

# Phonotactic and Metrical Influences on Adult Ratings of Spoken Nonsense Words

*Michael S. Vitevitch*

*Paul A. Luce*

Department of Psychology and Center for Cognitive Science

*Jan Charles-Luce*

Department of Communicative Disorders and Sciences

and

Center for Cognitive Science

*David Kemmerer*

Department of Linguistics and Center for Cognitive Science

State University of New York at Buffalo

Buffalo, NY 14260-4110

## ABSTRACT

This study examined the phonological intuitions of adults by having them rate the phonological "goodness" of nonsense words. Subjects were asked to use a ten point scale to rate how "English-like" each stimulus, which was presented auditorily, sounded. The stimuli were phonotactically legal (in English) bisyllabic, CVCCVC nonsense words that varied in their phonotactic probability and primary stress placement. Subjects rated highly probable phonotactic stimuli as more "English-like." In addition, stimuli with the primary stress on the first syllable were judged more "English-like" than stimuli with the primary stress on the second syllable. No interaction between phonotactic probability and stress was found. Our results show that subjects have consistent and reliable intuitions regarding phonotactic configurations and stress patterning, demonstrating that fairly detailed probabilistic segmental and suprasegmental information resides in memory for form-based lexical representations.

## Introduction

*Phonotactics* refers to the probability that a phonetic segment will be followed or preceded by a particular segment, as well as the probability that a given segment will occupy a particular position within a syllable or word.

A study by Brown and Hildum (1956) demonstrated that the phonotactic constraints of English have demonstrable effects on word perception. They found that identification of phonotactically legal monosyllables was significantly better than phonotactically illegal monosyllables, despite the fact that both of these sequences were nonsense words in English. These results suggest that subjects' identification of the degraded stimuli was influenced by their knowledge of phonotactic constraints on possible phonetic sequences in English.

Eukel (1980) investigated subjective frequency ratings of words and the possible mediating effects that phonotactic information may have on word frequency effects. Subjects in Eukel's study were presented with a list of real words and nonsense words and asked to estimate the frequency of each item. Subjects were able to estimate the frequency of words within one order of magnitude of their objective frequency counts. More interesting,

Eukel found that there was significant agreement among subjects' frequency judgments for the nonsense word items. These judgments were also found to correlate highly with Greenberg and Jenkins' (1964) metric for measuring similarity among phonetic patterns of spoken words. Basically, the Greenberg and Jenkins' metric indicates the degree to which a given phonetic pattern is similar to other patterns in the lexicon and thus constitutes an indirect measure of the probabilities of phonetic sequences. The finding that subjective frequency judgments of nonsense words is correlated with Greenberg and Jenkins' metric of similarity suggests that phonotactic information is stored, in one form or another, in lexical memory for form-based representations.

To further examine the representation of phonotactic information in the mental lexicon, we presented subjects with spoken nonsense words and asked them to rate how English-like they judged the items to be. All the items consisted of two nonsense syllables, each of which varied on the dimensions of phonotactic probability and primary stress placement. If subjects have access to phonotactic information stored in memory, we would expect to find evidence of this knowledge reflected in the subjective ratings. 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.

Previous studies of phonotactics have used only monosyllabic stimuli. In the current study, we used bisyllables to examine the possible additional influence of stress on subjects' judgments of phonological "goodness." Cutler and colleagues (Cutler, 1990; Cutler & Butterfield, 1992; Cutler & Carter, 1987; Cutler & Norris, 1988) have demonstrated that a high-proportion of words in English have primary stress on the first syllable. These researchers have argued that this regularity of stress patterning may enable speakers of English to adopt a strategy of lexical access based on stressed syllables, thus facilitating parsing of spoken words in the continuously varying speech stream. Several studies have also demonstrated that infants being raised in an English language environment prefer words with the primary stress on the first syllable (Jusczyk, Cutler, & Redanz, 1993a; Turk, Jusczyk, & Gerken, 1994). Perhaps stress

information, along with phonotactic information, is used to assist the processes of segmenting speech and recognizing words.

If subjects are indeed sensitive to phonotactic sequencing and stress patterns, one might expect to find an interaction of these two factors reflected in the subjects' ratings. In particular, phonotactics and stress may produce synergistic effects on subjective judgments, such that stress and phonotactic configurations combine in a non-additive fashion to produce exceptionally "good" or "bad" nonsense words. On the other hand, phonotactics and stress may have entirely independent effects on judgments of goodness, suggesting that subjects may be able to interrogate the two sources of information in memory separately in making their judgments.

Finally, our study differs from previous research on adults' judgments of phonotactics on one important dimension: We employed only phonotactically legal phonetic patterns in English. The present study attempted to demonstrate that adults are sensitive to phonotactic probabilities of nonsense words, even when these nonsense words are composed of phonotactic configurations that actually occur in English.

## METHOD

**Subjects:** Forty subjects from the University at Buffalo community were paid \$4 for their participation in this experiment. All subjects were native English speakers and reported no history of a speech or hearing disorder at the time of testing.

**Stimuli:** Two-hundred and forty nonsense syllables of varying phonotactic probability were created. 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 cooccurrence 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).

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 segment-to-segment cooccurrence probabilities, that is, CVC patterns with high probabilities of initial consonant-to-vowel and vowel-to-final consonant cooccurrences (e.g., /b/ followed by /æ/ and /æ/ followed by /p/ in the nonsense word , /bæp/).

Syllables that were considered low-probability patterns consisted of segments with low segment positional probabilities and low segment-to-segment cooccurrence probabilities. For example, the

low probability pattern /giθ/ ("geeth") has segments that are relatively rare in their respective positions and rarely cooccur. 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.

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.

Four lists of 120 bisyllabic nonsense words were then created by systematically combining the original 240 syllables. On each list, the 120 nonsense words appeared only once. Examples of the stimuli are presented in Table 1. Lists 1 and 2 were identical except for the placement of the stress. On one list the stress fell on the first syllable of the stimulus, while its phonetic match on the other list had the stress on the second syllable. Lists 3 and 4 differed from each other in the same manner. For each of the lists, half of the stimuli were strong-weak and half weak-strong.

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.

Table 1  
Examples of Stimuli

LIST 1	LIST 2	LIST 3	LIST 4
`taIs jaIb	taIs`jaIb	`jaIb taIs	jaIb`taIs

A high probability syllable, /taIs/, "tyce," and a low probability syllable, /jaIb/, "shibe," would be arranged in the following manner across the four lists.

The 480 stimuli were recorded by a trained phonetician (JC-L). All nonsense word 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 levels of three variables were examined: (1) first syllable phonotactic probability (high vs. low), (2) second syllable probability (high vs. low), and (3) stress pattern (primary stress on the first vs. second syllable).

**Procedure:** Subjects were tested individually or in pairs. Each subject 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. Subjects were presented auditorily with one of the stimulus items at a comfortable listening level. Subjects 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. Subjects 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 subject received one of the four lists of 120 randomly ordered stimuli. Each list was presented to 10 different subjects. Prior to the experimental trials each subject received 10 practice trials. These trials were used to familiarize the subjects with the task and were not included in the final data analysis.

## RESULTS

The mean ratings for each condition are shown in Figure 1.

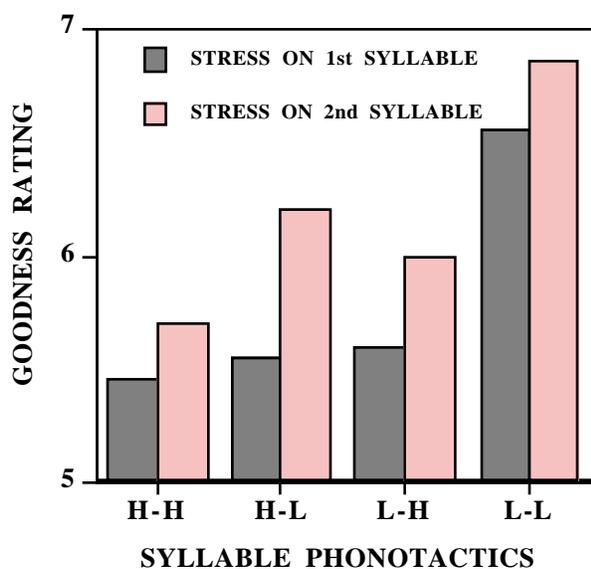


Figure 1. Goodness ratings as a function of syllable phonotactics and stress placement. Goodness rating is shown on the vertical axis. (The smaller the number, the more English-like the stimuli were judged.) Syllable phonotactics for the first and second syllables of the bisyllabic nonsense words are shown on the horizontal axis. High refers to high probability syllables; low refers to low probability syllables. Primary stress on the first syllable is represented by dotted bars; primary stress on the second syllable is represented by solid bars.

A 2 X 2 X 2 within-subjects analysis of variance was performed on the mean ratings for each condition for each of the 40 subjects. The variables and levels that were examined were: (1) first syllable phonotactic probability (high vs. low), (2) second syllable phonotactic probability (high vs. low), and (3) stress pattern (primary stress on the first vs. second syllable).

A main effect of stress pattern was found. Nonsense words with primary stress on the first syllable were judged more English-like than nonsense words having primary stress on the second syllable ( $F(1,39)=10.24, p<.01$ ).

In addition, a main effect of phonotactic probability was found for both the first syllable ( $F(1,39)=28.83, p=.0001$ ) and the second syllable ( $F(1,39)=38.97, p=.0001$ ). Stimuli containing high probability initial syllables were rated more English-like than stimuli containing low probability initial syllables; likewise, stimuli with high probability final syllables were rated as more English-like than stimuli with low probability final syllables.

To further interrogate that 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 for the purpose of these analyses.)

High-High stimuli were judged significantly more English-like than High-Low stimuli ( $F(1,39)=12.65, p=.001$ ). This difference is shown in Figure 1 by the difference between the two right-most sets of bars. High-Low stimuli were not judged significantly different than Low-High stimuli ( $F<1.0$ ), 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 ( $F(1,39)=95.68, p<.001$ ). This difference is shown in Figure 1 by the difference between the two left-most sets of bars.

## DISCUSSION

The results of the present experiment confirm that subjects 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, subjects 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 conjunction with previous research, these results provide strong evidence that subjects have access to information in memory regarding phonotactic probabilities. It remains unclear whether this information is derived from exemplars of form-based representations of spoken words or is instead abstract knowledge of the probabilistic phonotactic constraints of English. However, it is clear 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. Memory representations of spoken words appear to have richly (albeit probabilistically) constrained phonetic structures that can be revealed by subjects' reliable and systematic judgments of stimuli they have never before encountered.

A finding of further interest in the present study was that subjects consistently judged nonsense words with primary stress on the first syllable as more English-like than words with primary stress on the second syllable. This finding again reflects subjects' sensitivity to the probabilities of form-based representations in memory. As previously mentioned, Cutler and Carter (1987) demonstrated that most English words begin with strong syllables. In addition to probabilistic phonotactic information, subjects 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 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 subjects' judgments of phonological goodness. In particular, we were interested in determining if phonotactic probabilities might play a more important role for stressed syllables than for unstressed ones. This hypothesis was not confirmed: There was no interaction between phonotactic probability and stress, suggesting that subjects treated these two source of information separately in making their judgments.

Although our results demonstrate that subjects 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. Recent evidence (Auer, 1993) suggests that phonotactics may play a direct role in the speed and ease of processing of monosyllabic words. The present findings -- as well as those demonstrating infants' and young children's sensitivity to phonotactics and stress patterns (Jusczyk, Frederici, Wessels, Svenkerud, and Jusczyk 1993b; Jusczyk, Luce, and Charles-Luce, 1994; Messer, 1967) -- suggest that future research and modeling in spoken word recognition should carefully consider the role of segmental and suprasegmental probabilities in perceptual processing.

## 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.
- BROWN, R. W. and HILDUM, D. C. (1956) Expectancy and the perception of syllables. *Language*, **32**, 411-419.
- 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. and BUTTERFIELD, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*, **31**, 218-236.
- CUTLER, A. and CARTER, D.M. (1987). The predominance of strong initial syllables in English vocabulary. *Computer Speech and Language*, **2**, 133-142.
- CUTLER, A. and 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. *Journal of the Acoustical Society of America*, **68**, s33.
- GREENBERG and JENKINS (1964) Studies in the psychological correlates of the sound system of American English. *Word*, **20**, 157-177.
- JUSCZYK, P.W., CUTLER, A., and REDANZ, N. (1993a). 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., and JUSCZYK, A.M. (1993b). Infants' sensitivity to the sound pattern of native language words. *Journal of Memory and Language*, **32**, 402-420.
- JUSCZYK, P.W., LUCE, P.A. and CHARLES-LUCE, J. (1994) Infants sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, **33**, 630-645.
- MESSER, S. (1967) Implicit phonology in children. *Journal of Verbal Learning and Verbal Behavior*, **6**, 609-613.
- TURK, A.E., JUSCZYK, P.W., and GERKEN, L. (1994). *Do English-learning Infants Use Syllable Weight to Determine Stress?* Manuscript submitted for publication.

## Acknowledgments

This work is part of an ongoing project supported by NIH Grant No. DC-00879-01 to the State University of New York at Buffalo.