

RESEARCH ON SPOKEN LANGUAGE PROCESSING
Progress Report No. 23 (1999)
Indiana University

Neighborhood Density, the Tip-of-the-Tongue Phenomenon, and Aging¹

Michael S. Vitevitch and Mitchell S. Sommers²

*Speech Research Laboratory
Department of Psychology
Indiana University
Bloomington, Indiana 47405*

¹ This research was supported in part by NIH-NIDCD training grant T32 DC 00012 to Indiana University and the Brookdale Foundation (Washington University). The authors would like to thank Emily McCutcheon for her assistance in data collection and analysis, and Luis R. Hernandez for his programming assistance.

² Department of Psychology, Washington University, Campus Box 1125, St. Louis, MO 63130.

Neighborhood Density, the Tip-of-the-Tongue Phenomenon, and Aging

Abstract. A tip-of-the-tongue (TOT) elicitation task was used with younger adults in Experiment 1 and with older adults in Experiment 2 to examine the influence of word frequency, neighborhood density, and neighborhood frequency on the retrieval of phonological word forms from the lexicon. The results of Experiment 1 replicated the results of Harley and Bown (1998): More TOT states were elicited for words with low frequency and sparse neighborhood density (i.e., words with few similar sounding words). However, the results from Experiment 2 showed that in addition to word frequency and neighborhood density, neighborhood frequency, the mean frequency of phonological neighbors, also influenced lexical retrieval for speech production in older adults. Specifically, neighborhood-frequency interacted with word frequency and with neighborhood density. More TOT states were elicited for words with low neighborhood frequency and low word frequency. In addition, more TOT states were elicited for words with low neighborhood-frequency and sparse neighborhoods. These results demonstrate that the number of similar sounding words affects lexical retrieval in production as well as perception. Furthermore, the influence of these lexical characteristics on the process of retrieving word-forms during speech production changes with age. The results are discussed within the context of the Node Structure Theory (MacKay, 1987), a model of cognitive processing.

Introduction

Speech production is a rapid and efficient process.³ However, there are instances in which the fluent retrieval of a lexical item fails to occur. One example of failed retrieval occurs in tip-of-the-tongue (TOT) states. Tip-of-the-tongue states occur when only partial information associated with a word can be retrieved: information regarding the meaning or syntactic class of the word may be accessible, but not the complete phonological form of the word. The ability to access partial information often results in one having a “feeling of knowing” the word, despite being unable to retrieve all the information associated with the word.

Factors that Affect Lexicalization and TOTs

The process of retrieving a word form from lexical memory during speech production is called lexicalization. Several factors affect the speed and accuracy of lexicalization. The factors that affect the speed and accuracy of normal, unimpaired, lexicalization also affect instances of incomplete lexicalization, or TOT states. One factor that affects the speed and accuracy of normal lexicalization, as well as TOT states, is word frequency. Using a picture-naming task, Oldfield and Wingfield (1965) demonstrated that pictures of high-frequency words were named more quickly than were pictures of low-frequency words. Stemberger (1985; Stemberger & MacWhinney, 1986; see also Dell, 1990) found that phonological speech errors, such as spoonerisms (switching the initial phoneme of two nearby words, such as “darn bore” instead of “barn door”), perseverations (carrying-over a phoneme from a word produced earlier, such as “Ladies and lentlemen...” instead of “Ladies and gentlemen...”), and anticipations (producing a phoneme

³ This is about the only point that researchers agree on. Debates continue about the processes involved in speech production. For example, Garrett (1976) and Levelt (1989) take a modular approach to processing, whereas Dell (1986) and Harley (1984; 1993) argue for an interactive approach. There are also debates about where certain types of information are represented in the speech production system. For example, Roelofs, Meyer, and Levelt (1998) argue that syntactic information is accessed at the lemma level, whereas Caramazza and Miozzo (1998) claim that syntactic information is accessed at the lexeme level.

from an upcoming word, such as “pig pig” instead of “big pig”), occurred more often for low-frequency words than for high-frequency words. Finally, more TOT states occur and are experimentally induced in low frequency words than in high frequency words (Brown & McNeill, 1966; Burke, MacKay, Worthley, & Wade, 1991; Harley & Bown, 1998; cf. Yaniv & Meyer, 1987). Taken together, these findings suggest that the well-documented perceptual disadvantage of low-frequency words (i.e., poorer identification and slower response times) is paralleled by a similar disadvantage in production, as evidenced by greater difficulty with lexicalization.

Another factor that affects the speed and accuracy of both perception and lexicalization is neighborhood density. Neighborhood density refers to the number of words that are phonologically similar to a given target word (Luce & Pisoni, 1998). A rough measure of phonological similarity can be obtained by determining the number of new words that are created by the addition, deletion, or substitution of a phoneme in a target word.⁴ For example, the word “cat” has as neighbors the words “scat,” “at,” “hat,” “cut,” and “cap,” as well as other words. Words with many similar sounding words are said to have dense neighborhoods, whereas words with few similar sounding words are said to have sparse neighborhoods. Previous research has shown that words with sparse neighborhoods are recognized more quickly and more accurately than words with dense neighborhoods (Luce & Pisoni, 1998).

Previous investigations (Goldinger & Summers, 1989; Harley & Bown, 1998; Vitevitch, 1997a; 1997b) have also demonstrated that neighborhood density can affect speech production. For example, Goldinger and Summers (1989) found that neighborhood density influenced the voice onset time (VOT) for spoken words. VOT refers to the point in time at which vocal fold vibration starts, following the release of a closure (Crystal, 1992). Participants repeated word pairs that differed in the voicing of the initial consonants within the pair (e.g., *dutch-touch*) and that varied in neighborhood density across pairs. An acoustic analysis showed that the differences in VOT between the first word and the second word of the pairs were larger for word pairs with dense neighborhoods than for word pairs with sparse neighborhoods. These differences decreased across sessions for word pairs with sparse neighborhoods, but increased across sessions for word pairs with dense neighborhoods. Furthermore, Goldinger and Summers found that the interword interval, or the time between the offset of the first word and the onset of the second word within each minimal pair, varied with neighborhood density. The interword interval was greater for dense neighborhood word pairs than for sparse neighborhood word pairs. These results demonstrate that neighborhood density influences certain aspects of timing in speech production.

The accuracy of lexical retrieval in speech production is also affected by neighborhood density. Vitevitch (1997a) used tongue twisters containing words that had either dense or sparse neighborhoods to elicit phonological speech errors from participants. He found that more errors occurred in tongue twisters that contained words with sparse neighborhoods than with dense neighborhoods. These results suggest that neighborhood density influences speech production in demonstrable ways (see also Vitevitch, 1997b; Wright, 1997; cf. Jescheniak & Levelt, 1994). Specifically, neighborhood density appears to produce “supportive” or facilitative effects among words. That is, words with many similar sounding words (a dense neighborhood) are produced more accurately than words with few similar sounding words (a sparse neighborhood). These findings contrast with the competitive effect of neighborhood density typically observed in spoken word recognition: Words with dense neighborhoods are recognized more slowly and less accurately than words with sparse neighborhoods (Luce & Pisoni, 1998).

⁴ An alternate method used to measure “similarity” is to use phoneme confusion matrices as in Luce and Pisoni (1998) in the calculation of Neighborhood Probability Rules (NPRs). Both methods have been successfully used to demonstrate effects of neighborhood density on spoken word recognition.

Vitevitch (1997a; see also MacKay & Burke, 1990) speculated that the difference in neighborhood density effects found in speech production and spoken word recognition were due to the differences in the flow of information during speech production and spoken word recognition. That is, in spoken word recognition, acoustic-phonetic input activates many similar sounding words in memory (e.g., Luce & Pisoni, 1998). This candidate set must be winnowed down to a single item that will then retrieve semantic and syntactic information related to that word from the lexicon. Thus, more time will be required to winnow down the candidate set if there are many competitors (Luce & Pisoni, 1998). In contrast, speech production begins with a single conceptual representation that proceeds to activate a single lexical item and a single phonological word form (Levelt, 1989). That single phonological word form then activates the many sub-lexical units that it contains, such as syllables, phonemes, features, etc. Thus, a word that has components shared by many other words (a word with a dense neighborhood) will be able to spread activation along pathways between those components that are well traversed. A word that has components shared by few other words (a word with a sparse neighborhood) will have difficulty spreading activation along the pathways between components that are less traveled.

Additional evidence of neighborhood density affecting speech production can be found in the work of Harley and Bown (1998). They recently reported that more TOT states were elicited for words from sparse neighborhoods than for words from dense neighborhoods, and also suggested that neighborhood density played a “supportive” role in speech production (Harley & Bown, 1998). Although Harley and Bown (1998) accounted for their results in the context of extant interactive models of lexicalization (i.e., Dell, 1986; Harley, 1993), their results are difficult to clearly interpret because of confounding variables in their stimulus set.

Specifically, in two experiments that manipulated word frequency and neighborhood density, Harley and Bown (1998) attempted to induce TOT states experimentally using words that varied in length from one syllable (e.g., “act”) to five syllables (e.g., “chronological”). Word length was a variable that was not stringently controlled in their stimuli, and, unfortunately, proved to be a confounding variable. The results of their first experiment showed that more TOT states were reported for words that were low in frequency and that had few neighbors as defined by Coltheart-*N* (Coltheart, Davelaar, Jonasson, & Besner, 1977). Although the tip-of-the-tongue phenomenon is often described as an inability to retrieve a sound-based representation from the lexicon, Harley and Bown (1998) constructed their stimulus set using a metric of similarity based on orthographic similarity (Coltheart-*N*) instead of a metric based on phonological similarity. It should be noted, however, that when Harley and Bown analyzed the results from a reduced set of their stimuli based solely on phonological neighborhoods, their findings remained relatively unchanged.

However, when Harley and Bown performed a regression analysis on the data in Experiment 1, they found a significant effect of word length on TOT states: TOT states were more likely to occur with longer words than shorter words. Across the lexicon, short words tend to have denser neighborhoods than longer words (Bard & Shillcock, 1993; Pisoni, Nusbaum, Luce & Slowiaczek, 1985). Their results are further complicated by other relationships among word frequency, word length, and neighborhood density in the lexicon. For example, Zipf (1965) observed an inverse relationship between word length and word frequency in English: Short words are more common in English than long words. Also, Landauer and Streeter (1973) found a positive correlation between word frequency and neighborhood density: High frequency words tend to have denser phonological neighborhoods than low frequency words. Thus, it is unclear whether the results in Experiment 1 of Harley and Bown (1998) were due to neighborhood density or another related variable.

Harley and Bown (1998) attempted to control word length more precisely in their second experiment by using monosyllabic and disyllabic words (however, the trisyllabic word “audience” appears as a stimulus item in a low *N* condition) to examine the effects of word frequency and neighborhood density on TOT states. Although the word frequency and neighborhood density effects from Experiment 1 were replicated, a close examination of the stimuli in Experiment 2 reveals that word length was not entirely controlled. An analysis of the stimuli in appendix B of Harley and Bown (1998) shows that words with dense neighborhoods were still shorter than words with sparse neighborhoods. This is true when word length is measured in number of phonemes (dense words, mean = 3.17 phonemes; sparse words, mean = 5.07 phonemes; $F(1,58) = 54.15, p < .001$) and in number of syllables (dense words, mean = 1.06 syllables; sparse words, mean = 1.83 syllables; $F(1,58) = 8.82, p < .001$). Given the complex relationships among word length, word frequency, and neighborhood density, it is unclear how each of these individual factors affected TOTs in Harley and Bown (1998).⁵

Accounts of TOTs

Several hypotheses have been advanced to account for the occurrence of TOT states. One hypothesis states that similar sounding words interfere with—or “block”—the retrieval of the phonological word-form (Jones, 1989, Jones & Langford, 1987; Maylor, 1990; Woodworth, 1929). For example, Jones (1989) presented definitions to participants and primed them with a word that was semantically, phonologically, or both semantically and phonologically related to the target word. Jones (1989; see also Jones & Langford, 1987, and Maylor, 1990) found that more TOT states were elicited when a phonologically related prime was presented after hearing the definition of the target word. Jones (1989) interpreted these results as being consistent with the hypothesis that phonologically related words block the retrieval of the desired word-form.

An alternative explanation of TOTs claims that insufficient activation results in incomplete retrieval of the target word (Brown, 1991; Burke, MacKay, Worthley & Wade, 1991). According to this hypothesis, similar sounding words should act to aid rather than block the retrieval of word-forms. Evidence for this hypothesis comes from the work of Meyer and Bock (1992) and Perfect and Hanley (1992). Meyer and Bock (1992) and Perfect and Hanley (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, phonological primes did not interfere with the retrieval of the target word form; rather, phonological primes aided in the retrieval of the target word-form (Meyer & Bock, 1992; Perfect & Hanley, 1992).

The results of Harley and Bown (1998) also support the hypothesis that phonological similarity can serve to support lexicalization. They found more TOTs for words with sparse neighborhoods than for words with dense neighborhoods, suggesting that the more neighbors a word has, the more “support” it receives, and the more likely it will be correctly and completely retrieved. Harley and Bown (1998) accounted for their results by hypothesizing that the representation of an intended word was not fully activated because of insufficient amounts of supportive feedback between the lexeme level (which contains phonological information) and the lemma level (which contains semantic and/or syntactic information). They state that “[...]lemmas corresponding to phonological forms that have no or few close neighbours can receive little or no supporting activation from feedback between the phonological and lemma levels from

⁵ As in Experiment 1, Harley and Bown (1998) performed a regression analysis, but failed to find a relationship between word length and number of TOT states. However, the restricted range of word length in Experiment 2 (mostly mono- and disyllabic words, with one trisyllabic word) compared to the broader range of word length in experiment 1 (words with one to five syllables) may account for the non-significant regression.

related forms.” (Harley & Bown, 1998; pp. 163-164). Harley and Bown (1998; see also Harley & MacAndrew, 1992) further hypothesize that weak representations or random noise in the connections between the lemma and phonological representations may also contribute to TOT states.

Rather than being at the interface between the lemma and lexeme, as postulated by Harley and Bown (1998), the locus of TOT states may instead be at the interface between the lexeme (i.e., the phonological representation of the whole word) and sub-lexical representations. This hypothesis was postulated by Burke et al. (1991) within the context of the Node Structure Theory (NST), an interactive model of cognitive processing (see MacKay, 1987). Specifically, they state that “...TOTs result when phonological feature nodes receive insufficient priming to become activated.” (Burke et al., 1991, pp. 547). Insufficient activation between “word” and “phoneme” representations as postulated in NST can also explain the results of Harley and Bown (1998).

Node Structure Theory and TOTs

NST consists of a network of processing units, or nodes, organized hierarchically into semantic, phonologic, and motoric levels. Nodes are localist representations and are connected symmetrically—both bottom-up and top-down connections. The same network of nodes is involved in the perception and production of language (MacKay, 1987).

Two processes operate in NST: priming and activation. Priming is the sub-threshold excitation of a node that prepares it for activation. Activation is an all-or-none state in which the node has crossed a certain threshold.

Priming has several characteristics. It spreads in parallel to all connected nodes higher and lower in the hierarchical structure of nodes. Nodes can sum the priming that they receive simultaneously from several other nodes, or that they receive temporally across a single connection. Finally, transmission of priming becomes less efficient when a node has been satiated after prolonged and repeated activation.

In NST, activation is different from priming. For example, activation does not “spread” as it does in other network theories (e.g., McClelland & Rumelhart, 1981). Instead, activation proceeds sequentially and hierarchically through the network in a top-down and left-to-right manner. Activation must occur (i.e., the threshold must be crossed) in order to consciously retrieve the information associated with a node.

In addition to the nodes representing information at various levels, there are sequence nodes that connect nodes that share the same syntactic function or sequential privilege of occurrence in words and sentences. (This collection of similar nodes is referred to as a domain.) When a sequence node is activated, it multiplies the priming of all the nodes in a domain. The consequence of this multiplication of priming is that the node that initially had the most priming in that domain reaches threshold first and becomes activated. This domain specific activation accounts for the regularity found in many types of substitution errors. For example, nouns often substitute for nouns rather than verbs in a sentence, initial consonants often substitute for initial consonants rather than vowels or final consonants, etc. (e.g., Dell, 1986; Stemberger, 1985; MacKay, 1979).

Burke et al. (1991) suggest that TOT states arise in NST due to a deficit in transmission of priming across certain connections that are crucial for producing a target word. In a TOT state, semantic nodes become activated giving access to semantic information associated with that word. However, priming

does not spread sufficiently among phonological nodes, resulting in some phonological information not being activated and retrieved.

Transmission deficits may be caused by three factors: frequency of use, recency of use, and aging. The frequent activation of a node results in an increase in the rate and amount of priming that is transmitted across the connections of that node. Connections between less frequently activated nodes weaken with time, making the transmission of priming less efficient. This factor accounts for the frequency effects commonly found in speech error corpora: word and phoneme substitutions occur more often among low frequency items than among high frequency items (e.g., Stemmer, 1985, Stemmer & MacWhinney, 1986). Furthermore, object naming is faster for high frequency items than low frequency items (e.g., Oldfield & Wingfield, 1965), and TOT states tend to occur more often for low frequency than high frequency words (e.g., Burke et al., 1991).

Over time, connection strength between nodes decreases. If the connections become too weak, transmission of priming becomes less efficient. The longer a node has been “inactive,” the more decay occurs to the connections of that node. Evidence for this factor comes from a diary study by Burke et al. (1991) in which young and older participants recorded naturally occurring TOT states. Burke et al. (1991) found that a TOT state would more likely occur for a proper name the longer the duration since that acquaintance was contacted.

Finally, age weakens the strength of connections within the entire network (MacKay & Burke, 1990), reducing the rate and amount of priming being transmitted. This factor accounts for the general slowing of cognitive processes often associated with aging (e.g., Salthouse, 1985) and also the increased rate of TOT states seen for older compared with younger adults (Burke et al., 1991).

The neighborhood density effects observed by Harley and Bown (1998)—more TOTs for words with sparse neighborhoods—can also be accounted for within NST if one views neighborhood density as a frequency effect among the phonological constituents of a word. For example, a word with a dense neighborhood has many similar sounding neighbors, and, therefore, is comprised of segments that are common, or very frequent in the language. A word with a sparse neighborhood has fewer similar sounding neighbors, and, therefore is comprised of segments that are less frequent in the language. (See Vitevitch, Luce, Pisoni, & Auer (1999) for evidence of a positive correlation between neighborhood density and the frequency of segments and sequences of segments, also known as phonotactic probability.) Phonemes (represented by phonological nodes) that constitute words with dense neighborhoods receive more priming than phonological nodes that constitute words with sparse neighborhoods. The greater amount and rate of priming received by phonological nodes that constitute words in dense neighborhoods further strengthens those connections, whereas the connections between phonological nodes and words with sparse neighborhoods become weaker over time. Thus, the phonological information associated with a word with a dense neighborhood is more efficiently retrieved than the phonological information associated with a word with a sparse neighborhood. When the number of similar sounding words (i.e., neighborhood density) is viewed in terms of sub-lexical constituents, NST can also account for the neighborhood density effects observed by Harley and Bown (1998) without invoking additional feedback mechanisms between lexemes and lemmas.

NST, Neighborhoods, and Aging

In the current set of experiments, we examined the influence of neighborhood density on the number of TOT states elicited experimentally with stimuli that were controlled for word length. To

unambiguously demonstrate that neighborhood density, independent of word frequency and word length, affects TOT states, we used monosyllabic words with a consonant-vowel-consonant syllable structure that varied in word frequency, neighborhood density, and neighborhood frequency—a variable not manipulated by Harley and Bown (1998). (Neighborhood frequency is the mean frequency of the neighbors of a given target word). Also, given that the TOT phenomenon is the inability to retrieve phonological word forms from the lexicon, we used a metric of similarity based on phonological representations rather than orthographic representations as in Harley and Bown (1998).

A second goal of our investigations was to examine age differences in the effects of neighborhood density on lexicalization by eliciting TOT states from older adults in Experiment 2. Previous research (Sommers, 1996) reported that neighborhood density has greater effects on older than on younger adults. Therefore, Experiment 2 examined whether this age difference in the effects of neighborhood density on perception would have parallels in production. Work by Burke et al. (1991; see also Burke & Laver, 1990; MacKay & Burke, 1990; and Rastle & Burke, 1996) suggests that older adults have more problems than younger adults retrieving items from memory during language production. For example, older adults report more TOT states than younger adults. In Experiment 2, we examined how age modulates the effects of neighborhood density on the frequency of TOT states. In summary, we examined the influence of several lexical characteristics on speech production with a stringently controlled set of stimuli, and we also investigated how the influence of these characteristics may change over the life span by eliciting TOT states from young and older adults.

Experiment 1

The principle goal of Experiment 1 was to determine whether the effects of neighborhood density on TOT states reported by Harley and Bown (1999) could be replicated using a set of stimuli that were controlled for word length. Experiment 1 used a TOT elicitation task similar to that developed by Brown and McNeill (1966). In this task, participants are presented with a definition and must retrieve from memory the word that best matches the definition. Participants indicate whether they know the word (and produce it), don't know the word, or know the word but can't retrieve it (i.e., they are in a TOT state). The target words in this experiment were monosyllabic words that varied in word frequency, neighborhood density, and neighborhood frequency. Monosyllabic words were used to control for effects of word length on TOT susceptibility observed in Harley and Bown (1998).

Several predictions were made based on NST (Burke et al., 1991, MacKay, 1987) and the results of Harley and Bown (1998). Insufficient activation of phonological nodes would result in more TOT states for low frequency words than high frequency words. Based on the view that neighborhood density can be considered a frequency effect among the phonological constituents of a word, this same mechanism would also result in more TOT states for words with sparse neighborhoods than for words with dense neighborhoods. Finally, insufficient activation of phonological nodes would result in more TOT states for words with low-frequency neighbors than for words with high-frequency neighbors.

Method

Participants

Twenty-four native English-speaking adults were recruited from the Washington University community. None of the participants reported a history of a speech or hearing disorder and all received partial credit towards an Introductory Psychology class for their participation. The mean age of these

participants was 20.7 years ($SD = 1.9$). Mean years of education for the participants was 14.1 years ($SD = 1.4$). Data from one participant was excluded from all analyses because of failure to comply with experimental instructions.

Materials

One hundred twenty monosyllabic words consisting of a consonant-vowel-consonant syllable pattern were used as targets in the TOT elicitation task. Eight conditions, each containing fifteen words, were formed by orthogonally combining two factors of word frequency (HIGH and LOW), neighborhood density (SPARSE and DENSE), and neighborhood frequency (HIGH and LOW). The familiarity ratings (based on a scale from 1 “Don’t know the word” to 7 “Know the word”; Nusbaum, Pisoni, & Davis, 1984) did not differ across conditions ($F(1,112) = 2.02, p > .10$). The means and standard deviations for familiarity, word frequency, neighborhood density, and neighborhood frequency for the words in each condition are listed in Table 1.

High-frequency words (mean = 38.96 occurrences per million) had significantly higher frequencies of occurrence (based on values from the Kucera and Francis (1967) word counts) than low-frequency words (mean = 2.52 occurrences per million; $F(1,112) = 462.08, p < .001$). Neighborhood density was calculated by determining the number of words that could be created from a target word by adding, deleting, or substituting a phoneme. Words in the SPARSE NEIGHBORHOOD conditions (mean = 13.23 words) had significantly fewer neighbors, than the words in the DENSE NEIGHBORHOOD conditions (mean = 24.40 words; $F(1,112) = 247.17, p < .001$). Neighborhood Frequency, defined as the mean word frequency of all the neighbors of a target word, was also calculated using an on-line database. Words in the HIGH NEIGHBORHOOD FREQUENCY conditions (mean = 217.88 words per million) had neighbors with significantly higher values of word frequency than the neighbors of words in the LOW NEIGHBORHOOD FREQUENCY conditions (mean = 40.98 words per million; $F(1,112) = 255.36, p < .001$).

The questions for inducing TOT states were based on the definitions for each word found in the Webster’s New Collegiate Dictionary (1979). A pilot study using another group of participants (a young and an old adult group) determined if the target word was an appropriate answer to the question. Each question and associated target word was presented to participants for a rating of how well the word

CONDITION	Familiarity	Word Frequency	Neighborhood Density	Neighborhood Frequency
High Frequency Dense Neighborhood High Neigh. Freq.	6.92 (.19)	52.33 (67.88)	25.6 (4.51)	273.05 (403.49)
High Frequency Dense Neighborhood Low Neigh. Freq.	6.81 (.32)	31.73 (31.92)	22.6 (2.47)	51.05 (51.95)
High Frequency Sparse Neighborhood High Neigh. Freq.	6.89 (.17)	37.86 (27.98)	14.0 (2.59)	170.35 (180.30)
High Frequency Sparse Neighborhood Low Neigh. Freq.	6.80 (.42)	33.93 (26.71)	14.0 (3.25)	36.32 (38.92)
Low Frequency Dense Neighborhood	6.51	3.86	26.3	325.66

High Neigh. Freq.	(.56)	(2.79)	(5.09)	(273.09)
Low Frequency Dense Neighborhood Low Neigh. Freq.	6.68 (.37)	3.80 (2.62)	23.1 (3.63)	40.99 (48.21)
Low Frequency Sparse Neighborhood High Neigh. Freq.	6.20 (.94)	1.14 (.36)	12.3 (4.75)	102.47 (368.20)
Low Frequency Sparse Neighborhood Low Neigh. Freq.	6.60 (.37)	1.26 (.45)	12.6 (4.15)	35.55 (85.26)

Table 1. Mean familiarity, word frequency, neighborhood density, and neighborhood frequency values for the eight conditions of target words in the TOT elicitation task (standard deviations are in parenthesis).

answered the question (1 = “Does not answer the question at all,” 4 = “Answer acceptable, but there is a better word,” 7 = “Answers the question very well”). Any question that participants rated as not being appropriately answered by the target word (a mean rating of 5 or below) was modified until additional pilot study deemed the target word an appropriate response to the question.

Each target word and question pair had three additional foils that were of the same word class and were semantically similar. The foils were derived from the same source as the questions.

Procedure

The procedure followed that used by Burke et al. (1991). Participants were given the vocabulary sub-test of the Weschler Adult Intelligence Scale. They then heard a description of the TOT state from Brown and McNeill (1966), and were guided through a practice session. The practice session consisted of four questions that were similar to those used in the experimental session of the TOT elicitation task. The TOT-inducing questions were presented on an IBM-compatible computer. For each question, participants typed their responses on the computer keyboard.

A flow-chart description of the TOT elicitation task, adapted from Burke et al. (1991), is presented in Figure 1. For each question, three options were initially presented to the participants: K if they knew the answer, D if they didn’t know the answer and T if the answer was on the tip of their tongue. After providing the initial response (K, D or T) for each TOT-inducing question, participants were asked to rate how familiar they were with the word in question on a scale from 1 (“unfamiliar”) to 7 (“very familiar”). As in Burke et al. (1991), participants then rated how certain they were that they could recall the word in question on a scale from 1 (“uncertain”) to 7 (“certain”).

If participants had initially responded K (they knew the answer), they were asked to type the response to the question. If they were correct, appropriate feedback was given and the next trial was initiated with the presentation of a new TOT-inducing question. If the participant responded with an incorrect answer, they were given multiple choices from which to select a response. If they selected the correct option, appropriate feedback was given. If they selected an incorrect option, they were provided with the correct response and a new trial began.

If the initial response was D (they didn't know the answer), participants immediately received multiple options from which to select (after answering the questions regarding familiarity and likelihood of recall for the word). Appropriate feedback was again provided for each response before a new trial was initiated.

If the participant indicated that they were in a TOT state by initially selecting T, a number of other questions followed the two rating estimates. Participants were asked to provide, if possible: the initial sound of the word, the final sound of the word, the number of syllables in the word, and any similar sounding words that persistently came to mind. As in the "Don't Know" response, they were given multiple options to select from with the additional option of "None of the above," and received appropriate feedback.

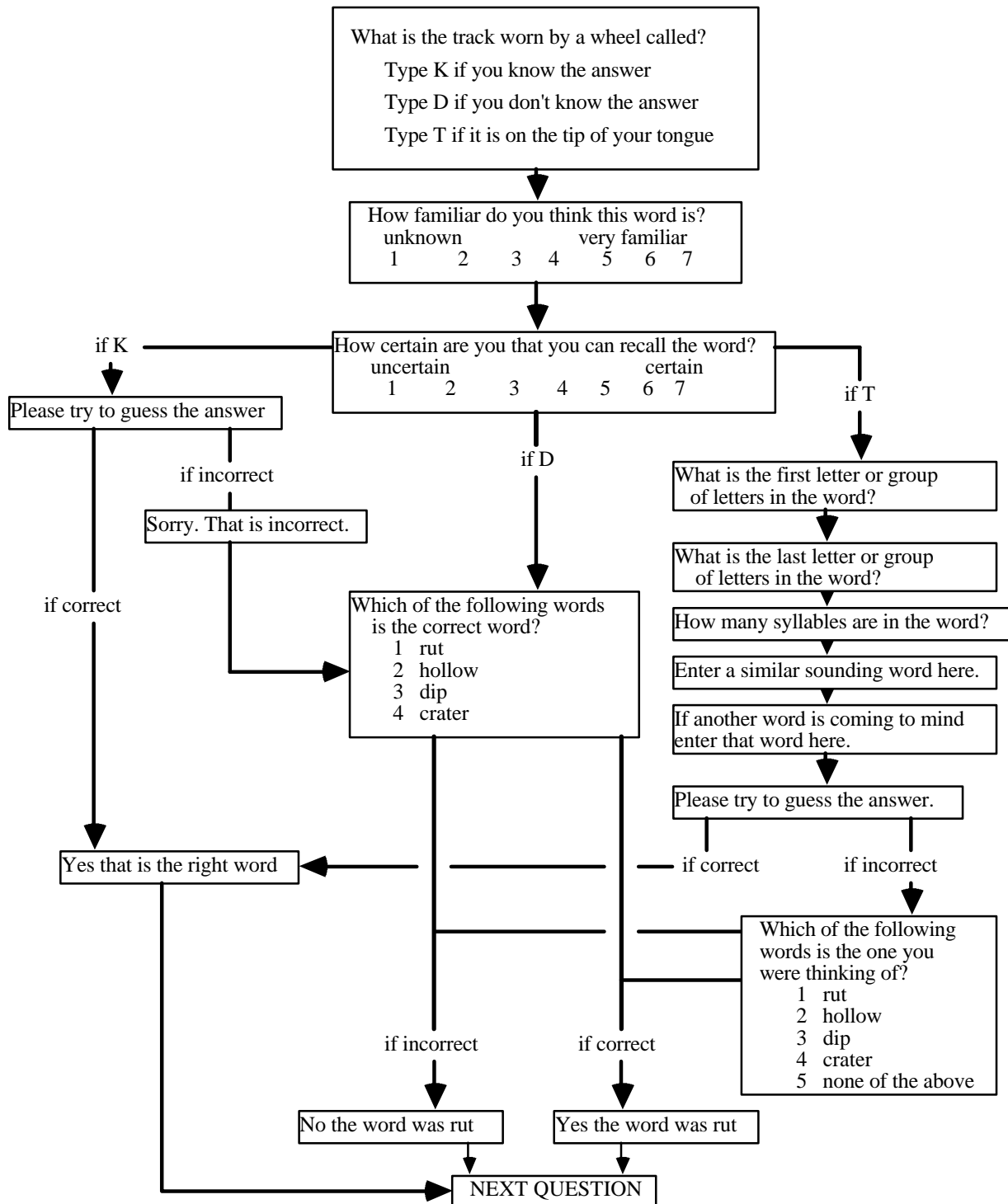


Figure 1. Flow chart description of the TOT elicitation task.

A brief practice session preceded the experiment. For the first question, participants were told to select K and to answer the questions that followed. For the second question, participants were told to select D and to answer the questions that followed. For the third question, participants were told to select T and to answer the questions that followed. For the last practice question, participants were allowed to select the option that was appropriate for their present state. This was done to familiarize the participants with all the possible types of questions that they might encounter. Upon completion of the practice session, participants began the experimental session of the TOT elicitation task and proceeded at their own pace. Participants were tested individually and received the 120 TOT-inducing questions in a different random order. As in Burke et al. (1991), participants could not backtrack to earlier questions, and the computer scored only the first three letters of an answer to minimize errors due to misspellings.

Results

Separate repeated-measures Analyses of Variance (ANOVAS) were performed on each type of response across the eight conditions formed by the orthogonal combination of two levels of three variables (word frequency, neighborhood density, and neighborhood frequency). Analyses by participants (F_1) and items (F_2) are reported. Additional analyses were also performed on the “interlopers,” (i.e., words that persistently came to mind instead of the desired word), and the other types of partially retrieved information.

Familiarity Rating

Participants were asked to rate how familiar they were with each word they were being questioned about on a scale from 1 (“Unknown”) to 7 (“Very Familiar”). There were no significant differences in either of the analyses (by F_1 and F_2) across the eight conditions in the familiarity ratings. The mean familiarity rating across the eight conditions (formed by crossing two levels of three variables) was 4.6 (range of 4.3 to 5.1 across conditions), suggesting that all the words were at least somewhat familiar to the participants.

Recall Rating

Participants were also asked to rate how certain they were that they could recall the word they were being questioned about on a scale from 1 (“Uncertain”) to 7 (“Certain”). Again, there were no significant differences in either of the analyses (by F_1 and F_2) across the eight conditions in the certainty ratings. The mean certainty rating across the eight conditions was 4.4 (range of 4.1 to 4.8 across conditions). This finding suggests that none of the stimulus variables (word frequency, neighborhood density, or neighborhood frequency) affected listeners' confidence in their ability to recall appropriate answers.

Total “Know” Responses

“Know” responses were analyzed in several ways. The first analysis examined the total number of “know” responses made. The second analysis examined the number of “know” responses that were actually correct. The final analysis examined the number of “know” responses that were initially incorrect when the participant did not have any options to choose from, but which were correctly selected once response alternatives were provided.

For the total number of “know” responses, a main effect of word frequency was found by participants ($F_1(1,22) = 5.74, p < .05$), but not by items ($F_2 < 1$), such that slightly more “know” responses were made to high frequency words (mean = 10.40) than to low frequency words (mean = 9.81). A significant interaction between neighborhood density and neighborhood frequency was also found by participants ($F_1(1,22) = 16.57, p < .001$) but not by items ($F_2(1,112) = 2.93, p > .10$). Slightly more “know” responses were made if the word was high in both neighborhood density and neighborhood frequency (mean = 10.41) or low in both neighborhood density and neighborhood frequency (mean = 10.87) than if one variable was high and the other variable was low (high neighborhood frequency/sparse neighborhood mean = 9.63; low neighborhood frequency/dense neighborhood mean = 9.52). Finally, a significant interaction was found by participants ($F_1(1,22) = 6.72, p < .05$) but not by items ($F_2 < 1$) between neighborhood density, word frequency, and neighborhood frequency. The means for this interaction can be found in the top portion of Table 2.

Younger Adults

	High Frequency				Low Frequency			
	Dense Neighborhood		Sparse Neighborhood		Dense Neighborhood		Sparse Neighborhood	
	Hi NHF	Lo NHF	Hi NHF	Lo NHF	Hi NHF	Lo NHF	Hi NHF	Lo NHF
Know Responses	11.0 (2.2)	9.2 (2.6)	10.0 (2.0)	11.3 (1.5)	9.7 (2.0)	9.8 (2.2)	9.2 (9.2)	10.4 (2.3)
Responses that were correct	6.0 (2.0)	6.1 (1.3)	4.7 (2.3)	7.0 (1.9)	4.9 (1.7)	6.5 (2.1)	4.6 (2.0)	3.5 (2.2)
Correct selection from multiple options	4.2 (1.7)	2.0 (1.2)	4.6 (2.2)	3.3 (1.8)	3.6 (1.9)	3.1 (2.0)	4.0 (1.9)	4.2 (1.9)
Don't know response	3.6 (2.3)	5.6 (2.6)	4.7 (2.1)	3.5 (1.5)	5.1 (2.0)	4.9 (2.1)	4.9 (2.1)	4.0 (2.1)
Correct response	2.6 (1.8)	3.7 (2.0)	3.4 (1.9)	2.0 (1.1)	3.7 (1.7)	4.2 (2.1)	3.6 (1.8)	3.0 (1.6)
TOT Reported	.26 (.54)	.13 (.34)	.21 (.67)	.17 (.38)	.13 (.45)	.22 (.42)	.83 (.93)	.57 (1.12)
Resolved	3	1	0	0	1	3	6	5

Table 2. Means for each type of response for young adults. Standard deviations are in parenthesis. *Note.* NHF = neighborhood frequency. The value for “TOT Reported” is the mean across 23 participants, whereas the value for “Resolved” is the raw number of resolved TOT states per condition.

Correctly Answered “Know” Responses

For the number of “know” responses that were actually correct, participants responded with the correct word more often for words with dense neighborhoods (6.0 correct responses) than for words with sparse neighborhoods (4.9 correct responses; $F_1(1,22) = 24.68, p < .001$; $F_2(1,112) = 4.65, p < .05$). No other differences were significant in the participant or items analyses (all F 's < 1). The means across each condition can be found in the top portion of Table 2.

“Know” Responses with Options

Of the “Know” responses that participants did not initially provide the correct word for, participants were relatively accurate in selecting the correct word from among the options displayed. There were no significant differences among the eight conditions in the participant or items analyses (all F 's < 1) for the number of correct selections made when participants selected from the multiple options presented. These means are also displayed in the top portion of Table 2.

Total “Don’t know” Responses

The means for the number of “don’t know” responses are shown in the middle of Table 2. Similar to the “know” responses, the “don’t know” responses were analyzed in several ways. First, the total number of “don’t know” responses were analyzed. When a “don’t know” response was made, participants were presented with several choices to select from. We then analyzed the number of correct selections made from the multiple choices.

The main effect of neighborhood density for the number of “don’t know” responses made was significant by participants ($F_1(1,22) = 5.06, p < .05$) but not by items ($F_2 < 1$). Participants were slightly more likely to respond “don’t know” for target words from dense neighborhoods (mean = 4.8) than for target words from sparse neighborhoods (mean = 4.2). Significant (by participants only) interactions were also observed between neighborhood density and neighborhood frequency ($F_1(1,22) = 17.04, p < .001$) and between neighborhood density, word frequency, and neighborhood frequency ($F_1(1,22) = 10.06, p < .01$). For the two-way interaction, slightly more “don’t know” responses were made to words with dense neighborhoods and either low neighborhood frequency (mean = 5.3) or high neighborhood frequency (mean = 4.3) than for words from sparse neighborhoods and either low neighborhood frequency (mean = 3.7) or high neighborhood frequency (mean = 4.8). The means for the three-way interaction are displayed in Table 2. However, none of these effects were significant by items (all F 's $< 2.6, p > .10$).

Correct “Don’t know” Responses

When participants made a “Don’t know” response, they were presented with multiple options from which to choose. No significant differences (by either participants or items) were found for the number of “Don’t know” responses for which the correct choice was selected (all F 's < 1).

Total TOT Responses

The bottom portion of Table 2 displays the mean number of TOT responses as a function of word frequency, neighborhood density, and neighborhood frequency. Several analyses were performed on the TOT responses. First, we analyzed the total number of TOT responses. A main effect of word frequency

was found ($F_1(1,22) = 11.30, p < .01$; $F_2(1,112) = 4.98, p < .05$) such that more TOT states were elicited for low frequency words (mean = .43) than for high frequency words (mean = .20). A main effect of neighborhood density was also found ($F_1(1,22) = 10.40, p < .01$; $F_2(1,112) = 5.93, p < .05$) such that more TOT states were elicited for words with sparse neighborhoods (mean = .45) than for words with dense neighborhoods (mean = .18). No main effect of neighborhood frequency was found ($F < 1.2$ for both participant and item).

The number of reported TOTs as a function of word frequency and neighborhood density are displayed in Figure 2. A significant interaction between word frequency and neighborhood density ($F_1(1,22) = 16.72, p < .001$; $F_2(1,112) = 5.93, p < .05$) was found. This interaction was due to more TOT states being elicited for words that had low frequency and sparse neighborhoods (the striped bar on the right; mean = .70) than for words in the other conditions. Pair-wise comparisons between words that had low frequency and sparse neighborhoods and the other conditions support this finding: 1) words that had low frequency and dense neighborhoods (the clear bar on the right; mean = .17; $F(1,22) = 33.49, p < .001$), 2) words that had high frequency and sparse neighborhoods (the striped bar on the left; mean = .20; $F(1,22) = 30.70, p < .001$), 3) words that had high frequency and dense neighborhoods (the clear bar on the left; mean = .20; $F(1,22) = 30.71, p < .001$). No other differences nor interactions were significant (all F 's < 1).

Across participants, the range of TOTs reported went from a high of 18 to a low of 1 reported TOT. The average rate of TOT responses in Experiment 1 was 3%. This value is considerably lower than the 13.2% and 19.9% reported in Harley and Bown (1998) and the 10.9% reported in Burke et al. (1991). The lower percentage of elicited TOTs in the current study may be due to the shorter length of the words used in this experiment. All of the stimuli in the present experiment were monosyllabic words, whereas other studies (e.g., Jones, 1989; Jones & Langford, 1987; Burke et al., 1991; Harley & Bown, 1998) have typically used bisyllabic or multisyllabic words. Recall that Harley and Bown (1998) found that the frequency of TOTs varied as a function of word length such that TOTs were more likely among longer words.

“Subjective” versus “Objective” TOTs

Additional analyses were conducted on the reported TOT states to examine participants' ability to resolve the TOT states and determine how much partial information was available in those TOTs that were not resolved. A distinction is often made between “subjective” TOT states, defined as the items self-reported by participants to be TOTs (Jones & Langford, 1987) and “objective” TOT states, defined as those TOT states that were correctly resolved. Nineteen of the 58 reported TOT states were resolved correctly without having to receive multiple options to select from (25 were guessed correctly when the participants received multiple options to select from, and 14 were guessed incorrectly by participants when presented with multiple options). A Chi-square analysis was used to compare the distribution of resolved TOT states to the distribution of reported TOT states. No difference between the two distributions was found ($\chi^2 4.12, p = .76$). Thus, the pattern of “subjective” TOT states was similar to the pattern of “objective” TOT states.

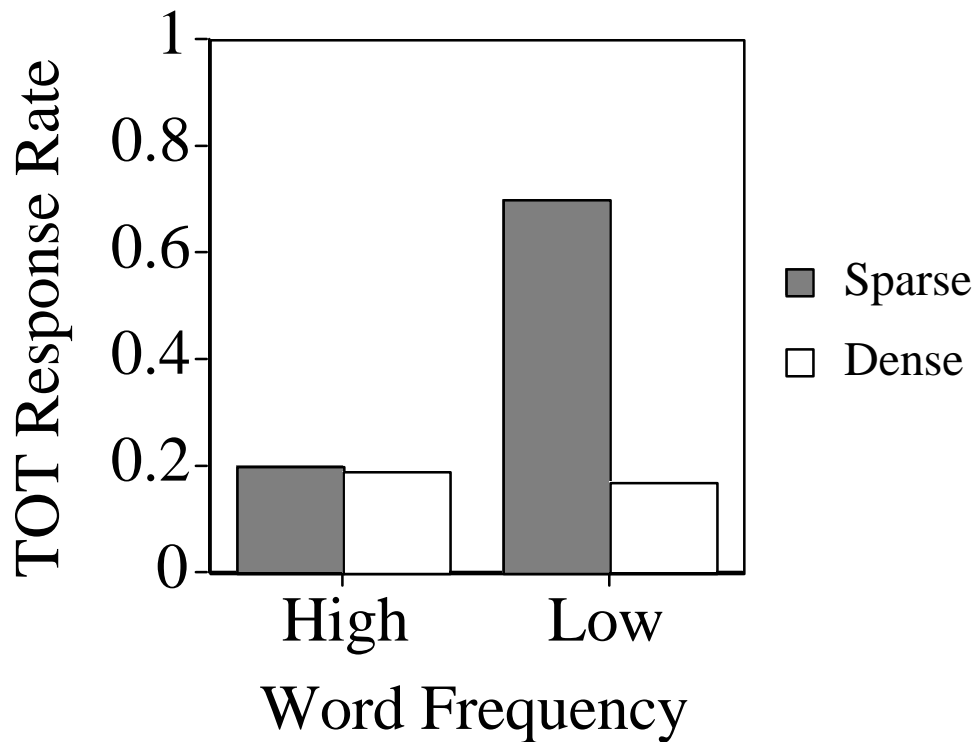


Figure 2. The number of TOTs for young adults as a function of word frequency and neighborhood density.

Analysis of the Interlopers

Among the 19 correctly resolved TOT states 10 interlopers (interfering words) were reported and all were either semantically or phonologically related to the target: Seven of the words were semantically related, 1 word was phonologically related, and 2 words were both semantically and phonologically related. We compared the frequency counts, familiarity ratings, neighborhood density, and neighborhood frequency of the interlopers and the targets. Information on 6 of the 10 interlopers was obtained from the same lexical database used to create the stimulus set. A paired *t*-test found no difference in the familiarity of the targets (mean = 6.4) and interlopers (mean = 6.8; $t(1,10) = 1.32, p = .21$) and no difference in the neighborhood frequency counts of the targets (mean = 73.56) and interlopers (mean = 18.53; $t(1,10) = 1.48, p = .17$). There was also no difference in the frequency counts (from Kucera & Francis, 1967) of the targets (mean = 12.50) and interlopers (mean = 18.83; $t(1,10) < 1$). This finding failed to replicate the results of Reason and Lucas (1983) who found that interlopers tended to be rated more frequent than target words.

One possible account of the differences between our experiment and the diary study of Reason and Lucas (1983) may be due to the methodology used in each study. As Reason and Lucas (1983) discuss, cognitive diary studies are subject to several types of biases in the selection and recording of events in the

diary. Perhaps only those interlopers that were more frequent were the interlopers that were reported by participants in the Reason and Lucas (1983) diary studies.

Another possible explanation for the differences between our findings and those of Reason and Lucas (1983) is the number of interlopers analyzed in each study. In our study we examined six interlopers, whereas Reason and Lucas (1983) examined 40 interlopers in their first diary study and 22 interlopers in their second diary study. The sample size of interlopers in our experiment may have been too small to provide adequate power to detect a difference in our statistical analysis.

There are also differences between the two studies in the way that word frequency was assessed. Reason and Lucas (1983) asked for subjective ratings of word frequency, whereas we used objective word counts (i.e., Kucera & Francis, 1967). Moreover, we used a *t*-test to quantitatively assess the difference in word frequency, whereas Reason and Lucas (1983) interpreted the differences in word frequency ratings qualitatively. (See Vitevitch, 1997 for a comparison of qualitative (e.g., chi-square) vs. quantitative (e.g., ANOVA) assessments of a speech error corpus, and the differences that may result in interpretation from the different assessment methods.) However, qualitative (chi-square) analysis performed on the current data failed to reveal significant effects for word frequency and neighborhood frequency (all $p > .10$), arguing against differences in the type of analyses as the primary factor responsible for the differences between the two studies.

However, a significant difference was found in the neighborhood density values for the interlopers (mean = 7.2 neighbors) and the targets (mean = 25.1 neighbors; $F(1,10) = 18.56, p < .01$). This finding suggests that when an interloper interfered with a participant's ability to correctly retrieve a target item, the interloper was likely to reside in a lower-density neighborhood than the target. Harley and Bown (1998) reported a similar finding in their study. It is difficult to clearly interpret this result in the current experiment because the interlopers were longer (3 words were 1 syllable long, 6 words were bisyllabic, and 1 word was trisyllabic) than the target words (all were monosyllabic words). Recall that Pisoni, Nusbaum, Luce and Slowiaczek (1985) found an inverse relationship between word length and neighborhood density: Short words tend to have many similar sounding items, whereas longer words tend to have fewer similar sounding items. The difference in word length between the targets and the interlopers may reflect a conscious search strategy adopted to resolve the TOT: if one is having a difficult time retrieving a lexical item (i.e., one is in a TOT state), it must be because the item is extremely "unusual" in some way. Such conscious awareness may result in the search for and retrieval of longer words, which are less common (i.e., "unusual") than monosyllabic words in English (Zipf, 1965). Conscious search strategies may also over-ride any affects of automatic processes, such as frequency-based biases, which may account for the non-significant difference in word-frequency between targets and interlopers observed in the present experiment.

Reports of Partial Information

Participants were also asked to report any partial information regarding the word they were attempting to retrieve. Participants correctly reported the number of syllables in the target word 5 out of 6 times, correctly reported the first letter of the target word 6 out of 7 times, and correctly reported the last letter of the target word 6 out of 6 times.

Discussion

The results from the current experiment replicate and extend the results of several previous studies examining factors that may influence the rate of TOTs (Brown & McNeill, 1966; Burke et al., 1991; Harley & Bown, 1998). Specifically, our findings indicated that TOT states were more likely for words with low-frequency and sparse neighborhoods. Several hypotheses have been advanced to account for the role of similar sounding words in TOT states. One hypothesis states that similar sounding words interfere with the retrieval of the lexeme, or phonological word form (Jones, 1989, Jones & Langford, 1987; Maylor, 1990; Woodworth, 1929). However, the current results—more TOT states for words with few rather than many similar sounding words—do not support this hypothesis. Instead, our results support the hypothesis that similar sounding words aid in the retrieval of intended representations (Brown, 1991; Burke et al., 1991; Meyer & Bock, 1992; Perfect & Hanley, 1992). Moreover, the finding of significantly more correctly answered “Know” responses for words with dense neighborhoods (6.0 correct answers) than for words with sparse neighborhoods (4.9 correct answers) further suggests that the number of similar sounding neighbors aids the retrieval of the intended word-form.

Our results are also consistent with predictions derived from NST regarding the influence of frequency and density on the probability of TOTs (Burke et al., 1991; MacKay, 1987). In NST, TOT states are the result of insufficient activation of the intended representations, particularly phonological nodes. Representations fail to be sufficiently activated because the connections between nodes transmit priming less efficiently. Transmission deficits result from the nodes not being used very often (frequency of use), not being used lately (recency of use), or because of aging. Because less common (low frequency) words do not receive priming as efficiently as more common (high frequency words), we predicted that more TOT states would occur for low frequency words than for high frequency words. The results of the current experiment support that prediction, replicating several other studies examining the role of word-frequency in lexicalization (e.g., Brown & McNeill, 1966; Burke, MacKay, Worthley, & Wade, 1991; Harley & Bown, 1998).

We also described neighborhood density in terms of the frequency of the sub-lexical constituents of a word. Specifically, words with dense neighborhoods are comprised of phonemes that are frequent and shared by many words, whereas words with sparse neighborhoods are comprised of phonemes that are less frequent and shared by few words (Vitevitch, Luce, Pisoni, & Auer, 1999). Phonemes (represented by phonological nodes in NST) that constitute words with dense neighborhoods receive priming more frequently and more recently than phonological nodes that constitute words with sparse neighborhoods. The difference in the frequency and recency of priming further strengthens the connections between words with dense neighborhoods and their constituent phonemes over time, whereas the connections between phonological nodes and words with sparse neighborhoods become weaker over time. Thus, the phonological information associated with a word with a dense neighborhood is more efficiently retrieved than the phonological information associated with a word with a sparse neighborhood. This hypothesis is supported by the finding that more TOT states were elicited for words with sparse neighborhoods than for words with dense neighborhoods (see also Harley & Bown, 1998). Additional support for the hypothesis that phonological information associated with dense rather than sparse neighborhoods is more readily available comes from the interesting observation that more of the interlopers that were either phonologically related or were both phonologically and semantically related occurred for target words with dense (66%) rather than sparse neighborhoods (33%).

Given the frequency-based effects (i.e., word-frequency and the frequency of the phonemes making up those words) on lexicalization that were observed in the present experiment, it is somewhat surprising that the frequency of the neighbors—defined as neighborhood frequency—did not appear to have an influence on the number of TOT states elicited. The effects of neighborhood frequency may have been

overshadowed by the more dominant influences of word frequency and neighborhood density. Although neighborhood frequency effects were not significant in the current experiment, the manipulation of this lexical characteristic is an important extension of Harley and Bown (1998). Both experiments by Harley and Bown failed to take this variable into consideration and neither manipulated nor controlled neighborhood frequency in their stimuli.

Of further interest is how the effects of these lexical characteristics on the process of lexicalization change over the life span. Work by Sommers (1996) has shown that the influence of neighborhood density on lexical retrieval in speech perception changes with age. Specifically, Sommers (1996) found that older adults were less accurate than younger adults at identifying words with many similar sounding neighbors, even under conditions producing similar identification performance for words with few similar sounding neighbors. One question that we wanted to address in the present study was whether the age-related changes in neighborhood density found in speech perception would also be found in speech production?

The findings of Burke et al. (1991) suggest that aging does affect speech production: older adults have more TOT states than younger adults. They attributed this increase in TOT states with age to an age-linked transmission deficit that decreased the availability of partial (phonological) information. Maylor (1990) found similar difficulties in accessing phonological information in older adults. Is this increase in TOT states with age differentially influenced by neighborhood density? Can this change with age also inform us about the nature of the decline associated with aging? To examine these questions we used the same stimuli from Experiment 1 to elicit TOT states in elderly adults.

Experiment 2

Previous studies have found more TOT states among elderly adults than among younger adults (Burke et al., 1991; Burke & Laver, 1990; MacKay & Burke, 1990; Rastle & Burke, 1996). Other research (Sommers, 1996; Sommers & Danielson, 1999) has demonstrated age-related declines in the ability to access the lexicon in spoken-word recognition. In Experiment 2, we wanted to examine how differential access to the lexicon in younger and older adults would affect the frequency of TOTs in these two populations.

NST proposes that transmission of priming between nodes becomes less efficient as a function of age if "...other characteristics of the nodes are assumed to be equal...especially their history of prior practice and their recency of activation." (pg. 222, MacKay & Burke, 1990). This principle generally explains why young adults can retrieve information more efficiently than can older adults. However, what happens if the history of the nodes in young and older adults is not equal, as in the case of words with dense and sparse neighborhoods? Recall that neighborhood density can be viewed as a frequency effect among the sub-lexical components of a word (Vitevitch, Luce, Pisoni & Auer, 1999). Thus, words with dense neighborhoods have sub-lexical components that have been activated more frequently than the sub-lexical components of words with sparse neighborhoods.

Furthermore, the difference in the frequency with which components of words with dense neighborhoods versus those with sparse neighborhoods are activated increases with age. To illustrate this point imagine that at time t_1 word x has a frequency of 2 occurrences per million and word y has a frequency of 100 occurrences per million, a difference of 98 occurrences per million. If one is exposed to a million words a year, in ten years, time t_{10} , word x will have been heard 20 times and word y will have been heard 1,000 times. The difference between words x and y from time t_1 to time t_{10} is now 980 occurrences per million, much greater than the original difference. Thus, even though overall vocabulary, as measured

by the vocabulary sub-test of the WAIS, for example, may not have increased, the same person at time t_{10} will have been exposed to more instances of a particular word than at time t_1 . This difference in exposure to words may also be observed in different people that are matched on all relevant variables, such as IQ or vocabulary size, but that differ in age.

The connections to the components of words that have a higher exposure will also have been activated and strengthened over time. With the increased linguistic exposure as a function of age, words with dense neighborhoods and strong connections to sub-lexical components may become at least partially “insulated” from the effects of age-related transmission deficits. In contrast, words with sparse neighborhoods will have much weaker connections to sub-lexical components and will be more susceptible to age-related transmission deficits. We predict that age-related transmission deficits will result in more TOT states for older adults than for younger adults, replicating the results of Burke et al. (1991). We further predict that the major source of TOT states in older adults will be for words with sparse neighborhoods. We hypothesize that the differences between older and younger adults is a result of differential exposure to words as a function of age, and the differential influence of age-related transmission deficits on the sub-lexical components of words with dense and sparse neighborhoods. To examine how lexicalization may be affected by normal aging, we used the same stimuli that were used in Experiment 1 and presented them to elderly participants.

Method

Participants

Twenty-four native English-speaking adults were recruited from the Washington University community. All subjects reported no history of a speech or hearing disorder and received \$20 for their participation. The mean age of these participants was 70.3 years ($SD = 4.9$). Mean WAIS vocabulary scores for this group of older adults did not differ significantly from vocabulary scores for younger adults ($F < 1.2$).

Materials and Procedure

The same materials and procedure used in Experiment 1 were used in the current experiment.

Results

Analyses similar to those in Experiment 1 were performed in the current experiment. In addition, the data from the two experiments were combined to compare the performance of younger and older adults. In analyses comparing the performance of younger and older adults Experiment (Experiment 1 versus Experiment 2) was treated as a between-participants factor.

Familiarity Rating

As with the young adults, there were no significant differences in familiarity ratings ($F < 3.14$, $p > .08$ for both participants and items) across the eight conditions. The mean familiarity rating across all conditions for older adults was 5.2 (range of 4.8 to 5.4 across conditions), suggesting that all the words were at least somewhat familiar to the participants. There was a significant difference in the familiarity ratings given by the younger and older adults by participants ($F_1(1, 45) = 5.45$, $p < .05$) but not by items

($F_2 < 1$). Overall, older adults tended to give higher familiarity ratings than younger adults. No other main effects or interactions were significant (all F 's < 1).

Recall Rating

The mean certainty rating for older adults regarding their ability to recall the word they were being questioned about was 5.2 (range of 4.7 to 5.5 across conditions). Certainty ratings did not vary significantly across the eight conditions. A main effect of age was found in the recall ratings by participants ($F_1(1, 45) = 15.50, p < .001$) but not by items ($F_2 < 1$). Overall, older adults gave higher certainty ratings than younger adults. No other main effects or interactions were significant (all F 's < 1).

Total “Know” Responses

As with the younger adults, “Know” responses from the older adults were analyzed in several ways. The first analysis examined the total number of “know” responses. The second analysis examined the number of those “know” responses that were actually correct. The final analysis examined the number of “know” responses that were initially incorrect when the participant did not have any options to choose from, but which were correctly selected when options were provided. Mean number of “Know” responses as a function of word frequency, neighborhood density, and neighborhood frequency are displayed in the top panel of Table 3.

For the older adults, a main effect of word frequency was found for the total number of “Know” responses by participants ($F_1(1,23) = 5.72, p < .05$), but not by items ($F_2 < 1$). There were slightly more “Know” responses to high-frequency words (mean = 11.61) than to low-frequency words (mean = 11.21). No other main effects or interactions were significant (all $F < 1$) for the number of “Know” responses.

In comparing the younger and older adults, a main effect of age was found for the total number of “know responses” ($F_1(1,45) = 3.94, p < .05$; $F_2(1,224) = 23.07, p < .001$) such that older adults responded “know” more often (mean = 11.42) than younger adults (mean = 10.10), replicating one of the findings of Burke et al. (1991).

Correctly Answered “Know” Responses

For older adults, there were no main effects or interactions that were significantly different in both the participants and item analyses for the number of “Know” responses that were actually answered correctly. There were also no main effects nor interactions significant by both participants and items when the number of “Know” responses that were actually answered correctly by the older adults was compared to the number of “Know” responses that were actually answered correctly by the younger adults (all F 's < 1).

“Know” Responses with Options

Of the “Know” responses that the older participants did not initially provide the correct word for, they were relatively accurate in selecting the correct word from among the options displayed. There were no significant differences among the eight conditions for the number of correct selections made from the multiple options presented to the participants (all F 's < 1). Furthermore, there were no differences between the younger and older adults in the number of “Know” responses that they answered correctly when presented with multiple options (all F 's < 1).

Total “Don’t know” Responses

The mean number of “don’t know” responses as a function of word frequency, neighborhood density, and neighborhood frequency are displayed in the middle panel of Table 3. “Don’t know” responses for older adults did not vary across the eight conditions (all F 's < 1). A main effect of age was found for the number of “Don’t know” responses ($F_1(1,45) = 5.65, p < .05; F_2(1,224) = 15.09, p < .001$) such that older adults responded “Don’t Know” fewer times (mean = 3.0) than younger adults (mean = 4.6), also replicating one of the findings of Burke et al. (1991). No other main effects or interactions across the age groups were significant by both analyses.

Older Adults

	High Frequency				Low Frequency			
	Dense Neighborhood		Sparse Neighborhood		Dense Neighborhood		Sparse Neighborhood	
	Hi NHF	Lo NHF	Hi NHF	Lo NHF	Hi NHF	Lo NHF	Hi NHF	Lo NHF
Know Total	11.7 (3.3)	11.5 (3.2)	11.1 (2.8)	12.0 (3.0)	11.2 (3.2)	11.4 (3.1)	11.5 (3.0)	10.6 (2.7)
Responses that were correct	6.9 (2.8)	7.4 (2.3)	5.1 (2.5)	7.8 (1.7)	6.8 (2.1)	6.6 (2.2)	7.6 (1.9)	4.1 (2.6)
Correct selection from multiple options	4.1 (2.4)	2.8 (1.9)	4.9 (2.2)	2.5 (1.3)	3.6 (1.8)	3.9 (1.9)	3.1 (1.6)	4.1 (2.3)
Don’t know Total	2.6 (3.1)	3.0 (3.0)	3.4 (2.9)	2.5 (3.0)	3.2 (3.2)	2.8 (3.1)	3.2 (2.9)	3.2 (2.4)
Correct response	1.9 (2.2)	2.1 (2.4)	2.5 (1.9)	1.8 (1.8)	2.7 (2.9)	2.0 (2.5)	2.2 (2.8)	2.3 (1.8)
TOT Reported	.58 (1.02)	.42 (.65)	.38 (.57)	.46 (1.10)	.50 (.78)	.67 (.87)	.29 (.69)	1.0 (1.51)
Resolved	4	8	5	8	8	9	3	8

Table 3. Means for each type of response for old adults. Standard deviations are in parenthesis. *Note.* NHF = neighborhood frequency. The value for “TOT Reported” is the mean across 24 participants, whereas the value for “Resolved” is the raw number of resolved TOT states per condition.

Correct “Don’t know” Responses

For the older adults, there was no difference among the conditions for the number of “Don’t know” responses answered correctly from among the multiple options presented (both F ’s < 1). However, when the number of “Don’t know” responses answered correctly by the older adults was compared to the number of “Don’t know” responses answered correctly by the younger adults, a significant difference was found by subjects ($F_1(1,45) = 4.81, p < .05$) but not by items ($F_2 < 1$). Younger adults tended to answer more of the “Don’t know” responses correctly (mean = 3.3) when presented with multiple options than the older adults (mean = 2.1). This marginal difference may reflect different criteria for selecting the “don't know” option, as is indicated by a reduced likelihood to select the “don't know” response for older adults. The adoption of different “error criteria” by young and older adults is also consistent with NST (MacKay & Burke, 1990).

Total TOT Responses for Older Adults

The mean number of TOT responses for older adults as a function of word frequency, neighborhood density, and neighborhood frequency are shown in the bottom panel of Table 3. The main effect of word frequency was marginally significant by participants but not by items ($F_1(1,23) = 3.54, p = .07$; $F_2(1,112) = 2.00, p = .15$). There tended to be more TOT states elicited for low-frequency words (mean = .62) than for high-frequency words (mean = .46). Main effects of neighborhood density and neighborhood frequency were not significant (all $F < 1$).

However, an interaction between neighborhood frequency and neighborhood density was significant by participants ($F_1(1,23) = 5.08, p < .05$) and marginally significant by items ($F_2(1,112) = 3.13, p = .08$). For words from sparse neighborhoods, more TOTs were observed for words with low neighborhood-frequency (mean = .75) than for words with high neighborhood-frequency (mean = .33; $F(1,23) = 10.18, p < .01$). For words from dense neighborhoods, however, neighborhood frequency did not influence the number of TOTs (means for each = .54). No other differences were significant (all F ’s < 1).

There was also an interaction between word frequency and neighborhood frequency ($F_1(1,23) = 8.12, p < .01$; $F_2(1,112) = 4.51, p = .05$). More TOT states were elicited for words that had low word- and neighborhood-frequency (mean = .85) than for words that had high word- and neighborhood-frequency (mean = .48; $F(1,23) = 9.13, p < .01$). These results are displayed in Figure 3. There was no difference between low frequency words (mean = .40) and high frequency words (mean = .48) with high neighborhood frequency ($F < 1$). No other differences nor interactions were significant (all F ’s < 1). The mean number of “TOT” responses for each condition is displayed in Table 3. TOT responses for older participants ranged from a high of 20 to a low of one.

“Subjective” versus “Objective” TOTs for Older Adults

Additional analyses were conducted to examine participants’ ability to resolve TOT states and to determine how much partial information was available when TOT states could not be resolved. Fifty-three of the 106 reported TOT states (“subjective TOT”) were resolved correctly (“objective TOT”) without having to receive multiple options to select from (36 were guessed correctly when the participants received multiple options to select from, and 17 were guessed incorrectly by participants when presented with multiple options). As in Experiment 1, a Chi-square analysis was used to compare the distribution of resolved (“objective”) TOT states in the older adults to the distribution of reported (“subjective”) TOT states in the older adults. No difference between the two distributions was found ($X^2 4.41, p = .73$), suggesting that the pattern of “subjective” TOT states is similar to the pattern of “objective” TOT states.

Among the 53 correctly resolved TOT states, 18 interlopers (interfering words) were reported: 13 of the words were semantically related, 0 words were phonologically related, 4 words were both semantically and phonologically related, and 1 word bore no obvious relationship to the target word. We compared the frequency counts, familiarity ratings, neighborhood density, and neighborhood frequency of the interlopers and the targets. Information on 8 of the 18 interlopers was available from the same lexical database used to create the stimulus set. Independent t-tests found no difference in the familiarity of the targets and interlopers ($t(1,14) < 1$), no difference in the frequency counts (from Kucera & Francis, 1967) of the targets and interlopers ($t(1,14) < 1$), and no difference in the neighborhood frequency counts of the targets and interlopers ($t(1,14) = 2.24, p = .15$). As in experiment 1, a chi-square analysis was also done, but no significant differences were found ($p > .10$).

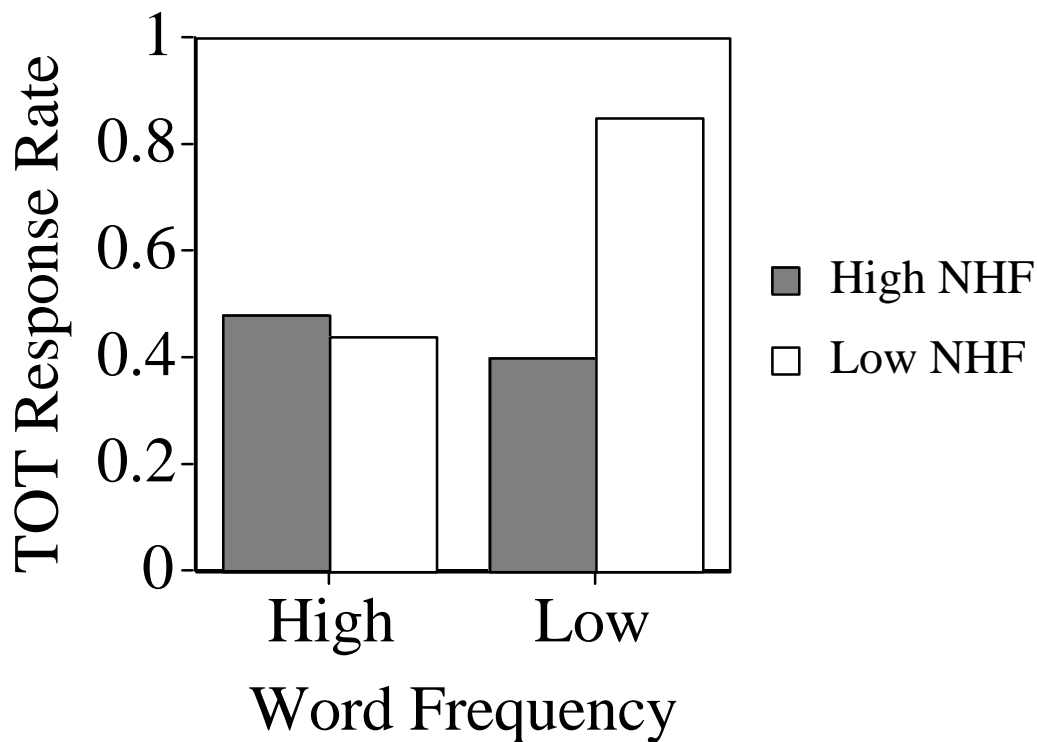


Figure 3. The number of TOTs for old adults as a function of word frequency and neighborhood frequency.

However, a significant difference was found in the neighborhood density values for the interlopers (mean = 18.0 neighbors) and the targets (mean = 29.6 neighbors; $t(1,14) = 2.81, p < .05$). This finding replicates the results of Harley and Bown (1998) and of Experiment 1, indicating that interlopers tended to be from sparser neighborhoods than their targets. Only one of the interlopers given by an older adult was a word with two syllables, and therefore would naturally have few similar sounding words (Pisoni, Nusbaum, Luce & Slowiczek, 1985). The difference in neighborhood density between the interlopers and the targets was still significant ($t(1,12) = 2.28, p < .05$) when this item (and the related target word) was excluded

from the analysis. As in Experiment 1, the difference in neighborhood density between the targets and the interlopers may be due to a conscious search strategy adopted to resolve the TOT. That is, participants become consciously aware that they are having difficulty retrieving a word from the lexicon, decide the item must be “unusual” in some way, and search out “lexical hermits,” or words with very few phonological neighbors. Adopting such conscious search strategies may mask natural biases used to select word-forms in normal processing.

Reports of Partial Information for Older Adults

Older participants were asked to report any partial information regarding the word they were attempting to retrieve. Participants correctly reported the number of syllables in the target word 8 out of 9 times, reported information regarding the first letter of the target word only once (and did so correctly), and did not report any information regarding the last letter of the target word. This result is interesting given the fact that younger adults always got the last letter correct.

Age differences in Total TOT states

To examine age-related differences in TOTs, we combined the TOT data from Experiments 1 and 2. A main effect of age was not significant by participants ($F_1(1,45) = 1.78, p = .19$), but was significant by items ($F_2(1,224) = 9.41, p < .01$). There was a tendency for older adults (mean = .54) to report more TOT responses than younger adults (mean = .32). This finding replicates the results of Burke et al. (1991). There was a marginally significant main effect of neighborhood density by participants ($F_1(1,45) = 4.08, p = .05$) but not by items ($F_2(1,224) = 2.56, p = .11$). There was a tendency for more TOT states to be reported for words with sparse neighborhoods (mean = .50) than for words with dense neighborhoods (mean = .36). A main effect of word frequency was found ($F_1(1,45) = 12.63, p < .001$; $F_2(1,224) = 6.42, p < .05$), such that more TOT states were reported for low frequency words (mean = .53) than for high frequency words (mean = .33).

A marginally significant interaction was found between age and neighborhood density by participants ($F_1(1,45) = 4.08, p = .05$) but not by items ($F_2(1,224) = 2.56, p = .11$). There was a tendency for more TOTs to be reported for words with sparse neighborhoods than for words with dense neighborhoods for young adults, and no difference between the two conditions for older adults.

A marginally significant interaction was found between age and neighborhood frequency ($F_1(1,45) = 4.04, p = .05$; $F_2(1,224) = 3.49, p = .06$). Older adults tended to report more TOTs for words that had low neighborhood frequency (mean = .64) than for words with high neighborhood frequency (mean = .43). This trend was reversed for younger adults; more TOTs were reported for words that had high neighborhood frequency (mean = .36) than for words with low neighborhood frequency (mean = .27).

A significant interaction was found between neighborhood density and word frequency ($F_1(1,45) = 5.62, p < .05$; $F_2(1,224) = 4.55, p < .05$). Figure 4 displays TOT response rates (combined across older and younger adults) as a function of word frequency and neighborhood density. As shown in the Figure, neighborhood density did not significantly influence the probability of TOTs for high frequency words. In contrast, for low frequency words significantly more TOTs were observed for words from sparse neighborhoods (mean = .68) than for words from dense neighborhoods (mean = .38; $F(1,45) = 8.68, p < .01$).

The interaction between neighborhood frequency and word frequency was significant by participants ($F_1(1,45) = 4.64, p < .05$), and marginally significant by items ($F_2(1, 224) = 2.56, p = .11$). There was a tendency for more TOTs to be reported for low frequency words than for high frequency words. This difference was greater for words with low neighborhood frequency than for words with high neighborhood frequency.

The interaction between neighborhood frequency, word frequency, and age was also significant by participants ($F_1(1,45) = 4.64, p < .05$), and marginally significant by items ($F_2(1, 224) = 2.56, p = .11$). Again, there was a tendency for more TOTs to be reported for low frequency words than high frequency words. For younger adults, this difference was approximately equal for words with both high and low neighborhood-frequency. However, for older adults the difference between the number of TOTs reported for high and low frequency words was much greater for words with low neighborhood frequency than high neighborhood frequency.

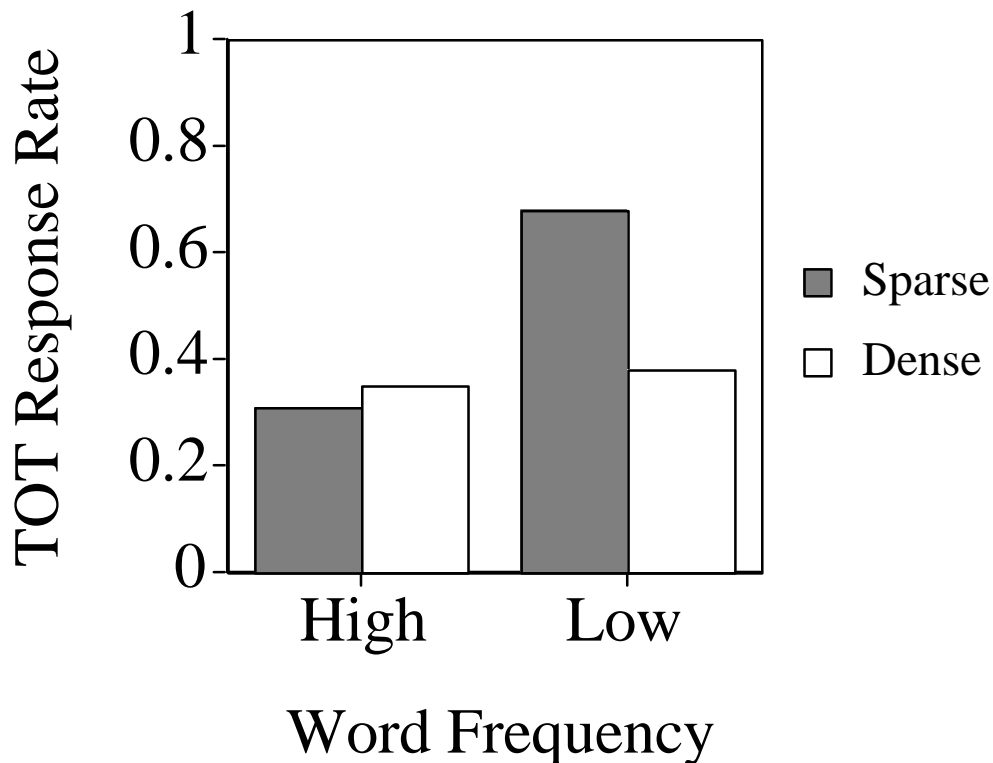


Figure 4. Overall mean number of TOT responses (younger and older adults combined) as a function of word frequency and neighborhood density.

Finally, an interaction was found between neighborhood density, neighborhood frequency, and age ($F_1(1,45) = 5.28, p < .05$). This interaction was marginally significant by items ($F_2(1, 224) = 3.01, p =$

.08). The trend shows that younger adults consistently reported more TOTs for words with sparse neighborhoods than for words with dense neighborhoods, regardless of neighborhood frequency. In contrast, older adults tended to report more TOTs for words with sparse neighborhoods only for words with low neighborhood frequency. This tendency was slightly reversed in older adults for words with high neighborhood frequency.

Discussion

The results of Experiment 2 replicate and extend findings of Burke et al. (1991; see also Burke & Laver, 1990; MacKay & Burke, 1990; and Rastle & Burke, 1996). For example, the current study found that more TOT states were elicited from older adults than from younger adults. These results are consistent with the transmission deficit account of TOT states (e.g., Bock & Levelt, 1994; Burke et al., 1991; Jescheniak & Levelt, 1994) within the NST (Burke et al., 1991). With age, the connections between nodes weaken, decreasing the rate and amount of priming transmitted between nodes. The decrease in the amount of priming between nodes in older adults results in the unsuccessful activation and retrieval of information associated with those nodes.

The “transmission deficit” hypothesis is also supported by the difference in availability of partial (phonological) information in younger and older adults. In the current experiment, older adults reported less partial information than did the younger adults. This observation is also consistent with the findings of Maylor (1990) who found that older adults had decreased availability of partial (phonological) information.

Of greater importance is the differential influence of neighborhood density with age. The results of Experiment 1, examining the influence of neighborhood density on TOT states in young adults, replicated the findings of Harley and Bown (1998): More TOT states were elicited for words with few neighbors than for words with many neighbors. However, older adults did not evidence significant main effects of neighborhood density as predicted. Instead, the process of lexicalization in older adults seemed to be differentially influenced by neighborhood frequency. Specifically, more TOT states were elicited in older adults for words with low neighborhood frequency and either sparse neighborhoods or low word frequency. No such interactions were observed in younger adults.

Although our predictions regarding the interaction between neighborhood density and aging were not observed in the results of Experiment 2, the findings of the present experiment can still be accounted for within NST (Burke et al., 1991; Burke & Laver, 1990; MacKay, 1987; MacKay & Burke, 1990). Within NST the efficiency of transmitting priming between nodes decreases with aging. Furthermore, the ability to access phonological information (i.e., phonological nodes) specifically decreases with age (Burke et al., 1991; Maylor, 1990). This accounts for the overall increase in the number of TOTs reported by older adults compared to younger adults.

Age-related transmission deficits may not be fully counter-acted by the stronger connections that arise from differential exposure rates as a function of age to words with dense neighborhoods versus words with sparse neighborhoods as we had initially predicted. Although phonemes shared by many words (i.e., phonemes in a word with a dense neighborhood) generally are “protected” from TOTs and other speech errors in younger adults, older adults activate and retrieve phonological information less successfully overall than younger adults. To compensate for the transmission deficits and weaker connections that result from aging, MacKay and Burke (1990) predicted that older adults rely on other sources of priming. Specifically they state that:

“...age differences are likely to be pronounced when a node critical to a task receives priming from only a single source or connection within the network. Age-linked transmission deficits are very likely to affect performance in such a task because no other sources of priming will be able to offset the reduced priming across that critical connection.” (MacKay & Burke, 1990, pg. 251).

One of the other sources of priming that older adults may use to compensate for age-related transmission deficits that affect a single source of priming is the frequency-based priming transmitted by the neighboring word nodes. A word with low neighborhood frequency has neighboring words with a mean word frequency that is low, whereas a word with high neighborhood frequency has neighboring words with a mean word frequency that is high. Neighboring words, like the target word, prime the phonological nodes they are connected to as a function of their frequency via the symmetrical connections in NST. Connections that are frequently primed are maintained and become stronger over time, whereas less frequently primed connections become weaker over time. Phonological nodes connected to a word with neighbors that have low frequency would, therefore, receive less priming from those neighboring words than nodes connected to a word with neighbors that have high frequency. The interaction of neighborhood frequency with word frequency and with neighborhood density in older but not younger adults supports the “single-source factor” hypothesis proposed by MacKay and Burke (1990) and the age-related transmission deficit hypothesis. That is, older adults will compensate for the decrements in priming from one source or connection by relying on additional sources or connections for priming. The current results suggest that older adults may be sensitive to neighborhood frequency as an alternative source of priming in speech production, verifying the predictions of MacKay and Burke (1990) regarding the “single-source factor.”

General Discussion

The results from Experiment 1 replicated and extended the results of Harley and Bown (1998): More TOT states were elicited (from young adults) for words with sparse neighborhoods than for words with dense neighborhoods. More importantly, we demonstrated this effect independent of word length, a factor that was not controlled for in Harley and Bown (1998), and with stimuli that varied in neighborhood density based on a phonological rather than orthographic similarity metric. Furthermore, we manipulated another variable—neighborhood frequency—that Harley and Bown (1998) neither manipulated nor controlled. Although no significant effects of neighborhood frequency were observed in young adults, demonstrable influences of this variable were observed in Experiment 2 among older adults.

The results of the current set of experiments provide crucial insight into the process of lexicalization and how it may change over the life span. Specifically, older, but not younger adults showed a significant influence of neighborhood frequency on the number of TOT states elicited. The influences of a single source of priming interacting with age-related transmission deficits led MacKay and Burke (1990) to predict that older adults may adopt additional sources of priming to compensate for their less efficient processing. The results of Experiment 2 verify the prediction that older adults may adopt a form of compensation in order to maintain “normal” processing. That is, because of age-related transmission deficits, older adults are unable to obtain sufficient priming from neighborhood density and are therefore forced to rely on additional sources of priming such as those available from neighborhood frequency. The absence of an effect of neighborhood density for older adults may indicate a shift in weighting functions for these two sources of priming, with neighborhood frequency having greater weights for older than for younger adults. Although clearly speculative at this point, the proposal of age-related shifts in weighting functions for neighborhood density and neighborhood frequency are consistent with the current findings.

More definitive conclusions will need to await further research on factors that can influence the weighting functions for different sources of priming.

In addition to “adopting” alternative sources of priming, older adults (and to some extent, younger adults) may adopt conscious search strategies to resolve TOT states. Indeed, Reason and Lucas (1984) found that the only internal strategy used with any success to resolve TOT states was the generation of similar sounding words. Such conscious search-strategies through the lexicon may account for our finding (also observed by Harley & Bown, 1998) of interlopers having sparser neighborhood density, or fewer similar sounding words, than the targets. One might reason that if one is having a difficult time retrieving a lexical item (i.e., is in a TOT state), it must be because the item is extremely “rare,” or “unusual” in some way. This may result in a conscious search strategy for unique items in memory. Indeed, Burke et al. (1991) found that TOTs commonly occurred for proper names, which are highly unique items. If a similar search strategy were adopted for words, “lexical hermits,” or words that have very few phonological neighbors, may be retrieved, accounting for the finding that interlopers have sparser neighborhood density than targets. Harley and Bown (1998) provide a similar explanation for their findings. Such conscious search strategies may also over-ride biases that arise as a result of automatic processes, such as frequency and word-length biases.

Furthermore, the results from the current set of experiments adds to a growing body of literature suggesting that “interlopers,” phonologically related words that continually come to mind instead of the intended target word, support rather than block the retrieval of the target word (Harley & Bown, 1998; Meyer & Bock, 1992; Perfect & Hanley, 1992). If interlopers were to block the retrieval of words, one would predict that a word with many similar sounding words, a word with a dense neighborhood, would elicit more TOT states than a word with few similar sound words, a word with a sparse neighborhood. Exactly the opposite was observed (see also Harley & Bown, 1998), suggesting that similar sounding words support the retrieval of the target word. A word with a dense neighborhood receives more support than a word with a sparse neighborhood and, therefore, will most likely be completely retrieved.

One of the conclusions from our data—the number of similar sounding words (i.e., neighborhood density) acts to support rather than block the retrieval of word forms from lexical memory—coincides with the conclusions of Harley and Bown (1998). However, our explanation for why this is observed differs from the explanation posited by Harley and Bown. We (along with Burke et al., 1991; MacKay & Burke, 1990) propose that TOT states result from insufficient activation at the interface between word-forms and sub-lexical representations. In contrast, Harley and Bown propose that insufficient feedback between the lemma level (which represents semantic and/or syntactic information) and the lexeme level (which represents the phonological word-form) is responsible for a TOT state.

Although the data from the present experiments (and from Harley & Bown, 1998) can not directly distinguish between these two accounts, we propose that TOTs are due to insufficient feedback between word-forms and sub-lexical representations. Identifying the locus of TOTs at the interface between word-forms and sub-lexical representations allows us to easily account for our results in an extant model of cognitive processing, namely NST (Burke et al. 1991, MacKay, 1987), without having to posit modifications or additional assumptions to a model of speech production. In postulating that the locus of TOTs is at the interface between lemmas and lexemes, Harley and Bown may have to propose an additional mechanism to account for the decreased ability of older adults to specifically retrieve phonological information as observed in Experiment 2 (see also Burke et al., 1991; Maylor, 1990). It is unclear how such a difference across the life span could be accounted for if the locus of TOTs is at the interface between lemmas and lexemes. Further empirical work must ultimately establish the locus of TOTs.

The results from the current experiment also address whether the speech production system could best be described as a modular or an interactive system. We accounted for our results in the context of an interactive model of cognitive processing, the Node Structure Theory (Burke et al. 1991, MacKay, 1987). It is not entirely clear how a modular system that posits independent processing could account for the current results. More TOT states for words with few rather than many similar sounding words seems counter-intuitive to a strict modular account of speech production. Such accounts argue that processing at one level must be completed before processing at another level can proceed (e.g., Garrett, 1976; Levelt, 1989). Competition among many similar sounding words at the word-form level would seem to slow rather than speed processing, much as it does in spoken word recognition (e.g., Luce & Pisoni, 1998). Furthermore, it is unclear how multiple candidates in a modular model might be activated at the word-form level given that many models of speech production, including interactive models, do not posit lateral connections between word-forms (e.g., Dell, 1986; Garrett, 1976; Harley, 1993; Levelt, 1989). It is dubious how multiple word-form candidates might be activated in a modular system if a lemma, or semantic representation, activates only the lexeme, or word-form representation, connected to it without allowing feedback from phonological units as in an interactive model of speech production.

A critical finding from Experiment 2 is that we demonstrated that word frequency and neighborhood density are not the only variables that influence lexicalization (see Harley & Bown, 1998). Rather, we have demonstrated that several variables—word frequency, neighborhood density, and neighborhood frequency—affect lexicalization differently at various points in the life span. Additional work on lexicalization must be done to further examine the locus of TOTs and how the processes that operate during speech production change across the life span. All of the results from the current experiments are consistent with a well-specified and tested model of cognitive processing, namely the Node Structure Theory (MacKay, 1987; MacKay & Burke, 1990). Our results also verified a prediction of NST regarding the interaction of priming from a single source and age-related transmission deficits, suggesting that NST is a useful model for understanding the changes in the speech production system as a function of age. Additional work in the context of NST may provide insight into how changes in speech production may differ from the changes that occur in speech perception across the life span and shed light on the differences between the two cognitive systems.

References

- Bard, E. & Shillcock, R. (1993). Competitor effects during lexical access: Chasing Zipf's tail. In G. Altmann & R. Shillcock (Eds.), *Cognitive Models of Speech Processing* (pp. 235-275). Hove, Sussex: Erlbaum.
- Brown, A.S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin*, **109**, 204-223.
- Brown, R. & McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior*, **5**, 325-337.
- Burke, D.M. & Laver, G.D. (1990). Aging and word retrieval: Selective age deficits in language. In E.A. Lovelace (ed.) *Aging and Cognition: Mental Processes, Self-Awareness and Interventions*. (pp. 281-300) North Holland: Elsevier Science Publishers.
- Burke, D.M., MacKay, D.G., Worthley, J.S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language*, **30**, 542-579.

- Caramazza, A. & Miozzo, M. (1998). More is not always better: a response to Roelofs, Meyer, and Levelt. *Cognition*, **69**, 231-241.
- Coltheart, M., Davelaar, E., Jonasson, J.T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.) *Attention and Performance*, VI, pp. 535-555. London: Academic Press.
- Crystal, D. (1992). *An Encyclopedic Dictionary of Language and Languages*. London: Blackwell.
- 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.
- Garrett, M.F. (1976). Syntactic processes in sentence production. In R. Wales & E. Walker (Eds.) *New approaches to language mechanisms* (pp. 231-256). Amsterdam: North-Holland.
- Goldinger, S.D. and Summers, V.W. (1989) Lexical neighborhood in speech production: A first report. In *Research on Speech Perception Progress Report No. 15* Bloomington, IN: Speech Research Laboratory, Indiana University.
- Harley, T.A. (1984). A critique of top-down independent levels models of speech production: Evidence from non-plan-internal speech errors. *Cognitive Science*, **8**, 191-219.
- 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.
- Harley, T.A. & MacAndrew, S.B.G. (1992). Interactive models of lexicalization: Some constraints from speech error, picture naming, and neuropsychological data. In J. Levy, D. Bairaktaris, J. Bullinaria & D. Cairns (Eds.), *Connectionist Models of Memory and Language*, pp. 378-383. London; UCL Press.
- 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, & Cognition*, **20**, 824-843.
- Jones, G.V. (1989). Back to Woodworth: Role of interlopers in the tip of the tongue phenomenon. *Memory & Cognition*, **17**, 69-76.
- Jones, G.V. & Langford, S. (1987). Phonological blocking in the tip of the tongue state. *Cognition*, **25**, 115-122.

- Kucera, H. & Francis, W.N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- 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.
- Levelt, W.J.M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Luce, P.A. & Pisoni, D.B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, **19**, 1-36.
- MacKay, D.G. & Burke, D.M. (1990). Cognition and aging: A theory of new learning and the use of old connections. In T.M. Hess (ed.) *Aging and Cognition: Knowledge Organization and Utilization*. (pp. 213-263) North-Holland: Elsevier Science Publishers.
- MacKay, D.G. (1987) *The Organization of Perception and Action: A Theory for Language and other Cognitive Skills*. New York: Springer-Verlag.
- Maylor, E. A. (1990). Age, blocking and the tip of the tongue state. *British Journal of Psychology*, **81**, 123-134.
- McClelland J.L. & Rumelhart, D.E. (1981) An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, **88**, 375-407.
- Meyer, A.S. & Bock, J.K. (1992). The tip of the tongue phenomenon: Blocking or partial activation? *Memory and Cognition*, **20**, 715-726.
- Nusbaum, H.C., Pisoni, D.B. & Davis, C.K. (1984). Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words. In *Research on Speech Perception Progress Report No. 10*. Bloomington, IN: Speech Research Laboratory, Indiana University.
- Oldfield, R.C. & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Psychology*, **17**, 273-281.
- Perfect, T.J. & Hanley, J.R. (1992). The tip of the tongue phenomenon: Do experimenter-presented interlopers have any effect? *Cognition*, **45**, 55-75.
- Pisoni, D.B. Nusbaum, H.C., Luce, P.A., & Slowiaczek (1985) Speech perception, word recognition and the structure of the lexicon. *Speech Communication*, **4**, 75-95.
- Rastle, K.G. & Burke, D.M. (1996). Priming the tip of the tongue: Effects of prior processing on word retrieval in young and older adults. *Journal of Memory and Language*, **35**, 586-605.
- Reason, J. & Lucas, D. (1984). Using cognitive diaries to investigate naturally occurring memory blocks. In J.E. Harris & P.E. Morris (Eds.) *Everyday Memory Actions and Absent-Mindedness*. (pp. 53-70). London: Academic Press.

- Roelofs, A., Meyer, A.S., & Levelt, W.J.M. (1998). A case for the lemma/lexeme distinction in models of speaking: comment on Caramazza and Miozzo (1997). *Cognition*, **69**, 219-230.
- Salthouse, T.A. (1985). *A theory of cognitive aging*. Amsterdam: North-Holland.
- Sommers, M.S. (1996). The structural organization of the mental lexicon and its contribution to age-related declines in spoken-word recognition. *Psychology and Aging*, **11**, 333-341.
- Sommers, M.S. & Danielson, S. M. (1999). Inhibitory processes and spoken word recognition in young and older adults: The interaction of lexical competition and semantic context. *Psychology and Aging*, **14**, 458-472.
- Stemberger, J.P. (1984) Structural errors in normal and agrammatic speech. *Cognitive Neuropsychology*, **1**, 281-313.
- Stemberger, J.P. (1985) The reliability and replicability of speech error data: A comparison with experimentally induced errors. In *Research on Speech Perception Progress Report No. 11*, Bloomington, IN: Speech Research Laboratory, Indiana University.
- Stemberger, J.P. & MacWhinney, B. (1986). Frequency and the lexical storage of regularly inflected forms. *Memory and Cognition*, **14**, 17-26.
- Vitevitch, M.S. (1997a). Tongue twisters reveal neighborhood density effects in speech production. In *Research on Spoken Language Processing Progress Report No. 21*. Bloomington, IN: Speech Research Laboratory, Indiana University.
- Vitevitch, M.S. (1997b). The neighborhood characteristics of malapropisms. *Language and Speech*, **40**, 211-228.
- Vitevitch, M.S., Luce, P.A., Pisoni, D.B., & Auer, E.T. (1999). Phonotactics, neighborhood activation and lexical access for spoken words. *Brain and Language*, **68**, 306-311.
- Webster's New Collegiate Dictionary (1979). Springfield, MA: G. & C. Merriam Company.
- Woodworth, R.S. (1929). *Psychology* (2nd rev. ed.) New York: Holt.
- Wright, R. (1997). Lexical competition and reduction in speech: A preliminary report. In *Research on Spoken Language Processing Progress Report No. 21*. Bloomington, IN: Speech Research Laboratory, Indiana University.
- Yaniv, I. & Meyer, D.E. (1987). Activation and metacognition of inaccessible stored information: Potential bases for incubation effects in problem solving. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **13**, 187-205.
- Zipf, G.K. (1965). *The psycho-biology of language: An introduction to dynamic philology*. Cambridge, Mass., M. I. T. Press.

