Journal of Experimental Psychology:
Learning, Memory, and Cognition
2004, Vol. 30, No. 2, 514–529

0278-7393/04/$12.00 DOI: 10.1037/0278-7393.30.2.514

Sublexical and Lexical Representations in Speech Production: Effects of Phonotactic Probability and Onset Density

Michael S. Vitevitch, Jonna Armbrüster, and Shinying Chu
University of Kansas

Phonotactic probability, neighborhood density, and onset density were manipulated in 4 picture-naming tasks. Experiment 1 showed that pictures of words with high phonotactic probability were named more quickly than pictures of words with low phonotactic probability. This effect was consistent over multiple presentations of the pictures (Experiment 2). Manipulations of phonotactic probability and neighborhood density showed only an influence of phonotactic probability (Experiment 3). In Experiment 4, pictures of words with sparse onsets were named more quickly than pictures of words with dense onsets. The results of these experiments provide additional constraints on the architecture and processes involved in models of speech production, as well as constraints on the connections between the recognition and production systems.

Models of the recognition and production of speech generally agree that phonological segments, more holistic word forms, and semantic information are represented and used in the processing of spoken language (e.g., McClelland & Elman, 1986; Norris, McQueen, & Cutler, 2000; Vitevitch & Luce, 1999, for perception and Dell, 1986; MacKay, 1987; Stemberger, 1985, for production). Given the similar representations in the recognition and production systems, it is perhaps not surprising that a variable that influences the recognition of a word also influences the production of a word. Consider, for example, the frequency with which a word occurs in the language as measured by an objective word count such as that provided by Kučera and Francis (1967). In recognition, words that occur often in the language are recognized more quickly and accurately than words that occur less often in the language (cf. Savin, 1963; Vitevitch, 2002a). Similarly, in speech production, words that occur often in the language are produced more quickly and accurately than words that occur less often in the language (Dell, 1988, 1990; Oldfield & Wingfield, 1965; Stemberger & MacWhinney, 1986).

Michael S. Vitevitch, Jonna Armbrüster, and Shinying Chu, Department of Psychology, University of Kansas.

This research was supported by Research Grant DC 04259 from the National Institute of Deafness and Other Communication Disorders, National Institutes of Health. Jonna Armbrüster and Shinying Chu received additional support from Undergraduate Research Awards from the University of Kansas.

We thank Gary Dell, Lori James, Don MacKay, and Holly Storkel for many helpful suggestions and comments. We also thank Revital Berkovith for her assistance in collecting the data in Experiment 4. The data in Experiment 1 were used to fulfill part of the requirements for an honors thesis for Jonna Armbrüster, and the data in Experiment 2 were used to fulfill part of the requirements for an honors thesis for Shinying Chu in the Psychology Department at the University of Kansas.

Correspondence concerning this article should be addressed to Michael S. Vitevitch, Spoken Language Laboratory, Department of Psychology, 1415 Jayhawk Boulevard, University of Kansas, Lawrence, KS 66045. E-mail: mvitevit@ku.edu

Work examining the perception of spoken language has found that phonotactic probability influences the speed and accuracy with which spoken stimuli are recognized (e.g., Vitevitch & Luce, 1998, 1999; Vitevitch, Luce, Pisoni, & Auer, 1999). *Phonotactic probability* refers to the frequency with which phonological segments and sequences of phonological segments occur in words in the English language (Jusczyk, Luce, & Charles-Luce, 1994). Certain segments, such as word initial /s/, and sequences of segments, such as /s^/, are more common than other segments, such as word initial /j/, and sequences of segments, such as /ji/. Vitevitch, Luce, Charles-Luce, and Kemmerer (1997; see also Vitevitch, Pisoni, Kirk, Hay-McCutcheon, & Yount, 2002) found that adult listeners rated nonsense words that contained common segments and sequences of segments (i.e., nonwords with high phonotactic probability) as being more like words in English than nonsense words that contained less common segments and sequences of segments (i.e., nonwords with low phonotactic probability). Vitevitch et al. (1997) also found that listeners repeated auditorily presented nonwords with high phonotactic probability more quickly than they repeated nonwords with low phonotactic probability, suggesting that adults are sensitive to phonotactic information and use it in the recognition of spoken words (see also Vitevitch, 2003; Vitevitch & Luce, 1998, 1999). To further examine the similarity between the perception and production of spoken words, and to provide additional constraints on models of each process, in the present set of experiments we examined the influence and locus of phonotactic probability effects in speech production.

In previous research, Motley and Baars (1975) found that phonotactic probability influenced the accuracy with which speech was produced using the spoonerisms of laboratory-induced predisposition (SLIP) task (Baars, 1992; Motley & Baars, 1976). In the SLIP technique, phonological speech errors are elicited by activating two incompatible speech plans. The competition between the two incompatible speech plans increases the likelihood of one's making a speech error when prompted to produce a verbal response. Competition among speech plans is accomplished by instructing participants to repeat to themselves word pairs that are rapidly presented on a computer screen. In each word pair, the initial phoneme of the first word is, for example, a /p/, and the initial phoneme of the second word is, for example, a /b/, such as *push–big, pig–bull,* or *pin–ban,* which strongly activates a *p–b* speech plan. Occasionally, participants are cued (by a tone or visual cue) to say a word pair out loud. In the word pair that must be produced, the initial phoneme of the first word is a /b/, and the initial phoneme of the second word is a /p/, such as *beach–palm.* This creates competition between the *p–b* speech plan that was activated by the preceding word pairs and the *b–p* speech plan that must be used to correctly produce the cued word pair. The competition between the two speech plans may result in the production of *peach–balm,* an induced speech error, instead of the intended *beach–palm* (note that other errors may result as well; cf. Motley, 1986; Sinsabaugh & Fox, 1986).

Motley and Baars (1975) used this task with specially constructed nonsense words that differed in how commonly the segments and sequences of segments (transitions between the initial consonant and subsequent vowel) occurred in the language and found that the segments and sequences of segments with lower probability tended to be replaced by the more common segment (Experiment 2) and by the more common sequence of segments (Experiment 3). That is, a spoonerism, or a speech error, was more likely to occur in an item with low rather than high phonotactic probability, suggesting that probabilistic phonotactic information of the language influences speech production (see also Levitt & Healy, 1985).

More recently, Dell, Reed, Adams, and Meyer (2000) demonstrated that the influence of phonotactic information is not limited to the constraints imposed by the language. In a multiple-session experiment, participants repeated sequences of syllables, such as "feng keg hem nes." In the experimental sessions, the probability with which certain sounds appeared in certain positions varied. For example, in the language, sounds like /f/ and /s/ can appear in the onset or the coda of a word, but in the context of the experiment, they appeared only in the onset or only in the coda. The errors that were elicited not only reflected language-wide phonotactic constraints, but also reflected the phonotactic constraints present in the experimental sessions. For example, if /f/ only appeared in the onset position during the experiment, it tended to erroneously occur only in onset positions (even though it can legally occur in the language in the coda position), further demonstrating that local (in the context of the experiment) and global (in the context of the language) phonotactic information influences the accuracy of speech production.

To further examine how phonotactic information influences speech production, we performed the following studies. In their previous studies of phonotactic information in speech production, Motley and Baars (1975) and Dell et al. (2000) elicited speech errors from participants by using sequences of nonsense words. The present studies differ from these previous studies in several important ways. First, in the present experiment we used real words in English that varied in phonotactic probability rather than specially constructed nonsense words, as in previous studies (see Luce & Large, 2001, and Vitevitch, 2003, for effects of phonotactic probability in real words rather than nonwords in spoken word recognition).

Also in contrast to previous studies, in the present experiments, we used a single-item picture-naming task rather than an error elicitation task such as SLIP or the rapid repetition of tongue-twister–like syllable sequences. Although a great deal has been learned from analyses of speech errors in perception and production (e.g., Bond, 1999; Dell, 1986, 1988; Fay & Cutler, 1977; MacKay, 1987; Shattuck-Hufnagel, 1979; Vitevitch, 1997, 2002a, 2002c; Vitevitch & Sommers, 2003), Levelt, Roelofs, and Meyer (1999) argued that speech errors tell us very little about the online processes involved in normal speech production. Rather, tasks that measure reaction time, like picture naming (Oldfield & Wingfield, 1965), provide more insight into the time course of speech production. Clearly, both dependent measures (reaction times and error rates) provide important and converging pieces of evidence regarding the processes and representations involved in speech production. By considering both the results from the picture-naming task used in the present set of experiments and the results obtained in previous studies that relied on error elicitation techniques, a more complete understanding of the influence of phonotactic probability on speech production can be obtained.

Finally, studies of spoken word recognition suggest that phonotactic information is stored in sublexical representations roughly corresponding to segments or sequences of segments rather than being an emergent property of the word forms stored in the lexicon (cf. McClelland & Elman, 1986; Vitevitch, 2003, Vitevitch & Luce, 1998, 1999). Few researchers have attempted to localize the source of phonotactic information in speech production. Note that Vitevitch et al. (1999) found that among consonant–vowel–consonant words, phonotactic probability is positively correlated with *neighborhood density,* or the number of words that sound similar to a target word (Luce & Pisoni, 1998). That is, words comprised of segments and sequences of segments with high probability tend to have many phonological neighbors. This relationship occurs because segments and sequences of segments become highly probable by occurring in many words (and/or because those words occur in the language quite often); words with the same segments and sequences of segments are phonologically similar, and therefore a positive relationship exists between phonotactic probability and neighborhood density. Given the positive correlation between phonotactic probability and neighborhood density, it is possible that influences of phonotactic information in speech production could either emerge from the word forms stored in the lexicon (e.g., McClelland & Elman, 1986) or be represented independently among sublexical or segmental representations (e.g., Vitevitch, 2003; Vitevitch & Luce, 1998, 1999). Therefore, in the present experiments, we will also attempt to determine the locus of phonotactic effects in speech production.

Experiment 1

In this experiment, we used a picture-naming task to examine whether phonotactic probability influenced the speed with which words are produced (rather than the accuracy with which words are produced, e.g., Dell et al., 2000; Motley & Baars, 1975). The reaction time measures obtained from this task provide converging evidence of the influence of phonotactic probability in speech production, as well as insight into the time course of that influence. In the picture-naming task, a black-and-white line drawing appeared on a computer screen, and participants named the object depicted as quickly and accurately as possible. The time from the onset of the object to the initiation of the participant's response (typically measured by a voice-activated switch interfaced to a timer) was then recorded. In this experiment, pictures of real, monosyllabic words in English varying in phonotactic probability but controlled for a number of other variables (including familiarity, word frequency, neighborhood density, syllable structure, number of phonemes, and the same initial phonemes) were used as stimuli. The same definition and method of estimating phonotactic probability that was used in previous studies of spoken-language research (e.g., Jusczyk et al., 1994; Storkel, 2001; Storkel & Rogers, 2000; Vitevitch & Luce, 1998, 1999) was used in the current set of experiments to facilitate comparison with more recent research on the influence of phonotactic information on processing (cf. Motley & Baars, 1975).

Recall that words that occur often in the language are produced more quickly and accurately than words that occur less often in the language (Dell, 1988; Oldfield & Wingfield, 1965). We made a similar prediction for words containing sounds and sequences of sounds that vary in the frequency with which those segments and sequences occur in the language. That is, words with high phonotactic probability will be named more quickly and accurately than words with low phonotactic probability. Indeed, Motley and Baars (1975) have already demonstrated that stimuli with high phonotactic probability are produced more accurately than stimuli with low phonotactic probability. The present experiment was, therefore, designed to examine how phonotactic probability influences the time course of processing during speech production.

Method

Participants. Twenty-three students from the Introductory Psychology pool of research participants at the University of Kansas took part in the experiment in exchange for partial course credit. All of the participants were native speakers of English, and they reported no history of speech or hearing disorders. None of the participants who took part in this experiment took part in any of the other experiments reported here.

Materials. Twenty-eight line drawings (Snodgrass & Vanderwart, 1980), half of which illustrated words with high phonotactic probability, and the other half of which illustrated words with low phonotactic probability, were used as stimuli in the present experiment. All of the words used in this (and subsequent experiments reported here) were monosyllabic words with a consonant–vowel–consonant syllable structure. Phonotactic probability was calculated with the same two measures that have been used extensively in other studies of phonotactic probability (e.g., Jusczyk et al., 1994; Storkel, 2001; Storkel & Rogers, 2000; Vitevitch & Luce, 1998, 1999; Vitevitch et al., 1997). These two measures are the sum of the positional segment probability and the sum of the biphone probability for the three segments and two biphones in each word. Both measures of phonotactic probability were used in the present studies simply to maintain consistency with earlier reports. At the present time, we do not make any claims regarding the influence or representation of biphones in models of speech production.

The mean probabilities for the sum of the segments in the given positions of the word were .206 (SEM = .01) for the items with high phonotactic probability and .153, SEM = .01; $F(1, 26) = 15.98$, $p < .01$ for the items with low phonotactic probability. The mean probabilities for the sum of the biphones were .013, SEM = .0020, for the items with high phonotactic probability (SEM = .00), and .005, SEM = .00; $F(1, 26) = 24.02$, $p < .01$ for the items with low phonotactic probability. The summed values for the segments and biphones in this experiment are comparable to the summed values for the segments and biphones reported in the spoken word recognition studies of Vitevitch and Luce (1999). The difference in the magnitudes of the segment and biphone probabilities reflects the fact that there are many more biphones than segments.

Similar to that in the study by Motley and Baars (1975), the transitional probability between the initial consonant (C_1) and the subsequent vowel (V) also differed between the two conditions, $F(1, 26) = 8.91$, $p < .01$; note that the transitional probability is equal to the biphone occurrence for C_1 and V. The transitional probability between (or biphone occurrence for) C_1 and V had a mean value of .006, SEM = .01 for words with high phonotactic probability and a mean value of .002, SEM = .00 for words with low phonotactic probability.

Although the differences in phonotactic probability for the two conditions were significantly different, the differences in familiarity ratings (which were obtained from Nusbaum, Pisoni, & Davis, 1984), word frequency, neighborhood density, and neighborhood frequency were not, all $Fs(1, 26) < 1$. Words with high phonotactic probability had a mean familiarity rating of 6.9, SEM = .01, a mean log-frequency of 1.6, SEM = .16, a mean neighborhood density value of 20.7, SEM = 1.55, and a mean log-neighborhood frequency of 1.3, SEM = .06. Words with low phonotactic probability had a mean familiarity rating of 6.9, SEM = .02, a mean log frequency of 1.3, SEM = .16, a mean neighborhood density value of 17.6, SEM = 1.62, and a mean log neighborhood frequency of 1.3, SEM = .07. Furthermore, the same number of initial consonants appeared in each condition (4 /b/, 1 /tʃ/, 1 /d/, 3 /k/, 3 /p/, 1 /s/, and 1 /w/).

The stimuli in the present experiment were not only controlled with regard to the phonological characteristics of the stimuli (as described above), but they were also controlled with regard to perceptual, syntactic, and conceptual characteristics. To evaluate the perceptual complexity of the stimuli, we used an objective measure of picture complexity. Forsythe, Sheehy, and Sawey (2003) used several objective methods to assess the complexity of images and found that they correlated relatively well with subjective ratings of complexity. In the present case, we used the size (in kilobytes) of the picture files saved in .pict file format. Data stored in .pict file format are stored in a vector format, which is essentially a collection of instructions (stored as vectors) on how to draw an object. These instructions are invariant with regard to the resolution or size of the object. However, a more complex picture will require a more complex set of instructions in order to draw the picture and will therefore require more kilobytes to store that information. To objectively evaluate the complexity of the pictures in each condition, we compared the size of the files in the two conditions. Pictures of words with high phonotactic probability had a mean file size of 7.7 kilobytes (SEM = .66), and pictures of words with low phonotactic probability had a mean file size of 6.8 kilobytes (SEM = .50). This difference was not statistically significant, $F(1, 26) = 1.07$, $p = .31$, suggesting that the pictures in each condition were of comparable complexity.

Previous research has found that syntactic class also influences how quickly pictures are named. For example, D'Amico, Bentrovato, Gasparini, Costabile, and Bates (2002) found that child and adult speakers of Italian named pictures of words depicting nouns more quickly than they named pictures of words depicting verbs (see also Vitevitch & Chu, 2003). All of the pictures in the present set of stimuli depicted nouns, thereby controlling for possible differences among the pictures based on syntactic class as observed by D'Amico et al.

Finally, previous research has shown that normal and patient populations name pictures of living, or natural, objects more quickly and accurately than pictures of nonliving objects or artifacts (e.g., Kiefer, 2001; Laws, Leeson, & Gale, 2002; Takarae & Levin, 2001), suggesting that conceptual characteristics also influence latencies in picture-naming tasks. To control the conceptual characteristics of the stimuli, approximately equal numbers of living and nonliving objects were used in each condition. For words with high phonotactic probability, there were 5 natural objects and 9 artifacts. For words with low phonotactic probability, there were 3 natural objects and 11 artifacts. A two-way chi-square analysis confirmed that there was no difference between the two conditions with regard to the number of natural objects and artifacts, $\chi^2 = (1, N = 28) = 0.70$, *ns*. The stimulus words, as well as the stimulus characteristics, are listed in Appendix A.

Procedure. Participants studied a booklet that, on each page, contained the stimulus picture and the monosyllabic word that identified that picture. When participants were confident that they could use the given label for each picture, they were seated in front of an iMac running PsyScope 1.2.2 (Cohen, MacWhinney, Flatt, & Provost, 1993), which controlled stimulus randomization and presentation, and collection of response latencies. A headphone-mounted microphone (Beyer-Dynamic DT109) was interfaced to a PsyScope button box that acted as a voice key with millisecond accuracy. A typical trial proceeded as follows: The word "READY" appeared in the center of the monitor for 500 ms. One of the 28 randomly selected stimulus pictures was then presented and remained visible until a verbal response was initiated. Response latency, measured from the beginning of the stimulus, was triggered by the onset of the participant's verbal response. Another trial began 1 s after a response was made. Responses were also recorded on high-quality audiotape for later accuracy analyses. No picture was presented more than once.

Results and Discussion

In all of the analyses that follow, participants (F_1) and items (F_2) were treated as random factors. Although the appropriateness of treating as a random factor stimulus items that have not been randomly selected is an issue of debate (cf. Clark, 1973, 1976; Cohen, 1976; Hino & Lupker, 2000; Keppel, 1976; Raaijmakers, Schrijnemakers, & Gremmen, 1999; Smith, 1976; Wike & Church, 1976), convention in psycholinguistic research dictates that such analyses be reported. Regardless of the outcome of the items analyses, theoretical interpretation in this article will be based only on the results from the analyses treating participants as a random factor. As such, estimates of effect size will only be reported for analyses treating participants as a random factor. The estimate of effect size that will be reported throughout this article is *PV*, or the proportion of variance explained by the dependent variable (Murphy & Myors, 1998). For reference, $PV = .01$ is considered a small effect, $PV = .10$ is considered a medium effect, and $PV = .25$ is considered a large effect.

The tape-recorded responses of each participant were scored for accuracy. Only accurate responses were included in the analyses of response latency. Errors included responses that were words other than the given label (e.g., responding with *hat* instead of *cap*), and responses in which the participant had improperly triggered the voice key (e.g., "cough," "uh," etc.). A significant effect of phonotactic probability was found, $F_1(1, 22) = 8.3$, $p < .01$, $PV = .27$; $F_2(1, 26) = 4.7$, $p = .04$. Participants responded to words with high phonotactic probability more quickly (677 ms; $SEM = 13.53$) than they did to words with low phonotactic probability (699 ms; $SEM = 14.69$).

There was no difference in overall error rates between the two sets of words (both $Fs < 1$), suggesting that participants did not sacrifice speed for accuracy in making their responses. Words with high phonotactic probability were correctly responded to 95.4% ($SEM = .01$) of the time, and words with low phonotactic probability were correctly responded to 93.3% ($SEM = .01$) of the time. There was also no difference between words with high and low phonotactic probability when the two types of errors (mislabeling vs. improper triggering of the voice key) were analyzed separately (both $Fs < 1$).

As predicted, words with high phonotactic probability were named more quickly than words with low phonotactic probability in the picture-naming task, suggesting that phonotactic information does influence the time course of processing in speech production. These results complement the findings of Dell et al. (2000) and Motley and Baars (1975), who used error elicitation tasks and found that phonotactic probability influenced the accuracy of speech production. Only by examining the patterns of (naturally occurring and elicited) speech errors and using tasks that measure online processing can we obtain a complete understanding of the representations and processes involved in the production of spoken words.

The results of the present experiment also complement the results of Experiments 4 and 5 in Vitevitch (2002c). In those experiments, words varying in neighborhood density but controlled for phonotactic probability (as well as a number of other variables) were presented in a picture-naming task. Words with dense neighborhoods (i.e., many phonologically related words) were produced more quickly than words with sparse neighborhoods (i.e., few phonologically related words). In the present experiment, words varying in phonotactic probability but controlled for neighborhood density (as well as a number of other variables) were presented in a picture-naming task. Words with high phonotactic probability were produced more quickly than words with low phonotactic probability. Observing effects of phonotactic probability while neighborhood density was controlled (and vice versa) was crucial for ruling out possible confounds (given the correlation between phonotactic probability and neighborhood density in the lexicon; Vitevitch et al., 1999) and for localizing the source of phonotactic effects in speech production.

Although we ruled out a possible confound between phonotactic probability and neighborhood density by controlling neighborhood density in this experiment, other confounding variables may have influenced participants' responses, making interpretation of the present results difficult. In the picture-naming task, two confounds that are quite common, but quite easy to address, include influences related to the articulation of the stimulus items and influences related to how well the picture depicts the stimulus item. There are several techniques that can be used to rule out influences related to the articulation of the stimulus items in the picture-naming task. One method involves delaying when the response is made, giving the speaker sufficient time (from 1000–1600 ms) to fully retrieve all aspects of a word before articulation is made. If differences between conditions persist despite the delay, articulation may have been involved (e.g., Jescheniak & Levelt, 1994). Note, however, that this method of ruling out effects of articulation involves the interpretation of a null result.

An alternative approach to ruling out influences of articulation in the picture-naming task is to replace the naming response with a manual response. Instead of measuring response latency from the onset of the picture to the onset of the vocal response, response latency can be measured from the onset of the picture to the onset of the push of a response button. After pressing the response button to indicate they have retrieved the word, participants then say the name of the object out loud in order to have their response scored for accuracy. In this case, if the differences between conditions remain, despite the removal of the voice-activated response, it can be inferred that the differences were not likely the result of articulation (e.g., Vitevitch, 2002c).

Perhaps the simplest way to rule out possible influences of articulation in the picture-naming task is to carefully balance the stimuli such that equal numbers of stimulus words have the same initial phoneme in each condition. In this approach, any influence of articulation on the picture-naming responses would be roughly equivalent in each condition. Recall that equal numbers of words in each condition in the present experiment started with the same initial phoneme, making it unlikely that the results of the present experiment were due to influences related to the articulation of the stimulus items. Although possible influences of articulation were ruled out for the present experiment, possible influences regarding differences in the degree to which the picture depicts the stimulus item may still have influenced the results. We performed Experiment 2 to rule out this possible confound.

Experiment 2

In the present studies examining the influence of phonotactic probability on speech production, we were able to rule out possible influences of articulation by having equal numbers of words in each condition begin with the same phoneme. The stimuli in Experiment 1 were also balanced on a number of other variables to minimize the influence of visual processing and object identification. Despite these precautions, it is still possible that the results of Experiment 1 were due to a systematic difference in the degree to which the pictures illustrated the intended lexical items. That is, the words with high phonotactic probability may have described the pictures used to elicit a response in the picture-naming task better than did the words with low phonotactic probability. To rule out this possibility, we performed a repeated picture-naming task in the present experiment using the same stimuli that we used in Experiment 1 (see Experiment 3 in Vitevitch & Sommers, 2003, for an alternative method of ruling out this confound). Using the same stimuli from Experiment 1 also allowed us to establish the generality of the finding from that experiment by systematically replicating the effect with a different sample of participants (e.g., Wike & Church, 1976).

If the results of Experiment 1 were due to the manipulation in phonotactic probability, then we should again observe an effect of phonotactic probability (in the first presentation of the stimuli as well as across all of the presentations) in the repeated picture-naming task used in the present experiment. We should also observe a main effect of repetition such that (both sets of) words should be produced more quickly with repeated presentation because the word was just recently retrieved (e.g., MacKay, 1987). However, if there indeed were systematic differences in the degree to which the pictures illustrated the intended lexical items, then repeated presentation of the pictures in the present experiment should quickly strengthen the relationship between the pictures and the lexical items. The difference in naming time between the high and low phonotactic conditions should then decrease with each repetition until the pictures in each condition are produced with equivalent speed (at or before the tenth repetition). That is, if a systematic difference in the degree to which the pictures illustrated the intended lexical items produced the results observed in Experiment 1, then phonotactic probability should interact with picture repetition in the present experiment.

Method

Participants. Forty students from the same population sampled in Experiment 1 took part in this experiment.

Materials. We used the same 28 line drawings and equipment that we used in Experiment 1 in Experiment 2.

Procedure. The same procedure used in Experiment 1 was used in Experiment 2. The only exception was that the 28 pictures were repeated 10 times each. More specifically, the 28 pictures were sampled without replacement until all 28 pictures had been presented. Once the 28 pictures had all been presented, they were randomized in a different order and presented again without replacement. This process continued until each picture had been presented 10 times. The experiment, therefore, consisted of 280 trials.

Results

The tape-recorded responses of each participant were scored for accuracy as in Experiment 1. The reaction times for words with high and low phonotactic probability across each repetition are displayed in Figure 1.

For the first presentation of the stimuli in the repeated picture-naming task, an effect of phonotactic probability was found, $F_1(1, 39) = 4.73, p = .03, PV = .11$; $F_2(1, 26) = 4.07, p = .05$. For the first presentation of the stimuli, participants responded to words with high phonotactic probability more quickly (701 ms; SEM = 18.55) than they did to words with low phonotactic probability (719 ms; SEM = 22.03). A significant main effect of phonotactic probability was also found across all ten presentations of the stimuli, $F_1(1, 39) = 7.66, p < .01, PV = .16$; $F_2(1, 260) = 3.15, p = .07$. Across all presentations, participants responded to words with high phonotactic probability more quickly (677 ms; SEM = 18.09) than they did to words with low phonotactic probability (685 ms; SEM = 18.33). A significant main effect of repetition was also found, $F_1(9, 351) = 5.15, p < .01, PV = .11$; $F_2(9, 260) = 2.71, p < .01$, such that with repetition, participants named pictures more quickly. Most important, no interaction between

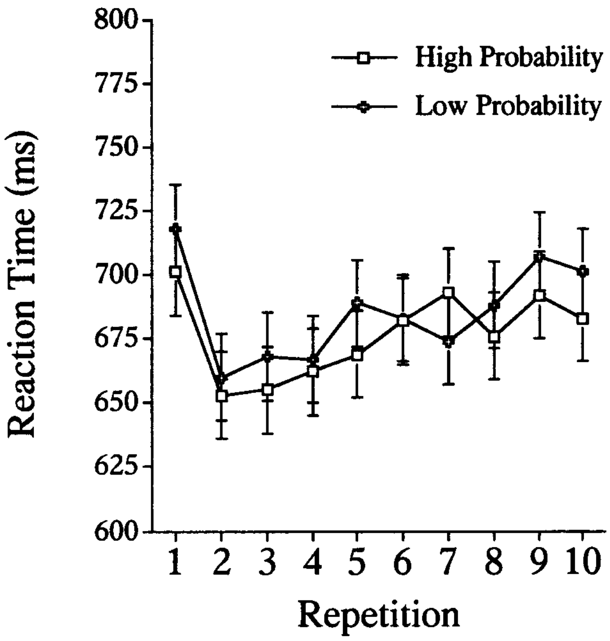

Figure 1. Mean reaction times (in milliseconds) from the repeated picture-naming task in Experiment 2. Error bars represent standard errors.

phonotactic probability and repetition was observed, $F_1(9, 351) = 1.30$, $p = .23$; $F_2(9, 260) < 1$, suggesting that the influence of phonotactic probability was consistent across repetitions.

There was no difference in error rates between the two sets of words for either the first presentation of the stimuli or across all of the presentations (all $Fs < 1$; across all presentations, high phonotactic probability = 97.7% correct and low phonotactic probability = 96.7% correct). There was also no interaction between phonotactic probability and repetition with regard to error rates (both $Fs < 1$). These results suggest that participants did not sacrifice speed for accuracy in making their responses.

Discussion

As in Experiment 1, words with high phonotactic probability were named more quickly than words with low phonotactic probability. This was true when just the first presentation of the stimulus items was examined and when all presentations of the stimulus items were examined. As predicted, a recency effect (MacKay, 1987) was also observed wherein words were retrieved and produced more rapidly as a result of recently being retrieved. Most important, the influence of phonotactic probability remained constant throughout the present experiment (i.e., there was no interaction between phonotactic probability and repetition). The robust influence of phonotactic probability over time suggests that there was no systematic difference in the degree to which the pictures illustrated the intended lexical items; this ruled out this possible confound as an account of the observed results.

The effects of phonotactic probability we observed in Experiments 1 and 2 can easily be accounted for in an interactive model of speech production (e.g., Dell, 1986; Stemberger, 1985; see also MacKay, 1987; Rapp & Goldrick, 2000). Although specific interactive models differ in the details of implementation, Dell (1988) succinctly described how his interactive model would account for findings previously observed by Levitt and Healy (1985) and Motley and Baars (1975) that resembled the presently observed processing advantage for words with high phonotactic probability in speech production: "The model produces this effect because phonemes that are present in many words, particularly in common words, receive additional input as activation reverberates between words and phonemes" (Dell, 1998, pp. 134). The additional input that common phonemes, or segments with high phonotactic probability, receive enables them to be retrieved more quickly and accurately than rare phonemes, or segments with low phonotactic probability (see also Dell, 1986). Thus, in interactive models of speech production, the influence of phonotactic probability arises from different amounts of activation between representations of word forms and phonological segments.

A strictly feedforward model of speech production, such as WEAVER++ (Levelt et al., 1999), could also account for the observed effects of phonotactic probability. In WEAVER++, there are phonological segments, phonological word forms, and a mental syllabary, or a store of syllables that are commonly used in the language. Syllables frequently used in the language are stored in and quickly retrieved from the syllabary. Syllables in the language that are used less frequently are assembled using the segmental information provided in the phonological representation—this is a process that is slower than direct retrieval. The difference between the time it takes to directly retrieve a syllable-sized representation from the syllabary and the time it takes to assemble a syllable-sized representation from the constituent segments could account for the difference in speed (and accuracy) in producing words varying in phonotactic probability (see Levelt & Wheeldon, 1994). This assumes, however, that the high-probability stimuli used in the present study are the items that are stored in the mental syllabary.

Alternatively, WEAVER++ could be modified such that phonotactic information (i.e., the frequency with which the segments and sequences of segments occur in the language) is encoded within the representations at the segmental level. Common segments, like common words, could have lower activation thresholds than rare segments (e.g., Jescheniak & Levelt, 1994) or different verification times as a function of frequency (e.g., Roelofs, 1997). However, an additional level of representation may need to be added to account for the influence of transitional probabilities. Unfortunately, there are few details regarding the contents of the mental syllabary in WEAVER++, so it is unclear which approach might be more successful in accounting for the results of the present experiments, or what consequence modifying the model such that phonotactic information is explicitly encoded in the segmental level might have for accounts of other speech production phenomena such as the influence of neighborhood density on speech production (e.g., Vitevitch, 1997, 2002c).

The results of Experiments 1 and 2 clearly demonstrate that phonotactic probability influences speech production when neighborhood density is controlled. Similarly, the results of Experiments 4 and 5 in Vitevitch (2002c) demonstrated that neighborhood density influenced speech production when phonotactic probability was controlled. Although examining the influences of phonotactic probability and neighborhood density independently of each other is important for increasing our knowledge of the processes and representations involved in speech production (and for ruling out possible confounds), a more complete understanding of the speech production system can be obtained if we examine the influence of each variable in the context of the other during processing. Therefore, we performed Experiment 3 to examine how stimuli varying in both phonotactic probability and neighborhood density influenced speech production.

Experiment 3

Recall that Vitevitch et al. (1999) demonstrated that phonotactic probability and neighborhood density are positively correlated in consonant–vowel–consonant words such that high-probability segments tend to occur in words with many phonological neighbors. Vitevitch (2002c) examined the influence of neighborhood density on speech production while phonotactic probability was controlled (see also Vitevitch & Sommers, 2003). These studies found that words with a dense neighborhood, or many phonological neighbors, were produced more quickly and accurately than words with sparse neighborhoods, or few phonological neighbors (see also Gordon, 2002; Gordon & Dell, 2001; Harley & Bown, 1998; James & Burke, 2000; Vitevitch, 1997). In Experiments 1 and 2, we examined the influence of phonotactic probability on speech production while neighborhood density was controlled and found that words with high phonotactic probability were produced more quickly than words with low phonotactic probability. To better understand how these two correlated variables influence speech production, we orthogonally manipulated phonotactic probability and neighborhood density in this study.

Despite the correlation between phonotactic probability and neighborhood density, Luce and Large (2001) were able to successfully manipulate the two variables in a study of spoken word recognition. We did the same in the present experiment, yielding four sets of words (with corresponding line drawings) such that some of the words were comprised of common segments and sequences and had many (or few) neighbors, and other words were comprised of rare segments and sequences and had many (or few) neighbors. As in Experiments 1 and 2, a number of other variables were controlled to minimize the influence of other extraneous and confounding variables on picture-naming latencies.

Method

Participants. Twenty-five students from the same population sampled in Experiments 1 and 2 took part in this experiment.

Materials. Forty-four line drawings either from or similar to those in Snodgrass and Vanderwart (1980) were used as stimuli in the present experiment. We formed four conditions by orthogonally manipulating phonotactic probability and neighborhood density (high phonotactic probability–dense neighborhood, high phonotactic probability–sparse neighborhood, low phonotactic probability–dense neighborhood, and low phonotactic probability–sparse neighborhood, with 11 pictures in each condition). In each condition, the same number of words had the same initial phoneme (one word each started with /b, d, g, h, l, n, p, r, v/, and two words each started with /m/), thereby ruling out the possibility that any observed differences were due to articulation. Phonotactic probability was calculated as in the previous experiments. As in the spoken word-recognition literature (e.g., Greenberg & Jenkins, 1964; Landauer & Streeter, 1973; Luce & Pisoni, 1998) and previous studies of neighborhood density in speech production (e.g., Vitevitch, 1997, 2002c), neighborhood density was estimated by the number of words that were similar to a target based on the addition, deletion, or substitution of a single phoneme in the target item. For example, in the word /kæt/ [cat], the words /skæt/ [scat], /æt/ [at], /hæt/ [hat], /kʌt/ [cut], and /kæp/ [cap], as well as other words (with a familiarity rating of 6 or higher) found in the computer-readable version of the Webster's Pocket Dictionary, which contains approximately 20,000 words (Nusbaum et al., 1984), would be considered neighbors. Familiarity ratings (contained in the database of Nusbaum et al., 1984) of 6 or higher were used so that the stimuli and the estimate of neighborhood density were based on words that were familiar to most of the participants.

In the high phonotactic probability–dense neighborhood condition, the mean probability for the sum of the segments was .164 (SEM = .00), the mean probability for the sum of the biphones was .008 (SEM = .00), and the mean neighborhood density was 23.9 neighbors (SEM = 1.10). In the high phonotactic probability–sparse neighborhood condition, the mean probability for the sum of the segments was .168 (SEM = .01), the mean probability for the sum of the biphones was .009 (SEM = .00), and the mean neighborhood density was 16.2 neighbors (SEM = .86). In the low phonotactic probability–dense neighborhood condition, the mean probability for the sum of the segments was .119 (SEM = .01), the mean probability for the sum of the biphones was .004 (SEM = .00), and the mean neighborhood density was 23.3 neighbors (SEM = .78).

In the low phonotactic probability–sparse neighborhood condition, the mean probability for the sum of the segments was .112 (SEM = .01), the mean probability for the sum of the biphones was .003 (SEM = .00), and the mean neighborhood density was 15.3 neighbors (SEM = 1.01). Statistical analyses confirmed that the high phonotactic probability conditions were significantly higher in phonotactic probability than were the low phonotactic probability conditions, sum of the phones: $F(1, 40) = 87.96$, $p < .01$; sum of the biphones: $F(1, 40) = 18.5$, $p < .01$. Statistical analyses also confirmed that the dense neighborhood conditions had significantly more neighbors than the sparse neighborhood conditions, $F(1, 40) = 67.88$, $p < .01$. Also note that the neighborhood density values in Experiment 3 are comparable to the values in previous experiments manipulating neighborhood density in speech production (e.g., Vitevitch, 2002c; Vitevitch & Sommers, 2003). Furthermore, there were no significant main effects for the other variable being manipulated or any interactions between the two variables in any of the analyses, all $Fs(1, 40) < 1$.

Although the differences in phonotactic probability and neighborhood density for the four conditions were significantly different, the differences in familiarity ratings, word frequency, and neighborhood frequency were not, all $Fs(1, 40) < 1$. In the high phonotactic probability–dense neighborhood condition, the mean familiarity rating was 6.9 (SEM = .10), the mean log frequency was 1.1 (SEM = .18), and the mean neighborhood log frequency was 1.0 (SEM = .06). In the high phonotactic probability–sparse neighborhood condition, the mean familiarity rating was 6.9 (SEM = .03), the mean log frequency was .99 (SEM = .19), and the mean neighborhood log frequency was 1.0 (SEM = .10). In the low phonotactic probability–dense neighborhood condition, the mean familiarity rating was 6.9 (SEM = .07), the mean log frequency was .85 (SEM = .16), and the mean neighborhood log frequency was .8 (SEM = .05). In the low phonotactic probability–sparse neighborhood condition, the mean familiarity rating was 6.9 (SEM = .04), the mean frequency was 1.0 (SEM = .23), and the mean neighborhood log frequency was 1.0 (SEM = .10).

As in Experiments 1 and 2, the stimuli in the present experiment were also controlled with regard to perceptual, syntactic, and conceptual characteristics. We again objectively evaluated the complexity of the pictures in each condition by comparing the size of the files in the four conditions. Pictures in the high phonotactic probability–dense neighborhood condition had a mean file size of 30 kb (SEM = 6.52). Pictures in the high phonotactic probability–sparse neighborhood condition had a mean file size of 36 kb (SEM = 6.32). Pictures in the low phonotactic probability–dense neighborhood condition had a mean file size of 34.1 kb (SEM = 9.65). Pictures in the low phonotactic probability–sparse neighborhood condition had a mean file size of 38.5 kb (SEM = 7.85). These differences were not statistically significant, all $Fs(1, 40) < 1$, the smallest $p = .52$, suggesting that the pictures in each condition were of comparable complexity.

In the present set of stimuli, four words were verbs and the rest were nouns. A one-way chi-square analysis confirmed that there was no difference between the four conditions with regard to the number of verbs in each condition, $\chi^2 = (3, n = 44) = 2.00$, *ns.* The number of living and nonliving objects in each condition also did not differ across conditions, as confirmed by a two-way chi-square analysis, $\chi^2 = (3, n = 44) = 3.27$, *ns.* The stimulus words, as well as the stimulus characteristics, are listed in Appendix B.

Procedure. The same procedure and equipment that was used in Experiment 1 was used in the present experiment.

Results

The tape-recorded responses of each participant were scored for accuracy as in Experiment 1. The reaction times for words varying in phonotactic probability and neighborhood density are displayed in Figure 2.

A significant main effect of phonotactic probability was found, $F_1(1, 24) = 31.58, p < .01, PV = .56; F_2(1, 40) = 3.94, p = .05$, such that participants responded to words with high phonotactic probability more quickly (829 ms, SEM = 14.5) than they did to words with low phonotactic probability (904 ms, SEM = 20.80). Neither the main effect of neighborhood density nor the interaction between phonotactic probability and neighborhood density were statistically significant (all $Fs < 1$).

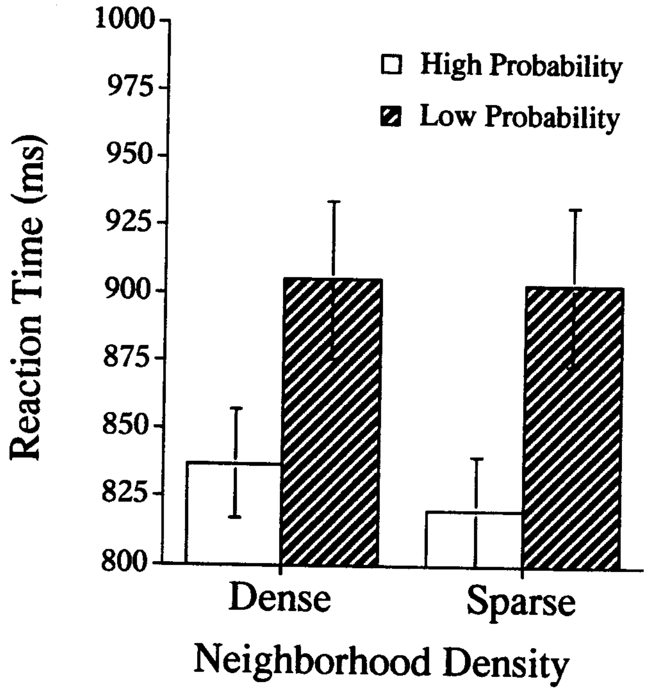

Figure 2. Mean reaction times (in milliseconds) from the picture-naming task in Experiment 3. Error bars represent standard errors.

For the accuracy rates, a significant main effect of phonotactic probability was also found, $F_1(1, 24) = 12.17, p < .01, PV = .56$; $F_2(1, 40) = 4.96$, $p = .03$, such that participants responded to words with high phonotactic probability more accurately (91.1%, $SEM = .01$) than they did to words with low phonotactic probability (84.5%, $SEM = .02$). As with the response times, neither the main effect of neighborhood density nor the interaction between phonotactic probability and neighborhood density were statistically significant (all $Fs < 1$) for the accuracy rates.

Discussion

The results of the present experiment, in which phonotactic probability and neighborhood density were both manipulated, showed that words with high phonotactic probability were produced more quickly and accurately than words with low phonotactic probability. No difference or interaction with neighborhood density was found. Although the results of the present experiment may appear to be unexpected—especially in the context of previously demonstrated effects of neighborhood density in speech production (e.g., Gordon, 2002; Gordon & Dell, 2001; Harley & Bown, 1998; James & Burke, 2000; Vitevitch, 1997, 2002c; Vitevitch & Sommers, 2003)—these results can be accounted for by interactive models of speech production in a fairly straightforward manner and without additional assumptions.

To account for neighborhood density effects in speech production, Vitevitch (2002c) suggested that phonologically related words (i.e., neighbors) facilitated the retrieval of a target word form by interacting via shared constituent phonological segments. Because of the difference in the number of neighbors, target words with dense neighborhoods receive greater amounts of activation from neighboring words via the shared phonological segments than target words with sparse neighborhoods. The amount of activation that targets receive from different numbers of neighbors results in words with dense neighborhoods being produced more quickly and accurately than words with sparse neighborhoods.

The account of the influence of phonotactic probability in Experiments 1 and 2 in the context of an interactive model also appealed to activation interacting between word forms and phonological segments. Phonemes with high phonotactic probability are constituents of many (and common) words, whereas phonemes with low phonotactic probability are constituents of few (and less common) words. The different amounts of activation received by common versus rare phonological segments from the word forms of which they are constituents accounted for the difference in the speed and accuracy of producing those items. How can the same mechanism in an interactive model of speech production—activation interacting between word forms and phonological segments—account for neighborhood density effects when phonotactic probability is controlled (Vitevitch, 2002c), phonotactic probability effects when neighborhood density is controlled (as in Experiments 1 and 2), and the appearance of only phonotactic effects when phonotactic probability and neighborhood density are both manipulated (as in Experiment 3)?

Consider first how phonotactic probability and neighborhood density operate in spoken word recognition. Vitevitch and Luce (1998, 1999; see also Vitevitch, 2003) suggested that effects of phonotactic probability arose when sublexical representations (something resembling phonological segments) dominated the process of spoken word recognition, whereas effects of neighborhood density arose when lexical representations (i.e., lexemes or word forms) dominated the process of spoken word recognition. In spoken word recognition—in which the goal is to recognize words not strings of phonemes—lexical representations, in general, tend to dominate the processing of sublexical representations, resulting in the observation of neighborhood density rather than phonotactic probability effects.

Similarly, one type of representation may dominate processing in speech production. In the case in which neighborhood density was manipulated but phonotactic probability was controlled (Vitevitch, 2002c), activation among phonological segments was equivalent, enabling activation among word forms to dominate and to determine how quickly and accurately a word form would be produced. In the case in which phonotactic probability was manipulated but neighborhood density was controlled (Experiments 1 and 2), activation among word forms was equivalent, enabling activation among phonological segments to dominate and determine how quickly and accurately a word form would be produced. In the case of Experiment 3, in which phonotactic probability and neighborhood density were both manipulated, the dynamics of speech production suggest that phonological segments should, in general, dominate processing and determine how quickly and accurately a word form would be produced. All current models of speech production postulate that holistic word forms activate constituent phonological segments (which in turn may activate constituent phonetic features, which in turn may activate the gestural scores that yield those phonetic features, etc.), so it is plausible that phonological segments can dominate processing of word forms. Furthermore, Harley (1993) demonstrated that processing among word forms dominated processing among lemmas (semantic or conceptual representations) in an interactive model, suggesting that the same mechanism may also allow phonological segments to dominate word forms in an interactive model of speech production.

Feedforward models of speech production, such as WEAVER++ (Levelt et al., 1999), may also account for the results of Experiment 3. In the Discussion section of Experiment 2, we offered two approaches for how a feedforward model might account for effects of phonotactic probability in speech production. One approach suggested that stimuli with high phonotactic probability might be stored and quickly retrieved from the mental syllabary, whereas stimuli with low phonotactic probability would have to be assembled using the segmental information provided in the phonological representation, which is a slower process. The difference between the time it takes to directly retrieve a representation from the syllabary and the time it takes to assemble the constituent segments would account for the difference in speed and accuracy in producing words varying in phonotactic probability.

However, the words in Experiment 3 did not just differ in phonotactic probability; they also differed in neighborhood density, or the number of words that sound similar to that word. The additional manipulation of neighborhood density in Experiment 3 may introduce a potential problem for a syllabary-based account of the present experiment. Recall that the constituent segments of stimuli with low phonotactic probability are assembled rather than retrieved directly from the syllabary. Assembling the segments for the low probability words should take the same amount of time regardless of whether the low probability word has a dense or sparse neighborhood (however, see Landauer & Streeter, 1973, for evidence that the set of phonemes that comprise words with dense neighborhoods is different than the set of phonemes that comprise words with sparse neighborhoods).

In the case of stimuli with high phonotactic probability, retrieval from the syllabary should differ as a function of neighborhood density because selection of items in the mental syllabary is based on the Luce choice rule. Selection based on the Luce choice rule suggests that high probability stimuli with dense neighborhoods (i.e., many similar neighbors, which are presumably stored in the syllabary as well) should be retrieved from the syllabary more slowly than high probability stimuli with sparse neighborhoods (i.e., few similar neighbors). Note that this was not the pattern of results that was observed in the present Experiment 3. It is unclear how problematic the results of the present experiment are for a syllabary-based account given the lack of detail regarding the point that demarcates high-frequency syllables of a language from low-frequency syllables, and the lack of detail about exactly which syllables are actually stored in the syllabary.

The other alternative offered for a feedforward model to account for the results of Experiments 1 and 2 might account for the results of Experiment 3. It was suggested that a feedforward model that encoded phonotactic probability in the representations at the segmental level might be able to account for the observed effects of phonotactic probability. Note that a decision would have to be made regarding how the frequency of the segments should be represented, either as nodes with different activation thresholds or as representations that require different amounts of time during the verification phase of production. It is unclear whether one method of representing phonotactic probability would be more beneficial than the other.

It is also unclear what effects such modifications to the model might have on the ability of the model to account for other speech production phenomena. Consider, for example, that such modifications would only enable the model to account for the results of Experiments 1 and 2 in which neighborhood density was controlled, and Experiment 3, in which neighborhood density effects were not present. Such modifications would most likely not help the model account for the cases in which effects of neighborhood density or of *neighborhood frequency,* defined as the mean frequency of occurrence for the neighbors, are observed (e.g., Vitevitch, 2002c; Vitevitch & Sommers, 2003). Recall that interactive models of speech production appeared to be able to account for all of these situations with the same mechanism, namely activation interacting between word forms and phonological segments. If information represented at the level of word forms and at the level of phonological segments interacts to influence speech production, then we should observe other influences on speech production as a result of information represented at these two levels interacting. To further examine the influence of word forms interacting with phonological segments on speech production, we manipulated a different variable—one that explicitly relies on the interaction of information stored at both levels of representation—in Experiment 4.

Experiment 4

Although we believe that the results of Experiment 3, in conjunction with the results of Vitevitch (2002c) and Vitevitch and Sommers (2003), pose a serious challenge to feedforward models of speech production, we wished to further explore the interaction of information represented at the level of word forms and phonological segments in another experiment. As we did in Experiment 3, we examined lexical and phonological information in Experiment 4. However, in Experiment 4, we probed lexical and phonological information by manipulating a single variable previously examined in spoken word recognition but not in speech production, namely *onset density* (Vitevitch, 2002b). Onset density refers to the proportion of neighbors that share the same initial phoneme as the target word; note that both lexical and phonological information must be available and consulted to assess a stimulus item in terms of this variable.

To illustrate the concept of onset density, consider the words /mæs/ [mass] and /sæd/ [sad]. The word /mæs/ [mass] has as neighbors—based on the computational metric—the words /mIs/ [miss], /mæd/ [mad], /mæn/ [man], and /pæs/ [pass]. The word /sæd/ [sad] has as neighbors the words /bæd/ [bad], /fæd/ [fad], /læd/ [lad], and /sæk/ [sack]. The words /mæs/ and /sæd/ have more words as neighbors; however, for presentation purposes, only these four neighbors have been listed. Note that three of the four neighbors of the word /mæs/ have the same initial phoneme as the target word /mæs/ (the phoneme /m/), whereas one of the four neighbors of the word /sæd/ has the same initial phoneme as the target word /sæd/ (the phoneme /s/). Despite an equal number of neighbors, the word /mæs/ has a greater proportion of neighbors that have the same initial phoneme as the target word (75%), whereas the word /sæd/ has a smaller proportion of neighbors that share the initial phoneme as the target word (25%). Words with a high proportion of neighbors sharing the onset of the target word are said to have a *dense onset,* whereas words with a low proporconceptual representations) in an interactive model, suggesting that the same mechanism may also allow phonological segments to dominate word forms in an interactive model of speech production.

Feedforward models of speech production, such as WEAVER++ (Levelt et al., 1999), may also account for the results of Experiment 3. In the Discussion section of Experiment 2, we offered two approaches for how a feedforward model might account for effects of phonotactic probability in speech production. One approach suggested that stimuli with high phonotactic probability might be stored and quickly retrieved from the mental syllabary, whereas stimuli with low phonotactic probability would have to be assembled using the segmental information provided in the phonological representation, which is a slower process. The difference between the time it takes to directly retrieve a representation from the syllabary and the time it takes to assemble the constituent segments would account for the difference in speed and accuracy in producing words varying in phonotactic probability.

However, the words in Experiment 3 did not just differ in phonotactic probability; they also differed in neighborhood density, or the number of words that sound similar to that word. The additional manipulation of neighborhood density in Experiment 3 may introduce a potential problem for a syllabary-based account of the present experiment. Recall that the constituent segments of stimuli with low phonotactic probability are assembled rather than retrieved directly from the syllabary. Assembling the segments for the low probability words should take the same amount of time regardless of whether the low probability word has a dense or sparse neighborhood (however, see Landauer & Streeter, 1973, for evidence that the set of phonemes that comprise words with dense neighborhoods is different than the set of phonemes that comprise words with sparse neighborhoods).

In the case of stimuli with high phonotactic probability, retrieval from the syllabary should differ as a function of neighborhood density because selection of items in the mental syllabary is based on the Luce choice rule. Selection based on the Luce choice rule suggests that high probability stimuli with dense neighborhoods (i.e., many similar neighbors, which are presumably stored in the syllabary as well) should be retrieved from the syllabary more slowly than high probability stimuli with sparse neighborhoods (i.e., few similar neighbors). Note that this was not the pattern of results that was observed in the present Experiment 3. It is unclear how problematic the results of the present experiment are for a syllabary-based account given the lack of detail regarding the point that demarcates high-frequency syllables of a language from low-frequency syllables, and the lack of detail about exactly which syllables are actually stored in the syllabary.

The other alternative offered for a feedforward model to account for the results of Experiments 1 and 2 might account for the results of Experiment 3. It was suggested that a feedforward model that encoded phonotactic probability in the representations at the segmental level might be able to account for the observed effects of phonotactic probability. Note that a decision would have to be made regarding how the frequency of the segments should be represented, either as nodes with different activation thresholds or as representations that require different amounts of time during the verification phase of production. It is unclear whether one method of representing phonotactic probability would be more beneficial than the other.

It is also unclear what effects such modifications to the model might have on the ability of the model to account for other speech production phenomena. Consider, for example, that such modifications would only enable the model to account for the results of Experiments 1 and 2 in which neighborhood density was controlled, and Experiment 3, in which neighborhood density effects were not present. Such modifications would most likely not help the model account for the cases in which effects of neighborhood density or of *neighborhood frequency,* defined as the mean frequency of occurrence for the neighbors, are observed (e.g., Vitevitch, 2002c; Vitevitch & Sommers, 2003). Recall that interactive models of speech production appeared to be able to account for all of these situations with the same mechanism, namely activation interacting between word forms and phonological segments. If information represented at the level of word forms and at the level of phonological segments interacts to influence speech production, then we should observe other influences on speech production as a result of information represented at these two levels interacting. To further examine the influence of word forms interacting with phonological segments on speech production, we manipulated a different variable—one that explicitly relies on the interaction of information stored at both levels of representation—in Experiment 4.

Experiment 4

Although we believe that the results of Experiment 3, in conjunction with the results of Vitevitch (2002c) and Vitevitch and Sommers (2003), pose a serious challenge to feedforward models of speech production, we wished to further explore the interaction of information represented at the level of word forms and phonological segments in another experiment. As we did in Experiment 3, we examined lexical and phonological information in Experiment 4. However, in Experiment 4, we probed lexical and phonological information by manipulating a single variable previously examined in spoken word recognition but not in speech production, namely *onset density* (Vitevitch, 2002b). Onset density refers to the proportion of neighbors that share the same initial phoneme as the target word; note that both lexical and phonological information must be available and consulted to assess a stimulus item in terms of this variable.

To illustrate the concept of onset density, consider the words /mæs/ [mass] and /sæd/ [sad]. The word /mæs/ [mass] has as neighbors—based on the computational metric—the words /mIs/ [miss], /mæd/ [mad], /mæn/ [man], and /pæs/ [pass]. The word /sæd/ [sad] has as neighbors the words /bæd/ [bad], /fæd/ [fad], /læd/ [lad], and /sæk/ [sack]. The words /mæs/ and /sæd/ have more words as neighbors; however, for presentation purposes, only these four neighbors have been listed. Note that three of the four neighbors of the word /mæs/ have the same initial phoneme as the target word /mæs/ (the phoneme /m/), whereas one of the four neighbors of the word /sæd/ has the same initial phoneme as the target word /sæd/ (the phoneme /s/). Despite an equal number of neighbors, the word /mæs/ has a greater proportion of neighbors that have the same initial phoneme as the target word (75%), whereas the word /sæd/ has a smaller proportion of neighbors that share the initial phoneme as the target word (25%). Words with a high proportion of neighbors sharing the onset of the target word are said to have a *dense onset,* whereas words with a low proportion of neighbors sharing the onset of the target word are said to have a *sparse onset.* Research in spoken word recognition found that words with sparse onsets were recognized more quickly than words with dense onsets despite having the same total number of neighbors (Vitevitch, 2002b). Looking at onset density in speech production (with overall neighborhood density and phonotactic probability, as well as a number of other variables, held constant) provides another opportunity to evaluate the influence of lexical and phonological information in speech production.

Method

Participants. Twenty-eight students from the same population sampled in the previous experiments took part in this experiment.

Materials. Twenty-six line drawings, either from Snodgrass and Vanderwart (1980) or similar in artistic style to those pictures, were used as stimuli in Experiment 4. Thirteen pictures were in each condition. One set of words had a proportion of neighbors that shared the initial phoneme of the target word that was greater than 50%. This set, referred to as the *dense-onset condition,* had a mean proportion of 74.4% (*SEM* = .01) of the neighbors with the same initial phoneme as the target word. The other set of stimuli contained words that had less than 50% of the neighbors sharing the initial phoneme of the target word. This set, referred to as the *sparse-onset condition,* had a mean proportion of 43.4% (*SEM* = .01) of the neighbors sharing the initial phoneme of the target word. The difference in the proportion of neighbors sharing the initial phoneme of the target word between the sparse- and dense-onset conditions was significantly different, $F(1, 24) = 393.5$, $p < .01$.

Both sets of words were equivalent in word familiarity, $F(1, 24) < 1$, as measured by a 7-point scale (Nusbaum et al., 1984). Words in the dense-onset condition had a mean familiarity of 6.89 (*SEM* = .05). Words in the sparse-onset condition had a mean familiarity of 6.91 (*SEM* = .05), indicating that all the stimuli used were highly familiar words to native speakers of English.

The two sets of words were also equivalent in word frequency, $F(1, 24) < 1$, as measured by log-transformed Kučera and Francis (1967) word counts. Words in the dense-onset condition had a mean log frequency of 1.13 (*SEM* = .18). Words in the sparse-onset condition had a mean log frequency of 1.09 (*SEM* = .19).

Using the same computational metric we used in Experiments 1–3, we found that the number of neighbors did not differ significantly between the two sets of words, $F(1, 24) < 1$. Words in the dense-onset condition had a mean neighborhood density of 20.6 words (*SEM* = 1.17). Words in the sparse-onset condition had a mean neighborhood density of 22.6 words (*SEM* = 1.15).

The two sets of stimuli did not differ in neighborhood frequency, $F(1, 24) < 1$—the mean frequency of occurrence for the neighbors—as measured by log-transformed Kučera and Francis (1967) word counts. The mean neighborhood log frequency for words in the dense-onset condition was 1.04 (*SEM* = .05). The mean neighborhood log frequency for words in the sparse-onset condition was 1.03 (*SEM* = .10).

Phonotactic probability was also equivalent between the two conditions, $F(1, 24) < 1$ (for the sum of the phones and sum of the biphones). In the dense-onset condition, the mean probability for the sum of the segments was .154 (*SEM* = .01), and the mean probability for the sum of the biphones was .007 (*SEM* = .00). In the sparse-onset condition, the mean probability for the sum of the segments was .164 (*SEM* = .01), and the mean probability for the sum of the biphones was .008 (*SEM* = .00).

As in Experiments 1–3, the stimuli in the present experiment were also controlled with regard to perceptual, syntactic, and conceptual characteristics. We again objectively evaluated the complexity of the pictures in each condition by comparing the size of the files in the two conditions. Pictures in the dense-onset condition had a mean file size of 30 kb (*SEM* = 5.19), and pictures in the sparse-onset condition had a mean file size of 38 kb (*SEM* = 5.96). This difference was not statistically significant, $F(1, 24) = 1.18$, $p = .29$, suggesting that the pictures in both conditions were of comparable complexity.

In the present set of stimuli, only one picture (in the sparse-onset condition) depicted the verb form of a word (*weep*). All other pictures depicted the noun form of the words. The number of living and nonliving objects in each condition also did not differ across conditions, as confirmed by a two-way chi-square analysis, $\chi^2 = (1, n = 26) = 1.53$, *ns*. The stimulus words, as well as the stimulus characteristics, are listed in Appendix C.

Finally, equal numbers of words in each set contained the following phonemes in the initial position: /d, k (3), l, m (3), p, r, s, t, w/. Controlling the phonemes in the initial position of the words in each set ensured that possible differences in reaction time between the two sets were due to the manipulation of onset density and not to differences in the phonemes used in the initial position between the two sets of words. Using the same number of initial phonemes in each condition also ensured that any observed differences were not due to articulatory differences.

Procedure. The same procedure and equipment that was used in Experiment 1 was used in Experiment 4.

Results

The tape-recorded responses of each participant were scored for accuracy as in Experiment 1. A significant effect of onset density was found, $F_1(1, 27) = 17.25$, $p < .01$, $PV = .39$; $F_2(1, 24) = 4.84$, $p = .03$ such that participants responded to words with sparse onsets more quickly (819 ms; *SEM* = 15.69) than they did to words with dense onsets (866 ms; *SEM* = 13.39). For the accuracy rates, a significant effect of onset density was also found, $F_1(1, 27) = 4.39$, $p = .04$, $PV = .14$; $F_2(1, 24) = 5.41$, $p = .02$, such that participants responded to words with sparse onsets more accurately (85%; *SEM* = .03) than they did to words with dense onsets (80%; *SEM* = .03).

Discussion

The results of Experiment 4 showed that words with a sparse onset, or a low proportion of neighbors sharing the onset of the target word, were responded to more quickly (and accurately) than words with a dense onset, or a high proportion of neighbors sharing the onset of the target word. Recall that the words varying in onset density were equivalent with regard to overall neighborhood density, phonotactic probability, and a number of other variables.

Consider first how an interactive model of speech production—in which activation between word forms and phonological segments interacts—might account for the results of Experiment 4. In the case of a word with a dense onset, such as /mæs/ [mass], activation of the target word form will spread to and partially activate the constituent phonological segments (/m/ /æ/ /s/). Those phonological segments will spread activation back to other word forms that contain those segments, such as /mIs/ [miss], /mæd/ [mad], /mæn/ [man], and /pæs/ [pass]. Retrieval of the initial phoneme (/m/) will occur quickly and accurately because many word forms are supporting the activation of that phonological segment. However, retrieval of the following /æ/ and /s/ of the target word /mæs/ will be impaired because many other segments (/I/ /d/ /n/) have become highly activated via the neighboring word forms. These other phonological segments will compete with the intended segments (/æ/ and /s/) for encoding and placement in the syllable frame, thereby slowing the overall production of words with dense onsets.

In contrast, retrieval (encoding and placement) of the segments that follow the onset will not be impaired in words with sparse onsets, such as /sæd/ [sad]. Activation spreading from the target word to the constituent segments (/s/ /æ/ /d/) and back to neighboring word forms like /bæd/ [bad], /fæd/ [fad], /læd/ [lad], and /sæk/ [sack] will result in many words supporting the retrieval of the segments that follow the onset (/æ/ and /d/ in this example). As a consequence, words with sparse onsets will be produced more quickly and accurately than words with dense onsets. Work by Sevald and Dell (1994) further supports this hypothesis (see also Sevald, Dell, & Cole, 1995; Yaniv, Meyer, Gordon, Huff, & Sevald, 1990; Wheeldon, 2003). Sevald and Dell found that participants produced more repetitions of word sequences during a fixed response time when the word sequences had the same segments that followed the onset (e.g., *pick tick*) than when the word sequences had the same onset (e.g., *pick pin*). In the present experiment, however, competition among the segments for the syllable frame positions that follow the onset came from neighboring word forms that were activated concurrently in the lexicon rather than from words that were read sequentially. It is important to note that the same mechanism—activation between word forms and phonological segments interacting—was used to account for the results of the present experiments that investigated phonotactic probability (Experiments 1 and 2), neighborhood density (Experiment 3; see also Vitevitch, 1997, 2002c; Vitevitch & Sommers, 2003), and onset density (Experiment 4) in speech production.

Now consider how a feedforward model of speech production, like WEAVER++ (Levelt et al., 1999), might account for the effects of onset density observed in Experiment 4. Levelt et al. (pp. 7–8) discuss how a visually or auditorily presented prime might influence naming latencies for picture targets. They suggest that primes, such as *escape*, that are in the same phonological cohort as the target *escort* (i.e., they share the same word-initial phonemes, /ɛsk/ in this case), will facilitate naming of the target word *escort* by partially activating the shared phonological segments. Presumably, target words with many words in the word-initial cohort will be facilitated more than target words with fewer words in the word-initial cohort. This prediction is exactly the opposite of the results observed in the present experiment, in which onset density, a variable similar to a cohort estimate of similarity, was manipulated: Words with sparse onsets were responded to more quickly (and accurately) than words with dense onsets. Also note that no priming words were presented in Experiment 4, so it is unclear whether the mechanism that Levelt et al. described would even apply to the simple picture-naming task employed in this experiment.

Consider instead the other approaches that we described to possibly account for the results of Experiments 1–3 in the context of a feedforward model of speech production, namely direct retrieval of items from the mental syllabary and assembly of items based on information in the phonological form. Recall that the stimuli in the present experiment were controlled with regards to word frequency, neighborhood density, and phonotactic probability (as well as a number of other variables). Given that neighborhood density and phonotactic probability were equivalent between the two stimulus conditions, either all of the stimuli would be directly retrieved from the syllabary, or all of the stimuli would be assembled using information in the phonological form. In the case in which all of the stimulus items are directly retrieved from the syllabary (with selection based on the Luce choice rule), there would be no difference in naming latency for the two sets of words because they have equivalent overall neighborhood densities. In the case in which all of the stimulus items are assembled using information in the phonological form (and assuming less common sequences take longer to assemble than more common sequences), there again would be no difference in the speed with which the two sets of words are assembled because they have equivalent phonotactic probability. However, the results of Experiment 4 clearly showed that words with sparse onsets were responded to more quickly (and accurately) than words with dense onsets. It is, therefore, unclear how a feedforward model of speech production, such as WEAVER++ (Levelt et al., 1999), would account for the results of Experiment 4.

General Discussion

Experiments 1–4 examined the influence of phonotactic probability, neighborhood density, and onset density on speech production. Although these variables have been examined in studies of spoken word recognition (e.g., Vitevitch & Luce, 1998, 1999; Vitevitch, 2002a, 2002b), comparatively fewer studies have examined the influence of these variables on speech production. Further examination of these variables in speech production may provide additional and important insights regarding the constraints that govern models of speech production. For example, both feedforward and interactive models of speech production appeared to be able to account for the results of Experiments 1 and 2, in which words with high phonotactic probability were produced more quickly than words with low phonotactic probability. In the case of an interactive model of speech production, it was suggested that activation between word forms and phonological segments interacted to produce the effects of phonotactic probability. In the case of a feedforward model of speech production, we offered two possible accounts of the results in Experiments 1 and 2. One approach suggested that high probability stimuli were directly retrieved from the mental syllabary, whereas low probability stimuli were assembled segment by segment (based on the information in the phonological form). An alternative approach suggested that phonotactic information could be directly represented in the phonological segments.

In Experiment 3, we orthogonally manipulated phonotactic probability and neighborhood density. We again observed that words with high phonotactic probability were produced more quickly than words with low phonotactic probability. However, neighborhood density did not seem to interact or influence speech production in this experiment (cf. Vitevitch, 2002c; Vitevitch & Sommers, 2003). The interactive model of speech production again appealed to the same mechanism that accounted for the results of Experiments 1 and 2—interactive activation between word forms and phonological segments—to account for the results of Experiment 3. However, the syllabary-based approach offered as a possible account of Experiments 1 and 2 for a feedforward model of speech production did not appear as elegant an account of Experiment 3 as the approach that suggested that phonotactic probability was encoded in the segmental representations.

In Experiment 4, we manipulated a different variable, namely onset density (Vitevitch, 2002c). Words varied in the number of neighbors that shared the initial phoneme as the target word despite having the same overall number of neighbors, phonotactic probability, word frequency, neighborhood frequency, and the same initial phonemes in both conditions. We found that words with sparse onsets were produced more quickly than words with dense onsets. The results of previous research (e.g., Sevald & Dell, 1994) supported the hypothesis that the same mechanism in an interactive model that accounted for the results of Experiments 1–3 could also account for the results observed in Experiment 4. In the case of the feedforward model, neither the syllabary-assembly approach nor the encoding of phonotactic information in the segmental representations approach appeared to adequately account for the results of Experiment 4. The failure of these two approaches does not imply that an account of the results in Experiment 4 does not exist in the context of a feedforward model. Rather, we hope that these empirical observations will stimulate the necessary computer simulations and motivate modifications to models of speech production as appropriate.

The results of this set of experiments may also be accounted for in a variety of other ways in other models. Mixed models, like the restricted interaction account proposed by Rapp and Goldrick (2000), which incorporates feedforward activation from the semantic to the word-form level but which allows word forms and phonological segments to interact, would most likely be able to account for the results of the present set of experiments. Similarly, models such as node structure theory (e.g., James & Burke, 2000), which allows priming energy to spread in a bidirectional manner but which only allows activation to proceed in one direction, would also most likely be able to account for the results of these experiments. By examining variables such as phonotactic probability, neighborhood density, and onset density in spoken word recognition and speech production, we can gain a better understanding of the two systems independently as well as a better understanding of the interaction between the two systems.

References

Baars, B. J. (1992). A dozen competing-plans techniques for inducing predictable slips in speech and action. In B. J. Baars (Ed.), *Experimental slips and human error: Exploring the architecture of volition* (pp. 129–150). New York: Plenum Press.

Bond, Z. S. (1999). *Slips of the ear: Errors in the perception of casual conversation.* New York: Academic Press.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12,* 335–359.

Clark, H. H. (1976). Discussion of Wike and Church's comments. *Journal of Verbal Learning and Verbal Behavior, 15,* 257–266.

Cohen, J. (1976). Random means random. *Journal of Verbal Learning and Verbal Behavior, 15,* 261–262.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, & Computers, 25,* 257–271.

D'Amico, S., Bentrovato, S., Gasparini, M., Costabile, D., & Bates, E. (2002, July). *Timed picture naming in Italian-speaking children and adults: Differences between nouns and verbs.* Paper presented at the joint conference of the Ninth International Congress for the Study of Child Language and the Symposium on Research in Child Language Disorders, Madison, WI.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93,* 283–321.

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language, 27,* 124–142.

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes, 5,* 313–349.

Dell, G. S., Reed, K. D., Adams, D. R., & Meyer, A. S. (2000). Speech errors, phonotactic constraints, and implicit learning: A study of experience in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1355–1367.

Fay, D., & Cutler, A. (1977). Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry, 8,* 505–520.

Forsythe, A., Sheehy, N., & Sawey, M. (2003). Measuring icon complexity: An automated analysis. *Behavior Research Methods, Instruments, & Computers, 35,* 334–342.

Gordon, J. K. (2002). Phonological neighbourhood effects in aphasic speech errors: Spontaneous and structured contexts. *Brain and Language, 82,* 113–145.

Gordon, J. K., & Dell, G. S. (2001). Phonological neighbourhood effects: Evidence from aphasia and connectionist modeling. *Brain and Language, 79,* 21–23.

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

Harley, T. A. (1993). Phonological activation of semantic competitors during lexical access in speech production. *Language and Cognitive Processes, 8,* 291–309.

Harley, T. A., & Bown, H. E. (1998). What causes a tip-of-the-tongue state? Evidence for lexical neighbourhood effects in speech production. *British Journal of Psychology, 89,* 151–174.

Hino, Y., & Lupker, S. J. (2000). Effects of word frequency and spelling-to-sound regularity in naming with and without preceding lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 166–183.

James, L. E., & Burke, D. M. (2000). Phonological priming effects on word retrieval and tip-of-the-tongue experiences in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1378–1392.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 824–843.

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.

Keppel, G. (1976). Words as random variables. *Journal of Verbal Learning and Verbal Behavior, 15,* 263–265.

Kiefer, M. (2001). Perceptual and semantic sources of category-specific effects: Event-related potentials during picture and word categorization. *Memory & Cognition, 29,* 100–116.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University.

Landauer, T. K., & Streeter, L. A. (1973). Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. *Journal of Verbal Learning and Verbal Behavior, 12,* 119–131.

Laws, K. R., Leeson, V. C., & Gale, T. M. (2002). The effect of 'masking' on picture naming. *Cortex, 38,* 137–148.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1–38.

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

Appendix A

Stimulus Items and Characteristics From Experiments 1 and 2

Word	Condition	Fam.	Freq.	Density	NHF	Σ Phone.	Σ Bipho.	Nat./Art.
bat	High	7.0	1.25	33	1.46	0.19	0.011	Artifact
bell	High	7.0	1.27	22	1.48	0.19	0.011	Artifact
bear	High	7.0	1.93	21	1.74	0.20	0.010	Natural
bus	High	7.0	1.54	17	1.21	0.16	0.007	Artifact
chair	High	7.0	1.81	18	1.65	0.16	0.008	Artifact
doll	High	6.9	1.00	9	1.17	0.18	0.008	Artifact
cap	High	7.0	1.43	24	1.13	0.20	0.016	Artifact
cat	High	7.0	1.36	27	1.50	0.23	0.018	Natural
car	High	7.0	2.43	20	1.23	0.23	0.032	Artifact
pen	High	7.0	1.25	18	1.23	0.25	0.019	Artifact
pear	High	7.0	1.76	23	1.54	0.23	0.012	Natural
pig	High	7.0	0.90	13	1.17	0.19	0.008	Natural
sun	High	7.0	2.44	22	1.58	0.23	0.011	Natural
well	High	7.0	2.95	23	1.23	0.16	0.011	Artifact
ball	Low	7.0	2.04	20	1.42	0.14	0.004	Artifact
bed	Low	7.0	2.10	20	1.54	0.16	0.006	Artifact
boat	Low	7.0	1.85	26	1.34	0.16	0.006	Artifact
boot	Low	7.0	1.11	24	1.11	0.13	0.003	Artifact
chain	Low	7.0	1.69	12	1.49	0.13	0.003	Artifact
duck	Low	6.8	0.95	22	0.91	0.14	0.004	Natural
cake	Low	7.0	1.11	21	1.40	0.17	0.003	Artifact
kite	Low	7.0	1.00	16	1.81	0.19	0.003	Artifact
cup	Low	7.0	1.65	13	1.45	0.16	0.005	Artifact
peach	Low	7.0	0.47	19	1.22	0.12	0.002	Natural
purse	Low	7.0	1.14	14	1.23	0.18	0.006	Artifact
pipe	Low	7.0	1.30	12	0.94	0.15	0.003	Artifact
seal	Low	7.0	1.23	24	1.46	0.20	0.005	Natural
watch	Low	7.0	1.90	4	1.61	0.04	0.002	Artifact

Note. Fam. = familiarity; Freq. = frequency of occurrence ($\log_{10}$); NHF = neighborhood frequency ($\log_{10}$); Σ Phone. = sum of the phoneme probabilities; Σ Bipho. = sum of the biphone probabilities; Nat./Art. = classification as either a naturally occurring object or an inorganic artifact. The picture for *bat* was that of a baseball bat, not the flying mammal.

(*Appendixes continue*)

Appendix B

Stimulus Items and Characteristics From Experiment 3

Word	Density Condition	Phono. Condition	Fam.	Freq.	Density	NHF	Σ Phone.	Σ Bipho.	Nat./Art.
bun	Dense	High	6.7	1.00	30	1.29	0.18	0.009	Artifact
duck	Dense	High	6.7	0.95	25	0.81	0.14	0.004	Natural
gun	Dense	High	7.0	3.07	20	1.15	0.16	0.007	Artifact
heal	Dense	High	6.9	1.04	29	1.12	0.14	0.004	Natural
lamb	Dense	High	7.0	0.84	27	0.74	0.16	0.009	Natural
map	Dense	High	7.0	1.11	20	0.88	0.17	0.014	Artifact
moon	Dense	High	7.0	1.77	21	1.13	0.17	0.003	Natural
net	Dense	High	6.9	1.53	26	1.40	0.16	0.007	Artifact
pearl	Dense	High	7.0	0.95	21	0.98	0.18	0.004	Artifact
ring	Dense	High	7.0	1.69	23	1.04	0.15	0.020	Artifact
vine	Dense	High	7.0	0.60	21	1.04	0.15	0.007	Natural
bug	Dense	Low	7.0	0.60	26	0.80	0.10	0.004	Natural
dime	Dense	Low	7.0	0.60	20	0.86	0.13	0.004	Artifact
goat	Dense	Low	7.0	0.77	26	0.91	0.14	0.005	Natural
hug	Dense	Low	7.0	0.47	21	0.60	0.09	0.003	Natural
leaf	Dense	Low	7.0	1.07	25	0.89	0.08	0.003	Natural
mud	Dense	Low	7.0	1.50	20	0.62	0.13	0.004	Natural
mug	Dense	Low	7.0	1.00	21	0.48	0.11	0.005	Artifact
nail	Dense	Low	7.0	0.77	26	0.81	0.12	0.004	Artifact
peach	Dense	Low	7.0	0.47	22	1.05	0.12	0.002	Natural
rose	Dense	Low	6.8	1.93	26	1.06	0.11	0.003	Natural
vail	Dense	Low	6.2	1.17	24	0.86	0.12	0.004	Artifact
bib	Sparse	High	6.8	0.30	13	1.25	0.17	0.006	Artifact
doll	Sparse	High	6.9	1.00	16	0.87	0.18	0.008	Artifact
gas	Sparse	High	7.0	1.99	19	0.86	0.18	0.010	Artifact
hitch	Sparse	High	6.7	0.69	19	1.43	0.14	0.005	Artifact
lawn	Sparse	High	7.0	1.17	19	1.41	0.14	0.003	Natural
mop	Sparse	High	7.0	0.47	16	0.80	0.15	0.008	Artifact
mouse	Sparse	High	7.0	1.00	14	0.92	0.14	0.001	Natural
neck	Sparse	High	7.0	1.90	13	0.73	0.15	0.009	Natural
pig	Sparse	High	7.0	0.90	19	0.87	0.19	0.008	Natural
rib	Sparse	High	7.0	0.00	19	0.77	0.17	0.019	Natural
van	Sparse	High	7.0	1.50	12	1.80	0.19	0.016	Artifact
badge	Sparse	Low	6.9	0.69	13	1.03	0.14	0.006	Artifact
dove	Sparse	Low	7.0	0.60	16	0.63	0.12	0.002	Natural
ghoul	Sparse	Low	6.6	1.00	17	0.90	0.12	0.002	Artifact
hawk	Sparse	Low	7.0	1.14	18	0.50	0.10	0.002	Natural
look	Sparse	Low	7.0	2.60	17	1.21	0.09	0.001	Artifact
maze	Sparse	Low	7.0	0.77	19	1.15	0.10	0.003	Artifact
merge	Sparse	Low	6.9	1.00	11	0.65	0.09	0.002	Artifact
knife	Sparse	Low	6.7	1.88	8	1.45	0.07	0.002	Artifact
page	Sparse	Low	7.0	1.82	16	1.15	0.12	0.003	Artifact
roach	Sparse	Low	7.0	0.30	18	1.09	0.10	0.002	Natural
vase	Sparse	Low	7.0	0.60	16	1.47	0.13	0.004	Artifact

Note. Phono. = phonotactic; Fam. = familiarity; Freq. = frequency of occurrence ($\log_{10}$); NHF = neighborhood frequency ($\log_{10}$); Σ Phone. = sum of the phoneme probabilities; Σ Bipho. = sum of the biphone probabilities; Nat./Art. = classification as either a naturally occurring object or an inorganic artifact.

Appendix C

Stimulus Items and Characteristics From Experiment 4

Word	Condition	Fam.	Freq.	Density	NHF	Σ Phone.	Σ Bipho.	Nat./Art.
doll	Dense	6.9	1.00	16	0.87	0.18	0.008	Artifact
calf	Dense	6.5	1.04	19	0.86	0.19	0.013	Natural
cape	Dense	6.9	1.30	26	0.95	0.15	0.003	Artifact
curl	Dense	7.0	0.30	23	0.85	0.19	0.002	Artifact
leaf	Dense	7.0	1.07	25	0.89	0.08	0.003	Natural
math	Dense	7.0	0.60	15	1.23	0.14	0.011	Artifact
moss	Dense	6.7	0.95	15	0.93	0.15	0.001	Natural
mice	Dense	7.0	1.00	22	1.23	0.17	0.003	Natural
peach	Dense	7.0	0.47	22	1.05	0.12	0.002	Natural
rim	Dense	6.5	0.69	26	0.95	0.19	0.023	Artifact
seam	Dense	6.9	2.37	22	1.37	0.18	0.004	Artifact
tool	Dense	7.0	1.61	23	1.35	0.14	0.003	Artifact
wife	Dense	7.0	2.35	15	0.92	0.07	0.002	Natural
dot	Sparse	7.0	1.11	26	1.05	0.17	0.005	Artifact
kite	Sparse	7.0	1.00	20	1.52	0.19	0.004	Artifact
lip	Sparse	7.0	1.25	29	0.80	0.16	0.011	Natural
map	Sparse	7.0	1.11	20	0.88	0.17	0.014	Artifact
mop	Sparse	7.0	0.47	16	0.80	0.15	0.008	Artifact
mare	Sparse	6.7	1.20	23	1.70	0.20	0.013	Natural
mug	Sparse	7.0	1.00	21	0.48	0.11	0.005	Artifact
pear	Sparse	7.0	1.76	26	1.40	0.23	0.012	Natural
rock	Sparse	6.6	1.87	23	0.67	0.16	0.004	Artifact
sash	Sparse	6.5	0.47	20	0.63	0.18	0.006	Artifact
tail	Sparse	7.0	1.65	30	1.09	0.14	0.004	Artifact
wall	Sparse	7.0	2.20	23	1.32	0.11	0.005	Artifact
weep	Sparse	7.0	1.14	18	1.07	0.08	0.002	Artifact

Note. Fam. = familiarity; Freq. = frequency of occurrence ($\log_{10}$); NHF = neighborhood frequency ($\log_{10}$); Σ Phone. = sum of the phoneme probabilities; Σ Bipho. = sum of the biphone probabilities; Nat./Art. = classification as either a naturally occurring object or an inorganic artifact.

Received May 14, 2003
Revision received August 20, 2003
Accepted September 12, 2003 ■