English speakers' sensitivity to phonotactic patterns

Rebecca Treiman

Wayne State University

Brett Kessler

Stanford University

Stephanie Knewasser

Wayne State University

Ruth Tincoff

Johns Hopkins University

Margo Bowman

Wayne State University*

Keywords: blending, English, phoneme (distributions), nonwords (word-likeness judgments), onsets, phonotactics, rimes, syllables, syllable structure

In M. B. Broe & J. B. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V: Acquisition and the lexicon* (pp. 269-282). Cambridge, England: Cambridge University Press. 2000.

* This research was supported by NSF Grant SBR 94-08456. We are grateful to the staff and children at the Walt Disney Elementary School for their cooperation, and to Kira Rodriguez for helpful comments on a draft of the manuscript.

X.1 Introduction

Linguists have often observed absolute restrictions in the patterning of phonemes in languages. Certain phonemes and sequences of phonemes commonly occur in a language whereas other phonemes and sequences of phonemes do not occur. For example, English does not allow /h/ or /uŋ/ at the ends of syllables. Across languages, it appears that absolute restrictions are more likely to involve VC (vowel-consonant) sequences than CV (consonant-vowel) sequences (e.g. Booij, 1983; Fudge, 1969, 1987; Selkirk, 1982). Such asymmetry has been taken to suggest that the vowel and final consonant form a subunit of the syllable, the *rime*. English, however, has relatively few absolute constraints. There has been a debate about whether there is enough imbalance in phonotactic constraints to support the postulation of a rime unit (Clements & Keyser, 1983; Fudge, 1987).

Kessler and Treiman (1997) recently broadened the scope of investigation to include *probabilistic* constraints, or situations in which a sequence of phonemes occurs less often than expected given the frequencies of the individual phonemes. When probabilistic constraints are considered, the case is greatly strengthened for a preponderance of constraints involving VC sequences as compared to those involving CV sequences. For instance, a number of English VCs occur significantly less often than expected based on the frequencies of their component phonemes. Sequences such as /æl/ and /os/, although not illegal, are relatively rare. In contrast, few CVs occur less often than expected by chance. Kessler and Treiman interpreted their results to support the view that vowels and final consonants in English have a special bond.

In the present study, we investigated language users' sensitivity to probabilistic VC constraints. On one view, the patterns uncovered by Kessler and Treiman (1997) are the detritus of historical all-or-nothing rules. If so, present-day users of English may not judge newly coined words that follow the statistical patterns as sounding more natural than words that do not. On the other hand, people may have developed a sensitivity to the distributional patterns of their language regardless of the historical causes of these patterns. We carried out three experiments to address this issue.

Our experiments focused on the distinction between more frequent and less frequent VCs. To assess people's implicit knowledge of rime frequency, we developed quadruplets of nonsense syllables. Two syllables in each quadruplet, the *High* syllables, contained more frequent VCs than the other two syllables, the *Low* syllables. Importantly, the High and Low syllables in each quadruplet contained the same phonemes. For example, one quadruplet had the High syllables /rup/ and /nɝk/ and the Low syllables /ruk/ and /nɝp/. Participants in Experiment 1 rated these syllables for the degree to which they sounded like English words. If participants rate High syllables as more word-like than Low syllables, we can conclude that they are sensitive to rime frequency.

X.2. Experiment 1

X.2.1 Method

There were 10 quadruplets of nonsense syllables based on five sets of four rimes each. One such set is /up/, /ɝk/, /uk/, and /ɝp/. Within each set, both High rimes (/up/ and /ɝk/ in the example) were more common than both Low rimes (/uk/ and /ɝp/). Before performing frequency counts, we collapsed the distinction between the vowels of *spa*, *cot*, and *caught* that is made in the Random House Dictionary, representing all three vowels as /ɑ/. This was done because speakers of American English do not always make these distinctions. The High rimes were more frequent than the Low rimes by both types and tokens (weighting each word by its frequency in Francis and Kučera, 1982). All High rimes were in the top half of rimes in frequency and all Low rimes were in the bottom half according to both the unabridged Random House Dictionary (1987) and the smaller Merriam-Webster Pocket Dictionary (Nusbaum, Pisoni, & Davis, 1984). The small number of rimes in the experiment reflects the fact that relatively few rimes fit our stringent constraints. The Appendix lists the rimes that we used. It was necessary to repeat two High rimes and two Low rimes within the experimental sets. Two quadruplets of CVC (consonant-vowel-consonant) syllables were constructed from each rime set by combining the rimes with different onsets. There were 40 test stimuli in all, comprising 10 quadruplets or 5 octuplets. The test syllables were arranged in a random order such that no syllables with the same rime were adjacent.

Another 10 nonwords were constructed for the practice phase of the experiment. Five of them, such as /trɪn/, were phonologically legal in English. The other five, such as /tlɪn/, were illegal. The practice syllables formed pairs that differed in a single phoneme. The first two stimuli in the practice list were /tlɪn/ and /trɪn/; the other eight practice stimuli followed in a random order.

The practice and test stimuli were tape recorded by a female native speaker of Midwestern American English. She repeated each stimulus three times, with a one-second pause between repetitions. She then paused 10 seconds before saying the next stimulus. The list was recorded in a sound-proof booth and played to participants using a Marantz tape recorder (Model PMD221) and Audio-Technica headphones (ATH-610).

Each participant in this and the following experiments was tested individually in a quiet room. The participant was told that he or she would hear “made-up words” and would rate the degree to which each one “sounded like it could be a real English word.” The participant was told to use a rating of 1 for a nonsense word that “doesn’t sound at all like an English word,” 7 for a nonsense word that “sounds very much like an English word,” and intermediate numbers for nonwords that fell between the two extremes. The experimenter played the first taped practice item, /tlɪn/, and said that this syllable does not sound much like an English word and should be rated as 1. The experimenter then played the second practice item, /trɪn/. She stated that this item sounds like it could be a real English word and so should receive a 7. The experimenter then reviewed the instructions and answered any questions. The participant rated the remaining practice items without help from the experimenter and proceeded to the test items.

Twenty-eight adults who were enrolled in courses at Wayne State University or who had recently graduated from college participated. All the participants in this and the following experiments were native speakers of English who reported no history of speech or hearing problems. None participated in more than one experiment of the present series.

X.2.2 Results

We first examined the data for the practice syllables to determine whether participants gave higher ratings to legal syllables than illegal syllables. On the eight practice syllables for which the experimenter did not provide the correct answer, all but two participants gave higher ratings to the legal syllables than to the illegal syllables. The results of these two participants were omitted from the analyses. These individuals may not have understood the directions, may not have paid attention to the task, or may have been swayed by knowledge of a second language.

The mean rating for the High test stimuli was 4.56, as compared to 3.80 for the Low stimuli. This difference was significant whether assessed by one-tailed t tests across subjects (t = 8.22, $p < .001$, $df = 25$), across the 20 pairs of High and Low stimuli that differed by a single phoneme ($t = 5.73$, $p < .001$, $df = 19$), across the 10 quadruplets of High and Low stimuli ($t = 5.51$, $p < .001$, one tailed, $df = 9$), or across the 5 octuplets of High and Low stimuli ($t = 5.31$, $p < .005$, $df = 4$). The reliable difference between High and Low stimuli suggests that adults consider nonwords with common rimes to be more word-like than nonwords with less common rimes.

A more stringent way of evaluating the data is to ask how many participants rated *both* High syllables in a quadruplet as more word-like than both Low syllables. Assume that participants had no actual preferences and that they assigned ratings evenly between 1 and 7. In this case, we can calculate that the probability of finding higher ratings for both High syllables than for both Low syllables in a given quadruplet would be .034. With 26 participants evaluating each quadruplet, about one (0.88) would be expected to show the "right" values for a given quadruplet by chance. By the binomial distribution, we would need to find at least four such participants before we could conclude that the deviation from chance is statistically significant at $p < .05$. Similarly, five or more participants would have to show the "right" values for an octuplet to achieve significance. Of the five octuplets, three showed significant results. This is highly unlikely to have occurred by chance ($p < .001$ by a binomial test). Three of the ten quadruplets showed significance, which is also unlikely to have occurred by chance ($p < .001$ by a binomial test). The results thus suggest that there is a statistically significant preference for both High syllables over both Low syllables. However, the effect is not proven to obtain for all stimulus sets.

Another way of asking whether participants are sensitive to rime frequency is to determine how well measures of the frequency of rimes and of other units predict participants' word-likeness ratings. We performed regression analyses to compare the predictive power of rime frequency with that of other variables -- the frequency of the CV, the frequency of the C_C (initial consonant-final consonant combination), and the frequencies of the initial consonant, vowel, and final consonant. Frequencies were taken from the list of monomorphemic monosyllabic words in the Random House Dictionary (1987). Both type measures and token measures (weighting each word by the natural log of its frequency in Francis and Kučera, 1982 and assigning a frequency of 0.3 to words that did not occur in that corpus[1]) were used. We counted a consonant as shared if it occurred as a singleton in the appropriate part of the syllable. (When we ran the analyses allowing clusters, we obtained very similar results to those reported.)

Simultaneous multiple regression analyses using the six predictors listed above accounted for 42.6% of the variance in word-likeness ratings using type measures and 47.4% using token measures. VC frequency made a significant contribution in both analyses ($p < .05$ for both). Participants gave higher word-likeness ratings to nonwords with more frequent rimes than to nonwords with less frequent rimes. The frequency of the CV also made a significant contribution ($p < .05$ by types; $p < .005$ by tokens). In the type analysis only, there were significant inverse relationships for vowel frequency and final consonant frequency ($p < .05$ for both). That is, once overall rime frequency was taken into account, participants gave lower ratings to syllables with more frequent vowels and those with more frequent codas.

We also carried out stepwise regression analyses. After determining which variable accounted for the most variance, we tested the remaining variables to see which should be entered next, continuing until all variables were included. In both type and token analyses, the

[1] We assigned a frequency of 0.3 to words that did not occur in the corpus so as not to assume that these words do not occur at all in English.

first variable to enter was CV frequency. VC frequency was second to enter the token analysis and fourth to enter the type analysis, after its components (vowel and coda) were added as negative predictors. VC frequency made a sizable contribution when added to the regressions -- 13 percentage points in the token analysis and 9 percentage points by types. We cannot conclude from these results that the frequency of the CV is more important than the frequency of the VC in explaining adults' word-likeness ratings. Because the onsets of the stimuli were selected randomly, whereas the VCs were systematically manipulated, there was more variability in the CV frequencies than in the VC frequencies. We *can* conclude that the frequency of the VC makes a significant contribution to the judged word-likeness of a syllable, above and beyond the separate contributions of the vowel and coda.

X.2.3 Discussion

The results of Experiment 1 suggest that adults are sensitive to some of the phonotactic patterns documented by Kessler and Treiman (1997). They implicitly know that some rimes, such as /ɝk/, are more probable than other rimes, such as /ɝp/. These results do not reflect the frequencies of the phonemes in the rimes because the syllables with more common rimes and the syllables with less common rimes contained the identical phonemes. The regression results further support the claim that adults are sensitive to rime frequency. The frequency of the VC made a significant contribution to the judged word-likeness of a nonword even after the frequencies of the vowel and coda were taken into account. The results thus show that adults make distinctions within the class of nonwords. They judge nonwords with more common rimes as "better" than those with less common rimes.

X.3 Experiment 2

The first goal of Experiment 2 was to replicate the findings of Experiment 1 using a different task. In the nonword comparison task of Experiment 2, participants heard pairs of syllables such as /nɝk/ (a High syllable) and /nɝp/ (a Low syllable). They judged which syllable sounded more like it could be a real English word. A second goal of Experiment 2 was to study children as well as adults. If people become sensitive to the frequencies of rimes in their language as a result of exposure to and use of the language, then even children might show this sensitivity.

X.3.1 Method

The test pairs of Experiment 2 were derived from the stimuli of Experiment 1 by pairing the two syllables in each quadruplet that differed in just the final consonant. (For one pair of stimuli, a different onset was used in Experiment 2.) The 20 test pairs were arranged in a random order such that no two adjacent pairs included syllables with the same rime. The order in a pair (High-Low or Low-High) was randomly chosen. The practice stimuli were derived from those of Experiment 1 by pairing the phonologically legal and illegal syllables that differed in a single phoneme. The first practice pair was /trɪn/-/tlɪn/; the other four practice pairs followed in a random order. The stimuli were recorded as in Experiment 1. For each stimulus pair, the speaker said, "Word A is ___; word B is ___," and then pronounced the stimuli as a pair two times. After a 15 second pause, the speaker announced the number of the next pair and proceeded as above.

Participants were told that they would hear pairs of "made-up words" and would decide which nonsense word in each pair "sounds more like it could be a real English word." The participants were asked to focus on the sounds of the nonwords in making their decisions. The experimenter played the first practice pair, /trɪn/-/tlɪn/, and stated that /trɪn/ sounds more like a word because no English words begin with /tl/. The participant was provided with a sheet on which two responses, A and B, were listed for each item. The experimenter directed the participant to circle A for the first practice item. The remaining practice items were played and the participant responded by circling A or B, with no help from the experimenter. The experimenter then reviewed the instructions and answered any questions before proceeding to the test items.

Second and fourth grade students at an elementary school in a Detroit suburb participated with parental permission. Parents reported that none of the children had a history of speech or hearing problems and that all had English as their native language. The adults were students at Wayne State University. Table X.1 shows the number of participants at each level and, for children, the mean ages.

Table X.1
Information on participants in Experiment 2 and results of the experiment (Standard deviations in parentheses)

	Group		
	Second grade	Fourth grade	College
Number of participants	32	28	26
Mean age of participants	8 years, 1 month	9 years, 11 months	[a]
Mean number of correct responses on practice trials for which experimenter did not provide correct answer, all participants (max. = 4)	3.28 (0.58)	3.46 (0.51)	3.77 (0.43)
Number (and percentage) of participants who performed perfectly on practice trials	11 (34%)	13 (46%)	20 (77%)
Mean number of correct responses in test for participants who performed perfectly on practice trials (max. =20)	11.09 (2.91)	11.62 (2.36)	12.70 (2.79)

[a] The college students were not asked to give their ages

X.3.2 Results

We first examined participants' responses on the practice pairs, each of which contrasted a legal syllable with an illegal syllable. Table X.1 shows the mean number of correct responses on the practice trials for which the experimenter did not provide the correct answer. The children made fewer correct responses than the adults, as confirmed by an analysis of variance by subjects using the factor of group (second grade, fourth grade, college student) ($F(2,83) = 6.48$, $p < .005$). The fact that the experimenter demonstrated the correct answer on only one practice trial may have contributed to the children's relatively poor performance, as this may not have given them sufficient opportunity to grasp the nature of the task.

When examining examine performance on the test trials, we focused on those participants who reliably distinguished legal and illegal syllables in the practice phase. Table X.1 shows the mean number of correct (i.e., High syllable) responses on the test trials for those participants who performed perfectly on the practice trials. An analysis of variance by subjects showed no significant difference among the second graders, fourth graders, and adults ($F(2,41) = 1.42$, $p = .25$). Similarly, an analysis of variance by stimulus pairs failed to find a significant group difference ($F(2,38) = 1.66$, $p = .20$). Thus, although the adults tended to perform better than the children on the test pairs, this tendency was *not* statistically reliable.

As Table X.1 shows, the mean number of correct responses on the test trials exceeded 10 out of 20 for all three groups of participants. Pooling across all participants who scored perfectly on the practice items, performance on the test items was significantly better than expected by chance. This held true whether tested by one-tailed *t* tests by subjects ($t = 4.80$, $p < .001$, $df = 43$), by stimulus pairs ($t = 2.74$, $p < .01$, $df = 19$), by quadruplets ($t = 2.58$, $p < .05$, $df = 9$), or by octuplets ($t = 2.82$, $p < .05$, $df = 4$).

X.3.3 Discussion

The first goal of Experiment 2 was to determine whether the sensitivity of adults to rime frequency that was demonstrated in Experiment 1 would be seen in a different task. It was. The adults tended to judge /nɝk/ as more word-like than /nɝp/, for example, in line with the fact that more English words end with /ɝk/ than with /ɝp/. This result must reflect rime frequency rather than phoneme frequency because the more common and less common rimes used in the experiment contained the identical phonemes. Although there were several uncontrolled factors -- the frequencies of words in people's own mental lexicons, the recency of their exposure to particular words, and the featural similarity between a nonword and a particular known word -- participants' responses were affected to a significant degree by rime frequency.

The second goal of Experiment 2 was to determine whether children are also sensitive to rime frequency. Although the second and fourth grade children found our nonword comparison task somewhat difficult, they performed well above chance on the practice items for which the experimenter did not provide the correct answer. This is consistent with the results of previous studies that show that children and even infants distinguish between legal and illegal syllables of their language (Friederici & Wessels, 1993; Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993; Messer, 1967). Moreover, those children who errorlessly distinguished between legal and illegal syllables in the practice phase of Experiment 2 performed at an above-chance level in the test phase. Indeed, these children did not perform significantly worse than the college students. These results suggest that an implicit knowledge of probabilistic patterns involving rimes is present by second grade in at least some children.

X.4 Experiment 3

In our final experiment, we asked whether rime frequency influences performance in a blending task. This task is of interest because it has been used to study the internal structure of syllables. In the blending task, participants combine two stimuli to form a single item. For example, participants might hear /pɑb/ and /gæf/. One way to combine these syllables is to blend the onset of the first syllable, /p/, with the rime of the second syllable, /æf/. This approach yields /pæf/, a C/VC blend. Another possible response joins the onset and vowel of the first syllable, /pɑ/, with the coda of the second syllable, /f/. The result is /pɑf/, a CV/C blend. Speakers of English generally prefer blends that divide syllables at the onset/rime boundary (i.e., C/VC blends) over other types of blends (Treiman, 1983, 1986; Kubozono, 1995).

In Experiment 3, college students performed a blending task with CVC syllables. Our main question was whether the preference for C/VC blends would be modulated by rime frequency. Thus, half the test pairs involved CVCs with relatively frequent rimes. The other half used CVCs with less frequent rimes. If rimes differ in cohesiveness, C/VC blends should be more common for syllables with High rimes than those with Low rimes.

X.4.1 Method

There were 40 pairs of syllables. In 20 of the pairs, such as /hɝk/-/jɪg/, both syllables contained High rimes. In the other 20 pairs, such as /hɝp/-/jɪʤ/, both syllables had Low rimes. Each pair permitted either a C/VC or a CV/C blend. The stimuli were based on the High and Low rimes of the previous experiments. Onsets were added to minimize the number of stimulus syllables, C/VC blends, and CV/C blends that were real words. The stimuli were divided into two equal-sized lists. The two lists involved the same sets of rimes. For example, /ɝk/ and /ɪg/ occurred in the pair /hɝk/-/jɪg/ in List A and the pair /ðɪg/-/sɝk/ in List B. The rimes were combined with different onsets in the two lists. The order of the rimes within a pair differed across the two lists. Two practice items constructed for which the stimuli, the C/VC responses, and the CV/C responses were all real words.

The experimenter explained that she would pronounce two words on each trial and that the participant should combine them into one new monosyllabic word "by taking part of the first word, starting from the beginning, followed by part of the second word." The experimenter pronounced the words of the first practice pair and the participant repeated each word. The experimenter said that "one possible answer is ___," giving either the C/VC blend or the CV/C blend. Half of the participants heard the C/VC blend as the answer for the first practice item and the other half heard the CV/C blend. The participant repeated the answer after the experimenter. The second practice pair was then presented. The experimenter gave the other type of blend as an answer for the second practice pair. Before presenting the test items, the experimenter explained that they would be similar to the practice items except that they would involve nonsense words rather than real words. The experimenter said that there might be more than one way of combining the syllables and that the participant should give the first answer that came to mind.

For each test trial, the experimenter said the first nonword and the participant repeated it. If the participant repeated the stimulus incorrectly, the experimenter said it again and the participant repeated it again, up to a maximum of three times. This procedure was repeated for the second nonword of the trial. The experimenter said the two syllables as a group twice and then asked for the participant's response. If the response was two syllables long, began with a portion of the second nonword rather than a portion of the first nonword, or began with a non-initial part of the first nonword, the experimenter repeated the relevant part of the instructions and requested another answer. Otherwise, the experimenter made no comment on the response and proceeded to the next item.

The participants were 35 Wayne State University students. Half received List A before List B and the other half had the reverse order. There was a short break between the two lists. The order of the test items within a list was randomly chosen for each participant.

X.4.2 Results

Responses were scored as C/VC, CV/C, or "other." For example, the C/VC response to the pair /hɝk/-/jɪg/ is /hɪg/ and the CV/C response is /hɝg/. An "other" response is one such as

/heg/ that does not fit either category. Table X.2 shows the mean number of responses of each type for High and Low stimuli.

Table X.2
Mean number of responses of various types in blending task of Experiment 3 (Standard deviations in parentheses)

Rime type	C/VC responses	CV/C responses	"Other" responses	Repetition errors
High	17.00 (2.72)	0.40 (0.91)	2.60 (2.52)	1.23 (1.31)
Low	16.31 (2.77)	0.86 (1.14)	2.83 (2.56)	1.94 (1.61)

The great majority of responses were C/VC blends. The number of "other" responses did not differ significantly as a function of rime frequency (t = .93 across subjects, *ns*, df = 34; t = .78 across pairs of High and Low stimuli, *ns*, df = 19; t = 1.27 across quadruplets, *ns*, df = 9; t = 1.49 across octuplets, *ns*, df = 4). However, the proportion of non-"other" responses that fell into the C/VC category (i.e., the proportion of C/VC responses relative to the total of C/VC and CV/C responses) was significantly greater for stimuli with High rimes than for stimuli with Low rimes by one-tailed t tests (t = 2.99, p < .005, df = 34 across subjects; t = 3.32, p < .005, df = 19 across stimulus pairs; t = 2.97, p < .01, df = 9 across quadruplets; t = 2.23, p < .05, df = 4 across octuplets).

The difference between High and Low stimuli in the proportion of non-"other" responses that were C/VC blends, although significant, was small. For High stimuli, 97.6% of all non-"other" responses were C/VC blends. The figure was 95.0% for Low stimuli, a difference of only 2.6%. One reason that the difference was so small is that C/VC responses became more common for Low pairs as the experiment progressed. Figure X.1 depicts the proportion of C/VC responses relative to all non-"other" responses as a function of the order of the stimulus in the experiment. The results for High and Low stimuli are presented as a function of the quartile in which the stimulus occurred -- first (i.e., within the first five High stimuli in the random order or within the first five Low stimuli), second, third, or fourth. For items in the first quartile, the proportion of C/VC responses relative to all non-"other" responses was clearly higher for High stimuli than for Low stimuli. Indeed, 96.2% of the non "other" responses for the very first High pair of the experiment were C/VCs, as compared to 83.3% for the first Low pair. As the experiment progressed, the difference between High and Low pairs diminished, with more responses to Low stimuli falling into the C/VC category.

Figure X 1. Proportion of C/VC responses relative to all non "other" responses in Experiment 3 as a function of the position of High and Low stimuli in the experiment.

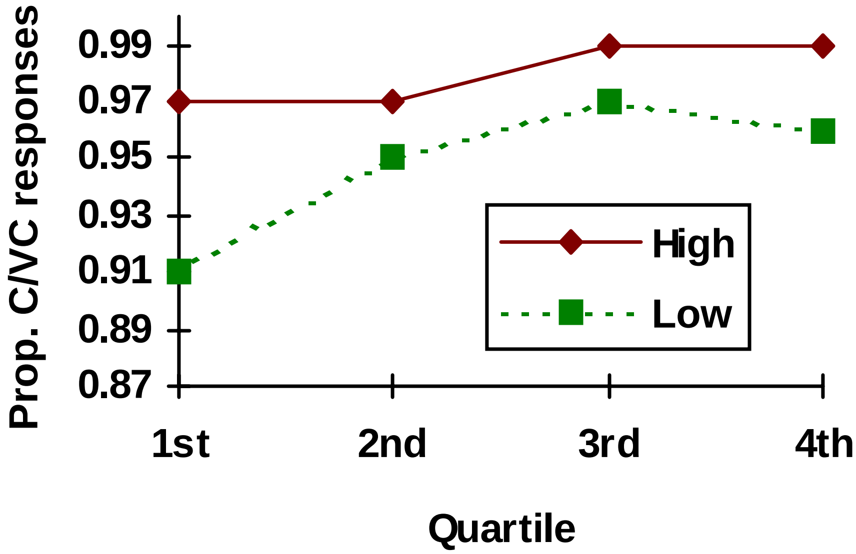

Table X.2 shows that participants made more repetition errors for stimuli with Low rimes than for stimuli with High rimes. The difference was significant by one-tailed *t* tests by subjects ($t = 1.93$, $p < .05$, $df = 34$), stimulus pairs ($t = 2.27$, $p < .025$, $df = 19$), and stimulus quadruplets ($t = 2.21$, $p < .05$, $df = 9$), and was marginally significant by octuplets ($t(4) = 1.86$, $p < .10$, $df = 4$).

X.4.3 Discussion

Syllables with more common and less common rimes behave somewhat differently in the blending task. Although the great majority of responses to all types of syllables joined the onset of the first syllable with the rime of the second syllable (C/VC responses), responses that broke up the rime (CV/C responses) were more common for syllables with less frequent rimes than syllables with more frequent rimes. This was particularly true during the early trials of the experiment.

These results are important because they speak to a debate about how the blending task is performed. Treiman (1983, 1986) interpreted the observed preference for breaks between the onset and vowel of a monosyllable to suggest that English syllables have the two primary constituents of onset and rime. More recently, an alternative explanation has been suggested. People are said to align the stimuli at some salient point -- in English, the vowels of the stressed syllables -- and switch from one stimulus to the other at that point (Beckman, 1995, Davis, 1989; Pierrehumbert & Nair, 1995).. According to this align-and-switch view, blends should not be affected by the nature of the stressed vowel and what comes after it. However, our results show that blends *are* affected to a small but significant extent by the properties of the VC, specifically its frequency. The results suggest that all rimes are cohesive units, but that common rimes are even more cohesive than less common rimes.

Although people usually repeated the syllables correctly, they tended to make more errors on syllables with less common rimes than syllables with more common rimes. This finding concurs with previous reports that the phonotactic structure of a stimulus affects repetition in adults and children (Beckman & Edwards, this volume; Gathercole, 1995; Gathercole, Willis, Emslie, & Baddeley, 1991; Vitevitch, Luce, Charles-Luce, & Kemmerer, 1997). It appears that syllables with uncommon VCs are harder to remember, to pronounce, or both than syllables with more common rimes.

X.5 General Discussion

In phonology, as in other areas of language, linguists have typically focused on all-or-nothing rules. Here, we have taken a more probabilistic approach. We reviewed findings showing that certain VC sequences, although not illegal, are less common than expected by chance (Kessler & Treiman, 1997). We then presented experimental evidence showing that speakers of English have an implicit knowledge of VC frequency. Thus, the adults in Experiment 1 gave significantly higher word-likeness ratings to syllables such as /nɝk/, which have relatively common rimes, than syllables such as /nɝp/, which have less common rimes. Further evidence for adults' sensitivity to rime frequency comes from Experiment 2, where participants tended to rate syllables such as /nɝk/ as "better" than syllables such as /nɝp/ when the syllables were presented in pairs. Previous studies have shown that adults distinguish between legal and illegal syllables (e.g. Brown & Hildum, 1956). Our findings go beyond the earlier results by demonstrating that adults also make distinctions *within* the class of legal syllables (see also Vitevich et al., 1997).

We also found a sensitivity to rime frequency among at least some children. In Experiment 2, those second and fourth graders who consistently picked legal syllables over illegal syllables in the practice phase tended to prefer nonwords with more frequent rimes over nonwords with less frequent rimes in the test phase. By second grade -- the youngest age tested here -- at least some children seem to have a sensitivity to rime frequency that is comparable to that of adults. Previous researchers have reported that preschoolers (Messer, 1967) and even infants (Friederici & Wessels, 1993; Jusczyk et al., 1993) can distinguish between phonologically legal and illegal syllables. Other results suggest that 9-month-old infants make distinctions within the class of legal syllables (Jusczyk, Luce, & Charles-Luce, 1994), although it is not clear from those results whether infants are sensitive to phoneme frequency, phoneme co-occurrence frequency, or both. We can rule out phoneme frequency as a factor in the present study because the High and Low stimuli contained the same phonemes.

Our results suggest that the special role of rimes in language processing (e.g., Treiman, 1983; 1986; Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995) may derive, in part, from the correlations between vowels and final consonants in the phonotactics of English. Supporting this view, we found in Experiment 3 that rime frequency influences adults' performance in the blending task. C/VC blends were slightly but significantly more common for syllables with more frequent VCs than for syllables with less frequent VCs. Apparently, common VCs form especially cohesive rimes that are most likely to behave as units.

In future work, it will be necessary to examine speakers' sensitivity to units other than rimes. Do speakers of English show a special sensitivity to the rime portion of syllables, or would a similar sensitivity to phoneme co-occurrence be found for any sequence of phonemes? The regression analyses of Experiment 1 suggest that adults' word-likeness ratings are influenced by the frequency of the CV as well as by the frequency of the VC. However, because we did not directly manipulate the frequency of the CV unit, we can draw no strong conclusions at this time.

The present findings underline the need for a detailed understanding of the statistical patterns in languages as a basis for understanding language behavior. We have shown that speakers of English pick up certain subtle statistical patterns in their language. Although these patterns may be, in part, the by-products of historical changes, present-day language users show a sensitivity to the patterns. Moreover, English speakers' tendency to treat rimes as units may reflect their experience with the correlations between vowels and final consonants that exist in their language. A similar sensitivity to statistical patterns may exist in other areas of language (e.g. Kelly, 1992; Treiman et al., 1995), and in other languages.

References

Beckman, M. E. (1995). On blending and the mora: Comments on Kubozono. In B. Connell & A. Arvaniti (Eds.), *Phonology and phonetic evidence: Papers in laboratory phonology IV* (pp. 157-167). Cambridge, England: Cambridge University Press.

Beckman, M. E., & Edwards, J. (this volume). Lexical frequency effects on young children's imitative productions.

Booij, G. E. (1983). Principles and parameters in prosodic phonology. *Linguistics, 21*, 249-280.

Brown, R., & Hildum, D. C. (1956). Expectancy and the perception of syllables. *Language, 32,* 411-419.

Clements, G. N., & Keyser, S. J. (1983). *CV phonology: A generative theory of the syllable*. Cambridge, MA: MIT Press.

Davis, S. (1989). On a non-argument for the Rhyme. *Journal of Linguistics, 25*, 211-217.

Francis, W. N., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.

Friederici, A. D., & Wessels, J. M. I. (1993). Phonotactic knowledge of word boundaries and its use in infant speech perception. *Perception and Psychophysics, 54,* 287-295.

Fudge, E. C. (1969). Syllables. *Journal of Linguistics, 5*, 253-286.

Fudge, E. (1987). Branching structure within the syllable. *Journal of Linguistics, 23,* 359-377.

Gathercole, S. E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory and Cognition, 23*, 83-94.

Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1991). The influences of number of syllables and wordlikeness on children's repetition of nonwords. *Applied Psycholinguistics, 12,* 349-367.

Jusczyk, P. W., Friederici, A. D., Wessels, J. M. I., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants' sensitivity to the sound patterns 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.

Kessler, B., & Treiman, R. (1997). Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language, 37,* 295-311.

Kelly, M. H. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review, 99*, 349-364.

Kubozono, H. (1995). Perceptual evidence for the mora in Japanese. In B. Connell & A. Arvaniti (Eds.), *Phonology and phonetic evidence: Papers in laboratory phonology IV* (pp. 141-156). Cambridge, England: Cambridge University Press.

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

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words. *Research on Speech Perception Progress Report No. 10*. Bloomington, IN: Indiana University.

Pierrehumbert, J., & Nair, R. (1995). Word games and syllable structure. *Language and Speech*, *38,* 77-114.

The Random House dictionary of the English language (2nd ed., unabridged). (1987). New York: Random House.

Selkirk, E. O. (1982). The syllable. In H. Van der Hulst & N. Smith (Eds.), *The structure of phonological representations (part II)* (pp. 337-383). Dordrecht, the Netherlands: Foris.

Treiman, R. (1983). The structure of spoken syllables: Evidence from novel word games. *Cognition, 15*, 49-74.

Treiman, R. (1986). The division between onsets and rimes in English syllables. *Journal of Memory and Language, 25*, 476-491.

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General*, *124*, 107-136.

Vitevitch, M. S., Luce, P. A., Charles-Luce, J., & Kemmerer, D. (1997). Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech*, *40*, 47-62.

Appendix

Sets of High and Low rimes used in experiments

Set 1:

High rimes: /up/, /ɝk/

Low rimes: /uk/, /ɝp/

Set 2:

High rimes: /ɪg/, /eʤ/

Low rimes: /ɪʤ/, /eg/

Set 3:

High rimes: /ɪm/, /ɑb/

Low rimes: /ɪb/, /ɑm/

Set 4:

High rimes: /uz/, /ɝk/

Low rimes: /uk/, /ɝz/

Set 5:

High rimes: /ɪʧ/, /ɑb/

Low rimes: /ɪb/,/ɑʧ/