an effect (i.e. speeded response latencies for back vowel nouns and front vowel verbs) is simply a distributional artifact of the language (response latency differences present only in high frequency words) or a general processing strategy.

18.2. Methods

18.2.1. Subjects

Twelve students (eight male, four female) attending Brown University volunteered to participate in the experiment. All were native speakers of American English with normal or corrected-to-normal visual acuity.

18.2.2. Stimuli

A total of 64 nouns and verbs was selected from the Brown Corpus. The stimuli are shown in appendix 18.1. The first set consisted of 32 high-frequency words (occurring more than 250 times per million) and the second set consisted of 32 low-frequency words (occurring between 30 and 50 times per million).

Within each set of words, all groups (front-vowel nouns, back-vowel nouns, front-vowel verbs, and back-vowel verbs) were matched for word frequency (Francis and Kučera 1982). In the high-frequency set, mean frequency of occurrence for front-vowel nouns, back-vowel nouns, front-vowel verbs, and back-vowel verbs was 452, 443, 442, and 437, respectively, with standard deviations of 179, 164, 153, 133, respectively. Low-frequency words were similarly matched. Mean frequency of occurrence was 42, 41, 39, and 44, respectively, with standard deviations of 5, 4, 10, 6, respectively.

Only monosyllabic, consonant-initial words were used (see Spoehr and Smith 1973). Mean letter frequency of initial consonants (Baddeley *et al.* 1960) and word length (Forster and Chambers 1973) were matched among all groups. Finally, only orthographically regular words (see Parkin 1982) were used in the experiment.

18.2.3. Procedure

All subjects were tested individually. They were told to respond as quickly and accurately as possible to each stimulus item. One half of the subjects were first given the High-Frequency List followed by the Low-Frequency List. The other half of the subjects were first given the Low-Frequency List followed by the High-Frequency List. After a short interval, the entire test procedure was repeated.

Following instructions, subjects were given a set of 20 practice items to introduce them to the procedure. A separate practice set accompanied each test set of words (High Frequency and Low Frequency). Practice items were not used in the experiment. For the test, each list (High Frequency, Low Frequency) consisted of 32 words – 16 nouns and 16 verbs. One half of each List contained front vowels and the remaining half contained back vowels. The 32 words were presented in a random order to each subject.

The experiment was controlled and presented on an Apple IIe computer. The stimuli, all in upper case letters, subtended a visual angle of approximately 2° horizontally and 0.5° vertically. Subjects were seated two feet in front of the display. All responses to the stimuli were made by pressing one of two clearly marked buttons on a control box placed in front of the subject. The entire experiment lasted approximately 20 minutes.

Subjects were instructed to identify each stimulus either as a noun or a verb. Each trial was completed when a subject moved the index finger of the preferred hand from a neutral location and pressed one of two equidistantly placed response buttons labeled "noun" or "verb."

Subjects were told that the stimuli were all familiar English words. They were further informed that, although some of the words could occur both as a noun and a verb, each stimulus was to be categorized on the basis of its more frequent usage.

Table 18.3 *Mean response latencies (in milliseconds) for each presentation of the test. Total number of errors is given in parentheses.*

First presentation							
High-frequency words				*Low-frequency words*			
Noun		Verb		Noun		Verb	
Front vowel	Back vowel	Front vowel	Back vowel	Front vowel	Back vowel	Front vowel	Back vowel
784	714	740	762	808	745	844	846
(3)	(1)	(10)	(5)	(3)	(5)	(7)	(9)
Second presentation							
High-frequency words				*Low-frequency words*			
Noun		Verb		Noun		Verb	
Front vowel	Back vowel	Front vowel	Back vowel	Front vowel	Back vowel	Front vowel	Back vowel
743	687	737	720	769	717	785	803
(5)	(3)	(5)	(2)	(3)	(3)	(7)	(8)

The stimuli were presented at a fixed rate. Initially, a fixation pattern "****" appeared at the center of the display screen. This pattern disappeared after 1.0 second, followed by a blank screen for 0.25 seconds. The stimulus word then appeared on the display screen and remained there until the subject responded. Reaction times were measured from the appearance of the stimulus on the screen until a key press was made. Following a response, the stimulus word disappeared, leaving a blank screen for three seconds until the next fixation pattern appeared. This entire sequence was repeated for every stimulus item.

18.3. Results

The mean latencies of noun/verb classifications are given in table 18.3. No errors were included in these averages, and all reaction times more than three standard deviations from the individual's mean were discarded. A four-way repeated measures ANOVA (Frequency List × Syntactic Class × Vowel Category × Trial) revealed a main effect for Frequency List (high frequency versus low frequency), [$F_{(1,11)} = 7.33$, $p < 0.02$], and Vowel Category (front vowel versus back vowel),

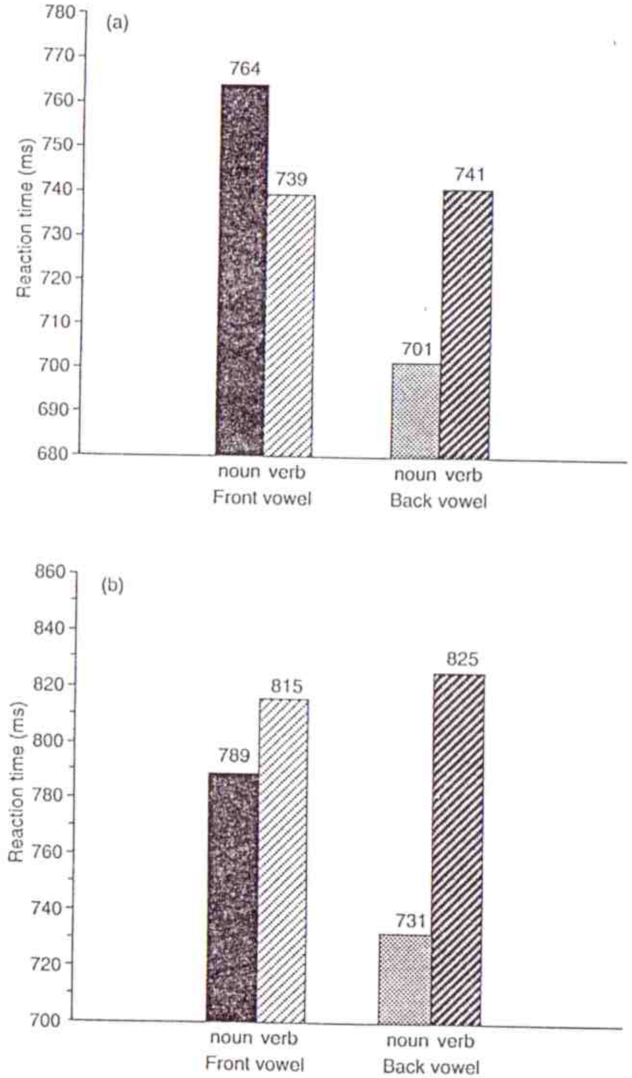

Figure 18.2. Response latencies (in milliseconds) for the classification of nouns and verbs as a function of vowel quality (front vowels vs. back vowels) in (a) high-frequency stimuli and (b) low-frequency stimuli.

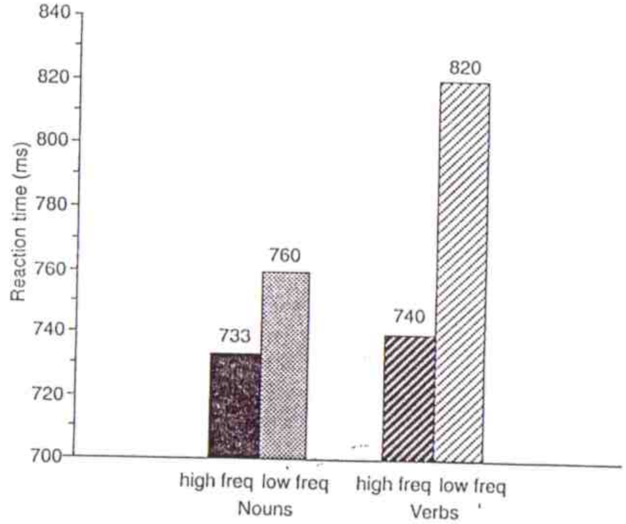

Figure 18.3. Response latencies (in milliseconds) for the classification of nouns and verbs as a function of frequency (high frequency vs. low frequency).

[$F_{(1,11)} = 24.49$, $p < 0.0004$]. High-frequency words (736 ms) were categorized faster than low frequency words (790 ms), and back-vowel words (750 ms) faster than front-vowel words (777 ms).

There was also a significant main effect for Trial, [$F_{(1,11)} = 26.06$, $p < 0.0003$]. Total mean response latencies were 780 ms on initial tests and 745 ms on repetition tests.

More importantly, however, only one significant interaction between factors was found. Specifically, there was a significant interaction between Syntactic Class (noun versus verb) and Vowel Category (front vowel versus back vowel), [$F_{(1,11)} = 11.54$, $p < 0.006$]. Nouns with back vowels (716 ms) were categorized significantly faster than nouns with front vowels (777 ms), and verbs with front vowels (776 ms) faster than verbs with back vowels (783 ms). This effect, moreover, was found for both high-frequency and low-frequency stimuli (see figures 18.2a and 18.2b).

Two interesting trends were also present in the data. Noun categorization (747 ms) was slightly faster than verb classification (780 ms), [$F_{(1,11)} = 3.40$, $p < 0.09$ (n.s.)]. In addition, a slight interaction was found between Frequency List (high frequency vs. low frequency) and Syntactic Class (noun vs. verb), [$F_{(1,11)} = 4.34$, $p < 0.06$ (n.s.)] (see figure 18.3). It seems that categorization latencies for frequency-matched nouns and verbs are approximately equal (733 ms and

740 ms, respectively) in commonly used, high-frequency words, but differ in low-frequency words where response latencies to verbs (820 ms) are slower than to nouns (760 ms).

The mean number of errors is shown in table 18.3. These totals include both mistakes in categorization (i.e. making a noun decision for a verb stimulus or vice versa) and trials in which latencies exceed three standard deviations of the subject's mean. A total of 79 errors were made, representing less than 5% of all responses. A four-way repeated measures ANOVA of errors (Frequency List × Syntactic Class × Vowel Category × Trial) revealed a main effect for Syntactic Class, [$F_{(1,11)} = 7.26$, $p < 0.02$]. There were significantly more errors for verbs (53) compared to nouns (26). In addition, there was a significant interaction of Frequency List × Syntactic Class × Trial, [$F_{(1,11)} = 9.16$, $p < 0.01$]. There were fewer mistakes for low-frequency nouns on the second trial. The error analysis is consistent with the analysis of mean response latencies in that both show that categorization of verbs, particularly low-frequency verbs, is more difficult (i.e. takes more time and results in more errors).

18.4. Discussion

The present study examines a form of sound symbolism in which syntactic form classes are distinguished in terms of the phonological classification of their vowels. A lexical search of American English revealed a striking pattern involving nouns and verbs and front and back vowels. In English, nouns were more likely to have back vowels rather than front vowels while verbs were more likely to have front vowels compared to back vowels.

The results of the psycholinguistic experiment showed that, in a noun/verb categorization task, significant processing differences exist between nouns and verbs depending on the front/back quality of the stressed vowel. Nouns with back vowels were categorized faster than nouns with front vowels. Conversely, verbs with front vowels were categorized faster than verbs with back vowels. These results are in accord with similar experiments run in the auditory modality investigating these phonological and syntactic class interactions. In both an auditory noun/verb categorization task and an auditory lexical decision task, Sereno and Jongman (1990) found that noun stimuli containing back vowels were processed faster than nouns with front vowels while verb stimuli containing front vowels were processed faster than verbs with back vowels.

A further result of the present study is that the facilitation of nouns with back vowels and verbs with front vowels occur for both high-frequency and low-frequency words. Recall that the lexical analysis of English showed that the phonological and syntactic class correspondences existed only in the high-frequency words. Low-frequency words in English, on the other hand, showed

equivalent distributions of front and back vowels in nouns and verbs. The psycholinguistic data, however, do not directly reflect this pattern. The present results show that processing differences exist for both high and low-frequency words. These psycholinguistic effects suggest the presence of a general processing strategy for both high and low-frequency words. This frequency-independent nature of the phonosyntactic correspondences was not found in the auditory tasks mentioned above. Sereno and Jongman (1990) found that in these auditory tasks the processing effects directly mimic the lexical distribution of nouns and verbs with front and back vowels in the language. The present data, however, seem to suggest the operation of a general processing strategy. Two tentative explanations are possible. First, the blocked presentation of the high and low-frequency stimuli in the present experiment may have encouraged the use of a general processing strategy. Second, the present experiment was presented visually. The visual presentation of the stimuli may have contributed some additional variance that may be affecting the pattern of results. One possible contributing factor may be the use of words with orthographic/phonological vowel clash, such as "wife," which is categorized phonologically as a back vowel but orthographically as a front vowel. However, it is not apparent how inclusion of such words could result in significant sound–syntax correspondences in both high and low-frequency stimuli. Further experimentation is clearly needed to resolve these issues.

Nevertheless, the present lexical and psycholinguistic data suggest that syntactic class membership and phonological information regarding the stressed vowel are important and influential variables in processing. It should be noted that the divisions between nouns and verbs and between front and back vowels are highly salient distinctions in language. The syntactic categories of noun and verb are universal categories in languages of the world (Hockett 1968; Sapir 1944). Moreover, the division between front and back vowels is a basic phonological contrast (e.g. Chomsky and Halle 1968; Jakobson *et al.* 1963), which can be easily characterized in terms of formant frequencies (e.g., Ladefoged 1993; Lieberman and Blumstein 1988). Specifically, front vowels have high second-formant frequencies which approximate third-formant levels, resulting in a concentration of energy in relatively higher frequencies. Back vowels, on the other hand, have low second-formant frequencies. The close proximity of the second and first formants, in this case, emphasizes lower frequencies of the spectrum. The perception of vowels has been shown, for the most part, to be based on the location of formant frequencies, especially the first and second formants (e.g. Carlson *et al.* 1975). Clearly, the categories of nouns and verbs and front and back vowels could be plausible candidates in perception.

The present data suggest that syntactic-form-class membership as well as phonological classification of vowels is explicitly coded in the lexicon. The interaction between syntactic class and phonological categorization of the stressed vowel suggests that this information may be used to structure the lexicon. Noun/verb categories and front/back vowel classification seem to be possible organizing principles of the lexicon.

In conclusion, a lexical search revealed that high-frequency words participate in a systematic skewed distribution in English. This distribution showed up clearly in terms of differences in processing times for both high and low-frequency words. In a noun/verb classification task, subjects were able to exploit differences that exist in commonly used words. It seems that a general processing strategy is used to access nouns and verbs in the lexicon based on vowel categorization.

Although the noun/verb classification task is rarely used in traditional psycholinguistic experiments, this task may be more appropriate for the analysis of linguistic processing. In contrast, the widely used lexical decision task, in which subjects are to decide whether a stimulus is a word or not, requires a decision which is not commonly made by a listener in normal sentence processing. Noun/verb classification data, therefore, may better reflect lexical orgaization. The present results suggest that there is a robust interaction between syntactic class and vowel categorization and that the lexicon may be organized on this basis.

Appendix 18.1. *Stimuli used in the experiment*

High-Frequency stimuli			
Front-vowel Nouns	Back-vowel Nouns	Front-vowel Verbs	Back-vowel Verbs
hand	school	leave	want
thing	house	feel	write
fact	group	keep	put
week	church	let	try
field	door	meet	move
street	month	reach	grow
friend	book	sit	serve
death	wife	wait	lose

Low-Frequency stimuli			
Front-vowel Nouns	Back-vowel Nouns	Front-vowel Verbs	Back-vowel Verbs
milk	clerk	mix	bind
fence	porch	swim	shut
gift	dirt	snap	pour
cat	bus	beg	quote
bench	bride	seize	hunt
prince	tongue	weigh	stir
sand	corn	melt	crawl
leaf	barn	cease	wipe

NOTES

* This research was conducted at Brown University while the author was supported by a National Science Foundation Graduate Fellowship.

1 The term "back vowel" in this paper refers to both back vowels and central vowels as defined in the phonetics literature (e.g. Ladefoged 1993).

REFERENCES

Baddeley, A., R. Conrad, and W. Thomson. 1960. Letter structure of the English language. *Nature* 186: 414–416.

Carlson, R., G. Fant, and B. Granström. 1975. Two-formant models, pitch, and vowel perception. In G. Fant and M. A. A. Tatham (eds.) *Auditory Analysis and the Perception of Speech*. London: Academic Press.

Chomsky, N. and M. Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.

Forster, K. and S. Chambers. 1973. Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior* 12: 627–635.

Francis, W. N. and H. Kučera. 1982. *Frequency Analysis of English Usage: Lexicon and Grammar*. Boston: Houghton Mifflin.

Greenberg, J. 1957. *Essays in Linguistics*. Chicago: University of Chicago Press.

1966. *Language Universals*. The Hague: Mouton.

1978. *Universals of Human Language*. Stanford: University Press.

Hockett, C. F. 1968. *The State of the Art*. The Hague: Mouton.

Jacobson, R., G. Fant and M. Halle. 1963. *Preliminaries to Speech Analysis*. Cambridge, MA: MIT Press.

Ladefoged, P. 1993. *A Course in Phonetics*. New York: Harcourt Brace Jovanovich.

Lieberman, P. and S. Blumstein. 1988. *Speech Physiology, Speech Perception, and Acoustic Phonetics*. Cambridge: University Press.

Lorge, I. 1949. *The Semantic Count of the 570 Commonest English Words*. New York: Bureau of Publications, Teachers' College at Columbia University.

Parkin, A. 1982. Phonological recoding in lexical decision: effects of spelling-to-sound regularity depend on how regularity is defined. *Memory and Cognition* 10(1): 43–53.

Sapir, E. 1944. *Language*. New York: Harcourt, Brace and World.

Sereno, J. A. 1983. Phonosyntactics: sound–syntax correspondences. MA thesis, Brown University.

Sereno, J. A. and A. Jongman. 1990. Phonological and form class relations in the lexicon. *Journal of Psycholinguistic Research* 19(6): 387–404.

Spoehr, K. and E. E. Smith. 1973. The role of syllables in perceptual processing. *Cognitive Psychology* 5: 71–89.

Ultan, R. 1978. Size–sound symbolism. In J. Greenberg (ed.) *Universals of Human Language*, vol. 2 – *Phonology*. Stanford: University Press, 525–568.

Weinreich, U. 1963. On the semantic structure of language. In J. Greenberg (ed.) *Universals of Language*. Cambridge, MA: MIT Press, 142–216.