LANGUAGE AND SPEECH, 1997, 40 (1), 47-62

Phonotactics and Syllable Stress: Implications for the Processing of Spoken Nonsense Words *

MICHAEL S. VITEVITCH
PAUL A. LUCE
JAN CHARLES-LUCE
DAVID KEMMERER
University at Buffalo

KEY WORDS

phonotactics

syllable stress

word recognition

ABSTRACT

Two experiments using bisyllabic CVCCVC nonsense words that varied in phonotactic probability and stress placement were conducted to examine the influences of phonotactic and metrical information on spoken word recognition. Experiment 1 examined participants' intuitions about the phonological "goodness" of nonsense words. Experiment 2 examined processing times for the same stimuli in a speeded auditory repetition task. The results of both studies provide further evidence that the phonotactic configuration and stress placement of spoken stimuli have important implications for the representation and processing of spoken words.

INTRODUCTION

The order and position of speech sounds in spoken words are highly constrained within a particular language. For example, in English only a subset of consonants may form syllable-initial clusters, and the order of consonants within clusters is severely restricted (see Clements & Keyser, 1983). The configuration of speech sounds within syllables and words is called *phonotactics*. Specifically, phonotactics accounts for the probability that a given phonetic segment will be followed or preceded by another particular segment. In addition, phonotactics refers to the probability that a given segment will occur in a specific position within a syllable or word.

---

*This research was supported in part by research grants R01 DC 0265801, R29 DC00957, and T32 DC 00036 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.

We would like to thank the two reviewers for their advice.

Portions of this work have been presented at the International Conference on Spoken Language Processing, Philadelphia, Pennsylvania, 1996.

Correspondence concerning this article should be sent to either of the first two authors at Language Perception Laboratory, Department of Psychology, University at Buffalo, Buffalo, NY 14260; or via e-mail to: *mikev@deuro.fss.buffalo.edu*

Phonotactic configuration may have important consequences for the representation and processing of spoken words. In particular, spoken words composed of common segments arranged in regular sequences may be processed more accurately and rapidly than words composed of less common segments and sequences. Whereas this finding alone is of interest, differential effects of phonotactics on recognition may also have important consequences for models of lexical processing. One important question relating to the representation and processing of phonotactic information concerns whether listeners have access to independent information in memory regarding phonetic segments and sequences, or whether all phonotactic effects emanate from individual representations of lexical form. An answer to this question could go far in deciding between models that postulate the existence of abstract phonetic and phonological information in memory, and those that postulate that all abstract phonological information is a by-product of conspiracies of individual representations of sound patterns in the lexicon (see, for example, Norris, 1994). In addition, understanding the role of phonotactics in recognition may provide insights into the listener's ability to segment words from the speech stream. Finally, the study of the processing and representation of phonotactic information may inform us not only about adult recognition capabilities but also about the development of the lexicon in infants.

Previous research has provided some evidence for the representation of phonotactic information in memory. In particular, several studies have demonstrated that phonotactic information has implications for perceptual processing and memory representations in infants and children. Work by Jusczyk, Frederici, Wessels, Svenkerud, and Jusczyk (1993) showed that 9-month-old infants are sensitive to the phonotactic patterns in their native language. Using the head turn preference procedure with Dutch and American infants, Jusczyk et al. (1993) found that infants listened longer to lists of words in their native language. Although Dutch and English have similar prosodic characteristics, the lists were constructed such that words from the non-native language violated the phonotactic constraints in the native language of the infant. Infants preferred listening to words in their native language over foreign words.

Jusczyk, Luce, and Charles-Luce (1994) also demonstrated that infants are sensitive to variations in the phonotactic probabilities of phonetic patterns *within* their native language. Jusczyk et al. used the head turn preference procedure with American 9-month-olds. They found that the infants preferred to listen to lists of monosyllabic nonsense words containing high probability phonotactic sequences over those containing low probability sequences. By nine months of age, infants are sensitive to the phonotactic probabilities of segments and segmental transitions within their native language. This suggests that phonotactic information is an early, fundamental property of the representations of spoken words in memory.

Infants may exploit their sensitivity to phonotactic information in learning their native language. Messer (1967) has shown that phonotactic information is also represented and used by children later in life. Messer presented three-year-old children with monosyllabic nonsense word pairs and asked them which member of a pair sounded more like an English word. He found that nonsense words with permissible phonetic sequences were judged as more like English words than nonsense words composed of illegal or very infrequent sequences.

Early sensitivity to phonotactics of the native language may arise from incipient lexical representations of form, or may be indicative of an independent source of knowledge regarding the probabilities of segments and sequences of segments. The demonstration that children or infants respond to phonotactics does not, unfortunately, provide evidence for the locus of possible effects of phonotactics in adults.

Adults also seem to be sensitive to phonotactic information. For example, Eukel (1980) found that frequency of occurrence judgments made on nonsense words (which, by definition, have little or no frequency in the language) are based on phonotactic information. Participants were presented with lists of real words and nonsense words and asked to estimate the frequency of each item. For the words, the subjective frequency ratings were within one order of magnitude of their objective frequency counts. More interestingly, Eukel found significant agreement among participants' frequency judgments for the *nonsense* words. Judgments for the nonsense words were also found to correlate highly with Greenberg and Jenkins' (1964) metric for measuring similarity among phonetic patterns of spoken words. Briefly, this metric indexes the extent to which a given phonetic pattern is similar to other patterns, thus providing an indirect measure of the probabilities of phonetic segments and sequences in the language (i.e., phonotactic probabilities). The finding that subjective frequency judgments of nonsense words are correlated with Greenberg and Jenkins' metric of similarity suggests that phonotactic information is stored and accessible, in one form or another, in lexical memory for form-based representations.

A study by Brown and Hildum (1956) demonstrated that phonotactic constraints have demonstrable effects on spoken word perception. They asked participants to identify monosyllabic items embedded in noise. These items were either: (1) real English words, (2) phonotactically *legal* nonsense words, or (3) phonotactically *illegal* nonsense words. Brown and Hildum found that both naive and phonetically sophisticated participants were best at identifying, or transcribing, the real words. In addition, they found that both groups were better at identifying phonotactically legal nonsense words than phonotactically illegal nonsense words, despite explicit instructions to expect illegal sound combinations. This suggests that participants' identification of degraded stimuli was influenced by their knowledge of phonotactic constraints on phonetic patterns. Drawing a parallel between semantic and phonetic context, Brown and Hildum argued that just as listeners may expect to find a particular noun in a particular semantic context, listeners may also expect certain phonetic segments to occur in certain phonotactically constrained slots in spoken syllables and words.

Auer (1993) has provided additional evidence of phonotactic influences on word recognition. In a number of auditory lexical decision and shadowing experiments, he found that phonotactic probabilities directly affected processing times for spoken consonant-vowel-consonant words. In particular, he found that words with high probability phonetic patterns were processed more rapidly than those with low probability patterns. Auer interpreted his results within the context of PARSYN, a connectionist model of spoken word recognition that explicitly encodes information regarding segmental probabilities and segment-to-segment transitions. His model successfully simulated processing time differences observed for spoken words varying on the probabilities of their phonotactic patterns. The results from studies involving infants, children, and adults suggest that phonotactic information may not be used just as a "bootstrap" into the language system, but may provide a continuous store of information throughout development.

To examine the representation of phonotactic information in the lexicon further, we asked participants to rate how "English-like" nonwords sounded. The bisyllabic nonwords were composed of initial and final syllables with high and low probability phonotactic patterns.

If participants have access to phonotactic information stored in memory, we would expect to find evidence of this knowledge reflected in the subjective "goodness" ratings of the nonsense words. That is, if phonotactic information resides in lexical memory — either by explicit rules, conspiracies of exemplars, or both — we expect subjective judgments to coincide with our objective measures of phonotactic information: Nonsense words with highly probable phonotactic patterns should be judged as more word-like than nonsense words with less probable patterns.

This work extends Brown and Hildum (1956), Eukel (1980), and other related work in two important ways. First, all the nonsense words in the present study are phonotactically legal items, unlike Eukel's stimuli, which were composed, in part, of illegal patterns in English. Thus, this study examined participants' sensitivity to *legal* patterns that varied in their segmental and sequential *probabilities*, rather than participants' sensitivity to the legality of segmental and sequential patterns.

Second, all the stimuli in the present study were bisyllabic items composed of consonant-vowel-consonant syllables in which either the first or second syllables received primary stress. Previous studies of phonotactics have used only monosyllabic stimuli. The manipulation of metrical stress enables us to determine if effects of phonotactic configuration vary with syllable stress. (For discussions of the role of stress in spoken word recognition and acquisition, see Cutler & Norris, 1988; and Jusczyk, Cutler, & Redanz, 1993.) If participants are indeed differentially sensitive to phonotactic sequencing as a function of syllable stress, one might expect to find these two factors reflected in the participants' ratings. In particular, phonotactics and stress may produce synergistic effects on subjective judgments, such that syllable stress may exaggerate and/or de-emphasize the importance of phonotactic configuration. For example, stressed syllables with highly probable phonotactic patterns may be judged to be exceptionally "good" patterns in English, whereas syllables with secondary stress and low probability patterns may be judged to be exceptionally "bad" patterns. On the other hand, phonotactic configuration may be independent of stress, suggesting that syllable stress does not mediate effects of probabilistic phonotactics.

EXPERIMENT 1

Method

Participants. Forty participants from the University at Buffalo community were paid for their participation in this experiment. All participants were native English speakers and reported no history of a speech or hearing disorder at the time of testing.

Materials. The two-hundred and forty nonsense syllables of varying phonotactic probability used in Jusczyk, Luce, and Charles-Luce (1994) were used in this experiment. These syllables were combined to form 120 bisyllabic nonsense words. No syllable was used more than once. The same two measures that were used to determine phonotactic probability in Jusczyk et al. (1994) were also used to define phonotactic probability in this experiment: (1) positional segment frequency (i.e., how often a particular phonetic segment occurs in a position in a word), and (2) biphone frequency (i.e., the segment-to-segment co-occurrence probability). These metrics were computed using an on-line version of Webster's Pocket Dictionary. This dictionary contains approximately 20,000 computer-readable phonetic transcriptions that were used to compute log-frequency weighted values for positional segment frequency and biphone frequency (see Auer, 1993). Given that we used frequency-weighted values in our computations, the segment and biphone statistics can be viewed as being based on *token* counts.

Syllables that were considered high-probability patterns consisted of segments with high segment positional probabilities. For example, in the high probability pattern, /kik/ ("keek"), the consonant /k/ is relatively frequent in initial position, the vowel /i/ is relatively frequent in the medial position, and the consonant /k/ is relatively frequent in the final position. In addition, a high probability phonotactic pattern had frequent biphone probabilities, that is, CVC patterns with high probabilities of initial consonant-vowel and vowel-final consonant co-occurrences (e.g., /b/ followed by /æ/ and /æ/ followed by /p/ in the nonsense word , /bæp/)

Syllables that were classified as low-probability patterns consisted of segments with low segment positional probabilities and low biphone probabilities. For example, the low probability pattern /giθ/ ("geeth") has segments that are relatively rare in their respective positions and rarely co-occur. Despite being relatively rare, none of the patterns formed were phonotactically illegal in English. Indeed, all segment positions and transitions in the stimuli occur in real English words. In addition, each of the five vowels used in the CVCs, /ʌ, aɪ, i, e, ɝ/, occurred in equal proportions in each of the syllable types.

The average segment probability was .1926 for the high-probability pattern list and .0543 for the low probability pattern list. The average biphone probability was .0143 for the high-probability list and .0006 for the low-probability list. The difference in the magnitudes of the segment and biphone probabilities reflects the fact that there are many more biphones than segments. Thus, biphones have a lower probability of occurrence overall than segments because the same total probability (i.e., 1.00) is divided among many more possible outcomes for the biphones than for the segments.

Four lists of 120 bisyllabic nonsense words were then created by systematically combining the original 240 syllables. (The full list can be found in the Appendix.) All resulting stimuli contained the same vowel in the first and second syllables. The 120 nonsense words appeared only once in each list. There were 15 stimuli in each of eight experimental conditions (2 stress placements × 4 phonotactic probability patterns), resulting in 120 stimuli per list. Examples of the stimuli are presented in Table 1.

Creation of four lists was necessitated by the orthogonal combination of syllable stress and syllable order: Lists 1 and 2 were identical except for the placement of the stress. On one list the primary stress fell on the first syllable of the stimulus, while its phonetic match on the other list had the primary stress on the second syllable. Lists 3 and 4 differed from each other in the same manner. The four lists also differed in terms of syllable order. Lists 1 and 2 had nonsense words with the syllables in one order. Lists 3 and 4 used the same syllables to form nonsense words, but reversed the order of the syllables from their matches on lists 1 and 2.

The 480 stimuli were recorded by a trained phonetician (JC-L). All nonsense word

TABLE 1

Examples of stimuli across the four lists

	List 1	*List 2*	*List 3*	*List 4*
CONDITION				
High-High	fʌlˈʧʌn	ˈfʌlʧʌn	ʧʌnˈfʌl	ˈʧʌnfʌl
High-Low	ˈlʌnðʌz	lʌnˈðʌz	saɪkˈgaɪb	ˈsaɪkgaɪb
Low-High	gaɪbˈsaɪk	ˈgaɪbsaɪk	ˈðʌzlʌn	ðʌzˈlʌn
Low-Low	ˈðaɪbʤaɪz	ðaɪbˈʤaɪz	ˈʤaɪzðaɪb	ʤaɪzˈðaɪb

stimuli were spoken in isolation. The stimuli were low-pass filtered at 4.8 kHz and digitized at a sampling rate of 10 kHz using a 12-bit analog-to-digital converter. All nonsense words were edited into individual files and stored on computer disk. Correct stress placement by the speaker was confirmed by measuring the amplitude of the vowel of each syllable using a digital wave-form editor.

Design: Two variables were examined: (1) syllable stress (initial primary stress vs. final primary stress) and (2) phonotactic probability (High-High vs. High-Low vs. Low-High vs. Low-Low).

Procedure: Participants were tested individually or in pairs. Each participant was seated in a booth equipped with a Microterm 5510 computer terminal and a pair of Telephonics headphones. The presentation of stimuli was controlled by a PDP 11/34 computer. All stimuli were presented in random order.

A typical trial proceeded as follows: A scale from 1, labeled "GOOD ENGLISH WORD," to 10, labeled "BAD ENGLISH WORD," appeared on the computer monitor. A prompt ("READY") then appeared on the monitor. Participants were presented auditorily with one of the stimulus items at a comfortable listening level. Participants were instructed to press one of the keys labeled 1 through 10 on the keyboard as quickly as possible. After recording the response, the computer began another trial. Participants were allowed a maximum of three seconds to respond before the computer automatically recorded a null response and presented the next trial. All responses were recorded by the PDP 11/34.

Each participant received one of the four lists of 120 randomly ordered stimuli. Each list was presented to 10 different participants. Prior to the experimental trials each participant received 10 practice trials. These trials were used to familiarize the participants with the task and were not included in the final data analysis.

Results

The mean ratings for each condition are shown in Figure 1. Ratings on a scale of 1 ("GOOD") to 10 ("BAD") are plotted on the *y* axis. Syllable phonotactic probability is represented on the *x* axis. "High-High" refers to nonsense words with high probability initial and final syllables, "High-Low" refers to nonsense words with high probability initial syllables and

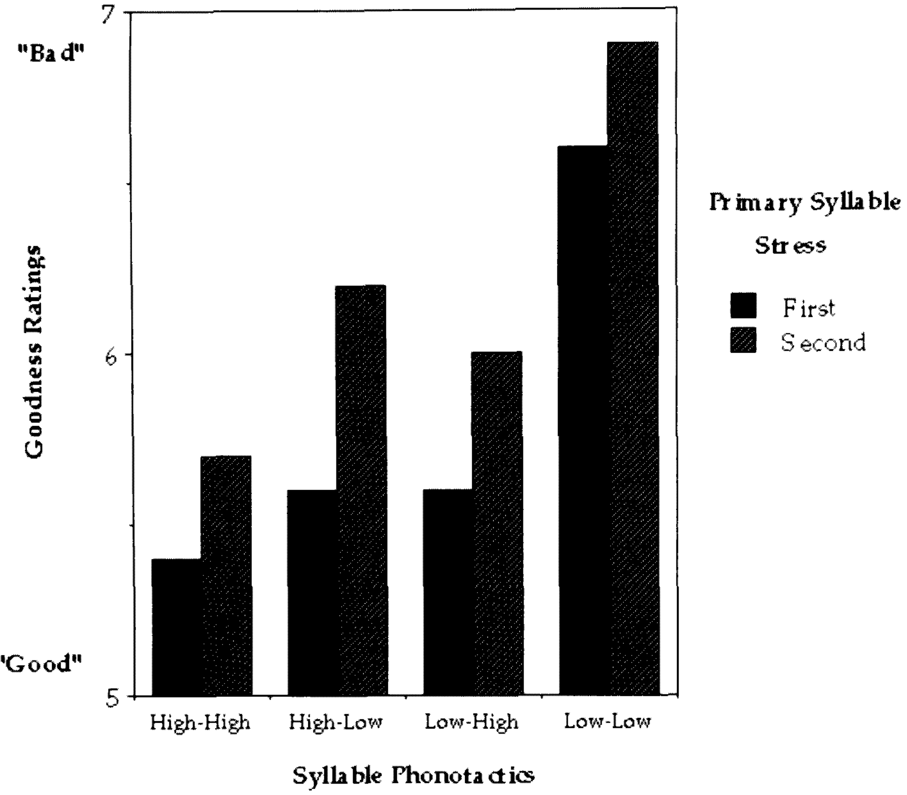

Figure 1

Averaged goodness-as-an-English-word ratings as a function of syllable phonotactics and stress placement. On the horizontal axis, first and second syllable probabilities are labeled "High" or "Low." Each bar represents 40 judgments.

low probability final syllables, "Low-High" refers to nonsense words with low probability initial syllables and high probability final syllables, and "Low-Low" refers to nonsense words with low probability initial and final syllables. The results are shown for stimuli with primary stress on the first syllable (solid bars) and for stimuli with primary stress on the second syllable (striped bars).

A 2 × 4 (syllable stress × phonotactic probability) within-subjects analysis of variance was performed on the mean ratings for each condition for each of the 40 participants. A main effect of syllable stress was found. Nonsense words with the primary stress on the first syllable were judged more English-like than nonsense words with the primary stress on the second syllable, $F_1(1, 39) = 10.24, p < .05$, and $F_2(1, 472) = 20.68, p < .0001$.

In addition, a main effect of phonotactic probability was obtained, $F_1(1, 39) = 31.26$, $p < .0001$, and $F_2(3, 472) = 30.81, p < .0001$. Stimuli containing two high probability syllables were rated most English-like. Stimuli containing two low probability syllables were rated least English-like.

The interaction between stress and phonotactics was not significant ($F < 1$, by subjects and items). The failure to find an interaction between syllable stress and phonotactics may have been due, in part, to the possibility that the vowels we used, particularly the tense vowels /aɪ, e, ɝ, i/, may be more typical of stressed than of unstressed syllables. In order to assess this possibility, we made separate calculations of log-frequency-weighted probabilities of each vowel occurring in syllables with primary stress and in syllables with secondary stress. We then rank-ordered the probabilities of the vowels for each stress type. The average rank of the vowels employed in the present study was 8.8 for stressed syllables (out of 18 vowels) and 9.0 for unstressed syllables. Thus, it does not appear that the vowels in the syllables with secondary stress were unusual or atypical compared to the vowels in syllables with primary stress. Nevertheless, although the present stimuli may have somewhat uncommon stress patterns because the vowel in the syllable receiving secondary stress is not reduced, the contrast under scrutiny (primary stress on the first vs. second syllable) should still be representative of the more extreme weak-strong contrast in English.

Finally, to further examine the data, we performed planned contrasts comparing the four phonotactic conditions (i.e., High-High, High-Low, Low-High, and Low-Low). Because stress did not interact with syllable phonotactics, we collapsed across stress placement (i.e., the four separate lists) for the purpose of these analyses.

High-High stimuli were judged significantly more English-like than High-Low stimuli, $F(1, 39) = 6.733, p < .02$. This difference is shown in Figure 1 by the difference between the two left-most sets of bars. High-Low stimuli were not judged significantly different than Low-High stimuli ($F < 1$), as shown by the middle two sets of bars in Figure 1. Finally, Low-High stimuli were judged significantly more English-like than Low-Low stimuli, $F(1, 39) = 50.89$, $p < .0001$. This difference is shown in Figure 1 by the difference between the two right-most sets of bars.

Discussion

The results of the rating experiment confirm that listeners have reliable intuitions about phonotactic probabilities in their language. When asked to judge whether a given bisyllabic nonsense word constitutes a "good" or "bad" English word, participants consistently responded in accordance with objective measures of phonotactic probability. In particular, nonsense words that were constructed to have highly frequent segments and segmental transitions were judged more "English-like" than nonwords with low probability phonotactic patterns. In addition, participants judged stimuli with stress on the first syllable to be more English-like than stimuli with second syllable stress.

It would be of interest to determine to what extent the obtained results of phonotactics were due to the separate contributions of the segment and biphone probabilities. Unfortunately, the present stimuli do not allow us to address this question. Indeed, the independent effects of segment and biphone probabilities may be quite difficult to assess because of the high correlation between the two. In our data, segment and biphone probabilities had a correlation of .86. Despite this strong correlation, we conducted a hierarchical multiple regression analysis on our rating data with segment and biphone probabilities as separate independent variables. There was no measurable independent contribution of either variable. Of course, our stimuli were not designed to test the two probabilities separately, so specifically designed future studies may be able to distinguish the effects of segments and biphones. However, the high correlation between the two probabilities we observed makes us appreciate the difficulty in selecting sufficient numbers of stimuli that vary orthogonally on these two variables.

One factor contributing to the rating results for the phonotactics may have been the co-occurrence probability of the segments *spanning* the syllable boundaries. Although we explicitly manipulated the within-syllable co-occurrence probabilities, we did not purposefully control for, or vary the probabilities of the segments across syllables. Therefore, we computed the biphone probabilities for the cross-syllable sequences for each condition. Indeed (and not surprisingly), the sequences in the High-High condition had the highest probability and the sequences in the Low-Low had the lowest probability. Analysis of variance revealed, however, that only the High-High sequences were significantly different from any of the other conditions. Thus, while the cross-syllable sequences may be contributing to the present results (a finding that is in no way inconsistent with our conclusions), cross-syllable probabilities are not sufficient for explaining the complete pattern of results observed, in particular, the finding that the Low-Low condition produced significantly lower ratings than the other three conditions.

Another factor that may contribute to the results are higher-order phonotactic probabilities, or the co-occurrence probabilities of nonadjacent segments. Our interest was primarily in the role of first-order phonotactic probabilities in spoken word processing. For work relating to higher order phonotactic probabilities, see Frisch, Broe, and Pierrehumbert (1995).

Although our results demonstrate that participants have access to fairly precise information in memory regarding probabilities of phonotactic configurations and stress patterns, we have yet to understand fully the implications of this information for on-line processing of spoken words. To date, there is little or no evidence demonstrating that phonotactic probabilities have predictable effects on on-line processing (as opposed to identification of degraded stimuli or goodness ratings) or processing time (see, however, Auer, 1993). Whereas listeners may have *access* to phonotactic information—as evidenced by our rating data—phonotactic information may have no important or demonstrable influence on processing. The aim of Experiment 2 was to attempt to determine if phonotactic probabilities and stress placement affect processing time and accuracy in a speeded auditory repetition task.

EXPERIMENT 2

In Experiment 2, we attempted to determine if measures of processing time coincide with subjective ratings of phonological goodness. In particular, we presented the same stimuli used in Experiment 1 to participants in an auditory repetition task (see Levelt & Wheeldon, 1994). Because the auditory repetition task has both perception and production components, we use the generic term "processing" throughout. We were interested in determining if participants' processing times would mirror the previously obtained subjective ratings, such that nonsense words judged as "good" would be processed more quickly.

We were furthermore interested in determining if effects of stress patterning would be evident in processing times. On the basis of our own previous work and that of Cutler and her colleagues (see Cutler, 1990; Cutler & Butterfield, 1992; Cutler & Carter, 1987; Cutler & Norris, 1988), we predicted that words containing initial syllables with primary stress would be responded to most quickly. Finally, we were interested in determining if phonotactics and stress might have interactive effects on processing times, despite their apparent independence in the data for the ratings. In short, we attempted to determine if the representation of phonotactics and stress patterning in memory demonstrated in previous work has implications for the processing of spoken language, such that highly probable sound patterns of spoken stimuli are processed more rapidly than stimuli with less probable patterns.

Method

Participants. Forty native English speakers from the University at Buffalo participated in partial fulfillment of course requirements. None reported a history of a speech or hearing disorder at the time of testing. None had participated in Experiment 1.

Materials. The same stimuli used in Experiment 1 were used in the present study.

Design. Two variables were examined: (1) phonotactic probability and (2) syllable stress.

Procedure. Participants were tested individually. Each participant was seated in a booth equipped with a Microterm 5510 computer terminal and a pair of Telephonics TDH-39 headphones equipped with a boom microphone that was positioned immediately in front of the participant's lips. The microphone was connected to a voice-key interfaced to a PDP 11/34 computer. The voice-key registered a response as soon as the participant began speaking. Presentation of stimuli and response collection was controlled by the computer.

A typical trial proceeded as follows: A prompt ("READY") appeared on the CRT. Participants were presented with one of the spoken stimulus items at a comfortable listening level. Participants then repeated the nonsense word as quickly and as accurately as possible into the microphone. Reaction times were measured by the computer from the onset of the stimulus to the onset of the participant's verbal response. After registering a response, the computer began another trial. Participants were allowed a maximum of three seconds to respond before the computer automatically recorded a null response and presented the next trial.

All responses were recorded on audio tape for accuracy analysis. Accuracy was assessed by listening to the participants' responses and comparing them to a written transcription of the stimuli. A response was scored as correct if there was an identical match on all segments of the stimulus.

Each participant received one of four randomly ordered lists of 120 stimuli. Prior to the experimental trials, each participant received ten practice trials. These trials were used to familiarize the participants with the task and were not included in the final data analysis.

Results. The mean reaction time for each condition is shown in Figure 2. Reaction time in milliseconds is plotted on the y axis. Syllable phonotactic probability is represented on the x axis. "High-High" refers to nonsense words with high probability initial and final syllables, "High-Low" refers to nonsense words with high probability initial syllables and low probability final syllables, and so on. The results are shown for stimuli with primary stress on the first syllable (solid bars) and for stimuli with primary stress on the second syllable (striped bars).

Two × four (syllable stress X phonotactic probability) within-subjects analyses were performed on latencies and accuracy. For the latencies, a significant main effect of stress pattern was observed, $F_1(1, 39) = 35.99, p < .0001$, and $F_2(1, 472) = 10.65, p < .002$. Stimuli with primary stress on the first syllable were responded to significantly more quickly than

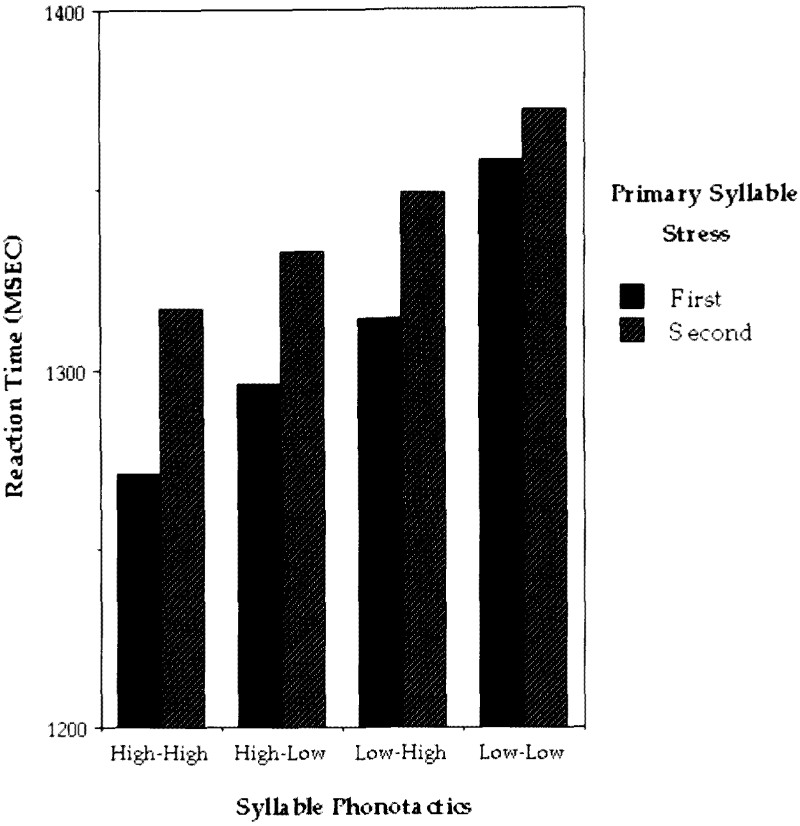

Figure 2

Averaged reaction times to repeat auditorily presented nonsense words as a function of syllable phonotactics and stress placement. On the horizontal axis, first and second syllable probabilities are labeled "High" or "Low."

stimuli with primary stress on the second syllable. A significant effect of phonotactic probability was also obtained, $F_1(3, 117) = 32.63, p < .0001$, and $F_2(3, 472) = 9.031, p < .0001$. Overall, highly probable patterns were responded to more quickly than less probable patterns. The interaction of stress pattern and phonotactic probability was significant in the subject analysis, $F_1(1, 39) = 3.16, p < .03$, but not in the item analysis, $F_2\ (3, 472) < 1$.

Planned contrasts comparing the four phonotactic conditions were also performed. (Lacking significant interaction between stress pattern and phonotactic probability in item analyses, we collapsed across stress for these analyses.) Participants had shorter latencies for High-High stimuli than for High-Low stimuli, $F(1, 39) = 7.50, p < .008$, or Low-High stimuli, $F(1, 39) = 25.40, p < .0001$. High-Low stimuli were also responded to significantly more quickly than Low-High stimuli, $F(1, 39) = 5.30, p < .03$. Both High-Low stimuli, $F(1, 39) = 46.29, p < .0001$, and Low-High stimuli, $F(1, 39) = 20.27, p < .0001$, were responded to faster than Low-Low stimuli. Finally, High-High stimuli were responded to more quickly than Low-Low stimuli, $F(1, 39) = 91.05, p < .0001$. In general, nonsense words with two high probability syllables were responded to most quickly; those with two low probability syllables were responded to most slowly. In addition, High-Low stimuli were responded to faster than Low-High stimuli, suggesting that the phonotactics of initial syllables may play

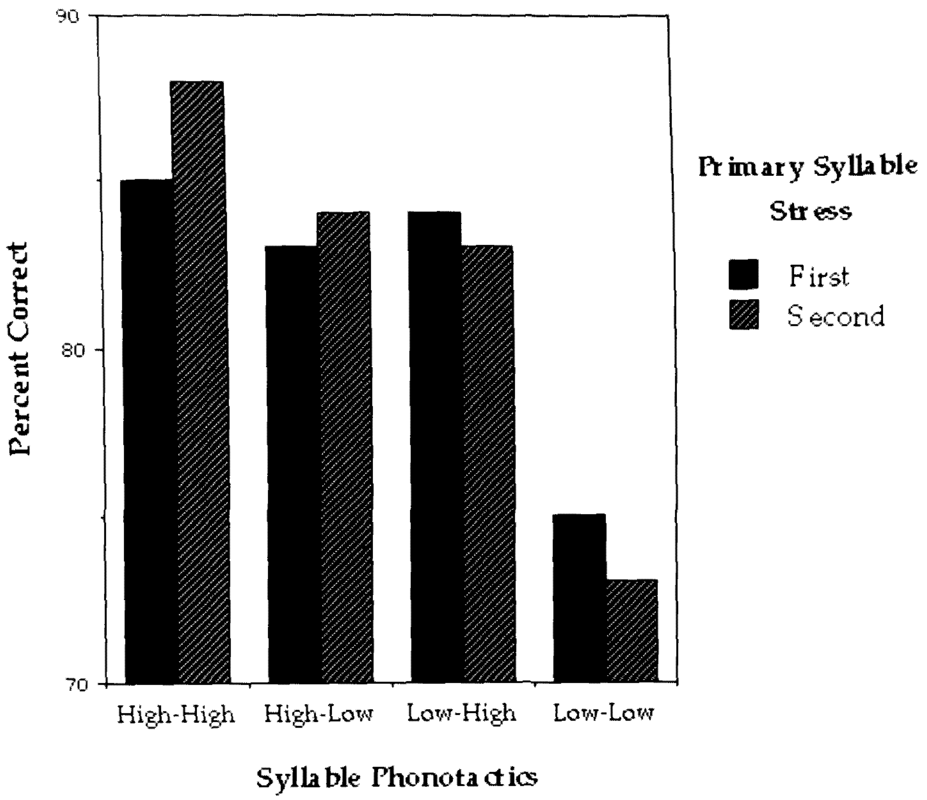

Figure 3

Averaged accuracy rates in naming auditorily presented nonsense words as a function of syllable phonotactics and stress placement. On the horizontal axis, first and second syllable probabilities are labeled "High" or "Low."

a stronger role in processing than the phonotactics of final syllables.[1]

The mean percentages correct for each condition are shown in Figure 3. We obtained a main effect of phonotactic probability only, F_1 (3, 117) = 17.80, $p < .0001$, and F_2 (3, 472) = 10.61,

---

[1] Latencies to repeat a word may reflect characteristics of the stimulus itself, producing contaminated measures of processing time. In particular, the voice key may be differentially sensitive to the phonetic characteristics of the initial segment of the word to be articulated. One means of evaluating these potential confounds is to present a delayed cue for the repetition response well after the stimulus word has been presented (see Balota & Chumbley, 1985). Any effects remaining in the delayed repetition condition can thus be attributed not to processing but instead to stimulus characteristics, such as differences in amplitude rise times of the initial segments of the stimuli. On the other hand, the failure to obtain effects in the delayed condition rules out the possibility that characteristics of the initial segments alone were responsible for response times in the immediate repetition condition.

We re-ran the repetition experiment with a delayed response cue of 1200 msec following offset of the stimulus. The delayed repetition experiment was identical in all other respects to the immediate repetition experiment. We obtained no significant effect of phonotactic condition on response times, $F(3, 117) = 1.129, p = .34$. We are therefore confident that the reaction times in the immediate repetition experiment are not attributable to differential sensitivity of the voice key to the initial segments of the stimulus items.

$p < .0001$. The effect of stress pattern and the interaction of stress pattern and probability were not significant, $F_1(1, 39) < 1$ and $F_2(3, 117) = 1.37$, $p > .2$, respectively. Planned contrasts comparing phonotactic probabilities revealed significant differences for the accuracy scores in three of the five comparisons: High-Low versus Low-Low, $F(1, 39) = 25.548$, $p < .0001$, Low-High versus Low-Low, $F(1, 39) = 25.08$, $p < .0001$, and High-High versus Low-Low, $F(1, 39) = 48.63$, $p < .0001$. In general, for the accuracy scores, the observed significant effects appear to have all been the result of participants' lower accuracy at shadowing the Low-Low stimuli. The important result for the present study, however, is the failure to observe any possible speed-accuracy tradeoffs in the data: When reaction times slowed, accuracy either remained statistically stable or dropped.

Discussion

The results of Experiment 2 join a growing body of evidence demonstrating the role of phonotactic information in the representation and processing of spoken words. In the first experiment, we demonstrated that participants have reliable intuitions about phonotactic probabilities in their language. When asked to judge whether a given bisyllabic nonsense word constituted a "good" or "bad" English word, participants in Experiment 1 consistently responded in accordance with objective measures of phonotactic probability: Nonsense words with highly frequent segments and sequences of segments were judged as "better" sounding English than nonwords with less frequent segments and sequences. Experiment 2 demonstrated that phonotactic probabilities affect not only subjective ratings but *reaction times* as well. Our results show that bisyllabic nonsense word stimuli composed of two high probability phonotactic patterns (High-High) were responded to most quickly and those with two low probability patterns (Low-Low) least quickly. Interestingly, we also found that nonsense words with high probability initial syllables and low probability final syllables (High-Low) were responded to more quickly than nonsense words with the reverse order of syllables (Low-High). This finding suggests that phonotactic probability may play a more important role earlier in the processing of spoken words.

In addition to our findings regarding phonotactic probability, we found that nonsense words with primary stress on the initial syllable were processed more quickly than those with primary stress on the final syllable. These findings are consistent with theories of spoken word recognition that ascribe priority to strong (i.e., stressed) syllables (Cutler, 1990; Cutler & Carter, 1987; Cutler & Norris, 1988). We also found that phonotactic probability and stress placement failed to interact (in either Experiments 1 or 2), suggesting that segmental and sequential probabilities have equivalent effects across stress patterns.

GENERAL DISCUSSION

In conjunction with previous research, these results provide evidence that participants have access to information in memory regarding phonotactic probabilities. Participants had consistent intuitions about phonotactic probabilities and stress patterns. These variables had demonstrable effects on processing times, as measured by the auditory repetition task. Participants were sensitive to fine-grained differences in probabilities of acceptable sequences and not simply grosser differences between legal and illegal patterns. Our results join with previous research on adults and children in implicating a significant role for probabilistic phonotactic information in the memory representations and perceptual processes involved in spoken word recognition and production.

The task before us now is to provide mechanistic accounts for effects of phonotactics on recognition. Earlier models of spoken word recognition (Marslen-Wilson, 1990; Morton, 1969), while attempting to account for other, fundamental findings in the literature, had little to say about the role of phonotactic configuration. More recently, however, recognition models in the connectionist tradition have provided potential mechanisms for accounting for effects of segmental and sequential regularity on processing time. TRACE (McClelland & Elman, 1986), for example, may be able to account for effects of phonotactics via conspiracies among lexical items sharing similar phonetic compositions. In TRACE, therefore, effects of phonotactics may be entirely top-down. In Norris' (1994) Shortlist model, however, phonotactic constraints are encoded at the phoneme level and do not emanate from lexical entries themselves. As Norris points out, the treatment of phonotactic effects in Shortlist harks back to earlier notions that phonotactic information constitutes independently represented knowledge and is not simply the result of overlapping phonetic patterns associated with form-based representations of lexical items. Nonetheless, it remains unclear whether the information that gave rise to the results in the present study is derived from form-based representations of spoken words (as happens in TRACE) or is instead abstract knowledge of the probabilistic phonotactic constraints of English (as in Shortlist). Both alternatives are also simultaneously possible. Abstract knowledge of phonotactic constraints may be an epiphenomenon of generalizations across form-based lexical representations.

Although the current research cannot distinguish among these interesting alternatives, we believe our results underscore the fact that any model purporting to explain spoken word recognition must account for probabilistic phonotactic effects. It seems certain that information regarding the probability—and not simply the legality or illegality—of a given phonetic sequence is, in one form or another, represented in memory. Lexical representations of spoken words appear to have richly (albeit probabilistically) constrained phonetic structures that can be revealed by participants' reliable and systematic judgments and on-line processing of stimuli they have never before encountered.

A finding of further interest in the present study was that participants consistently judged nonsense words with the primary stress on the first syllable as more English-like than words with the primary stress on the second syllable. This finding again reflects participants' sensitivity to the probabilities of form-based representations in memory. As previously mentioned, Cutler and Carter (1987) demonstrated that most English words have the primary stress on the first syllable. In addition to probabilistic phonotactic information, participants clearly have access to some form of information representing the likelihood of particular stress patterns, given that they consistently and reliably judge bisyllabic nonwords with the primary stress on the first syllable as constituting "better" English words.

We examined phonotactic probabilities and stress placement in tandem in an attempt to determine if these two types of phonetic information interact in participants' judgments of phonological goodness. In particular, we were interested in determining if phonotactic probabilities might play a more important role for words with the primary stress on the first syllable, than words with the primary stress on the second syllable. This hypothesis was not confirmed: There was no interaction between phonotactic probability and stress, suggesting that participants treated these two sources of information separately in making their judgments.

In summary, we believe the present research on effects of phonotactic probabilities may have important implications for the role of phonotactic information in memory, its consequences for on-line perceptual processing, and current and future modeling efforts in spoken word recognition.

Received: August 22, 1995; revised manuscript received: January 22, 1997, accepted: February 4, 1997

REFERENCES

AUER, E. T. (1993). *Dynamic processing in spoken word recognition: The influence of paradigmatic and syntagmatic states.* Unpublished doctoral dissertation, University at Buffalo, Buffalo, NY.

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 and Language*, **24**, 89 – 106.

BROWN, R. W., & HILDUM, D. C. (1956). Expectancy and the perception of syllables. *Language*, **32**, 411 – 419.

CHURCH, K. (1987). Phonological parsing and lexical retrieval. *Cognition*, **25**, 53 – 69.

CLEMENTS, G. N., & KEYSER, S. J. (1983). *CV Phonology: A generative theory of the syllable.* Cambridge, MA: MIT Press.

CUTLER, A. (1990). Exploiting prosodic probabilities in speech segmentation. In G.T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp.105 – 121). Cambridge: MIT Press.

CUTLER, A., & BUTTERFIELD, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*, **31**, 218 – 236.

CUTLER, A., & CARTER, D. M. (1987). The predominance of strong initial syllables in English vocabulary. *Computer Speech and Language*, **2**, 133 – 142.

CUTLER, A., & NORRIS, D. G. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, **14**, 113 – 121.

EUKEL, B. (1980). Phonotactic basis for word frequency effects: Implications for lexical distance metrics (Abstract from the Proceedings of the 100th Meeting of the Acoustical Society of America, Los Angeles, CA). *Journal of the Acoustical Society of America*, **68** (Suppl. 1), S33.

FRISCH, S., BROE, M., & PIERREHUMBERT, J. (1995). *The role of similarity in phonotactic constraints.* Unpublished Manuscript. Northwestern University. Evanston, IL.

GREENBERG, J. H., & JENKINS, J. J. (1964). Studies in the psychological correlates of the sound system of American English. *Word*, **20**, 157 – 177.

JUSCZYK, P. W., CUTLER, A., & REDANZ, N. (1993). Preference for the predominant stress patterns of English words. *Child Development*, **64**, 675 – 687.

JUSCZYK, P. W., FREDERICI, A. D., WESSELS, J., SVENKERUD, V. Y., & JUSCZYK, A. M. (1993). Infants' sensitivity to the sound pattern of native language words. *Journal of Memory and Language*, **32**, 402 – 420.

JUSCZYK, P. W., LUCE, P. A., & CHARLES-LUCE, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, **33**, 630 – 645.

LEVELT, W. J. M., & WHEELDON, L. (1994). Do speakers have access to a mental syllabary? *Cognition*, **50**, 239 – 269.

LUCE, P. A., PISONI, D. B., & GOLDINGER, S. D. (1990). Similarity neighborhoods of spoken words. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp.105 – 121). Cambridge: MIT Press.

MARSLEN-WILSON, W. D. (1990). Activation, competition, and frequency in lexical access. In G.T.M. Altman (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 148 – 172). Cambridge, MA: MIT Press.

McCLELLAND, J., & ELMAN, J. (1986). The TRACE model of speech perception. *Cognitive Psychology*, **18**, 1–86.

MESSER, S. (1967). Implicit phonology in children. *Journal of Verbal Learning and Verbal Behavior*, **6**, 609–613.

MORTON, J. (1969). Interaction of information in word recognition. *Psychological Review*, **76**, 165–178.

NORRIS, D. (1994). SHORTLIST: A connectionist model of continuous speech recognition. *Cognition*, **52**, 189–234.

TURK, A. E., JUSCZYK, P. W., & GERKEN, L. (1995). Do English-learning infants use syllable weight to determine stress? *Language and Speech*, **38**, 143–158.

APPENDIX

Stimulus materials for Experiments 1 and 2.

High probability syllables

fʌl ʧʌn mʌb sʌʃ tʌl sʌʤ hʌs ʤʌn dʌs sʌz sʌg kʌk sʌv ɹʌl sʌd lʌn pʌm bʌl pʌl sʌt mʌn sʌs sʌl kʌn taɪs daɪp vaɪɹ vaɪk baɪs faɪk ɹaɪɹ maɪd haɪs saɪb vaɪt ʤaɪn saɪv ʧaɪn saɪp saɪm gaɪn paɪt saɪs daɪt saɪk saɪl baɪn haɪn kis ʧin kik ɹig sig θin fik kit pim fis vin ɹiz bis siv dik nin hin bil dis dit fin ɹit ɹis ɹin vet ɹeb meb keb seb mep ges wes hes sep peb ɹem nes tes pep lel hen pem ked sed nen ten pek ses dɝs mɝn sɝz fɝt tɝt sɝg pɝv vɝn pɝb mɝs kɝm sɝp pɝd fɝs bɝs kɝn sɝd sɝl sɝm sɝk pɝn sɝt sɝn sɝs

Low probability syllables

ðʌʃ ðʌʤ jʌʃ ðʌʧ θʌʃ jʌʤ ʧʌʃ θʌʤ jʌʧ ʧʌʤ θʌʧ ʃʌʤ wʌʃ ʧʌʧ ʃʌʧ ðʌz wʌʤ ðʌg ðʌv jʌz wʌʧ ðʌd θʌz ʧʌz ðaɪð ʃaɪð ʤaɪð ʧaɪð gaɪð ðaɪz ðaɪb ʤaɪz ðaɪv ʃaɪb ʧaɪz waɪð ðaɪm naɪð kaɪð ðaɪp ʃaɪv ʧaɪb ʃaɪm ʃaɪp gaɪb ʤaɪm ʤaɪp faɪð ðiʃ ðið jiʃ giʃ ziʃ jið zið gið ðiθ jiθ ziθ giθ ðiʧ jiʤ ziʤ jiʧ ziʧ ʧiʃ ʧið giʤ ðig giʧ jig zig θeʒ θeð θeθ θeg ʤeʒ ʤeð ʧeʒ ʧeð ʤeθ ʃeʒ θeʤ ʃeð ʧeθ ʃeθ ʤeg ʧeg ʃeg ʤeʤ θez ʧeʤ veʒ ʃeʤ veð geʒ jɝz jɝθ ʃɝθ jɝg ʃɝz ʧɝθ ʧɝz ʃɝg jɝʧ ʧɝg nɝθ nɝz ʃɝʧ lɝθ ʃɝʤ θɝθ lɝz jɝp ʧɝʤ nɝg gɝg ʤɝθ θɝz jɝv