

Effects of Neighborhood Density on Adult Word Repetition

Skott E. Freedman¹ & Jessica A. Barlow²

¹ Department of Speech-Language Pathology and Audiology, Ithaca College, USA

² School of Speech, Language, and Hearing Sciences, San Diego State University, USA

Correspondence: Skott E. Freedman, Department of Speech-Language Pathology and Audiology, Ithaca College, 953 Danby Road, Ithaca, NY 14850, USA. Tel: 607-274-3731. E-mail: sfreedman@ithaca.edu

Received: January 9, 2013 Accepted: January, 2013 Online Published: March 25, 2013

doi:10.5539/ijel.v3n2p1

URL: <http://dx.doi.org/10.5539/ijel.v3n2p1>

Abstract

Presumable lexical competition has been found to result in higher perceptual accuracy for words with few versus many neighbors. Previous studies have typically only analyzed the lexical-semantic level, however. In order to also explore the possibility of phonological effects, a word repetition task was administered to 46 typical adults in which 80 stimuli differed only in neighborhood density. In contrast to previous studies, verbal responses were elicited in order to analyze productions holistically and segmentally at the phonological level. An additional error analysis examined differences in neighborhood density between target words and substitutions. Findings revealed that words with more neighbors *facilitated* recognition, and were more accurately repeated than those with fewer neighbors. When a target word was misperceived, its substitution tended to be higher in neighborhood density, unrelated to word frequency. In order to interpret these results, an account of lexical competition is re-visited with consideration of characteristics of the lexicon discovered using graph theory (Vitevitch, 2008).

Keywords: neighborhood density, word repetition, lexical competition, phonology

1. Introduction

1.1 Neighborhood Density

The notion that competition may exist between several candidates during perception dates back to the 1950's, when Selfridge (1959) proposed a structural network foundation dubbed "pandemonium architecture." Developed further, Jackson (1987) theorized that certain "dormant demons" in a network might be summoned to the spotlight depending on the demands of more "active" ones. Considering the lexicon itself is a network, many items in the lexicon may become activated during speech perception in addition to an intended target word (MacWhinney, 1988). Work by Vitevitch and colleagues has consistently shown that, potentially due to such lexical competition, words with fewer "competitors" are identified more quickly compared to forms facing more competition (e.g., Vitevitch & Luce, 1998, 1999). The effects of a variety of lexical and sublexical variables have been investigated in this respect, such as the frequency of a word's occurrence in a language (Cluff & Luce, 1990; Savin, 1963), the probability of a word's sounds occurring in a language (i.e., phonotactic probability; Vitevitch, Luce, Charles-Luce, & Kemmerer, 1997), and how phonologically similar words are to one another (Benki, 2003; Vitevitch & Luce, 1998, 1999). It is this last variable which forms the basis of the current study.

Neighborhood density (ND) has been operationalized as the number of meaningful words, or *neighbors*, that are present for a given word in a language by adding, substituting, or deleting a phoneme in any word position (Vitevitch & Luce, 1998, 1999; cf. Luce & Pisoni, 1998, for an alternative index of similarity related to phonetic overlap). Consider the English word "cat" /kæt/. By substituting the first sound, the words "mat" /mæt/, "bat" /bæt/, "sat" /sæt/ (and so on) would be considered neighbors. Similarly, "at" /æt/ is a neighbor of "cat" by deleting the initial segment, as are "can" /kæn/ and "cap" /kæp/, by substituting the final sound. A word with many neighbors, such as the above example of "cat", is said to reside in a *dense* neighborhood and have high ND. In contrast, the English word "sniff" /snɪf/, which has few neighbors (e.g., "stiff" /stɪf/, "snuff" /snʌf/), is considered to reside in a *sparse* neighborhood and have low ND.

1.2 Background Literature

Research focusing on ND has been conducted in English across a broad range of language processes, particularly

in the areas of phonological acquisition (Hogan, Bowles, Catts, & Storkel, 2011; Morrisette & Gierut, 2002), word learning (Storkel, Armbruster, & Hogan, 2006; Storkel & Lee, 2011), speech production (Gahl, Yao, & Johnson, 2011; Vitevitch, 2002), and most relevant to the present study, speech perception (Luce & Pisoni, 1998; Vitevitch & Luce, 1998, 1999). In one of the first investigations of its kind, Luce and Pisoni (1998) presented monosyllabic CVC words to adults in the presence of white noise. Stimuli varied in a number of variables, including ND, and were presented at one of three signal-to-noise ratios. Results indicated that words with high ND were repeated less accurately than words with low ND, arguably due to a greater number of competing lexical forms for words with high ND. Speech productions were interestingly not elicited in the study. Instead, participants were provided up to 30 seconds to type their perceptions. This time delay is potentially problematic given that Savin (1963) found that when participants were unsure of presented words, their incorrect responses tended to be higher in word frequency than the target stimuli. It is possible then that some participants in Luce and Pisoni (1998) may have eliminated initially perceived words based on assumed infrequencies in the language. Additionally, the response analysis in the study was restricted to the holistic level (that is, based on the whole word), with responses scored as either correct or incorrect. Consequently, other aspects of recognition that may have affected perception, such as featural accuracy at the segmental level, were not considered.

In a follow-up study designed to further explore the relationship between ND and perception, Benkí (2003) altered the signal-to-noise ratio across four conditions and presented CVC words and nonwords differing in ND. This time, however, there was no time limit at all in which participants had to respond. Consistent with Luce and Pisoni (1998), findings revealed that words with low ND facilitated perception to a greater degree than words with high ND. In the study, stimuli were unmatched for PP and word frequency. Given that words with high ND are composed of similar, frequently-occurring sound sequences, an important positive correlation exists between PP and ND (Vitevitch, Luce, Pisoni, & Auer, 1999). Namely, words with low ND tend to have low PP (e.g., “beige” /beɪʒ/), while words with high ND tend to have high PP (e.g., “sick” /sɪk/). As such, effects of PP could have influenced the results reported in Benkí (2003). This is especially relevant given consistently robust findings of PP (Storkel, 2001; Storkel & Rogers, 2000; Vitevitch et al., 1997).

Finally, Taler, Aaron, Steinmetz, and Pisoni (2010) conducted a sentence repetition task in which four conditions varied both in ND and word frequency. Unlike in Luce and Pisoni (1998) and Benkí (2003), participants were asked to *verbally* repeat the presented stimuli. Sentences containing words with low ND were more accurately repeated relative to those with high ND. The authors concluded that words with high ND cause greater competition in spoken word recognition than words with low ND. One limitation to the study relates to the manner in which stimuli were presented, though. The experimenter manually presented each stimulus; once a participant decided to respond, the next stimulus was then presented. This raises the aforementioned concern regarding a lack of time limit in speech recognition tasks, during which other cognitive and linguistic variables may also be impacting participants' judgments.

In summary, existing limitations of prior work include type of elicited response, confounding stimuli factors, and response analysis. The present study aims to address these limitations, but first a discussion of a theoretical account of how ND may operate in the lexicon is warranted.

1.3 Graph Theory

Based on the findings presented in 1.2, it appears that ND influences perceptual identification in the presence of noise, with an advantage for perceiving words with low ND versus high ND due to possible lexical competition. If words with high ND have more similarly-sounding forms than words with low ND, and such forms are all competing for recognition, words with low ND would be most accurately identified. Yet, a recent analysis of the lexicon using graph theory (Vitevitch, 2008) seems to contrast this notion. Graph theory derives from mathematics and has been typically used by computer scientists and physicists to examine the structure of complex systems (for review, see Albert & Barabási, 2002; Barabási, 2002). Two key terms necessary for understanding graph theory are *nodes* and *links*. Words are represented by nodes, and links are the relationships between two nodes (neighbors). One of the main advantages of graph theory is the ability to compare complex systems with one another based on a number of measurements. One of these measurements is termed the *degree of assortative mixing*, which refers to the probability that highly connected nodes (i.e., words) are connected to other nodes that are also highly connected (Newman & Park, 2003). In a network with *disassortative* mixing, nodes with many connections are typically connected to nodes with relatively few connections. In contrast, in a network with a *positive* degree of assortative mixing, nodes with many connections tend to associate with similarly connected nodes.

Following the assumption that entries in the lexicon are stored in lexical neighborhoods (Luce & Pisoni, 1998;

Storkel & Morrisette, 2002), Vitevitch (2008) visually graphed approximately 20,000 words from the 1964 *Merriam-Webster Pocket Dictionary* according to ND. Among other things, he found that the lexicon has a positive degree of assortative mixing (Vitevitch, 2008); therefore, words that are highly connected (i.e., high ND) tend to be neighbors to one another and likewise for words with low ND. Vitevitch (2008) proposed that if a target form is temporarily inaccessible in the lexicon such that activation levels do not reach a certain threshold, a nearby lexical candidate may be selected. In other words, partial information about a target form can still be accessed (e.g., the first two sounds, the final consonant). Perhaps the many neighbors of words with high ND offer such information to a greater degree than those with low ND. Previous studies seem to support this notion (MacKay & Burke, 1990; Vitevitch & Sommers, 2003). For example, MacKay and Burke (1990) found that during tip-of-the-tongue states, participants were able to provide more segmental detail about words with high versus low ND. In conclusion, contrary to results reported in previous studies of ND and repetition, an analysis of the lexicon using graph theory appears to predict an advantage for recognizing words with high ND more accurately than those with low ND.

In the event of misperceptions, words with high ND might also be perceived more often than words with low ND. Given the positive degree of assortative mixing in the lexicon, a word with high ND might be the first candidate selected during perceptual uncertainty due to more available phonetic information from its numerous high ND neighbors.

2. The Present Study

Despite repeated findings of a perceptual advantage for words with low ND versus high ND, it is still uncertain how ND *alone* may impact perception, that is, without co-investigated variables such as word frequency and PP. It is also unclear how ND may affect perception at the featural level as well as during perceptual errors, particularly since response analysis in prior research has only used a binary criterion: correct or incorrect. The current study will add to our understanding of the nature of the lexicon by considering perceptual accuracy beyond the holistic level as well as the nature of misperceptions. First, a segmental measure of perceptual accuracy can offer unique insight into how ND might affect word recognition at a more fine-grained level. For example, although an overall effect of ND may exist during word recognition, participants may nonetheless perceive similar amounts of featural accuracy between words with low and high ND. A word with low ND (e.g., *badge*) may be more accurately perceived than a word with high ND (e.g., *back*), yet it is possible that participants correctly perceived a word-initial bilabial stop in both words. Analyzing participants' approximations of target words using a featural analysis at the segmental level can offer such information. Based on previous work and accounts of lexical competition believed to occur during adult perceptual identification (Taler et al., 2010; Vitevitch & Luce, 1998), it is expected that adults will repeat words with low ND with a greater degree of accuracy than those with high ND, using both holistic and segmental analyses at the phonological level. If words with many neighbors compete with one another during repetition, then advantages for words with low ND (with considerable less competition) should be observed regardless of the type of response scoring used.

Second, when a word is misperceived, is its substituted form likely to have lower, equal, or higher ND? Very little research currently exists in this area. Previous investigations have found that when a target word cannot be accessed, participants with word-finding impairment substitute forms that are higher in ND (German & Newman, 2004). Similarly, when misperceptions occur in the present study, it is predicted that participants will produce words that are higher in ND than the target. If this is found, yet words with low ND are perceived best, such findings would suggest that lexical competition only operates during *accurate* word recognition, when participants are more confident of what they heard. When accuracy suffers, and participants become more unsure of a presented target, lexical facilitation might result in substituted forms that tend to be higher in ND. Words with high ND could be perceived to a greater degree during such moments given their high resemblance to many other forms in the lexicon, as predicted by an analysis of the lexicon using graph theory (Vitevitch, 2008). In this case, high ND words would be expected to occur more often as substitutes for target words than low ND words.

Lastly, in order to control for interference of other cognitive processes (e.g., reasoning) and variables such as word frequency, verbal repetitions must be elicited within a short period of time. Since ND is believed to be represented early on at the lexical phonological level (Goldrick & Rapp, 2007), requiring participants to verbally repeat presented stimuli as soon as possible helps to minimize the influence of other factors. Additionally, an adequate number of misperceptions must be obtained by adding background noise to the stimuli. This is warranted noting that previous repetition studies have not always found effects of ND in ideal listening conditions due to ceiling effects. For example, Vitevitch and Luce (Experiment 4; 1998) found that participants repeated words/nonwords with low ND and high ND with similar rates of high accuracy (>80%). Likewise, Luce

and Pisoni (1998) argued that, without noise, a repetition task might simply be one of phonetic perception with minimal consequences for recognition. Thus, a repetition task without noise might fail to reveal sufficient misperceptions to answer the current research question.

Results of this study may offer implications for clinical practice, such as expanding our understanding of how items in the lexicon interact with one another. If words compete with one another, then words with high ND may require better discrimination trials and increased naming practice for adults with aphasia. In turn, words with low ND, with relatively fewer competitors, might be more accurately perceived and named by clients with naming deficits. There might be academic implications for the study's findings as well, given the acoustic quality of classrooms. Instructors using tasks of repetition, for example in English-as-a-Second-Language classes, might benefit from understanding which words are more easily confused with one another.

3. Method

3.1 Participant Characteristics

Forty-six college students (44 females, 2 males) at San Diego State University were recruited to participate in exchange for extra course credit. The average age of individuals participating in the study was 23 years ($sd=3$; $range=21-33$). Participants completed an in-depth questionnaire which addressed their speech, language, hearing, and overall development. All participants were native speakers of English and reported no history of speech, language, hearing, or cognitive impairment.

3.2 Stimuli

Eighty English words, half of which had low ND, and the other half of which had high ND, were used as stimuli (see Appendix A: Stimuli). Monosyllabic English words consisting of 3-5 phonemes served as initial candidates, and this list was narrowed down further based on a variety of sublexical and lexical variables in order to balance the two ND conditions (described in detail below). Consistent with previous investigations (Newman & German, 2002; Vitevitch & Luce, 1998), ND was determined by the number of words that could be created by deleting, substituting, or adding a single phoneme to a target item. ND was calculated with the *Irvine Phonotactic Online Database* (IPhOD; Vaden & Halpin, 2005; <http://www.iphod.com>), an online database of 33,432 words which offers a variety of information about a word's lexical and sublexical properties.

Stimuli were divided at the median value for ND; all words below the median were considered to have low ND, while those above the median value were classified as containing high ND. These classifications are consistent with prior methods of determining low and high conditions for ND (e.g., Storkel et al., 2006). In the low ND condition, the mean ND was 6.8 neighbors ($sd=2.0$; $range=3-9$), and the mean ND in the high ND condition was 19.7 neighbors ($sd=7.3$; $range=12-43$). An independent-sample t-test confirmed that words in the low ND condition had significantly fewer neighbors than those in the high ND condition, $t(78)=10.81$, $p<.001$.

The word frequencies of each word's neighbors (i.e., neighborhood frequency) was also calculated for each condition. For words with low ND, the mean frequency-weighted ND was 6.39 ($sd=2.37$); for words with high ND, the mean frequency-weighted ND was 17.93 ($sd=7.15$). An independent-samples t-test confirmed that the high frequency-weighted ND condition had a significantly greater number of neighbors than the low ND condition, $t(78)=9.70$, $p<0.001$. Note that the means in each condition were similar to the original means using the traditional one-phoneme metric of ND. As such, confounding effects due to neighborhood frequency would be less likely to occur.

As described earlier, it was crucial to control for many other confounding variables that have shown to affect speech perception and production (Newman & German, 2002; Savin, 1963; Vitevitch et al., 2004; Vitevitch & Luce, 1998, 1999). Although stimuli differed in ND, they were not significantly different in any of the following factors (all $ps>.05$; see Appendix B: Stimuli Control), as calculated with the IPhOD (Vaden & Halpin, 2005; <http://www.iphod.com>):

- a) word frequency (how frequently a word occurs in a language),
- b) stress-weighted phonotactic probability (stress-weighted probability of a sound's occurrence and co-occurrence with other sounds in a language, as indexed by positional segment frequency and biphone frequency, respectively),
- c) word length (number of phonemes; number of syllables),
- d) grammatical class (nouns, verbs, or adjectives),
- e) phonological composition (e.g., number of word-initial and word-final consonant clusters, sound class),

- f) morphology (number of morphemes), and
- g) duration.

In addition to the above variables, it was necessary to control for sensitivity to various initial segments of stimuli. Benki (2003) found that the advantage of perceiving words with low ND was attributed to the initial two segments of overlap between target words and competitors. In order to avoid a confounding effect of word position, words with low and high ND were therefore matched for initial phonological segments in the present study. For instance, a word with low ND was “grudge”; a word with high ND was “grand”.

All stimuli were digitally recorded directly to a Roland Edirol R-09 digital recorder at a sampling rate of 44.1 kHz, using a high quality, bidirectional microphone. Stimuli were recorded in a sound-attenuated booth (IAC Controlled Acoustical Environment). A female native speaker of English, who speaks with a General American English dialect, read aloud each item at a normal speech rate, taking care to pronounce each form with a similar inflection. Differing prosodic intonations due to list-reading were further minimized by including “filler” words at the beginning and end of each word list. Stimuli were normalized using Adobe© Audition recording software, adjusting the overall level of each stimulus so that the amplitude peaks were similar, then spliced into individual 16-bit WAV files at a sampling rate of 44.1 kHz. Words were randomized using a true random number generator. A screening of the words was then conducted with a different native speaker of English, who identified single-word speech samples with 100% accuracy. This indicated that the recordings were intelligible and of high quality.

In order to dilute the listening signal, speech spectrum shaped noise was mixed with the stimuli at an appropriate signal-to-noise ratio (SNR, described below). Briefly, speech spectrum-shaped noise has a spectrum which approximates the average long spectrum heard in typical speech, and serves as a more realistic type of noise (versus white noise, for example) that occurs in everyday situations (Crandell, 1991; Plomp, 1986). Consistent with previous related tasks (e.g., Choi, Lotto, Lewis, Hoover, & Stelmachowicz, 2008), the noise began 50 milliseconds before the beginning of each word, and lasted for 50 milliseconds following the end of each word.

The next step involved identifying an appropriate SNR for the task. Three SNRs were piloted with native speakers of English: -2 dB, -3 dB, and -4 dB. Overall repetition accuracy was 90% at -2 dB, 30-40% at -3 dB, and 20% at -4 dB. Based on pilot testing, -3 dB was selected as the target SNR for this task, considering that other SNRs might yield accuracy levels that were either too low or too high. Furthermore, Benki (2003) and Taler et al. (2010) found a main effect of ND using a similar SNR.

3.3 Procedure

All participants attended one session and were tested individually. Participants were told that they would be hearing words over some background noise, and were instructed to repeat each word as quickly and accurately as possible. All participants were strongly encouraged to produce a response even when they were uncertain of a word; this ensured a sufficient number of repetition errors to analyze. A high-quality, bidirectional microphone was positioned in close proximity to the participant’s lips. Presentation of stimuli was controlled by Sound Studio© digital editing software and played over Sony MDR-7506 headphones at a comfortable listening level.

Each participant received three practice trials. These trials were used to familiarize participants with the nature of the task; practice trials were not included in the final data analysis. Pilot studies confirmed that three trials were sufficient to familiarize participants with the task. The research protocol was then administered. Stimuli presentation was randomized for each participant using a true random number generator. After each stimulus was presented, participants had exactly three seconds to repeat what they heard (Vitevitch & Luce, 1999). Otherwise, a “no response” was scored and the next word was automatically presented. By limiting the amount of time a participant could respond, cognitive-linguistic processes which may have intervened during the task (i.e., word frequency effects) were minimized. Note that this methodology differs from previous studies, where participants had up to 30 seconds to respond or even no time limit at all. All single-word speech samples were digitally recorded at a sampling rate of 44.1 kHz directly to a Roland Edirol R-09 digital recorder.

3.4 Dependent Measures

A participant’s responses were phonetically transcribed by the first author, a native English speaker and speech-language pathologist trained in English phonetics and phonology. Inter-rater transcription reliability was calculated for approximately 15% of each participant’s productions by a research assistant trained in English phonetic transcription. Mean point-to-point transcription agreement reached 98% between listeners (sd=3%; range=92%-100%).

Three dependent variables were measured for all responses: 1) binary perceptual accuracy, 2) segmental

perceptual accuracy, and 3) difference scores between targets and misperceptions.

The first analysis evaluated overall perceptual accuracy. This type of analysis, consistent with earlier studies of this nature (Luce & Pisoni, 1998), allowed for the possibility that some repetition errors could be related to a word's ND. Using a binary criterion (yes, no), participants' responses were marked as correct if all phonetic segments identically matched the target. Responses were otherwise marked as incorrect, including no responses. Dialectal variations were not penalized (e.g., /kɔlz/ and /kɔlz/ were accepted for "calls").

A second analysis considered perceptual accuracy with respect to featural properties of the sounds in target words. Note that this was a more fine-grained analysis relative to previous studies that examined effects of ND only at the holistic level. Following Edwards, Beckman, and Munson (2004), each consonant in a participant's repetition was coded for accuracy on a 3-point scale: place of articulation, manner of articulation, and voicing. Each vowel was also coded for accuracy on a 3-point scale: dimension (front, center, back), height (high, mid, low), and length (lax, tense). One point was awarded for each correct feature; thus, each phoneme could receive a maximum of 3 points. In the example of the word "wheat" being repeated as [wik], a participant would receive 3 points for target /w/ (1 point for correct voicing: voiced; 1 point for correct place: bilabial; 1 point for correct manner: glide), 3 points for target /i/ (1 point for correct dimension: front; 1 point for correct height: high; 1 point for correct length: tense), and 2 points for target /t/, produced as [k] (1 point for correct voicing: voiceless; 1 point for correct manner: stop; 0 points for incorrect place) for a total of 8/9 points. As before, dialectal and common variations were not penalized. Lastly, 1 point was deducted from a phoneme if an epenthetic segment (i.e., an additional sound not part of the target) occurred directly before it (or in the case of a word-final addition, directly after it). Inter-rater scoring reliability was calculated for approximately 15% of each participant's productions, including consonants and vowels, by a research assistant trained in English phonetics. Mean point-to-point scoring agreement reached 97% between scorers (sd=2%; range=92%-100%).

Finally, a third analysis explored the nature of a misperception (e.g., "chest" for "taste") with regards to ND. This comparison was conducted in order to investigate the possibility that erred perceptions might be lower or higher in ND than the target. Such findings could indicate an inhibitory or facilitory influence of ND on misperceptions. Using the previously mentioned database (Vaden & Halpin, 2005), ND was calculated for all repetitions regardless of accuracy. "No responses", nonwords (e.g., [flit] for "fruit"), or words not found in the database (e.g., "drudge") were excluded from such analyses. These responses accounted for 1.06%, 0.71%, and 0.82% of the data, respectively.

3.5 Analysis

Four accuracy scores were calculated for each participant: 1) binary accuracy for words with low ND, 2) binary accuracy for words with high ND, 3) segmental accuracy for words with low ND; and 4) segmental accuracy for words with high ND. Binary accuracy was determined by calculating how many words were correctly repeated out of the 40 possible targets in each condition. For the segmental analysis, each target was assigned a total possible number of points, with 3 points assigned per phoneme. Average scores for segmental accuracy were then calculated by dividing the total number of points each participant received in each condition by the total number of possible points in each condition. Accuracy scores for binary and segmental accuracy were converted to a proportion score and then examined as a function of condition. Proportions were arcsine-transformed to approximate a normal distribution. A separate analysis of individual means revealed that all participants performed similarly to the overall group (i.e., within two standard deviations of the mean). In order to conduct the error analysis, differences in ND were calculated between all targets and substitutions for each participant.

A paired samples t-test was completed on the transformed data for each dependent variable (binary accuracy, segmental accuracy, difference in ND) in order to investigate the influence of ND on word repetition. All statistical tests were conducted using Bonferroni adjusted alpha levels of .0167 per test (.05/3). The effect size for each analysis was calculated using Cohen's *d* (1988); sizes were deemed small (.2), medium (.5), or large (.8).

4. Results

4.1 Binary Perceptual Accuracy

The first analysis considered how ND might influence perceptual accuracy using a binary, holistic measure. The average accuracy rates for the binary analysis are displayed in Table 1. There was a main effect of ND, $t(45)=5.87, p<.001, d=.8$. Unexpectedly, words with high ND were more accurately repeated than words with low ND.

4.2 Segmental Perceptual Accuracy

Featural differences in responses were evaluated relative to the target. This was necessary to test effects of ND during word recognition beyond the holistic level. The average accuracy rates for the segmental analysis are provided in Table 1. A significant effect for ND was found, $t(45)=7.46$, $p<.001$, $d=.9$. Again, participants repeated words with high ND with greater accuracy than words with low ND.

Table 1. The mean percentage accuracy rates and standard deviations for word repetition by analysis and condition

| | Binary | Segmental |
|---------|--------------|--------------|
| Low ND | 33.21 (8.13) | 73.56 (5.24) |
| High ND | 40.27 (9.34) | 78.05 (4.44) |

As shown in Table 1, repetition accuracy was highest on words with high ND.

4.3 Target-Misperception Comparison

A final analysis compared the ND of misperceptions with those of target words. The means for targets and errors are presented in Table 2. The difference in ND was significant, $t(45)=19.19$, $p<.001$, $d=.8$. As predicted, erred repetitions (due to misperception) tended to be higher in ND than the target form, which is a novel finding.

Table 2. The means and standard deviations of ND for adult word repetition by response type

| Target Words | Substitutions |
|--------------|---------------|
| 13.24 (8.40) | 18.38 (1.82) |

As seen in Table 2, substitutions were higher in ND compared to the target words.

5. Discussion

The primary purpose of the current study was to determine the independent influence of ND on adult word recognition in a noisy condition, using both holistic and segmental analyses at the phonological level, as well as consider how ND may impact the nature of misperceptions. Importantly, a host of confounding factors were necessarily controlled for in the stimuli, notably PP and neighborhood frequency. Additionally, immediate verbal responses were elicited in order to reduce the likelihood of cognitive-linguistic processes impacting perceptual judgment.

5.1 Lexical Facilitation

Studies in the past have proposed that lexical competition can account for the observed perceptual advantages for words that are low in ND. If a word has many neighbors, greater competition might exist for recognition, hindering accurate perception. Yet in the present study, words with high ND were repeated most accurately. What can account for this finding?

Recall from Vitevitch's (2008) analysis of the lexicon using graph theory, that the lexicon has a positive degree of assortative mixing. Namely, words having a high number of connections (i.e., high ND) tend to be neighbors to one another; the same is true for words with low ND. Vitevitch (2008) hypothesized that in the event a target word does not receive sufficient activation levels, a similar candidate may be retrieved in its place. Following this line of thinking, partial phonological information of a target word could still be accessible such as the initial sound or medial vowel. Words with high ND have many more phonologically similar forms compared to words with low ND. Such information may therefore be even more accessible to the listener for words with high ND than low ND during degraded listening conditions. This might explain why words with high ND were more accurately repeated than those with low ND, which would become even more apparent using a fine-grained phonological analysis (i.e., segmental accuracy) versus a holistic analysis (Vitevitch & Luce, 1998, 1999).

Contrary to increased lexical competition (Luce & Pisoni, 1998; Vitevitch & Luce, 1998), results of the current study, especially considering the large effect sizes, suggest that there may be increased levels of *facilitation* in the lexicon during repetition tasks. Such possibilities have been demonstrated previously. Vitevitch (2002) found that more "slips of the tongue" were made on words with low ND versus high ND. He argued that during speech

production, multiple lexical forms are activated that facilitate, rather than compete with, one another. According to the current results, it seems that word repetition may also benefit from multiple activated lexical forms. The levels of activation that a target item receives from its neighbors, through shared phonological representations, might aid in final selection/repetition of the target word. Along the same lines, a phonologically similar word may be produced when the target item is not activated sufficiently.

Results obtained in the current study may differ from previous studies of this nature due to specific aspects of the task and stimuli. First, the nature of the task required immediate verbal responses to the stimuli, whereas prior research has allowed ample time for participants to respond. By requiring participants to verbally repeat stimuli quickly, this likely limited other variables from influencing perceptual judgments (e.g., word frequency; Savin, 1963). Second, the majority of previous studies have used CVC stimuli where stimuli differed not only in ND, but often other influential variables such as word frequency and PP. In order to better control for a variety of other influential factors, in particular the strong correlation between PP and ND, it was necessary in this study to expand the syllable structure of the stimuli. It is possible that the traditional one-phoneme metric of ND captures neighborhood structures of CVC stimuli differently than stimuli containing clusters. Lastly, differences in how ND itself has been defined could likely be responsible for the varying findings. Both Luce and Pisoni (1998) and Benkí (2003) did not operationalize ND using the one-phoneme metric, but by the *degree of phonetic overlap*. Dubbed “segmental confusability”, the neighbors of a target word were determined based on how similar individual segments were to one another. For example, a neighbor for /kʌt/ was /skɪd/; a neighbor for /dɔŋ/ was /tæg/. Note that under the more traditional one-phoneme metric, none of these word pairs would be considered neighbors. For example, /kʌt/ and /skɪd/ differ by three phonemes although both forms contain velar stops, lax vowels, and word-final alveolars. Still, it should be noted that, Luce and Pisoni (1998) and Benkí (2003) aside, nearly every other study of ND over the past decade has defined phonological similarity using a one-phoneme metric (Newman & German, 2002; Morrisette & Gierut, 2002; Storkel et al., 2006; Vitevitch, 2002, 2008; Vitevitch & Luce, 1998, 1999). As such, the present study defined ND in a manner that is more consistent with the literature. Past studies of perception using a traditional definition of ND were not conducted in noise; therefore, it is difficult to compare them with the current findings since differences in accuracy were not found in ideal listening conditions (Vitevitch et al., 2004; Vitevitch & Luce, 1998, 1999). Future work should attempt to replicate the present findings by using a traditional one-phoneme definition of ND during tasks of noise-induced repetition.

5.2 High ND Bias

Considering the present results are in direct contrast to prior work similar in nature, noting still the large effect sizes obtained in each analysis (0.8 or higher), yet another possibility is worth considering. Participants in the current study may implicitly have used a high ND-biased selection system where, in the event of uncertainty, substitutions were more likely to resemble other words. An explanation of why such a selection process was used could be related to signal degradation. Unlike clean signal conditions, under which participants can correctly perceive both low and high ND words (Vitevitch & Luce, 1998), the majority of stimuli in the current study (60-70%) were incorrectly perceived at the holistic level. Using a high ND-biased selection strategy would theoretically increase the chance of approximating a word more accurately versus producing a word that resembles few other words. By selecting responses from a pool of high ND words (i.e., due to overall higher lexical activation, for example), greater accuracy would be observed on stimuli high in ND versus low in ND. In other words, more items in such a pool would be similar to high ND stimuli relative to low ND stimuli. The same selection process could also explain why misperceptions of words were substituted with words higher in ND, even when targets were already high in ND.

5.3 Error Analysis

In addition to perceptual effects of ND, one of the central purposes of this study was to examine the nature of errors following misperceptions. Results indicated that participants were more likely to name words that were higher in ND than the target word. This was true for target words with low *and* high ND. Thus, even when a word with high ND was misperceived, a participant’s repetition tended to be even higher in ND. These findings are consistent with those of German and Newman (2004), who found that participants with word-finding impairment substituted words that are higher in ND than the target word. It seems possible that when a word is unknown (or when only partial phonemic detail is available), the substituted form has high levels of activation from many different neighbors. Words with high ND, receiving the greatest levels of activation from multiple forms, may be the most frequently selected candidates during such instances. Alternatively, the high ND-biased selection system mentioned in 5.2 could have resulted in these findings as well.

Lastly, it should be acknowledged that other factors may have influenced the results, such as word frequency. Prior work has showed that substituted words tend to be higher in word frequency than target forms (German & Newman, 2004). Vitevitch (2008) also cautions that word frequency may be correlated with ND (but only during the initial stages of lexical acquisition). In order to address this possibility, a post-hoc analysis was conducted comparing the word frequencies of targets with those of substitutions (Vaden & Halpin, 2005). No difference was found, $t(90)=0.15, p=.88$. Thus, although substitutions were higher in ND, they did not significantly differ in terms of word frequency.

5.4 Academic and Clinical Implications

In addition to theoretical implications, academic implications likewise exist for the present study. Consider first that speech spectrum-shaped noise was used in the present study, which is a realistic type of noise similar to that occurring in normal environments, including school classrooms (Crandell, 1991). Current findings revealed that words with high ND were more accurately perceived than those with low ND. Put a different way, words that do not phonologically resemble many other words in the lexicon are not repeated as accurately as those that do. This discrepancy may become important when considering classroom environments in schools and universities. When teaching subject areas that require auditory models and repetition (e.g., English as a second language), certain words may need better presentation. Otherwise, an individual's phonological representation and first productions of a word may be incorrect.

Populations outside of the classroom may also benefit from the current study. In therapeutic settings, words with high ND may be ideal targets for adults with naming deficits. Limited evidence has suggested that not only may the ND of a word impact naming errors, but that words with high ND can facilitate the naming process. Middleton and Schwartz (2010) found that an adult with aphasia demonstrated higher accuracy and fewer semantic errors while naming pictures corresponding to words with high ND versus low ND. Perhaps the facilitative nature of the lexicon, as found in the current study, helps to elevate levels of phonemic distinction during word retrieval. If words with high ND have a greater number of activated forms than words with low ND, many similarly-sounding forms may facilitate accurate selection of a target word. Clinically, speech-language pathologists could consider the ND of a target word and provide phonemic cues as necessary as a word-naming strategy. Studies in the future should further determine the interaction between naming and ND for adults with aphasia.

5.5 Limitations and Future Directions

While the current study offers insight into the influence of ND during word recognition, two limitations should be noted. First, given the relatively difficulty of the task (participants generally scored between 30-40% for overall accuracy), it would be informative to assess repetition when overall accuracy is higher. Given that no differences in repetition accuracy emerged during clean listening conditions (Vitevitch & Luce, 1998), it is predicted that effects of ND might be smaller if listening conditions were "cleaner." Second, the type of noise used in the present study may have influenced the findings. Although speech spectrum-shaped noise is more realistic than other types of noise (e.g., white noise), lab-manufactured noise is arguably more unrealistic compared to the natural acoustic environments of a classroom or restaurant. This would be an important factor to consider in future work.

Further investigation is warranted to determine if effects of ND on repetition are similar across development (preschoolers, adolescents), populations (typical, clinical), and different languages. For instance, current work in Spanish has yielded conflicting effects of ND. Vitevitch and Stamer (2006) reported an inhibitory effect of ND, while Baus, Costa, and Carreiras (2008) found a facilitatory effect. It is important to explore the extent to which influences of ND may be universal or language-specific. Additionally, it would be worthwhile to explore how using different definitions of ND can affect the outcome of a study. Given the conflicting findings of past and present studies (e.g., Luce & Pisoni, 1998), the manner in which phonological similarity is defined is a critical variable to consider. Finally, it would be informative to explore effects of ND on repetition in a more naturalistic setting, such as in classrooms or activity centers. This would help determine the robustness of ND in the presence of more realistic distractions, such as perturbations of pitch and noise in the surrounding environment, in a context full of reductions.

Acknowledgements

Thank you to all of the participants in the study, to Sarah Cercone for lending her voice to the project, and to Adam Jacobson and Erin Brown for lending their ears for analyses. We additionally thank Sonja Pruitt, Eric Baković, Rachel Mayberry, and Vic Ferreira, as well as Tiffany Hogan, Autumn McIlraith, and an anonymous reviewer for their comments on earlier aspects of this work. This research was supported in part by a National

Institute on Deafness and Other Communication Disorders training grant, an American Speech-Language-Hearing Foundation New Century Scholars Program Doctoral Scholarship, and the Sheila and Jeffrey Lipinsky Family Doctoral Scholarship (all awarded to the first author).

References

- Albert, R., & Barabási, A. L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74, 47-97. <http://dx.doi.org/10.1103/RevModPhys.74.47>
- Barabási, A. L. (2002). *Linked: The new science of networks*. Cambridge, MA: Perseus Publishing.
- Baus, C., Costa, A., & Carreiras, M. (2008). Neighbourhood density and frequency effects in speech production: A case for interactivity. *Language and Cognitive Processes*, 23, 866-888. <http://dx.doi.org/10.1080/01690960801962372>
- Benkí, J. R. (2003). Quantitative evaluation of lexical status, word frequency, and neighborhood density as context effects in spoken word recognition. *Journal of the Acoustical Society of America*, 113, 1689-1705. <http://dx.doi.org/10.1121/1.1534102>
- Choi, S., Lotto, A., Lewis, D., Hoover, B., & Stelmachowicz, P. (2008). Attentional modulation of word recognition by children in a dual-task paradigm. *Journal of Speech, Language and Hearing Research*, 51, 1042-1054. [http://dx.doi.org/10.1044/1092-4388\(2008\)076](http://dx.doi.org/10.1044/1092-4388(2008)076)
- Cluff, M. S., & Luce, P. A. (1990). Similarity neighborhoods of spoken two-syllable words: Retroactive effects on multiple activation. *Journal of Experimental Psychology*, 16, 551-563.
- Cohen, J. (1988). *Statistical power analysis for the behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Crandell, C. (1991). Classroom acoustics for normal-hearing children. Implications for rehabilitation. *Educational Audiology Monographs*, 2, 18-38.
- Edwards, J., Beckman, M. E., & Munson, B. R. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research*, 47, 421-436. [http://dx.doi.org/10.1044/1092-4388\(2004\)034](http://dx.doi.org/10.1044/1092-4388(2004)034)
- Gahl, S., Yao, Y., & Johnson, K. (2012). Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language*, 66, 789-806. <http://dx.doi.org/10.1016/j.jml.2011.11.006>
- German, D. J., & Newman, R. S. (2004). The impact of lexical factors on children's word finding errors. *Journal of Speech, Language, and Hearing Research*, 47, 624-636. [http://dx.doi.org/10.1044/1092-4388\(2004\)048](http://dx.doi.org/10.1044/1092-4388(2004)048)
- Goldrick, M., & Rapp, B. (2007). Lexical and post-lexical phonological representations in spoken production. *Cognition*, 102, 219-260. <http://dx.doi.org/10.1016/j.cognition.2005.12.010>
- Hogan, T. P., Bowles, R. P., Catts, H. W., & Storkel, H. L. (2011). The influence of neighborhood density and word frequency on phoneme awareness on 2nd and 4th grades. *Journal of Communication Disorders*, 44, 49-58. <http://dx.doi.org/10.1016/j.jcomdis.2010.07.002>
- Jackson, J. (1987). Idea for a Mind. *Siggart Newsletter*, 181, 23-26.
- Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19, 1-36. <http://dx.doi.org/10.1097/00003446-199802000-00001>
- MacKay, D. G., & Burke, D. M. (1990). Cognition and aging: A theory of new learning and the use of old connections. In T. Hess (Ed.), *Aging and cognition: Knowledge organization and utilization* (pp. 281-300). Amsterdam: North-Holland. [http://dx.doi.org/10.1016/S0166-4115\(08\)60159-4](http://dx.doi.org/10.1016/S0166-4115(08)60159-4)
- MacWhinney, B. (1988). Competition and teachability. In R. Schiefelbusch, & M. Rice (Eds.), *The Teachability of Language* (pp. 63-104). New York: Cambridge University Press.
- Morrisette, M. L., & Gierut, J. A. (2002). Lexical organization and phonological change in treatment. *Journal of Speech, Language, and Hearing Research*, 45, 143-159. [http://dx.doi.org/10.1044/1092-4388\(2002\)011](http://dx.doi.org/10.1044/1092-4388(2002)011)
- Newman, M. E. J., & Park, J. (2003). Why social networks are different from other types of networks. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 68, 036122.1-036122.8.
- Newman, R. S., & German, D. J. (2002). Effects of lexical factors on lexical access among typical language-learning children and children with word-finding difficulties. *Language and Speech*, 45, 285-317.

- <http://dx.doi.org/10.1177/00238309020450030401>
- Savin, H. B. (1963). Word frequency effects and errors in the perception of speech. *Journal of the Acoustical Society of America*, 35, 200-206. <http://dx.doi.org/10.1121/1.1918432>
- Selfridge, O. G. (1959). Pandemonium: A paradigm for learning. In R. Schiefelbusch, & A. M. Utley (Eds.), *Proceedings of the Symposium on Mechanisation of Thought Processes* (pp. 511-529). London: Her Majesty's Stationery Office.
- Storkel, H. L. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research*, 44, 1321-1337. [http://dx.doi.org/10.1044/1092-4388\(2001/103\)](http://dx.doi.org/10.1044/1092-4388(2001/103))
- Storkel, H. L., & Lee, S. (2011). The independent effects of phonotactic probability and neighborhood density on lexical acquisition by preschool children. *Language and Cognitive Processes*, 26, 191-211. <http://dx.doi.org/10.1080/01690961003787609>
- Storkel, H. L., & Morrisette, M. L. (2002). The lexicon and phonology: Interactions in language acquisition. *Language, Speech, and Hearing Services in Schools*, 33, 22-35. [http://dx.doi.org/10.1044/0161-1461\(2002/003\)](http://dx.doi.org/10.1044/0161-1461(2002/003))
- Storkel, H. L., & Rogers, M. A. (2000). The effect of probabilistic phonotactics on lexical acquisition. *Clinical Linguistics and Phonetics*, 14, 407-425. <http://dx.doi.org/10.1080/026992000415859>
- Storkel, H. L., Armbruster, J., & Hogan, T. P. (2006). Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech, Language, and Hearing Research*, 49, 1175-1192. [http://dx.doi.org/10.1044/1092-4388\(2006/085\)](http://dx.doi.org/10.1044/1092-4388(2006/085))
- Taler, V., Aaron, G. P., Steinmetz, L. G., & Pisoni, D. B. (2010). Lexical neighborhood density effects on spoken word recognition and production in healthy aging. *Journal of Gerontology: Psychological Sciences*, 65B(5), 551-560. <http://dx.doi.org/10.1093/geronb/gbq039>
- Vaden, K., & Halpin, H. (2005). *Irvine Phonotactic Online Dictionary, Version 1.3*. [Data file]. Retrieved from <http://www.iphod.com>
- Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 735-747. <http://dx.doi.org/10.1037/0278-7393.28.4.735>
- Vitevitch, M. S. (2008). What can graph theory tell us about word learning and lexical retrieval? *Journal of Speech Language Hearing Research*, 51, 408-422. [http://dx.doi.org/10.1044/1092-4388\(2008/030\)](http://dx.doi.org/10.1044/1092-4388(2008/030))
- Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in spoken word perception. *Psychological Science*, 9, 325-329. <http://dx.doi.org/10.1111/1467-9280.00064>
- Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, 40, 374-408. <http://dx.doi.org/10.1006/jmla.1998.2618>
- Vitevitch, M. S., & Sommers, M. S. (2003). The facilitative influence of phonological similarity and neighborhood frequency in speech production in younger and older adults. *Memory & Cognition*, 31, 491-504. <http://dx.doi.org/10.3758/BF03196091>
- Vitevitch, M. S., & Stamer, M. K. (2006). The curious case of competition in Spanish speech production. *Language and Cognitive Processes*, 21, 760-770. <http://dx.doi.org/10.1080/01690960500287196>
- Vitevitch, M. S., Armbruster, J., & Chu, S. (2004). Sublexical and lexical representations in speech production: Effects of phonotactic probability and onset density. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 514-529. <http://dx.doi.org/10.1037/0278-7393.30.2.514>
- Vitevitch, M. S., Luce, P. A., Charles-Luce, J., & Kemmerer, D. (1997). Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech*, 40, 47-62.
- 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. <http://dx.doi.org/10.1006/brln.1999.2116>

Appendices**Appendix A***Stimuli*

| Low ND Words | ND | High ND Words | ND |
|--------------|----|---------------|----|
| blind | 8 | blocks | 12 |
| brush | 8 | bride | 24 |
| caused | 5 | calls | 23 |
| charge | 7 | chest | 22 |
| cloth | 6 | clean | 19 |
| coined | 7 | coast | 16 |
| crawl | 7 | crowd | 12 |
| dance | 9 | dean | 42 |
| draw | 7 | dried | 16 |
| ends | 5 | ears | 30 |
| facts | 5 | failed | 22 |
| flesh | 9 | flight | 18 |
| fruit | 8 | freight | 19 |
| grudge | 3 | grand | 15 |
| joined | 5 | jet | 22 |
| lunch | 7 | luck | 27 |
| minds | 9 | meets | 16 |
| mixed | 7 | missed | 25 |
| parked | 7 | paint | 21 |
| plow | 9 | plays | 22 |
| pleased | 3 | please | 14 |
| plus | 6 | plot | 13 |
| proof | 3 | pride | 18 |
| proud | 8 | prize | 16 |
| rev | 8 | rights | 28 |
| scheme | 9 | scale | 13 |
| sharp | 6 | shares | 31 |
| silk | 9 | sick | 43 |
| sky | 9 | skin | 15 |
| slave | 5 | slight | 19 |
| speech | 3 | speed | 13 |
| spoke | 9 | spite | 20 |
| stem | 7 | stores | 15 |
| straw | 3 | strain | 12 |
| task | 9 | taste | 19 |
| tossed | 9 | tools | 16 |
| trend | 6 | trained | 12 |
| trot | 6 | trip | 19 |
| wished | 9 | waste | 15 |
| yarn | 6 | yard | 14 |

Appendix B*Stimuli Control*

| Variable | Means and Standard Deviations of Low Neighborhood Density Stimuli | Means and Standard Deviations of High Neighborhood Density Stimuli | Statistic |
|------------------------------|---|--|---------------------------|
| Word frequency | 45.2 (24.19) | 50.45 (16.11) | $t(78)=1.14$, $p=.26$ |
| Positional segment frequency | 0.23 (.05) | 0.23 (.05) | $t(78)=0.40$, $p=.69$ |
| Biphone frequency | .003 (.001) | .003 (.002) | $t(78)=0.58$, $p=.56$ |
| Number of phonemes | 4.05 (.50) | 4.00 (.51) | $t(78)=0.44$, $p=.66$ |
| Duration | 670 (102) | 704 (173) | $t(78)=1.06$, $p=.29$ |