The Influence of Phonological Similarity Neighborhoods on Speech Production

Michael S. Vitevitch
Department of Psychology, Indiana University.

Abstract

The influence of phonological similarity neighborhoods on the speed and accuracy of speech production was investigated with speech-error elicitation and picture-naming tasks. The results from 2 speech-error elicitation techniques—the spoonerisms of laboratory induced predisposition technique (B. J. Baars. 1992; B. J. Baars & M. T. Motley, 1974; M. T. Motley & B. J. Baars, 1976) and tongue twisters—showed that more errors were elicited for words with few similar sounding words (i.e., a sparse neighborhood) than for words with many similar sounding words (i.e., a dense neighborhood). The results from 3 picture-naming tasks showed that words with sparse neighborhoods were also named more slowly than words with dense neighborhoods. These findings demonstrate that multiple word forms are activated simultaneously and influence the speed and accuracy of speech production. The implications of these findings for current models of speech production are discussed.

Current models of spoken-word recognition treat as axiomatic the hypothesis that acoustic-phonetic input activates multiple phonological word forms that compete among each other, thereby affecting the speed and accuracy of lexical access during word recognition (e.g., Luce & Pisoni, 1998; Marslen-Wilson & Zwit-serlood, 1989; Norris, McQueen, & Cutler, 2000). In contrast, the influence of phonologically related words on the speed and accuracy of speech production is unclear. Evidence supports the hypothesis that words with similar forms compete with each other during speech production, as well as the hypothesis that formally similar words facilitate speech production. The present experiments attempted to better describe the nature of the activation among phonological word forms; do phonologically related representations compete among each other or facilitate processing at the word-form level during speech production?

During the retrieval of a phonological word form in speech production, phonologically similar words may block each other (Schacter, 1999; Woodworth, 1929) or compete with each other, as they do in models of spoken-word recognition. Using a tip-of-the-tongue (TOT) elicitation task, Jones (1989) presented definitions to participants and asked them to retrieve the word (i.e., the target) that fit the definition. Along with the definition, a prime that was semantically, phonologically, or both semantically and phonologically related to the target was presented. Jones (1989; see also Jones & Langford, 1987; Maylor, 1990) found that more TOT states were elicited when a phonologically related prime was presented after hearing the definition of the target. The increase in TOT states—or the decreased ability to retrieve the target word—in the context of a phonologically related prime suggests that phonologically related words compete with each other during speech production.

Work by Sevald and Dell (1994) also supports the hypothesis that formally related words compete during speech production. Sevald and Dell showed that speakers had slower production times to sequences of words with the same initial sounds (e. g., *cat, cab, can, cad*) than to sequences of words with different initial sounds (e. g., *cat, bat, mat, rat*). Together, these demonstrations of slower and less accurate speech production in the context of phonologically related words suggest competition among formally similar words during speech production.

Alternatively, phonologically similar words may facilitate the activation and retrieval of a lexical word form during speech production (e.g., A. S. Brown, 1991; Burke, MacKay, Worthley, & Wade, 1991). Meyer and Bock (1992) showed that the targets used by Jones (1989) differed across conditions in the susceptibility to TOT states. When targets with equal susceptibility to TOT states were used across conditions in a TOT-elicitation task, phonological primes did not interfere with the retrieval of the target word form; rather, phonological primes aided in, or facilitated, the retrieval of the target word form. The results of another TOT-elicitation task by James and Burke (2000) further support the hypothesis that phonologically related word forms facilitate retrieval. James and Burke presented participants with words like *indigent, abstract,* and *locate* and then presented the question "What word means to formally renounce a throne?" They elicited fewer TOT states when the target word, in this case *abdicate,* was preceded by phonologically related rather than unrelated words, suggesting facilitated retrieval of the phonologically similar target word.

Evidence from the cross-modal picture–word interference task also supports the idea that phonologically related words facilitate speech production (e.g., Costa & Sebastian-Galles, 1998; Jescheniak & Schriefers, 2001; Meyer, 1996; Schriefers, Meyer, & Levelt, 1990). For example, Jescheniak and Schriefers (2001) presented pictures that participants had to name while a word that was either phonologically related or unrelated to the picture was presented auditorily. Jescheniak and Schriefers found faster naming times when the auditorily presented words were phonologically related rather than unrelated to the to-be-named picture, suggesting that phonologically related words facilitate the process of speech production.

Note that the tasks used in the previous experiments relied on some form of sequential presentation of relevant stimuli, or priming. A word (presented visually or auditorily) was either related or unrelated to a subsequently to-be-produced item. Several studies have pointed out the limitations of priming methodologies (e.g., Bowles & Poon, 1985; Roediger, Neely, & Blaxton, 1983). Specifically, the relationship between the prime and the target may (consciously or unconsciously) induce task-specific strategies to use the prime as a cue to retrieve the target. The particular retrieval strategy that is induced may facilitate or inhibit the retrieval of the target and may not accurately reflect the strategy used during normal processing.

Rather than relying on priming methodologies, an alternative approach can be used to examine the influence of phonologically related words on speech production. Namely, words that vary in the number of formally related neighbors can be used as targets. The number of words that are phonologically similar to a target word is a variable that is commonly manipulated in studies of spoken-word recognition and is referred to as *neighborhood density* (e.g., Goldinger, Luce, & Pisoni, 1989; Luce & Pisoni, 1998; Vitevitch & Luce, 1998, 1999; Vitevitch, Luce, Pisoni, & Auer, 1999). A word with many similar sounding words has a *dense neighborhood,* whereas a word with few similar sounding words has a *sparse neighborhood*.

Previous studies examining neighborhood density in speech production have found evidence to suggest that phonologically similar words facilitate processing (Harley & Bown, 1998; Vitevitch, 1997; Vitevitch & Sommers, 2001). Vitevitch (1997) examined the neighborhood density characteristics of whole-word speech errors known as malapropisms (e.g., saying

octane instead of *octave*) that were collected by means of naturalistic observation (Fay & Cutler, 1977). The neighborhood density of the target and the error that was produced were compared with 10 samples of words of comparable length and syntactic class that were randomly sampled from a computer-readable version of *Webster's Pocket Dictionary,* which contains approximately 20,000 words (Nusbaum, Pisoni, & Davis, 1984). The results showed that targets and errors had fewer similar sounding words (i.e., sparser neighborhoods) than the words randomly sampled from the lexicon. It was hypothesized that words that have sparse neighborhoods do not receive sufficient activation to be accurately retrieved from the lexicon, resulting in a malapropism. In contrast, words with denser neighborhoods receive sufficient activation (from more phonological neighbors) and are, therefore, more likely to be accurately retrieved from the lexicon.

Similarly, Vitevitch and Sommers (2001; see also Harley & Bown, 1998), using the traditional TOT-elicitation task (i.e., no prime words were used; see R. Brown & McNeill, 1966), found that more TOT states were elicited for words with sparse rather than dense neighborhoods. As in Vitevitch (1997), it was hypothesized that words with many phonological neighbors receive sufficient amounts of activation from formally related neighbors to be completely retrieved from the lexicon. However, words with few phonological neighbors do not receive sufficient amounts of activation to be completely retrieved from the lexicon, resulting in a TOT state. Together, these studies support the hypothesis that phonologically related words facilitate the retrieval of, rather than compete with, target words during speech production.

Two points in the previous studies examining neighborhood density in speech production should be noted. First, Vitevitch (1997), Vitevitch and Sommers (2001), and Harley and Bown (1998) used methods that did not involve priming. That is, the influence of phonological similarity on speech production was not examined by manipulating the formal relationship between a word and a subsequently presented and to-be-named item. Rather, the ability of participants to retrieve words that had many neighbors was compared with the ability of participants to retrieve words that had fewer neighbors. Manipulating neighborhood density rather than the relationship between prime and target words may provide evidence regarding the influence of phonologically similar words on speech production that is less prone to task-specific strategies (e.g. Bowles & Poon, 1985; Roediger et al., 1983). Second, evidence from these studies examining neighborhood density in speech production suggests that phonologically related items facilitate speech production. To further examine the influence of phonologically related word forms, I manipulated neighborhood density in several experimental tasks in the present set of experiments. The results of these experiments converge on the idea that simultaneously activated word forms facilitate speech production, a finding that may be difficult for some models of speech production to account for.

Experiment 1

The spoonerisms of laboratory induced predisposition (SLIP) technique (Baars, 1992; Baars & Motley, 1974; Motley & Baars, 1976) was used to elicit phonological speech errors on words that varied in neighborhood density. The SLIP technique elicits phonological speech errors by activating two incompatible speech plans. The competition between the two incompatible speech plans increases the likelihood of making a speech error when prompted to produce a verbal response. Competition among speech plans is accomplished by instructing a participant 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, /p/, and the initial phoneme of the second word is, for example, /b/ (e.g., *push–big, pig–bull,* and *pin–ban*), strongly activating 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 /b/, and the initial phoneme of the second word is /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.*

Although other responses that differ from the intended word pair may also be produced, these responses are not counted as speech errors. Speech errors induced in elicitation tasks, such as the SLIP technique, are similar in kind to naturally occurring speech errors collected in various error corpora (Cutler, 1982; Ferber, 1991; Stemberger, 1992). Furthermore, the errors produced in the SLIP paradigm are not artifacts of proactive interference or confusions in short-term memory (cf. Motley, 1986, and Sinsabaugh & Fox, 1986). The SLIP technique is simply one of many competing-plans techniques (Baars, 1992; Bock, 1996) that have been used to examine phonological speech errors (Dell, 1984, 1986, 1990; Levitt & Healy, 1985; Shattuck-Hufnagel & Klatt, 1979; Stemberger & Treiman, 1986), syntactic representations (e.g., Bock, 1986), and idiom blends (Cutting & Bock, 1997) in speech production. The advantage of using such techniques includes the precise calculation, rather than estimation, of actual error probabilities (Motley & Baars, 1976). These techniques also allow for the control and manipulation of selected variables, such as word frequency, neighborhood density, and neighborhood frequency in the present experiment.

Dell (1988, 1990) and Stemberger and MacWhinney (1986) found more speech errors for rare rather than for common words in the English language, suggesting that word frequency affects speech production. Word frequency was manipulated in the present experiment in an attempt to replicate these results. *Neighborhood density* refers to the number of similar sounding neighbors a target word has and was manipulated to examine the influence that the simultaneous activation of multiple word candidates might have on speech production. Finally, *neighborhood frequency* refers to the mean frequency of occurrence of the neighbors (Luce & Pisoni, 1998). Luce and Pisoni (1998) have shown that this variable (along with density and frequency) influences the speed and accuracy of lexical retrieval during spoken-word recognition. Neighborhood frequency was manipulated in the present experiment to see if the frequency of similar sounding words also influenced lexical retrieval in speech production.

Method

Participants—In all the experiments reported, participants were native English speakers with normal or corrected-to-normal vision and no history of speech or hearing problems as determined by self-report. None of the participants who took part in a given experiment took part in any of the other experiments. In this experiment, 78 speakers from the State University of New York at Buffalo pool of introductory psychology students participated in partial fulfillment of a course requirement. The data from 2 participants were not included in the analyses because of technical problems that occurred during the experiment resulting in only part of the session being recorded.

Materials—The stimuli consisted of 112 consonant–vowel–consonant (CVC) words. The mean familiarity rating for the words was 6.84 on the basis of a 7-point subjective rating scale of familiarity that ranged from 1 (*don't know the word*) to 7 (*know the word and know its meaning*). All of the subjective familiarity ratings in the experiments presented here are based on Nusbaum, Pisoni, and Davis's (1984) study. The median value was used in all of the experiments to equally divide the words among the relevant conditions. In the present experiment, this resulted in eight orthogonal stimulus conditions (high vs. low frequency, dense vs. sparse neighborhood density, high vs. low neighborhood frequency) with 14 words per condition. Unless otherwise specified, all analyses were significant at $p < .05$. The initial consonants found in the eight conditions did not differ significantly across conditions, $\chi^2(126,$

$N = 112) = 146$. Words in the high-frequency conditions had a mean frequency of occurrence of 45.0 per million, and words in the low-frequency conditions had a mean frequency of occurrence of 3.0 per million, $F(1, 110) = 197$, $MSE = 380$. All of the frequency and neighborhood frequency counts in the experiments presented are based on Kuèera and Francis (1967).

As in the spoken-word recognition literature (e.g., Greenberg & Jenkins, 1964; Landauer & Streeter, 1973; Luce, 1986; Luce & Pisoni, 1998), neighborhood density was defined as the number of words that were similar to a target on the basis of the addition, deletion, or substitution of a single phoneme in the target item. For example, in the word *cat* [/kæt/], the words *scat* [/skæt/], *at* [/æt/], *hat* [/hæt/], *cut* [/kʌt/] and *cap,* [/kæp/], as well as other words found in the computer readable version of the *Webster's Pocket Dictionary* (with a familiarity rating of 6 or higher; Nusbaum, Pisoni, & Davis, 1984) would be considered neighbors. Familiarity ratings 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. The mean value for stimuli in the dense-neighborhood condition was 24.86 neighbors and in the sparse-neighborhood condition was 14.50 neighbors, $F(1, 110) = 176$, $MSE = 1,938$.

Neighborhood frequency is the mean frequency of the neighbors of the target word. Words in the high-neighborhood-frequency conditions had a mean frequency of occurrence of 19.0 per million, whereas words in the low-neighborhood-frequency conditions had a mean frequency of occurrence of 4.5 per million, $F(1, 110) = 274$, $MSE = 10$. The mean and standard errors of each variable for each condition are listed in Table 1.

The 14 stimulus words in each condition were grouped to form seven word pairs. Each word in the pair was similar in word frequency, neighborhood density, and neighborhood frequency and had minimal overlap of the consonants or vowels; the frequency with which overlap occurred did not differ across conditions ($F < 1$). Furthermore, when the initial consonants of each word in the pair were switched, either a real word or a pronounceable nonword was formed. There were no cases in which both words in the pair formed nonwords. Note that Baars, Motley, and MacKay (1975), among others, have suggested that there may be a bias in the speech production system to output lexical items and prohibit the output of nonlexical items. The data they used to support this claim, however, were based on elicited speech errors in which both words in the pair formed nonwords. It is unclear whether the so-called lexical bias affects word pairs in which only one item in the pair forms a nonword when switched. Furthermore, Dell (1986) found an interaction between lexical bias and output cue deadline. Specifically the lexical bias was present when participants were cued to produce a response at longer deadlines (700 and 1,000 ms), but it was not present when participants produced a response at a short deadline (500 ms). A deadline of 600 ms (relatively short compared with that in Dell's, 1986, study) was used in the present experiment to decrease the lexical bias in error output.

Additional words were paired to act as interference and filler word pairs. The interference word pairs were constructed according to the criteria described in Motley and Baars (1976): The first interference pair contained word-initial phonemes that were different from those of the targeted error, or spoonerism, but resembled the spoonerism in all other respects. The second interference pair differed from the spoonerism in the initial phoneme of the second word. The third interference pair differed from the spoonerism in the initial phonemes of both words. In all cases, the interference words were as similar in all other respects to the spoonerism as possible. By way of example, the target pair *name–life* had the following as interference pairs: *same-strife, lake-fife,* and *late-nine.* Across the two lists and across the eight conditions there were no differences in word frequency, neighborhood density, or neighborhood frequency for the interference pairs (all *F*s < 1). The filler word pairs contained words that were not part of

the stimulus or interference pairs. Filler pairs were randomly dispersed among the interference-stimulus groupings to complete the experimental list. In total, 486 word pairs were presented to participants: 56 stimulus pairs, 168 interference pairs, and 262 filler pairs.

Procedure—Participants were seated at a comfortable distance from a Macintosh Centris 650 computer with a 13-inch Macintosh monitor used for stimulus presentation. The participants were instructed to repeat to themselves each pair of words that appeared on the monitor and to be prepared to say some of the word pairs out loud when periodically cued by the computer.

Each participant received all of the word pairs with 112 pairs being cued for a response. Half of the cued-word pairs were the target stimuli (and are available from Michael S. Vitevitch on request, as are all the stimuli), and half were filler items that were cued to prevent participants from noticing a pattern in the stimuli. No word (or pair of words) was presented more than once.

Each participant received one of two lists that differed in the order of the stimulus (e.g., *name–life* on one list; *life–name* on the other). The word pairs were also presented in different pseudorandom orders (to maintain interference-stimulus groupings) on each list. Each word pair appeared in the center of the monitor for 900 ms. An interstimulus interval (consisting of a blank screen) of 100 ms separated each pair.

Periodically, participants were cued to repeat the previously presented pair of words by a string of eight question marks that appeared on the screen in the same location as the word pairs. The visual cue remained on the screen for 900 ms. A computer beep was presented 600 ms after the onset of the visual cue to encourage participants to respond rapidly. Participants were encouraged in the instructions to repeat the word pairs prior to the onset of the auditory cue if possible. The auditory cue was used only to encourage participants to repeat the word pairs quickly; it was not used as an exclusion criterion for responses.

After the offset of the visual cue to respond, 200 ms elapsed before another stimulus pair was presented. Responses from the participants were recorded on high quality audiotape, using a microphone (Shure 5755; Evanston, IL) and a tape recorder (Marantz, PMD221; Aurora, IL) to be scored at a later time. The number of filler pairs that occurred between each cued pair (whether the cued pair was a stimulus or filler item) in the practice session and in the experimental list ranged from 2 to 8 word pairs. The experimental session was preceded by a short practice session consisting of 20 pairs of words, 6 of which were cued for a verbal response.

Results

The recorded responses were examined for speech errors. Intrarater reliability (with at least 36 months passing between the initial and second coding) was very high (97.6%). Cases in which there was disagreement between the initial and second coding were resolved by an independent judge who was naive in regards to the nature of the experiment. Following the scoring conventions in Baars et al. (1975), a response was scored as a speech error if the response was either a complete or an incomplete reversal of the initial phonemes of the word pair. As per Baars et al. (1975), some responses to the stimulus items were not counted as speech errors. These responses included errors not involving initial consonants, producing a word not in the cued pair, failures to repeat any of the words in the cued pair, and errors in which participants misread or mispronounced a word (e.g., saying *lamb* for *lame* or *tone* for *ton*). These errors accounted for less than 1% of the responses made by participants. There was no difference across conditions or across lists (all *F*s < 1) for these other kinds of responses.

Because a set of highly controlled, nonrandomly selected stimuli that almost exhausted the pool of possible items was used in this and all the experiments that follow, only analyses of variance (ANOVAs) with participants as a random factor were conducted (J. Cohen, 1976; Hino & Lupker, 2000; Keppel, 1976; Raaijmakers, Schrijnemakers, & Gremmen, 1999; Smith, 1976; Wike & Church, 1976). Because there were no significant main effects or interactions for list/word order (all $Fs < 1$), all further analyses were collapsed across lists. The lack of a difference for list/word order also suggests that the order of the words (or the ordering of the initial segments in those words; see Levitt & Healy, 1985) was not a major factor in eliciting speech errors in this experiment.

The percentage of speech errors elicited for each condition is displayed in Table 2. A significant main effect of frequency was found, $F(1, 75) = 14.3$, $MSE = 3$; more speech errors were elicited for low-frequency word pairs (31.9%) than for high-frequency word pairs (16.4%). A significant main effect of neighborhood density was also found, $F(1, 75) = 15.3$, $MSE = 3$; more errors were elicited for word pairs from sparse neighborhoods (31.6%) than for word pairs from dense neighborhoods (16.8%). No difference in neighborhood frequency, nor any significant interactions were observed (all $Fs < 1$). The overall mean number of speech errors in this experiment (24.2%) was within the expected range of speech errors for this task (up to 30%; Motley & Baars, 1976).

Discussion

The results of the present experiment showed that more speech errors were elicited for word pairs with low rather than high frequency of occurrence, replicating analyses of speech-error corpora (Stemberger & MacWhinney, 1986) and studies using error-elicitation techniques (Dell, 1988, 1990). Although there was a significant effect of word frequency, no influence of neighborhood frequency was observed (cf. Vitevitch & Sommers, 2001).

More important, the results of the present experiment showed that more speech errors were elicited for word pairs with sparse rather than with dense neighborhoods. The influence of phonological similarity neighborhoods observed in the current experiment is consistent with the results of Vitevitch (1997) and Vitevitch and Sommers (2001; see also Harley & Bown, 1998). Recall that Vitevitch (1997) found that malapropisms tended to have sparser neighborhoods than comparable words randomly selected from the lexicon, and Vitevitch and Sommers (2001; see also Harley & Bown, 1998) found that more TOT states were elicited in college-age adults for words that had sparse rather than dense neighborhoods. Together, these results suggest that multiple word forms become activated during speech production. Furthermore, simultaneously active word forms facilitate processing in speech production rather than compete among each other. That is, lexical representations with many similar sounding neighbors receive a greater amount of activation than lexical representations with few similar sounding neighbors, supporting the more accurate retrieval of the target word form for words in dense neighborhoods. To further examine how the simultaneous activation of multiple word forms influences the accuracy of producing words, I selected another set of words varying in neighborhood density for use in a different phonological speech error elicitation task.

Experiment 2

Experiment 2 attempted to replicate the results of Experiment 1 with a different set of stimulus words and a different speech-error-elicitation task, namely the tongue twister task (Shattuck-Hufnagel, 1992). The tongue twister task, like the SLIP task, elicits speech errors from participants by activating competing speech plans (Baars, 1992; Bock, 1996). The tongue twister task differs from the SLIP task in that words are repeated rapidly rather than presented rapidly.

The stimuli used in Experiment 2 were even more rigorously controlled than were the stimuli in Experiment 1. Although there was no difference in the distribution of the initial segments across conditions in Experiment 1, the stimuli used in the present experiment had equal numbers of initial segments in each condition. Having equal numbers of words with the same initial phoneme in each condition rules out the possibility that the observed effects may be due to differences in the phonological segments in each condition. Furthermore, the stimuli in the present experiment had equivalent familiarity, word frequency, and neighborhood frequency, focusing only on the influences of neighborhood density in speech production.

Method

Participants—Twenty-eight speakers from the Indiana University pool of introductory psychology students participated in partial fulfillment of a course requirement.

Materials—Ten pairs of highly confusable target segments (from Experiment 2 of Shattuck-Hufnagel, 1992) were used to select CVC words for the tongue twisters in this experiment. Twenty tongue twisters, each containing four words of comparable neighborhood density, were created. Half of the tongue twisters consisted of words from sparse neighborhoods, and the other half consisted of words from dense neighborhoods.

Neighborhood density was computed as in Experiment 1. The median value was used to separate the words into stimuli with either dense or sparse neighborhoods. In the dense condition the mean number of neighbors was 23.9 words, and in the sparse condition the mean number of neighbors was 15.4 words. The difference between the dense and sparse neighborhood conditions was significant, $F(1, 78) = 143$, $MSE = 1,453$.

Although the stimuli differed in neighborhood density, the words in each condition did not differ in word frequency, $F(1, 78) = 2.0$, $MSE = 0.83$. The mean frequency of the items was 6.9 occurrences per million in the dense condition and 10.7 occurrences per million in the sparse condition. The two sets of words also did not differ in neighborhood frequency, $F(1, 78) = 1.2$, $MSE = 0.1$. The mean neighborhood frequency values of the items was 16.3 occurrences per million in the dense condition and 14.0 occurrences per million in the sparse condition. Finally, the words in each condition were also equivalent in subjective familiarity ratings, $F(1, 78) = 1.2$, $MSE = 0.2$ (dense condition: $M = 6.7$; sparse condition: $M = 6.8$).

Procedure—Participants were seated individually in a soundproof booth (IAC Model 402; Bronx, NY) equipped with a 13-inch monitor (Gateway 2000 Crystal Scan 1024 CRT; San Diego, CA) and a head-mounted microphone (Shure SM-98; Evanston, IL). A 486 PC computer presented a prompt ("Please repeat the following words six times in a row.") in the center of the monitor for 3 s. After the offset of the prompt, a tongue twister was randomly presented in the center of the monitor for 12 s. Participants were instructed to repeat the tongue twister six times as quickly as they could. The tongue twister remained on the monitor for the entire duration. Responses were recorded on a Sony (New York, NY) TCD-D8 tape recorder, using high quality audiotape, for later analysis. At the end of 12 s, the prompt was presented on the monitor, and a new trial began.

A practice session using five pseudo-tongue twisters, each composed of four randomly selected monosyllabic words not included in the stimulus set, familiarized the participants with the task. The responses from the practice session were not included in the final analyses.

Results

The recorded responses of each participant were examined for accuracy. Intrarater reliability (with at least 24 months passing between the initial and second coding) was very high (94.4%).

Cases in which there was disagreement between the initial and second coding were resolved by an independent judge who was naive in regards to the nature of the experiment. To maintain consistency with Experiment 1, I did not score the responses of the present experiment as perseverations, anticipations, or reversals of the initial phonemes but only as speech errors. For example, if the tongue twister was *peach balm bull pig* but the participant responded "beach balm bull big," (note that two words have incorrect initial phonemes) this repetition was scored as a single speech error.

Repetitions that were not correct but not counted as speech errors (less than 1% of the responses made by participants) included errors not involving initial consonants (e.g., *peep* for *peach*), substitutions of words other than those presented, failures to repeat one of the four words in the list, and errors in which participants misread a word (e.g., *made* for *mead* or *doze* for *dose*). There was no difference for these responses between conditions ($F < 1$).

For each density condition, there were 10 tongue twisters (repeated six times each). Thus, there were 60 opportunities to correctly repeat tongue twisters containing dense words and 60 opportunities to correctly repeat tongue twisters containing sparse words. A significant difference in the number of errors elicited between conditions was observed, $F(1, 27) = 16.8$, $MSE = 154$. More erroneous repetitions were elicited from tongue twisters containing words with sparse neighborhoods ($M = 12\%$, $SEM = 1$) than from tongue twisters containing words with dense neighborhoods ($M = 7\%$, $SEM = 0.9$). The overall mean percentage of speech errors in this experiment (9.5%) was comparable to the percentage of speech errors made in other tasks (Motley & Baars, 1976).

Discussion

In Experiment 2 a different error-elicitation task, different stimuli, and a sample of participants from a different university were used. The results, however, are the same as those in Experiment 1: More speech errors were elicited for words with sparse neighborhoods than for words with dense neighborhoods. The results of Experiment 2 are also consistent with the results of other studies investigating the role of neighborhood density in speech production (Harley & Bown, 1998; Vitevitch, 1997; Vitevitch & Sommers, 2001). These results suggest that multiple word forms become activated simultaneously during lexical access and influence the accuracy of speech production. Furthermore, the multiple word forms activated in memory facilitate rather than inhibit or compete during lexical access in speech production. Word forms with many neighbors receive a greater amount of activation than word forms with fewer neighbors, facilitating the accurate retrieval of words in dense neighborhoods. In Experiments 3–5, the influence of neighborhood density on the speed of lexical access during speech production was examined.

Experiment 3

A great deal has been learned from spontaneous and experimentally elicited speech errors. Indeed, many models of speech production were developed to account for speech-error data (e.g., Dell, 1986, 1988; Fay & Cutler, 1977; MacKay, 1987; Shattuck-Hufnagel, 1979). However, Levelt, Roelofs, and Meyer (1999) have argued that models of lexical access have always been conceived as process models of normal speech production. Their ultimate test … cannot lie in how they account for infrequent derailments of the process but rather must lie in how they deal with the normal process itself. RT studies, of object naming in particular, can bring us much closer to this ideal … [because] … object naming is a normal, everyday activity … [and] … reaction time measurement is still an ideal procedure for analyzing the time course of a mental process, (p. 2)

To meet the challenge that Levelt et al. (1999) have set for speech production research, I used an object-naming task (also known as a picture-naming task; Oldfield & Wingfield, 1965) in Experiment 3 to examine how monosyllabic words varying in neighborhood density influence the speed of lexical access during speech production. Given the facilitative effects of neighborhood density observed in the previous experiments, it was predicted that participants would more quickly name pictures illustrating words from dense rather than sparse neighborhoods.

Method

Participants—Thirty-four participants from the same population sampled in Experiment 2 took part in this experiment.

Materials—Forty-eight line drawings (Snodgrass & Vanderwart, 1980), half of which illustrated words from sparse neighborhoods and the other half of which illustrated words from dense neighborhoods, were used as stimuli in the present experiment. The words from sparse neighborhoods had significantly fewer neighbors ($M = 6.8$ words) than the words from dense neighborhoods ($M = 19.4$ words), $F(1, 46) = 107$, $MSE = 1,887$.

Although the difference in neighborhood density of the two conditions was significantly different, the differences in familiarity ratings, word frequency, and neighborhood frequency were not, all $Fs(1, 46) < 1$. Words from sparse neighborhoods had a mean familiarity rating of 6.9, a mean frequency of 38.5 occurrences per million, and a mean neighborhood frequency of 16.5 occurrences per million. Words from dense neighborhoods had a mean familiarity rating of 6.9, a mean frequency of 24.2 occurrences per million, and a mean neighborhood frequency of 17.6 occurrences per million. There was also no difference in the distribution of the initial phonemes used in each set of stimulus words, $\chi^2(13, N = 48) = 9.27$.

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 a Macintosh Quadra 950, with a 17-inch Macintosh monitor, that was running PsyScope 1.2.2 (J. D. Cohen, MacWhinney, Flatt, & Provost, 1993), which controlled stimulus randomization and presentation and the collection of response latencies. A headphone-mounted microphone (Beyer-Dynamic DT109, Heilbronn, Germany) 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 48 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

The tape-recorded responses of each participant were scored for accuracy. Only accurate responses were included in the repeated measures ANOVA for response latency. Errors included responses that were words other than the given label (e.g., responding with "sofa" instead of "couch") and improperly triggering the voice key (e.g., by coughing or saying "uh"). A significant main effect of neighborhood density was found, $F(1, 33) = 8.3$, $MSE = 8,768$. Participants responded to words from dense neighborhoods more quickly (716 ms, $SEM = 9$) than to words from sparse neighborhoods (739 ms, $SEM = 11$). There was no difference in error rates between the two sets of words ($F < 1$), suggesting that participants did not sacrifice speed for accuracy in making their responses. Words from dense neighborhoods were correctly

responded to 94.4% of the time, and words from sparse neighborhoods were correctly responded to 94.0% of the time.

Discussion

The results of the present experiment show that words from dense neighborhoods are produced more quickly than words from sparse neighborhoods, suggesting that multiple word forms do become simultaneously activated in memory and do influence the speed in addition to the accuracy of speech production. The results of the present experiment further suggest that multiple word forms activated simultaneously facilitate speech production.

The results of Experiments 1–3 suggest that many phonological neighbors facilitate the accurate and rapid retrieval of word forms. How might a model of speech production account for the facilitative effects of simultaneously activated phonologically related words on speech production? Current models of speech production can be generally classified into one of two types of models: interactive and feedforward models. An example of an interactive model of speech production is described in Dell (1986), and an example of a feedforward model of speech production is described in Levelt et al. (1999).

In Dell's (1986) interactive model of speech production (indeed, in most models of speech production) there are no lateral connections between representations within a level. Without lateral connections between similar sounding word forms, an interactive model of speech production can still account for the facilitative effects of neighborhood density in the following way. When the representation of a word form (*cat*) is partially activated by semantic information, the word form will partially activate the phonological nodes that constitute it (/k/, /æ/, /t/). (Note that in an interactive model other word forms may be partially activated by semantic information, but for ease of explication we only follow the activation of *cat*.) The activated phonological nodes (/k/, /æ/, /t/) will feed activation back to the word-form level to all the word forms that contain those phonemes (e.g., *hat, cut, cap*). Those phonologically related word forms will in turn send activation back down to the phonological nodes, thereby increasing the activation of those shared phonological nodes. The activation that is sent to the phonological nodes from similar sounding word forms will further activate those phonological nodes, which will in turn increase the activation of the target word that is composed of those phonemes. The higher levels of activation that the target word receives from similar sounding words via the shared phonological nodes will increase the probability that the target word (being the highest activated representation) will be selected.

Thus, in an interactive model of speech production, a target word with many phonological neighbors (i.e., a dense neighborhood) will receive activation from many similar sounding words via the shared phonological nodes. A target word with fewer neighbors (i.e., a sparse neighborhood) will receive activation from few similar sounding words via the shared phonological nodes. The difference in the number of similar words contributing to the activation that is sent to the target via the phonological nodes results in words with dense neighborhoods being produced faster and more accurately than words with sparse neighborhoods.

Levelt et al. (1999) described a different model of speech production, WEAVER++, that has a strictly feedforward architecture. That is, activation at the word-form level cannot spread "backward" to influence the activation of a lemma, nor can activation among phonological segments spread "backward" to influence the activation of word forms. (The only "feedback" in WEAVER++ is indirectly through the speech comprehension system, which is not considered feedback in the traditional sense.) Like most models of speech production, there are no lateral connections between representations within a level. Levelt et al. described how their model could—without lateral inhibitory connections—account for inhibitory effects of

phonologically similar words observed in some speech production tasks (e.g., Sevald & Dell, 1994). The mechanism they described involved the weighting of recently activated syllable nodes in the (Luce choice) decision rule. This resulted in inhibitory effects on subsequently produced words if they had similar syllable nodes. Note, however, that this weighting mechanism accounts only for inhibitory effects on subsequently presented words. This mechanism does not address the issue investigated in the present set of experiments, namely the influence of phonologically related words that are simultaneously activated during the retrieval and production of isolated words.

In section 5.2.1, Levelt et al. (1999) also discussed how WEAVER++ accounts for some facilitative effects observed in the literature. The facilitative effects they discussed, however, are facilitative effects among words that are semantically related. It is unclear whether the mechanisms described in section 5.2.1 of Levelt et al. would also apply to words that are phonologically related. Furthermore, given the strictly feedforward architecture of WEAVER ++ and the constraint that only selected lemmas become phonologically activated (Levelt et al., 1991), it is unclear if multiple word forms that are phonologically related can even become activated simultaneously. Levelt et al. (see also Roelofs, 1992) did suggest that multiple word forms may be activated in situations in which two (or presumably more) lemmas are equally activated and selected. However, given the arbitrary relationship between meaning and sound (e.g., Saussure, 1916/1966), it is unlikely that these semantically related representations would also be phonologically related (e.g., *sofa* and *couch*). In its present form, it is not clear that the strictly feedforward model, WEAVER++, can account for the facilitative effects of simultaneously activated and phonologically related words observed in Experiments 1–3.

WEAVER++ might be able to account for the effects observed in Experiments 1–3 if (a) the effects observed in Experiments 1–3 were due to phoneme frequency (i.e., *phonotactic probability*) rather than neighborhood density, and (b) the phonological nodes in the model were sensitive to the frequency with which those phonemes occur (nota bene, in its present form, frequency is encoded only at the word-form level in WEAVER++).

Taking each of these points in turn, phonotactic probability refers to the frequency that a particular segment or sequence of segments occurs in a given position in a word or syllable (Vitevitch, Luce, Charles-Luce, & Kemmerer, 1997). Vitevitch et al. (1999) found a positive correlation between phonotactic probability and neighborhood density. Common segments and sequences—those with high phonotactic probability—tend to be segments and sequences that are found in many words. Conversely, patterns with low-probability phonotactics typically occur in words with sparse phonological neighborhoods. Work by Dell, Reed, Adams, & Meyer (2000) suggested that the frequency with which segments occur in the language (and within the context of the experiment) influences the frequency of occurrence of certain errors elicited experimentally. It is, therefore, possible that the results of Experiments 1–3 (which did not control phonotactic probability) were the result of the difference in frequency among the phonological segments and sequences of segments, rather than the difference in the number of similar sounding words. If the results of Experiments 1–3 were indeed due to frequency differences among the phonological segments, WEAVER++ might be able to account for the present findings if it is modified to weight the activation of phonological segments as a function of their frequency of occurrence.

Given the role that phonological segments play in the interactive account of the present results, it is also important to rule out the possibility that the observed effects are due solely to activity at the level of phonological segments. Recall that in the interactive account activation spread from the partially activated word node via phonological segments to phonologically related word forms (and back again to increase the activation of the target). If the observed effects are due solely to activity at the level of phonological segments, then an interactive model may not

be required to account for the observed effects. However, if neighborhood density effects are still observed when the frequency of the phonological segments (i.e., phonotactic probability) is controlled, then only a model that allows activation to feedback from representations of phonological segments to representations of word forms can account for these findings. To better determine the locus of the neighborhood-density effect in speech production, and to adjudicate between an interactive and feedforward account of the findings, the picture-naming task was used in Experiments 4 and 5.

Experiment 4

In the present experiment the picture-naming task was used with stimuli that had equivalent phonotactic probability, word frequency, and neighborhood frequency but different neighborhood density, to ascertain whether the effects observed in Experiments 1–3 were due to neighborhood density. Given the language-wide correlation between phonotactic probability and neighborhood density and the effects of phonotactic constraints on speech production demonstrated by Dell et al. (2000; see also Motley & Baars, 1975), it is possible that the effects observed in Experiments 1–3 were due to differences in phonotactic probability and not due to differences in the number of words simultaneously activated in memory. To rule out the possibility that phonotactic probabilities were the source of the observed effects, this variable was controlled in the following experiments. Furthermore, only CVC words with the same initial segments were used as stimuli in each condition. Given that the same initial segments and the overall frequency of occurrence of the segments in the words are equivalent, no difference between the two conditions should be observed if differences among the phonological segments that constitute the words are the source of the effects observed in Experiments 1–3. Alternatively, if a difference in the number of simultaneously activated similar sounding words (i.e., neighborhood density) is the locus of the effects observed in Experiments 1–3, then effects of neighborhood density should be observed in the present experiment.

Method

Participants—Twenty-five participants from the same population sampled in Experiment 2 took part in this experiment.

Materials—Forty-eight monosyllabic words with a CVC syllable structure were used as labels for line drawings selected from Snodgrass and Vanderwart (1980) and Cycowicz, Friedman, Rothstein, and Snodgrass (1997). All lexical characteristics were assessed in the same manner as in Experiments 1–3. The words from sparse neighborhoods had a mean density of 11.72 neighbors, and the words from dense neighborhoods had a mean density of 21.38 neighbors. This difference was significant, $F(1, 46) = 103$, $MSE = 1.102$.

Although the difference in neighborhood density of the two conditions was significantly different, the differences in familiarity ratings, word frequency, and neighborhood frequency were not, Fs (1, 46) < 1. Words from sparse neighborhoods had a mean familiarity rating of 6.9, a mean frequency of 11.5 occurrences per million, and a mean neighborhood frequency of 21.5 occurrences per million. Words from dense neighborhoods had a mean familiarity rating of 6.9, a mean frequency value of 11.0 occurrences per million, and a mean neighborhood frequency of 20.5 occurrences per million.

To control the phonological segments used in the stimuli, I ensured that equal numbers of words in each stimulus condition contained the same initial phonemes. Furthermore, the phonotactic probabilities of the words in the two conditions were also equivalent. Phonotactic probability was calculated with the same two measures that have been used extensively in other studies of phonotactic probability (e.g., Jusczyk, Luce, & Charles-Luce, 1994; 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. The average probabilities for the segments in the given word positions for the items with sparse neighborhoods was .144, and for the items with dense neighborhoods it was .151, $F(1, 46) < 1$. The average probabilities for the biphones for the items with sparse neighborhoods was .005, and for the items with dense neighborhoods it was .006, $F(1, 46) < 1$.

Procedure—The equipment used in this experiment was the same as that used in Experiment 3.

Results

The tape recorded responses of each participant were scored for accuracy using the same criteria and type of analysis used in Experiment 3. A significant main effect of neighborhood density was found, $F(1, 24) = 15.9$, $MSE = 7,662$ such that words from dense neighborhoods were responded to more quickly (795 ms, $SEM = 12$) than were words from sparse neighborhoods (820 ms, $SEM = 12$). There was no difference in 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 from dense neighborhoods were correctly responded to 84.0% of the time, and words from sparse neighborhoods were correctly responded to 85.8% of the time. Note that the responses in the present experiment were overall less accurate that those in Experiment 3. A number of factors—including different stimuli and different participants—may account for the difference in error rates between the two experiments. What is important, however, is that there was no speed-accuracy trade off in the present results.

Discussion

The results of the present experiment again showed that words from dense neighborhoods were produced more quickly than words from sparse neighborhoods. The observation of a neighborhood-density effect in the present picture-naming experiment is important not only because it replicated the facilitative neighborhood density effect found in Experiment 3 with a different set of words and different participants but because the words in the present experiment were equivalent in phonotactic probability. In addition to having the same number of words in each condition with the same initial segments, the phonological segments that composed the words used in the present experiment were equivalent in their positional frequency and biphone frequency. By using words equivalent in these two measures of phonotactic probability, it was possible to rule out the possibility that the effects observed in the previous experiments were due to differences in the phonological segments that composed the words rather than the number of words simultaneously activated in memory. The results of the present experiment further suggest that multiple word forms become simultaneously activated during speech production and that these phonologically related word forms facilitate lexical access in speech production.

Experiment 5

The results of Experiment 4 clearly rule out the possibility that differences among the phonological segments composing the words in each condition were the only source for the effects observed in the previous experiments. The results of that experiment do not, however, rule out the possibility that words with dense neighborhoods were simply easier to articulate than words with sparse neighborhoods. That is, the previously observed results may not be due to differences in the time it takes to retrieve from the lexicon words varying in neighborhood density. Rather, the differences may be due to differences in the ease with which the

musculature involved in producing words varying in neighborhood density can be moved. To evaluate this possibility a modified picture-naming task was used.

The standard picture-naming task was modified in the following manner. Instead of using a voice key triggered by a vocal response to measure the response time, a buttonpress was used to collect response times. Participants were instructed to view the picture presented on the computer screen and retrieve the word used to label that picture. Participants were further instructed to press a button on a response box as soon as they retrieved the correct word and then say the name of the object out loud so that the accuracy of the response could be assessed. Cutler, Sebastian-Galles, Soler-Vilageliu, and van Ooijen (2000) and McQueen (1998), for example, have used similar modifications to standard word-recognition and word-segmentation tasks to show that the effects they observed were due to differences in lexical access and not to differences in articulation.

We hypothesized that if the results of Experiments 3 and 4 were due to differences in the ease of articulation between dense and sparse words, then no difference in response times should be observed in the present experiment because the speed with which articulation is initiated is not being measured. However, if the number of simultaneously activated neighbors does influence the speed with which lexical access occurs during speech production, then words with dense neighborhoods should still be responded to more quickly than words with sparse neighborhoods.

Method

Participants—Twenty-five participants from the same population sampled in Experiment 2 took part in this experiment.

Materials—The same stimuli used in Experiment 4 were used in the present experiment.

Procedure—The procedure and equipment used in Experiments 3 and 4 were also used in the present experiment with the following exception. Instead of using the voice key to measure reaction time, participants pressed a button on a response box with their dominant hand and then said the name of the object aloud so that the response could be scored for accuracy.

Results

The author scored the tape recorded responses of each participant for accuracy using the same criteria used in Experiments 3 and 4. In addition, responses were counted as incorrect if the participant responded with the name of the object before pressing the response button or failed to produce the name of the object after pressing the button. Repeated measures ANOVAs showed a significant main effect of neighborhood density, $F(1, 24) = 7.9$, $MSE = 5,640$. Participants responded more quickly to words from dense (662 ms, $SEM = 22$) rather than sparse neighborhoods (683 ms, $SEM = 23$). There was no difference in error rates between the two sets of words ($F < 1$), suggesting that participants did not sacrifice speed for accuracy in making their responses. Words from dense neighborhoods were correctly responded to 89.2% of the time and words from sparse neighborhoods were correctly responded to 89.7% of the time. As in Experiment 4, the responses in the present experiment were overall less accurate than those in Experiment 3 (but more accurate than those in Experiment 4). A number of factors —including the use of slightly different tasks and different participants—may account for the difference in error rates among the experiments. The more important thing to note about the accuracy results is that there was no speed-accuracy trade off.

Discussion

The results from the modified picture-naming task used in the present experiment showed that words from dense neighborhoods were responded to more quickly than were words from sparse neighborhoods. This result replicated the neighborhood density effect observed in Experiments 3 and 4 with a different sample of participants and a slightly different task. Replication of the neighborhood-density effect observed in Experiment 4 was especially important because the stimuli used in that experiment as well as the present experiment varied in neighborhood density but were controlled for the initial phonological segments, familiarity, word frequency, neighborhood frequency, and phonotactic probability of the stimuli. Keeping phonotactic probability constant between the two conditions was essential for ruling out the possibility that the effects observed were due only to differences in the phonological segments that made up the words, a variable closely related to neighborhood density, rather than to neighborhood density itself.

The results of the present experiment also ruled out the possibility that the effects observed in the previous experiments reported here were due to differences among the stimuli in the ease of articulation. The present experiment used a buttonpress instead of a voice key to measure response time. Observing an effect of neighborhood density with this modified task further suggests that differences in articulation between the two types of words were not the source of the differences in response times found in Experiments 3–5 (Cutler et al., 2000; McQueen, 1998). The results of the present experiment more strongly suggest that multiple word forms that are phonologically related are simultaneously activated and facilitate the retrieval of words during speech production.

General Discussion

The facilitative effects of simultaneously activated phonologically related word forms on lexical access in speech production that were demonstrated by means of both accuracy rates (see also Harley & Bown, 1998; Vitevitch, 1997; Vitevitch & Sommers, 2001) and response times (cf. Jescheniak & Levelt, 1994) in the present set of experiments are important for several reasons. First, manipulating the number of phonological neighbors for words presented in isolation provides a clearer picture of how phonologically related words affect lexical access during speech production than do experiments that rely on priming methods. Tasks that rely on priming or the sequential presentation of words and that vary the relationship between a presented item and a subsequently to-be-produced item are potentially prone to task-specific strategies that may not reflect normal processing (Bowles & Poon, 1985; Roediger, Neely, & Blaxton, 1983). Although all experimental tasks are prone to various task-specific strategies, including the tasks in the present set of experiments, the consistent findings across different words, different task demands, different dependent variables, and different participants clearly speaks to the reliability of facilitative effects of phonologically related words in speech production.

The consistent results of the present set of experiments also localized the source of these effects in the speech-production process. The modified picture-naming task in Experiment 5 showed that the observed effects were more likely the result of the lexical retrieval process than the result of the articulatory processes involved in speech production. Furthermore, the stimuli used in Experiments 4 and 5 that controlled phonotactic probability ruled out the possibility that the observed effects were due solely to processes at the level of phonological segments. Rather, the results of these and other experiments manipulating neighborhood density in speech production (Vitevitch, 1997; Vitevitch & Sommers, 2001) suggest that phonologically related word forms facilitate the activation of target word forms by their interaction with shared constituent phonological segments. That is, the observed effects do not appear to be the result

of processes among just word forms or among just phonological segments but are the result of word forms and phonological segments interacting with each another.

Although the results and conclusions of this set of experiments may appear incongruent with those of other studies (e.g., James & Burke, 2000; Sevald & Dell, 1994; Yaniv, Meyer, Gordon, Huff, & Sevald, 1990), they, in fact, are consistent within a broader view of the speech-production system. For example, James and Burke (2000) found fewer TOT states when targets were primed with words that shared phonological segments with the target (e.g., *pellet* primed the phonological segments /ɛl/ in *velcro*). They suggested that the phonologically related primes served to strengthen the connections between lexical items and their constituent phonological segments, making it easier for the target word, which shared many of these phonological segments, to be retrieved.

The findings of the present experiments, which examined simultaneous rather than sequential activation as in James and Burke (2000), can easily be accommodated by the same model used to account for the findings of James and Burke, namely, node structure theory (NST; MacKay, 1987). In NST there are bidirectional connections between phonological segments and lexical representations. (To be more precise, phonological segments connect to syllable nodes in NST, however, for ease of explication I assume some form of isomorphism between the syllable and lexical nodes for monosyllabic words like those used in the present experiments.) The bidirectional connections between levels of representation enable the model to account for the results of the present experiments with the same interactive mechanism described earlier. Thus, the same model that James and Burke used to account for results of sequential activation in their experiments can also account for the results of simultaneous activation in the present experiments.

The results of Sevald and Dell (1994; cf. Sevald, Dell, & Cole, 1995) or Yaniv et al. (1990) are also consistent with the results of the present experiment when viewed from a broader context in the speech-production system. Specifically, Sevald and Dell (1994) found competitive effects for the production of sequences of words with the same initial sounds (e.g., *cat, cab, can, cad* vs. *cat, bat, mat, rat),* suggesting that there was competition among phonemes for placement in the representation of the word frame. Yaniv et al. (1990) also found inhibitory effects when the vowels in pairs of CVC words were similar, suggesting that a lateral inhibitory mechanism may modulate the motor programming of vowels during speech production. Although the results of these studies propose inhibition-competition between similar representations, rather than facilitation as in the present set of experiments, these studies proposed such processes at different levels of representation (word frame and motor programming levels) than that investigated in the present experiments. Vitevitch and Luce (1998, 1999; see also Pitt & Samuel, 1995) found evidence for facilitation among phonological segments and competition among word forms in studies of spoken word recognition, suggesting that different processes may operate at different levels of representation. The speech-production system may also operate with different processes at different levels of representation. From this broader perspective, the present results and those of Sevald and Dell (1994) and Yaniv et al. (1990) are not at odds but serve to more precisely describe the speech production system.

The results of the present experiments are also consistent with interactive models of speech production. In discussing the work of James and Burke (2000) it was noted that the results of the present experiments were consistent with the predictions of NST, an interactive model of speech production. Furthermore, simulations by Gordon and Dell (2001) have produced facilitative effects of phonologically related neighbors in an interactive model of speech production using normal processing parameters.

The results of the present experiments are not, however, easily accounted for by strictly feedforward models (e.g., WEAVER++; Levelt et al., 1999). As discussed, it is unclear how multiple-word forms that are phonologically related can become simultaneously activated in the current instantiation of WEAVER++. In addition, the results of Experiment 4 and 5 ruled out the possibilities that the observed effects were due to differences in the phonological segments or to articulatory processing. These results further constrain the modifications that could be made to WEAVER++ to allow it to account for the data reported here.

Finally, the results of the present experiments investigating speech production demonstrated a facilitative effect of neighborhood density, which contrasts with the competitive effects of neighborhood density typically observed in investigations of spoken-word recognition (e.g., Luce & Pisoni, 1998). This finding may further guide modeling efforts in speech production and speech perception, especially those efforts that attempt to model the interface between the two processes (e.g., NST; MacKay, 1987).

References

Baars, BJ. A dozen competing-plans techniques for inducing predictable slips in speech and action. In: Baars, BJ., editor. Experimental slips and human error: Exploring the architecture of volition. New York: Plenum Press; 1992. p. 129-150.

Baars BJ, Motley MT. Spoonerisms: Experimental elicitation of human speech errors. Catalog of Selected Documents in Psychology 1974;4:118.Abstract obtained from Journal Supplement Abstract Service

Baars BJ, Motley MT, MacKay DG. Output editing for lexical status in artificially elicited slips of the tongue. Journal of Verbal Learning and Verbal Behavior 1975;14:382–391.

Bock JK. Syntactic persistence in language production. Cognitive Psychology 1986;18:355–387.

Bock JK. Language production: Methods and methodologies. Psychonomic Bulletin & Review 1996;3:395–421.

Bowles NL, Poon LW. Effects of priming in word retrieval. Journal of Experimental Psychology: Learning, Memory, and Cognition 1985;11:272–283.

Brown AS. A review of the tip-of-the-tongue experience. Psychological Bulletin 1991;109:204–223. [PubMed: 2034750]

Brown R, McNeill D. The "tip of the tongue" phenomenon. Journal of Verbal Learning and Verbal Behavior 1966;5:325–337.

Burke DM, MacKay DG, Worthley JS, Wade E. On the tip of the tongue: What causes word finding failures in young and older adults? Journal of Memory and Language 1991;30:542–579.

Cohen J. Random means random. Journal of Verbal Learning and Verbal Behavior 1976;15:261–262.

Cohen JD, MacWhinney B, Flatt M, Provost J. PsyScope: A new graphic interactive environment for designing psychology experiments. Behavioral Research Methods, Instruments, and Computers 1993;25:251–271.

Costa A, Sebastian-Galles N. Abstract phonological structure in language production: Evidence from Spanish. Journal of Experimental Psychology: Learning, Memory, and Cognition 1998;24:886–903.

Cutler, A. The reliability of speech error data. In: Cutler, A., editor. Slips of the tongue and language production. Berlin: Walter de Gruyter/Mouton; 1982. p. 7-28.

Cutler A, Sebastian-Galles N, Soler-Vilageliu O, van Ooijen B. Constraints of vowel and consonants on lexical selection: Cross-linguistic comparisons. Memory & Cognition 2000;28:746–755.

Cutting JC, Bock K. That's the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends. Memory & Cognition 1997;25:57–71.

Cycowicz YM, Friedman D, Rothstein M, Snodgrass JG. Picture naming by young children: Norms for name agreement, familiarity, and visual complexity. Journal of Experimental Child Psychology 1997;65:171–237. [PubMed: 9169209]

Dell GS. The representation of serial order in speech: Evidence from the repeated phoneme effect in speech errors. Journal of Experimental Psychology: Learning, Memory, and Cognition 1984;10:222–233.

Dell GS. A spreading-activation theory of retrieval in sentence production. Psychological Review 1986;93:283–321. [PubMed: 3749399]

Dell GS. The retrieval of phonological forms in production: Tests of predictions from a connectionist model. Journal of Memory and Language 1988;27:124–142.

Dell GS. Effects of frequency and vocabulary type on phonological speech errors. Language and Cognitive Processes 1990;5:313–349.

Dell GS, Reed KD, Adams DR, Meyer AS. Speech errors, phonotactic constraints, and implicit learning: A study of experience in language production. Journal of Experimental Psychology: Learning, Memory, and Cognition 2000;26:1355–1367.

Fay D, Cutler A. Malapropisms and the structure of the mental lexicon. Linguistic Inquiry 1977;8:505–520.

Ferber R. Slip of the tongue or slip of the ear? On the perception and transcription of naturalistic slips of the tongue. Journal of Psycholinguistic Research 1991;20:105–122. [PubMed: 2066909]

Goldinger SD, Luce PA, Pisoni DB. Priming lexical neighbors of spoken words: Effects of competition and inhibition. Journal of Memory and Language 1989;28:501–518.

Gordon JK, Dell GS. Phonological neighbourhood effects: Evidence from aphasia and connectionist modeling. Brain and Language 2001;79:21–23.

Greenberg JH, Jenkins JJ. Studies in the psychological correlates of the sound system of American English. Word 1964;20:157–177.

Harley TA, Bown HE. What causes a tip-of-the-tongue state? Evidence for lexical neighbourhood effects in speech production. British Journal of Psychology 1998;89:151–174.

Hino Y, Lupker SJ. 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 2000;26:166–183. [PubMed: 10696612]

James LE, Burke DM. 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 2000;26:1378–1392.

Jescheniak JD, Levelt WJM. Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. Journal of Experimental Psychology: Learning, Memory, and Cognition 1994;20:824–843.

Jescheniak JD, Schriefers H. Priming effects from phonologically related distractors in picture-word interference. Quarterly Journal of Experimental Psychology: Human Experimental Psychology 2001;54(A):371–382.

Jones GV. Back to Woodworth: Role of interlopers in the tip-of-the-tongue phenomenon. Memory & Cognition 1989;17:69–76.

Jones GV, Langford S. Phonological blocking in the tip of the tongue state. Cognition 1987;25:115–122. [PubMed: 3652647]

Jusczyk PW, Luce PA, Charles-Luce J. 'Infants' sensitivity to phonotactic patterns in the native language. Journal of Memory and Language 1994;33:630–645.

Keppel G. Words as random variables. Journal of Verbal Learning and Verbal Behavior 1976;15:263–265.

Kuèera, H.; Francis, WN. Computational analysis of present-day American English. Providence, RI: Brown University Press; 1967.

Landauer TK, Streeter LA. Structural differences between common and rare words: Failure of equivalence and assumptions for theories of word recognition. Journal of Verbal Learning and Verbal Behavior 1973;12:119–131.

Levelt WJM, Roelofs A, Meyer AS. A theory of lexical access in speech production. Behavioral and Brain Sciences 1999;22:1–38. [PubMed: 11301520]

Levelt WJM, Schriefers R, Vorberg D, Meyer AS, Pechmann T, Havinga J. The time course of lexical access in speech production: A study of picture naming. Psychological Review 1991;98:122–142.

Levitt AG, Healy AF. The roles of phoneme frequency, similarity, and availability in the experimental elicitation of speech errors. Journal of Memory and Language 1985;24:717–733.

Luce, PA. Neighborhoods of words in the mental lexicon. Indiana University; Bloomington: 1986. Doctoral dissertation

Luce PA, Pisoni DB. Recognizing spoken words: The neighborhood activation model. Ear and Hearing 1998;19:1–36. [PubMed: 9504270]

MacKay, DG. The organization of perception and action: A theory for language and other cognitive skills. New York: Springer-Verlag; 1987.

Marslen-Wilson WD, Zwitserlood P. Accessing spoken words: The importance of word onsets. Journal of Experimental Psychology: Human Perception and Performance 1989;15:576–585.

Maylor EA. Age, blocking and the tip of the tongue state. British Journal of Psychology 1990;81:123–134. [PubMed: 2364243]

McQueen JM. Segmentation of continuous speech using phonotactics. Journal of Memory and Language 1998;39:21–46.

Meyer AS. Lexical access in phrase and sentence production: Results from picture-word interference experiments. Journal of Memory and Language 1996;35:477–496.

Meyer AS, Bock JK. The tip-of-the-tongue phenomenon: Blocking or partial activation? Memory & Cognition 1992;20:715–726.

Motley MT. On replicating the SLIP technique: A reply to Sinsabaugh and Fox. Communication Monographs 1986;53:342–351.

Motley MT, Baars BJ. Encoding sensitivities to phonological markedness and transitional probability: Evidence from spoonerisms. Human Communication Research 1975;2:351–361.

Motley MT, Baars BJ. Laboratory induction of verbal slips: A new method for psycholinguistic research. Communication Quarterly 1976;24:28–34.

Norris D, McQueen JM, Cutler A. Merging information in speech recognition: Feedback is never necessary. Brain and Behavioral Sciences 2000;23:299–370.

Nusbaum, HC.; Pisoni, DB.; Davis, CK. Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words (Research on Speech Perception, Progress Report #10). Bloomington: Indiana University, Speech Research Laboratory; 1984.

Oldfield RC, Wingfield A. Response latencies in naming objects. Quarterly Journal of Experimental Psychology 1965;17:273–281. [PubMed: 5852918]

Pitt MA, Samuel AG. Lexical and sublexical feedback in auditory word recognition. Cognitive Psychology 1995;29:149–188. [PubMed: 7587137]

Raaijmakers JGW, Schrijnemakers JMC, Gremmen F. How to deal with "The language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. Journal of Memory and Language 1999;41:416–426.

Roediger HL, Neely JH, Blaxton TA. Inhibition from related primes in semantic memory retrieval: A reappraisal of Brown's (1979) paradigm. Journal of Experimental Psychology: Learning, Memory, and Cognition 1983;9:478–485.

Roelofs A. A spreading-activation theory of lemma retrieval in speaking. Cognition 1992;42:107–142. [PubMed: 1582154]

Saussure, Fd. Course in general linguistics. Baskin, W., translator. New York: McGraw Hill; 1966. (Original work published 1916)

Schacter DL. The seven sins of memory: Insights from psychology and cognitive neuroscience. American Psychologist 1999;54:182–203. [PubMed: 10199218]

Schriefers H, Meyer AS, Levelt WJ. Exploring the time course of lexical access in language production: Picture-word interference studies. Journal of Memory and Language 1990;29:86–102.

Sevald CA, Dell GS. The sequential cuing effect in speech production. Cognition 1994;53:91–127. [PubMed: 7805353]

Sevald CA, Dell GS, Cole JS. Syllable structure in speech production: Are syllables chunks or schemas? Journal of Memory and Language 1995;34:807–820.

Shattuck-Hufnagel, S. Speech errors as evidence for a serial order mechanism in sentence production. In: Cooper, WE.; Walker, ECT., editors. Sentence processing: Psycholinguistic studies presented to Merrill Garrett. Hillsdale, NJ: Erlbaum; 1979. p. 295-342.

Shattuck-Hufnagel S. The role of word structure in segmental serial ordering. Cognition 1992;42:213–259. [PubMed: 1582157]

Shattuck-Hufnagel S, Klatt D. The limited use of distinctive features and markedness in speech production: Evidence from speech error data. Journal of Verbal Learning and Verbal Behavior 1979;18:41–55.

Sinsabaugh BA, Fox RA. Reevaluating the SLIP paradigm: A research note. Communication Monographs 1986;53:335–341.

Smith JEK. The assuming-will-make-it-so fallacy. Journal of Verbal Learning and Verbal Behavior 1976;15:262–263.

Snodgrass JG, Vanderwart M. A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. Journal of Experimental Psychology: Human Learning and Memory 1980;6:174–215. [PubMed: 7373248]

Stemberger, JP. The reliability and replicability of naturalistic speech error data: A comparison with experimentally induced errors. In: Baars, BJ., editor. Experimental slips and human error: Exploring the architecture of volition. New York: Plenum Press; 1992. p. 195-215.

Stemberger JP, MacWhinney B. Frequency and the lexical storage or regularly inflected forms. Memory & Cognition 1986;14:17–26.

Stemberger JP, Treiman R. The internal structure of word-initial consonant clusters. Journal of Memory and Language 1986;25:163–180.

Storkel HL, Rogers MA. The effect of probabilistic phonotactics on lexical acquisition. Clinical Linguistics & Phonetics 2000;14:407–425.

Vitevitch MS. The neighborhood characteristics of malapropisms. Language and Speech 1997;40:211–228. [PubMed: 9509578]

Vitevitch MS, Luce PA. When words compete: Levels of processing in spoken word perception. Psychological Science 1998;9:325–329.

Vitevitch MS, Luce PA. Probabilistic phonotactics and spoken word recognition. Journal of Memory and Language 1999;40:374–408.

Vitevitch MS, Luce PA, Charles-Luce J, Kemmerer D. Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. Language and Speech 1997;40:47–62. [PubMed: 9230698]

Vitevitch MS, Luce PA, Pisoni DB, Auer ET. Phonotactics, neighborhood activation, and lexical access for spoken words. Brain and Language 1999;68:306–311. [PubMed: 10433774]

Vitevitch MS, Sommers MS. The role of phonological neighbors in the tip-of-the-tongue state. 2001Manuscript submitted for publication

Wike EL, Church JD. Comments on Clark's "The language- as-fixed-effect fallacy". Journal of Verbal Learning and Verbal Behavior 1976;15:249–255.

Woodworth, RS. Psychology. 2. New York: Holt; 1929.

Yaniv I, Meyer DE, Gordon PC, Huff CA, Sevald CA. Vowel similarity, connectionist models, and syllable structure in motor programming of speech. Journal of Memory and Language 1990;29:1–26.

Table 1

Mean Values by Condition for the Stimuli in Experiment 1

Variable	High frequency								Low frequency							
	Dense				Sparse				Dense				Sparse			
	High NF		Low NF		High NF		Low NF		High NF		Low NF		High NF		Low NF	
	M	SE	M	SE	M	SE	M	SE	M	SE	M	SE	M	SE	M	SE
Freq.	61.9	1.6	21.5	0.7	58.9	1.4	32.9	1.4	7.6	1.1	3.3	0.8	3.3	0.5	2.4	1.6
Density	23.71	0.8	20.43	1.1	14.07	1.1	11.64	1.1	22.43	1.5	21.50	0.9	14.56	1.1	12.36	1.1
NF	24.8	0.4	7.0	0.3	22.0	0.4	4.2	0.4	16.3	0.4	5.9	0.3	17.8	0.4	3.6	0.4

Note. Freq. = word frequency in occurrences per million; NF = neighborhood frequency in occurrences per million.

Table 2

The Rate of Speech Errors for Each Condition

	High frequency								Low frequency							
	Dense				Sparse				Dense				Sparse			
	High NF		Low NF		High NF		Low NF		High NF		Low NF		High NF		Low NF	
	%	SE	%	SE	%	SE	%	SE	%	SE	%	SE	%	SE	%	SE
	17.1	0.5	7.9	0.3	23.7	0.6	17.1	0.5	27.6	0.6	14.5	0.4	42.1	0.7	43.4	0.8

Note. NF = neighborhood frequency.