Original Article

Language and Speech

Using Network Science and Psycholinguistic Megastudies to Examine the Dimensions of Phonological Similarity

Language *and* Speech
1–32

DOI: 10.1177/00238309221095455
journals.sagepub.com/home/las
SAGE

Nichol Castro
Department of Psychology, The University of Kansas, USA; Department of Communicative Disorders and Sciences, University at Buffalo, USA

Michael S. Vitevitch
Department of Psychology, The University of Kansas, USA

Abstract
Network science was used to examine different dimensions of phonological similarity in English. Data from a phonological associate task and an identification of words in noise task were used to create a phonological association network and a misperception network. These networks were compared to a network formed by a computational metric widely used to assess phonological similarity (i.e., one-phoneme metric). The phonological association network and the misperception network were topographically more similar to each other than either were to the one-phoneme metric network, but there were several network features in common between the one-phoneme metric network and the phonological association network. To assess the influence of network structure on processing, we compared the influence of degree (i.e., neighborhood density) from each of the networks on visual and auditory lexical decision reaction times obtained from two psycholinguistic megastudies. The effect of degree differed across network types and tasks. We discuss the use of each approach to determine phonological similarity and a possible direction forward for language research through the use of multiplex networks.

Keywords
Phonological similarity, phonological associate, misperception, network science, one-phoneme metric

Corresponding author:
Nichol Castro, Department of Communicative Disorders and Sciences, University at Buffalo, 3435 Main Street, 122 Cary Hall, Buffalo, NY 14214, USA.
Email: nicholca@buffalo.edu

1 Introduction

Phonological similarity influences a wide range of cognitive processes including working/short-term memory (Goh & Pisoni, 2003; Vitevitch et al., 2012), word learning (Stamer & Vitevitch, 2012; Storkel, 2004), spoken-word recognition by monolinguals and bilinguals (Luce & Pisoni, 1998; Sommers, 1996; Vitevitch, 2002; Vitevitch & Luce, 1998, 1999; for a review see Vitevitch & Luce, 2016), spoken word production in typically developed adults and in people with certain language disorders (Gordon, 2002; Munson & Solomon, 2004; Scarborough, 2012; Scarborough & Zellou, 2013; Vitevitch, 1997; Vitevitch & Stamer, 2006), and reading (Yates et al., 2004). Although effects of phonological similarity on cognitive processing have been observed and replicated in each of these (and other) domains, and typically produce effect sizes (e.g., Cohen's *d*) that are medium to large in size, definitions of phonological similarity vary across and within different research domains. The present study of English words sought to examine three different definitions of phonological similarity, and how those different definitions might be reflected in processing. We used the mathematical techniques from network science to construct networks reflecting the different definitions of phonological similarity and compared the resulting network structures. We then used data from psycholinguistic megastudies to examine how the different network structures might influence language-related cognitive processes.

1.1 Definitions of phonological similarity

Consider first the definition of phonological similarity as found in the phonological similarity effect studied by working memory researchers. In the phonological similarity effect, it is typically observed that lists of letters, such as *c, b, d*, and *v*, are recalled less accurately than lists of letters, such as *c, r, m*, and *k* (e.g., Conrad & Hull, 1964). The names of the letters *c, b, d*, and *v* are often described as "sounding similar" to each other, whereas the names of the letters *c, r, m*, and *k* are often described as "sounding dissimilar" to each other. More recent research has instead defined these phoneme-sized units in terms of the *phonetic features* that comprise them and found that acoustically similar lists of phonemes (e.g., /pa-ta-ka/) were recalled less accurately than lists of dissimilar phonemes (e.g., /fa-na-ga/) when the response did not require overt articulation (i.e., saying the items out loud). However, when the response did require an overt articulation performance for acoustically similar and articulatorily similar lists of phonemes (e.g., /da-la-za/) were both recalled less accurately than lists of dissimilar phonemes (Schweppe et al., 2011). The work of Schweppe et al. (2011) hints that "phonological similarity" is not a simple or monolithic construct. From this work, we see that "phonological similarity" includes not only how phonemes sound (i.e., acoustic-phonetic features) but also how a speaker must move the articulators to produce the phonemes (i.e., articulatory features).

Furthermore, we see that the phonological similarity effect studied by working memory researchers has also been observed when the definition of phonological similarity is operationalized as phoneme overlap. Lists of words that share phonemes, such as *cat, fad, pan*, and *map*, are typically recalled less accurately than lists of words that do not share phonemes, such as *bar, kid, sun*, and *toe* (Baddeley, 1966). Thus, even within the well-studied phenomenon of the phonological similarity effect studied by working memory researchers, we see variability in the definition of phonological similarity.

Now consider how phonological similarity is defined in studies of various language-related processes. Studies of bilinguals often compare the processing of cognates, or words from two different languages that are similar semantically and phonologically, such as *vampire* (English) and *vampiro* (Spanish), to the processing of noncognate words such as *beehive* (English) and the

Spanish equivalent *colmena* (Gollan & Acenas, 2004). In studies of visual word recognition (also known as reading) we see words that are homophones (i.e., pronounced the same), but not homographs (i.e., spelled differently, such as *you* and *ewe*) being defined as phonologically similar (Lindell & Lum, 2008). We also find perfect rhymes (i.e., words that are phonologically identical from the final stressed vowel onward) and imperfect rhymes (i.e., where vowels/consonants that follow the final stressed vowel may differ) in studies of poetry and rhyme processing (Knoop et al., 2021), further increasing the variability in what it means for words to be phonologically similar.

In studies of spoken word recognition, we see that words composed of phonemes that share some phonetic features are "phonologically similar" enough to influence processing via form priming, despite none of the phonemes in the prime and target overlapping. For example, Luce et al. (2000) showed that words like *pill* would interfere with the speed and accuracy with which a word like *tear* was responded. Note that each phoneme in the words *pill* and *tear* share several phonetic features (e.g., /*p*/ and /*t*/ are both voiceless stops, but differ in place of articulation), resulting in the activation of *pill* (the prime in the form-priming task) competing with and therefore slowing down responses to the target word, *tear*. Thus, two words may be "phonologically similar" even when they do not have any phonemes in common.

Perhaps, the most common way that phonological similarity is operationally defined in studies of various language processes is with a variant of Hamming or Levenshtein distance, which are metrics commonly used in information theory and computer science to compare two strings of symbols or characters (e.g., Greenberg & Jenkins, 1964; Landauer & Streeter, 1973; Luce, 1986; Pisoni et al., 1985). Using this metric, two words are said to be phonologically similar if they differ by the addition, deletion, or substitution of a single phoneme. Thus, the word *cat* would be considered phonologically similar to the words *scat* (via one-phoneme addition); *_at* (via one-phoneme deletion); and *fat, cot*, and *cab* (via one-phoneme substitutions) in addition to other words (either in the same language or in another language; Vitevitch, 2012).

Although the one-phoneme metric is simple to implement and widely used, it is far from perfect and has been criticized for several reasons. One criticism is that the metric is not "cognitively grounded," meaning that the simple computational metric does not closely match perceptual distances (Luce, 1986; Wieling et al., 2014). That is, in the one-phoneme metric, each phoneme is considered to be equal to all other phonemes (i.e., a consonant can be substituted for a vowel, and it is equally likely that one consonant will be replaced by another consonant). However, when looking at the confusion matrices of phoneme perception in noise reported by Miller and Nicely (1955), for example, one does not see a consonant being perceived as a vowel, and some consonants are more similar to each other than others. Similar patterns of substitution tend to be observed in speech production error data as well (Stemberger, 1992). Furthermore, the data in the confusion matrices reported by Miller and Nicely (1955) show clear asymmetries in substitution. That is, /k/ may often be misperceived as /p/, but /p/ is misperceived as /k/ less so. The one-phoneme metric does not capture such differences.

Another criticism against the one-phoneme metric points to the importance of the position in which the phoneme change occurs in the word (e.g., Fricke et al., 2016; Turnbull & Peperkamp, 2017). In the one-phoneme metric, a phoneme change at the beginning of the word is equivalent to a phoneme change at the end of a word. However, previous studies have demonstrated that processing is influenced by the number of positions in a word that can be altered to produce another word (Vitevitch, 2007), that the onsets of words are important in processing (Marslen-Wilson & Zwitserlood, 1989; Vitevitch, 2002; cf., Connine et al., 1993), and that the ends of words are important in processing (De Cara & Goswami, 2002). The one-phoneme metric does not capture the different influences that phoneme changes may have in different positions of a word.

Increasingly, more complex metrics have been proposed that include more parameters to capture various dimensions of "phonological similarity" (e.g., Hahn & Bailey, 2005; Strand, 2014; Suárez et al., 2011). In the present study, rather than proposing yet another computational metric, we instead used the tools of network science to explore some of the dimensions that contribute to the concept of "phonological similarity" and to examine how those different dimensions of phonological similarity might influence the way in which representations of words are organized in the mental lexicon, that part of long-term memory that stores language-related information.

1.2 Network science terminology and measures

To examine different dimensions of phonological similarity, the present work was inspired in part by the work of several researchers who used the techniques from network science to examine different definitions of semantic similarity by considering the structure found among words that are "semantically similar" to each other (for review, see Baronchelli et al., 2013; Beckage & Colunga, 2015; Siew et al., 2019; see also https://cqllab.upc.edu/biblio/networks/). A *network* (sometimes referred to as a graph in Mathematics) is composed of *nodes*, which represent some sort of entity, and *edges* between nodes (sometimes referred to as "neighbors") to represent some sort of relationship between the connected nodes. Semantic networks have been defined in a variety of ways based on different definitions of semantic similarity. Consider, for example, the work of Steyvers and Tenenbaum (2005), where three different networks of "semantically similar" words were examined. In one network, nodes were the cue and target words from a semantic association task (Nelson et al., 2004) with connections between cues and their respective targets. In the second network, nodes were words from Roget's (1911) Thesaurus and they were connected if they shared at least one class in common. In the final network, nodes were words found in WordNet (Miller, 1995), and they were connected if they were antonyms (e.g., *tall* and *short*), hypernyms (*red* is a *color*), or meronyms (*cats* have *tails*). Despite defining semantic similarity in three different ways, Steyvers and Tenenbaum (2005) found several structural similarities among the three different types of semantic networks.

In contrast to the study of semantic networks, there has been only one prominently studied phonological network; one based on the one-phoneme metric (Vitevitch, 2008). That is, nodes represent words in the dictionary and two words are connected if they differ by only one phoneme through addition, substitution, or deletion (Vitevitch, 2008). In the present study, we compared the network structure that results from using the simple computational metric of phonological similarity, to networks formed by using behavioral data to determine the "phonological similarity" between words (e.g., a phonological associate task, Neergaard & Huang, 2019; a phonological verbal fluency task, Neergaard et al., 2019). By comparing the structure of networks formed from using three different definitions of "phonological similarity," we hoped to gain further insight into the various dimensions that contribute to phonological similarity.

Analysis of networks requires careful definition of nodes and connections between nodes, relying heavily on theoretical motivation. Whereas nodes may be more straightforward in the case of language networks (i.e., words), connections between nodes can vary not just in the type of similarity (e.g., semantic vs. phonological), but in directionality and weighting. In terms of directionality, edges can be directed (also called *arcs*) or undirected. For example, consider a free association network based on free association data, where a cue word elicits a response word from participants. Here, an undirected edge would indicate bidirectionality, so an edge between node A and node B would imply that A (as a cue word) leads to response B and that B (as a cue word) leads to response A. On the other hand, a directed edge would indicate a relationship with directionality. For example, if only A (as a cue word) leads to response B (and B as a cue word does not lead to response A), then a directed edge (often represented as an arrow) would extend from node A to node B.

In addition, edges could be represented with weighting. An *unweighted* edge means that the strength of the relationship between nodes is not represented (or is assumed to be equal), while a *weighted* edge means that values of some sort are associated with the edge to represent the strength of the relationship between nodes. In the case of a free association network, weighting could represent the frequency with which a particular cue-response pair was produced across participants, such that greater weighting represents a more common cue-response pair and weaker weighting represents a less common cue-response pair. Finally, *degree* is simply a count of how many edges a node has.

Once the nodes and edges in the system of interest have been identified, a web-like structure, or network, emerges. Various measurements can then be made of individual nodes, of groups of nodes, or of the overall network structure. Critically, the way in which a network is structured can have important implications for not just how processing occurs in the network (Strogatz, 2001), but also for the growth and stability of the network over time (Albert et al., 2000; Newman, 2003b). In what follows, we outline several measures commonly used to assess the structure of a network, and provide examples based on the previously studied one-phoneme metric network given our focus on capturing different dimensions of phonological similarity (Vitevitch, 2008).

One way to characterize the overall structure of a network is to consider the *location of nodes*, particularly in networks that are not fully connected and contain disconnected subnetworks, or *components*. In networks that are not fully connected, nodes can reside in one of three locations: the giant component (i.e., the largest connected component of the network), smaller components (called lexical islands in phonological networks; Vitevitch, 2008), or they may be isolates that are not connected to any other nodes in the network (called lexical hermits in phonological networks; Vitevitch, 2008). Previous analyses of words in an English phonological network constructed using the one-phoneme metric (Arbesman et al., 2010; Vitevitch, 2008) found that 34% of the word nodes resided in the giant component of the phonological network in contrast to many other real-world networks where 80%–90% of nodes reside in the giant component (Newman, 2001). Having a relatively small giant component with many smaller components may contribute to the resilience of the phonological network to node removal (Arbesman et al., 2010). Furthermore, the lexical processing of words differs based on their location in the phonological network; Siew and Vitevitch (2016) found that words located in lexical islands were more quickly recognized and accurately recalled than words located in the giant component, even when controlling for a number of other psycholinguistic variables.

Another way to evaluate the overall structure of a network is to assess whether it has a *small-world structure* (Kleinberg, 2000; Watts & Strogatz, 1998), whose name comes from the "six-degrees of separation," small-world phenomenon described by Milgram (1967). The idea here is that two randomly selected nodes in the network are connected via a small number of connections by taking advantage of key structural features, such as short-cut paths and high clustering. In the pioneering work of Watts and Strogatz (1998), two network measures were calculated to determine if a network had a small-world structure: average shortest path length and average clustering coefficient. *Average shortest path length* captures the average distance (i.e., number of edges) between two nodes in the network and *average clustering coefficient* captures the relative amount of interconnectivity among the neighbors of a node. A network is said to have a small-world structure if: (1) the average shortest path length is about the same as the average shortest path length of a comparably sized network whose connections are placed at random and (2) the average clustering coefficient is several orders of magnitude larger than the average clustering coefficient of a comparably sized network whose connections are placed at random.

More recently, a statistical measure called small-worldness has been developed (Humphries & Gurney, 2008) that considers similar ratio comparisons between average shortest path length and average clustering coefficient of observed and comparably sized random networks. Small-world networks are interesting because search and spreading processes are extremely efficient and rapid on such networks compared to comparably sized networks that are structured in some other way (Kleinberg, 2000). Thus, having a small-world structure in the phonological network may be important for quick and efficient word retrieval in daily communication.

In addition to the small-world structure of a network, researchers often assess the overall network to determine if it exhibits a scale-free structure. In a scale-free network, most nodes in the network have few connections, but a few nodes have many connections, leading to a degree distribution that follows a power-law function. (The *degree distribution* of a network can be visualized by the histogram of node degree, depicting how many nodes have degree = 0, degree = 1, degree = 2, and so on). Scale-free networks are interesting because they have been shown to be resilient to attempts to damage the network (Albert et al., 2000; Barabasi & Albert, 1999). In the context of language networks, the damage could represent degraded connectivity between words or eventual loss of words. Although the semantic networks examined by Steyvers and Tenenbaum (2005) exhibited scale-free characteristics, the degree distribution of the phonological network formed with the one-phoneme metric (Arbesman et al., 2010; Vitevitch, 2008) was instead best fit by an exponential or truncated-exponential function (see also Broido & Clauset, 2019), rather than a power-law function.

Mixing patterns in a network, or the way in which nodes tend to connect to other nodes in a network, is another way to characterize the overall structure of a network. For example, in a social network, people tend to be connected to others who have the same gender, race, or age. Other properties of nodes, like degree, can be assessed for mixing patterns. In a network that shows *assortative mixing by degree*, nodes with high degree tend to connect to other nodes with high degree (and low-degree nodes tend to connect to other low-degree nodes). In a network that shows *disassortative mixing by degree*, nodes with high degree tend to connect to other nodes with low degree (and vice versa; Newman, 2003a). No mixing by degree indicates no relationship between the degree of a node and the degree of connected nodes.

Vitevitch (2008; see also Arbesman et al., 2010; Turnbull & Peperkamp, 2017) found that the phonological network formed with the one-phoneme metric exhibits assortative mixing by degree. This assortative mixing pattern is not just a feature of the network structure, but is also observed in linguistic behavior; participants were more likely to respond with a word of the same degree in a variety of psycholinguistic tasks (e.g., hear a high degree word and respond with a similar sounding high degree word; Vitevitch et al., 2014). Networks with assortative mixing by degree are hypothesized to be more resilient to targeted node removal (e.g., removing nodes with the highest degree first) than networks with disassortative mixing by degree (Newman, 2003a). Arbesman et al. (2010) found in phonological networks across a variety of languages that the average path length of the network remained relatively consistent when removing nodes by degree (highest degree first), indicating that the networks remained well-connected. This feature of a phonological network may also have important implications for the resilience of language processes despite certain neurological diseases, stroke, or even healthy aging.

Finally, *community structure* refers to how groups of nodes connect to each other beyond the immediate neighborhood. Specifically, communities are groups of nodes that tend to connect to each other more than they connect to nodes in another community (Newman & Girvan, 2004; Ravasz & Barabási, 2003). Siew (2013) extracted the community structure of the giant component in the phonological network formed with the one-phoneme metric, and found 17 communities ranging in size from 31 to 697 nodes. Furthermore, Siew (2013) found that words within a community tended to share similar biphone sequences (e.g., the sequences /ɨn/, /bɨ/ and /ɜb/ as found in the words "urban," "turbine," and "bourbon") with each other as compared to words in other communities. The community structure of a network could impact the ease of word retrieval and other language processes (e.g., Vitevitch et al., 2021), but this remains to be tested.

2 Research question

In the present study, we sought to examine different dimensions of "phonological similarity" using network science to assess how the organization of words in the mental lexicon might vary as a function of the different aspects of phonological similarity that each definition captures, akin to what has been done in the study of semantic similarity networks (see Steyvers & Tenenbaum, 2005). Using network science to study phonological similarity is not new, but has primarily used the one-phoneme metric to define edges in the network (see Vitevitch, 2008). However, studies have pointed out that certain topological features of phonological networks derived from the one-phoneme metric can also be found in randomly generated phonological networks (Brown et al., 2018; Gruenenfelder & Pisoni, 2009; Stella & Brede, 2016a), increasing the need to understand the dimensions of phonological similarity and how those dimensions might influence network structure and processing.

In addition to using the one-phoneme metric to define phonological similarity, we defined phonological similarity using behavioral responses from human participants in two different behavioral tasks to shed light on how phonological similarity may be represented in the mind. To obtain behavioral responses from human participants that elicited "explicit" responses to define phonological similarity, we used the phonological association task in which participants report which word(s) in their lexicon sounds similar to a cue word we provided to them. This task is similar to the commonly used semantic associate task (Nelson et al., 2000), but instead requires participants to provide a phonologically related word rather than a semantically related word as a response (see Supplementary Material for details on the phonological association task and data; see also Neergaard & Huang, 2019).

To obtain behavioral responses from human participants that elicited "implicit" responses to define phonological similarity, we used misperceptions of words heard in noise (Felty et al., 2013). We assumed that when participants misperceived a word presented in noise, they would produce a response that was phonologically similar to the stimulus, rather than a response that was completely unrelated to the degraded stimulus. Thus, we can infer what responses are phonologically similar to the target word by considering what participants produce on error trials. In contrast to the phonological association task, which explicitly asked participants to produce a word in response to the cue, participants in the identification of words in noise task were simply asked to indicate what they heard, which resulted in many word responses that we assumed were "phonologically similar" to the cue word, but occasionally resulted in nonwords as well. Given the perceptual nature of the identification of words in the noise task, we elected to include both the word and nonword responses in the network that we constructed to more fully capture a "perceptual" dimension of "phonological similarity."

Using the tools of network science, we analyzed a phonological association network (PAN), a misperception network (MPN), and a network constructed from the one-phoneme metric to examine the different dimensions that may contribute to "phonological similarity." Furthermore, we used the data from two psycholinguistic mega-studies (Balota et al., 2007; Goh et al., 2020) to examine how the different definitions of phonological similarity might also influence language processing. In an exploratory analysis, we considered the influence of degree (referred to as neighborhood density in the psycholinguistic literature), as calculated from each of these three different phonological networks, on the existing megastudy data from an auditory and a visual lexical decision task.

3 Network analysis of phonologically similar words

Because behavioral studies have found that various structural characteristics of the one-phoneme metric network influence the perception, production, and learning of spoken words (e.g., Chan & Vitevitch, 2009, 2010; Goldstein & Vitevitch, 2014, 2017; Vitevitch & Castro, 2015), it is important to consider how different dimensions of "phonological similarity" might influence the overall organization of words stored in the mental lexicon and how those potentially different network structures might influence various language-related processes. To that end, we constructed networks of words that were "phonologically similar" using (1) responses from a phonological association task (see Supplementary Material for details on the phonological association task and data) that "explicitly" captured phonological similarity (henceforth the PAN), (2) misperceived responses of words heard in noise that "implicitly" captured phonological similarity (Felty et al., 2013; MPN), and (3) as a point of comparison a computational, one-phoneme metric using the phonological association data (see also Vitevitch (2008); henceforth, the one-phoneme metric network, 1PN]. Then, we used an information-theoretic approach, called Network Portrait Divergence (Bagrow & Bollt, 2019), to obtain a coarse overview of whether these networks differed in their topological structure and considered a more fine-grained view of these networks by directly comparing several network measures discussed previously.

3.1 Method

3.1.1 Network creation. Phonological association data were used to construct the PAN. We provide a brief review of the data, but further details are provided in the Supplementary Material. In the phonological association task, participants were shown a cue word on the computer screen and responded by typing up to three words that sounded similar to the cue word. This task has been used in other studies to assess phonological similarity, where cues are words (also called a neighbor generation task; Muneaux & Ziegler, 2004; Neergaard & Huang, 2019; Vitevitch et al., 2014, 2016), syllables (Neergaard & Huang, 2019), or nonwords (also called a word reconstruction task; Cutler et al., 2000; Luce & Large, 2001; Van Ooijen, 1996). Because a large database of phonological associates was not available for analysis in the present study, we obtained phonological associates to 9,371 cue words, representing a range of word lengths, frequencies, and parts of speech, which resulted in the generation of 77,451 cue-response pairs. We retained all response words that were produced, regardless of morphological similarity (e.g., walk, walks, and walked) or derivation (e.g., discern and discernment). Because the present study is interested in phonological similarity, but not semantic similarity, homophones were treated as a single wordform (i.e., node in the network). Homographs (e.g., dove the noun and dove the verb) were rare; in such cases, we determined the intended pronunciation based on the cue word eliciting the response, such that two different wordforms (i.e., nodes in the network) could be included if appropriate. Consistent with previous studies (e.g., Luce & Large, 2001; Vitevitch et al., 2014, 2016), the phonological associate data obtained in the present study had a majority of cue-response pairs that differed by one or two phonemes, and a positive correlation between cue word length and response word length (as measured by the number of phonemes). To create the network in the present study, the PAN contained nodes representing the cue and response words ($N = 20,615$) from the phonological association task. Edges were placed between cue-response pairs ($N = 56,747$).

The cue and response words/nodes from the phonological association task were also used to construct the 1PN. Note that there were 40 cue words shown to participants that did not elicit a response (see Supplementary Material), and were therefore not included in the 1PN ($N = 20,575$). Edges in the 1PN were placed between any two words that differed by one phoneme through addition, substitution, or deletion (Luce & Pisoni, 1998), similar to the construction of the phonological network reported in Vitevitch (2008). Note that the 1PN assumes any cue word responded to and any response made would exist in the mental lexicon. Thus, connections in the 1PN can be placed between any two words, regardless if they were both cue words, both response words, and participant-generated cue-response pairs.

The MPN was constructed using misperception data from Felty et al. (2013). We briefly review the data here and refer readers to the original publication for more details. Participants were told that they would hear English words over headphones and were asked to identify them by typing the word that they heard. There was a total of 1,428 cue words, representing a range of word lengths, frequencies, and parts of speech, presented in noise (i.e., six-talker babble at one of three signal-to-noise ratios) across participants, which resulted in a total of 67,952 trials. Of these trials, 55.6% were correct (i.e., the participant provided the correct word response), and of the incorrect responses, only 26.7% were nonwords (27.1% if including foreign words as nonwords). In addition, incorrect responses tended to be phonologically similar to the cue word as measured by the one-phoneme metric, with the majority of incorrect responses being one or two phonemes different (Felty et al., 2013). To create the network in the present study, only the incorrect data were considered. The MPN contained nodes representing the cue and incorrect response words and nonwords ($N = 16{,}457$), and edges were placed between cue-response pairs ($N = 20{,}116$).

3.1.2 Network analysis. Python was used to calculate Network Portrait Divergence (*D*; Bagrow & Bollt, 2019), which allowed us to assess how similar in structure two networks were to each other. For each network, a Network Portrait (Bagrow et al., 2008) is created, which is simply a matrix where the columns denote the number of nodes and the rows indicate the different path lengths between a pair of nodes from 1 (immediate connections) to the maximum diameter of the network. Several pieces of information can be obtained from the Network Portrait, including its degree distribution (i.e., the number of nodes that have each value of degree from 0 to the maximum node degree found in the network) and community structure, or subgroups of densely connected nodes. The amount of difference between networks can be quantified based on a comparison between two network portraits using Jensen–Shannon divergence. This measure of divergence (*D*) is determined by comparing the shortest path distributions encoded in each portrait, and is particularly sensitive to capturing topological features of the network. *D* ranges from 0 to 1, with values closer to 0 indicating that the networks are more similar to each other and values closer to 1 indicating that the networks are more dissimilar from each other. One of the benefits of this network comparison approach is the ability to compare networks that vary in size (i.e., number of nodes and/or edges) and that do not overlap in nodes.

The "igraph" package (Csardi & Nepusz, 2006) in R was used to compute several structural measures on the individual networks. Specifically, we examined the distribution of nodes in various locations of the network (giant component, island, or hermit); the average shortest path length, average clustering coefficient, and small-worldness (*S*; Humphries & Gurney, 2008) of the giant component of each network to assess small-world structure; fitting of a power-law function and exponential curve to the degree distribution of each network to assess the scale-free structure; Pearson correlations between the degree of a node and the degree of its connected nodes to assess mixing by degree patterns; and the Louvain method (Blondel et al., 2008) was used to determine the number of communities in the giant component of each network, along with the modularity (*Q;* Clauset et al., 2004) of those partitions. Table 1 provides a description of how these network measures are calculated. In addition to those measures of network structure, we assessed how the location of words changed or remained the same across the networks with different definitions of phonological similarity, and calculated the Pearson correlation of the degree and clustering coefficient of nodes across networks.

Table 1. Network Measure Description.

Network measure	Description
Node Location	The number and percentage of nodes residing in each of three possible locations were reported: the giant component, an island, or as a hermit. Recall that the giant component is the largest, connected group of nodes in the network. Islands are smaller, connected groups of nodes (also simply referred to as "components" in the network science literature). And, hermits are nodes that have no connection to any other node in the network (also referred to as "isolates" in the network science literature)
Community Structure	There are several ways to determine the community structure of a network, with the Louvain method commonly used (Blondel et al., 2008). In the Louvain method, modularity (or the partitioning of the network into communities) is optimized by considering the density of edges within a community compared to edges outside a community. This is an iterative process of placing nodes into different communities and assessing changes in modularity until a significant improvement in modularity is reached. Importantly, modularity values higher than .3 are indicative of significant community structure (Clauset et al., 2004).
Small-Worldness	Small-world structure occurs when a network has a similar average shortest path length and larger average clustering coefficient than a comparably-sized random network (Watts & Strogatz, 1998), as determined by standard network analysis convention where differences in values greater than 1.5 times in magnitude indicate a significant difference. In addition, a statistical measure of small-worldness was also calculated following Humphries and Gurney (2008), with the following equation: $$S = \frac{C_g / C_{\mathrm{rand}}}{L_g / L_{\mathrm{rand}}},$$ where C is a measure of network clustering based on transitivity for the network g and a comparably-sized random network rand, and L is the average shortest path length of the network g and a comparably-sized random network rand. S values greater than 1 indicate a small-world structure
Degree distribution	The degree distribution of a network is described by the best fitting line of the number of nodes with each unique value of degree on a log-log plot. Many networks exhibit power-law functions (Albert et al., 2000), but an exponential or truncated-exponential curve was identified in the phonological network of Vitevitch (2008; see also Arbesman et al., 2010). A network is said to have a scale-free structure if the degree distribution is the best fit by a power-law function
Mixing by degree	Recall that there are three mixing by degree patterns: assortative mixing, disassortative mixing, and no mixing (Newman, 2003a). A Pearson's correlation between the degree of a node and each of its neighbors was used to determine the mixing by degree pattern, such that a positive correlation indicates assortative mixing, a negative correlation indicates disassortative mixing, and zero correlation indicates no mixing. Assortative mixing by degree occurs when nodes with a high degree tend to be connected to other nodes with high degree. On the other hand, disassortative mixing by degree occurs when nodes with high degree tend to be connected to nodes with low degree. No mixing by degree indicates no pattern

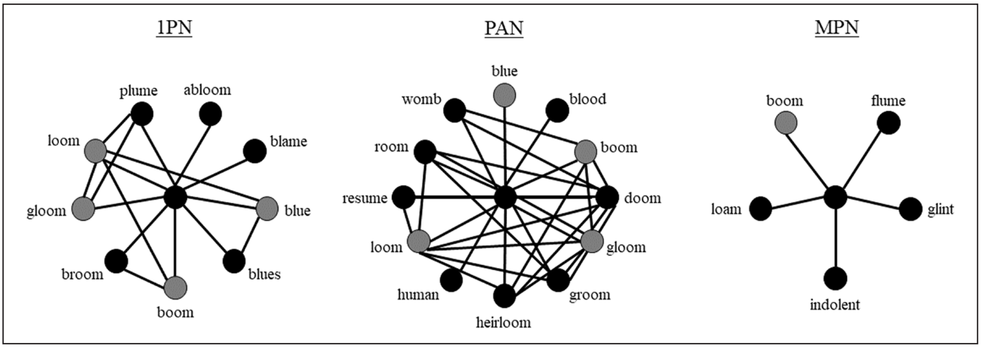

Figure 1. The 1-hop neighborhood of *bloom* in each network.
Note. The node for *bloom* is located in the center of each network. Gray nodes represent neighbors that are found in more than one of the networks, whereas black nodes represent neighbors found in only that specific network. 1PN: one-phoneme metric network; PAN: phonological association network; MPN: misperception network.

3.2 Results

3.2.1 Coarse topological structure. Using Network Portrait Divergence (Bagrow & Bollt, 2019), we conducted pairwise comparisons between the PAN, MPN, and 1PN to assess whether their topological structures were similar or not. Divergence (D) ranges from 0 to 1, with values closer to 0 indicating that the networks are more similar to each other and values closer to 1 indicating that the networks are less similar to each other. D was .599 when comparing the PAN and the MPN, .726 when comparing the PAN and the 1PN, and .769 when comparing the MPN and the 1PN.

To better understand the Network Portrait Divergence results, we looked at several network measures individually. In what follows, we compare the networks on each of several network measures, highlighting key similarities and differences. Figure 1 provides a visualization of the same word, *bloom*, and its 1-hop neighborhood in each of the networks. Table 2 summarizes the network values for the location of nodes (giant component, island, or hermit), community structure, small-world structure, degree distribution, and mixing by the degree of the PAN, 1PN, and MPN. In addition, we include previously reported network values in Vitevitch (2008) and Siew (2013) for the previously reported phonological network of Vitevitch (2008) for comparison.

3.2.2 Specific network structure features

Location of Nodes. The PAN and MPN had large giant components, such that the majority of the network was connected (98.2% and 97.2% of nodes, respectively). In contrast, the 1PN had a much smaller giant component (50.9% of nodes), with more nodes located in islands or as hermits than the PAN or MPN.

Community Structure. The Louvain method (Blondel et al., 2008) was used to determine the community structure of the giant component of each network. Modularity (Q) captures the quality of partitions in the network (Fortunato, 2010), with values greater than .3 indicative of significant community structure (Clauset et al., 2004). All three networks exhibited community structure: PAN = .86, MPN = .89, and 1PN = .68. In addition, we found fewer communities in the 1PN than in the PAN and MPN. The 1PN had a total of 37 communities, ranging in size from 6 to 1,054 nodes (M = 283.27, SD = 324.99). The PAN had a total of 70 communities, ranging in size from 8 to 1,060 nodes (M = 289.33, SD = 228.53). And, the MPN had a total of 98 communities, ranging in size from 57 to 454 nodes (M = 163.21, SD = 76.19).

Table 2. Network Structure Measures for the PAN, 1PN, MPN, and Phonological Network of Vitevitch (2008).

Network measure	PAN	1PN	MPN	Vitevitch (2008) network[a]
Network size	Nodes = 20,615 Edges = 56,747	Nodes = 20,575 Edges = 57,042	Nodes = 16,457 Edges = 20,116	Nodes = 19,340 Edges = 31,267
Location of nodes[b]	GC = 20,253 (98.2) Islands = 322 (1.6) Hermits = 40 (.2)	GC = 10,481 (50.9) Islands = 3,347 (16.3) Hermits = 6,747 (32.8)	GC = 15,995 (97.2) Islands = 460 (2.8) Hermits = 2 (.0)	GC = 6,508 (33.7) Islands = 2,567 (13.3) Hermits = 10,265 (53.1)
Small world structure[c]	Avg. path Len. = 9.80 Avg. C = .12 S = 724.79	Avg. Path Len. = 6.46 Avg. C = .16 S = 1157.50	Avg. Path Len. = 7.57 Avg. C = .007 S = 48.15	Avg. Path Len. = 6.05 Avg. C = .13 S = 1064.38
Scale-free structure[d]	P.L. RMSE = .64 Exp. RMSE = .03	P.L. RMSE = .12 Exp. RMSE = .02	P.L. RMSE = .03 Exp. RMSE = .11	P.L. RMSE = .09 Exp. RMSE = .01
Mixing by degree[e]	$r = .44$, $p < .0001$	$r = .67$, $p < .0001$	$r = -.02$, $p = .009$	$r = .62$, $p < .0001$
Community structure[f]	70 communities Mod. = .86	37 communities Mod. = .68	98 communities Mod. = .89	17 communities Mod. = .66

Note. 1PN: One-phoneme metric network; GC: giant component; MPN: Misperception network; PAN: phonological association network; RMSE: root-mean-square error.
[a]Measures reported were obtained from Vitevitch (2008) for all but community structure, which was obtained from Siew (2013).
[b]GC (proportion of nodes in parentheses).
[c]Average shortest path length (Avg. Path Len.), average clustering coefficient (Avg. C), and small-world-ness (S) from Humphries and Gurney (2008), where values greater than 1 indicate significant small-world structure.
[d]Comparison of the root-mean-square error (RMSE) of the Power-Law (P.L.) and Exponential (Exp.) functions to the degree distribution.
[e]Correlation between the degree of a node and the degree of each of its neighbors' degree.
[f]Modularity (Mod.) values above .3 indicate significant community structure (Clauset et al., 2004).

Small-World Structure. The conventional approach that considers the ratio of average shortest path length and average clustering coefficient of the observed and comparably sized random networks was used to determine if the networks exhibited small-world structure. The average shortest path length of the 1PN was 6.46. The PAN and MPN had larger average shortest path lengths, 9.80 and 7.57, respectively. Note that even controlling for network size differences across the three networks (e.g., a ratio of path length to network size), these numerical differences remain. The average clustering coefficient varied across the three networks, with the 1PN having the highest average clustering coefficient (.16), followed by the PAN (.12) and the MPN (.007). The markedly smaller average clustering coefficient found in the MPN may be driven in part by the inclusion of nonword nodes, which would never be cues, limiting the ability of connections to exist between neighbors of a cue word. Regardless, all networks displayed similar average path lengths and larger average clustering coefficients than comparably sized random networks, indicating significant small-world structure through the conventional approach.

Furthermore, we used a statistical approach (Humphries & Gurney, 2008) to assess the small-worldness of the networks. In this approach, small-worldness (S) values greater than 1 are indicative of a small-world structure. All of the observed networks in the present analysis had values greater than 1 (PAN=724.79, 1PN=1157.50, MPN=48.15), indicating a significant small-world structure.

Scale-Free Structure. The degree distribution of each network was examined to determine if the networks displayed a scale-free structure. The degree distribution data of each network was visualized with a log-log plot fitted with a power-law function and an exponential curve (see Figure 2). Networks whose degree distribution is a better fit with a power-law function would be indicative of having a scale-free structure, with the exponential curve serving as a commonly used contrast. We used root-mean-square error (RMSE) values to indicate the best fit (i.e., smaller values indicate better fit), following previous approaches (e.g., Vitevitch, 2008). Specifically, the degree distribution of the MPN was the best fit by the power-law function ($Y=.576x^{-2.085}$, RMSE=.03) than the exponential curve ($Y=.046e^{-.156x}$, RMSE=.11). In contrast, the PAN was best fit by the exponential curve ($Y=-.172e^{-.177x}$, RMSE=.03) than the power-law function ($Y=4.703x^{-2.616}$, RMSE=.64), as well as the 1PN ($Y=.103e^{-.109x}$, RMSE=.02 and $Y=1.126x^{-1.768}$, RMSE=.12, respectively). Thus, only the MPN would be considered as having a scale-free structure.

Mixing by Degree. We used a Pearson correlation between the degree of a node and the degree of its neighbors to determine the mixing by degree pattern of each network. The PAN and the 1PN both displayed assortative mixing by degree as indicated by positive correlations. The PAN had a Pearson correlation r (56,515)=.44, $p<.0001$, and the 1PN had a Pearson correlation r (54,648)=.67, $p<.0001$. Recall that assortative mixing by degree occurs when high-degree nodes tend to connect to other high-degree nodes, and low-degree nodes tend to connect to other low-degree nodes. In contrast, the MPN indicated a small, but significant, negative correlation indicative of disassortative mixing by degree, r (20,114)=−.02, $p=.009$. Recall that disassortative mixing by degree occurs when high-degree nodes tend to connect to low-degree nodes, and low-degree nodes tend to connect to high-degree nodes.

3.2.3 Analysis of individual nodes across networks. The previous network examinations focused on the topological features of the networks. In this set of analyses, we compared the network structure of an individual word across the PAN, 1PN, and MPN to provide additional information about how the different definitions of "phonological similarity" may influence the resulting network structure at the node level. Note that while the PAN and 1PN share almost all nodes (N=20,575), the MPN

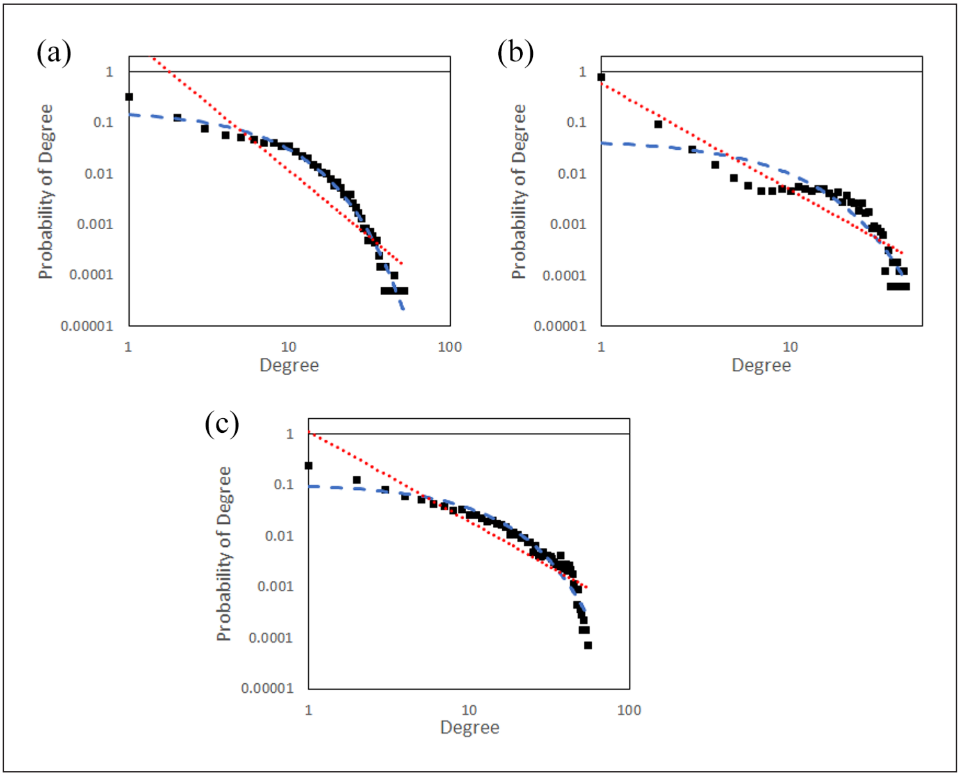

Figure 2. Log-log plot of the degree distribution for the three networks: PAN, 1PN, and MPN. (a) Phonological association network (PAN), (b) misperception network (MPN), and (c) one-phoneme metric network (1PN).

Note. The power-law function is represented by the red (dotted) line and the exponential curve is represented by the blue (dashed) line.

only overlaps in 6,481 words with the 1PN and PAN, limiting the dataset available for some of the following analyses.

Node Location across the PAN, 1PN, and MPN. First, we examined the location of a word in each network (giant component, an island, or as a hermit), as well as how that location may have differed (or remained the same) between networks. The location of nodes can vary from one network to the next in several ways: from the giant component to an island (or vice versa), from the giant component to a hermit (or vice versa), from an island to a hermit (or vice versa), or remain in the same location. We conducted this analysis for three specific, directional pairwise network comparisons: PAN → 1PN, MPN → 1PN, PAN → MPN. We acknowledge that comparisons can be bidirectional between the different networks, but we report these specific directional comparisons for simplicity. For each comparison, we considered only those nodes that overlapped between the networks being compared. Table 3 provides the proportion of nodes for each type of location change for the three pairwise network comparisons. One interesting, but perhaps not surprising, finding is the similarity in proportion of node location changes between the PAN to 1PN and the MPN to 1PN. In both cases, a large proportion of nodes located in the giant component of the PAN and MPN were found as islands or hermits in the 1PN, reflecting the more stringent definition of a one-phoneme metric

Table 3. Proportion of Nodes for Each Type of Location Change Between the PAN, the 1PN, and the MPN.

Type of Location Change	PAN to 1PN	MPN to 1PN	PAN to MPN
Giant component to island	3,280 (15.9%)	913 (14.0%)	115 (1.8%)
Giant component to hermit	6,605 (32.1%)	1,567 (24.2%)	2 (.3%)
Island to giant component	113 (.5%)	34 (.5%)	45 (.7%)
Island to hermit	142 (.7%)	51 (.8%)	0 (.0%)
Hermit to giant component	[a]	1 (<.01%)	8 (.1%)
Same location	10,435 (50.7%)	3,915 (60.4%)	6,311 (97.4%)

Note. Count (and percentage) of nodes in each category. The PAN and 1PN overlap in 20,575 words, whereas the MPN only overlaps in 6,481 words with the 1PN and PAN. 1PN: One-phoneme metric network; MPN: misperception network; PAN: phonological association network.
[a]The hermit nodes of the PAN were not included as nodes in the 1PN; therefore, no comparison can be made.

to determine similarity in the 1PN. In contrast, the majority of nodes remained in the same location when considering the PAN to MPN, in part an artifact of the naturally large giant components of these two networks.

Correlation of Node Degree across the PAN, 1PN, and MPN. We next considered the degree of a word in each network by conducting pairwise Pearson correlations. Given the reduced word overlap of the MPN with the PAN and 1PN, this analysis conducted correlations only using the overlapping 6,481 nodes. With the reduced dataset, the PAN had an average degree of 8.85 ($SD = 7.08$), ranging from 0 to 52. The 1PN had an average degree of 7.95 ($SD = 10.41$), ranging from 0 to 53. The MPN had an average degree of 3.51 ($SD = 5.08$), ranging from 0 to 37.

A Pearson correlation test indicated a positive correlation between the degree of a word in the PAN and the 1PN, r (6,479) = .61, $p < .0001$. However, a Pearson correlation test indicated no significant relationship between the degree of a word in the MPN and the 1PN, r (6,479) = .01 $p = .13$, nor between the degree of a word in the MPN and the PAN, r (6,479) = .001, $p = .89$.

Correlation of Node Clustering Coefficient across the PAN, 1PN, and MPN. Finally, we considered the clustering coefficient of a word in each network by conducting pairwise Pearson correlations. We again started with the reduced dataset, and then further removed words with a clustering coefficient of 0 (i.e., hermit words or words with no clustering) or 1 (i.e., words with full clustering), resulting in a total of 317 words in the following analysis. With this further reduced dataset, the average clustering coefficient of words in the PAN was .18 ($SD = .13$), ranging from .01 to .90. The average clustering coefficient of words in the 1PN was .27 ($SD = .09$), ranging from .04 to .81. The average clustering coefficient of words in the MPN was .11 ($SD = .14$), ranging from .002 to .66.

A Pearson correlation indicated a negative correlation between the clustering coefficient of a word in the PAN and the MPN, r (315) = −.13, $p = .019$. However, a Pearson correlation indicated no significant relationship between the clustering coefficient of a word in the PAN and the 1PN, r (315) = .01, $p = .82$, nor between the clustering coefficient of a word in the MPN and the 1PN, r (315) = −.01, $p = .85$.

3.3 Discussion

A structural comparison of different types of phonological networks provided a means to understand how different dimensions of phonological similarity might be represented in the mental lexicon. We analyzed an "explicit" phonological network using data from a phonological associate task (PAN; see Supplementary Material) and an "implicit" phonological network using data from misperceptions of words heard in noise (MPN; Felty et al., 2013), and compared these networks to a network based on the computational, one-phoneme metric (1PN; see also Vitevitch, 2008).

Overall, we found that the only similarity across all three networks was small-world structure, indicating that these phonological structures were not random, despite differences in sources of data and dimensions of phonological similarity assessed (for a discussion of limitations in this approach, see Turnbull & Peperkamp, 2017). The remaining network structure features that were examined differentiated the three networks either as a function of the demands of the behavioral tasks that were employed or as a dimension of phonological similarity. Recall that while both the PAN and MPN were constructed using participant-driven data, we proposed that the PAN would capture an "explicit" dimension of phonological similarity and the MPN would capture an "implicit" dimension of phonological similarity. In contrast, the 1PN was constructed using the computational, one-phoneme metric, which could be viewed as an objective, albeit strict, comparison for the PAN and MPN.

The organization of nodes across the networks, as captured by the analyses of node location and community structure, was likely influenced by the demands of the behavioral tasks that were employed and contributed to the higher similarity scores in Network Portrait Divergence between the PAN and MPN. In particular, we found that the PAN and MPN had much larger giant components with greater modularity in community structure than the 1PN. In addition, our consideration of node location change across networks supports the idea that weakly connected communities in the giant component of the PAN and MPN are likely to be represented as islands in the 1PN, leading to greater modularity in the PAN and MPN (i.e., weakly connected communities are easy to identify, thereby increasing modularity; Clauset et al., 2004). This topographical similarity between the MPN and PAN is most likely a consequence of the demands of the tasks used to generate the data (i.e., the association and misperception tasks). That is, participants were very likely to provide a response rather than no response, leading to a high rate of connections being placed in the network (i.e., few hermits and islands). Indeed, similar topographical structures are also commonly seen in semantic free association networks, where the network is either fully connected (i.e., no hermits) or the network has a very large giant component (e.g., 96% of nodes; Steyvers & Tenenbaum, 2005). Therefore, although we find similarities in the topographical structure of the PAN and MPN, these similarities should be interpreted cautiously as evidence of a common definition or dimension of phonological similarity. Rather the similar topographical structures in the PAN and MPN may have emerged due to the demands of the laboratory-based tasks used to elicit (and indeed requiring) responses.

The other network structure features examined in this section differentiated the PAN and MPN, with the 1PN often aligning in directional effects with the PAN. The PAN and 1PN exhibited assortative mixing by degree, whereas the MPN exhibited disassortative mixing by degree. The PAN and 1PN did not exhibit scale-free structure, whereas the MPN did exhibit scale-free structure. There was also a positive correlation between the degree of a node in the PAN and 1PN, whereas the degree of a node in the MPN did not correlate with the degree of a node in the PAN or 1PN. And finally, while the clustering coefficient of a node in the PAN and MPN did correlate, it was a negative correlation. These findings seem to indicate that the PAN may be capturing some features of the 1PN, perhaps related to the one-phoneme metric. Alternatively, though, it is important to remember that the MPN includes a small number of nonword nodes and limited overlap in real-word nodes across networks, which may influence what was observed. In what follows, we briefly speculate on what differences in scale-free structure and mixing by degree may mean for phonological representation.

First, consider scale-free structure. Finding scale-free structure in the MPN indicates that there are a small number of nodes that have many connections and a large number of nodes that have few connections. This suggests that phonological misperception, while significant for a small set of nodes (i.e., they are easily confusable with other similar sounding words/nonwords), is perhaps more likely to be idiosyncratic and/or nonexistent for a large number of words/nonwords. Furthermore, the targeted removal of a highly connected node in the MPN is likely to lead to a fractured network (Albert et al., 2000; Barabasi & Albert, 1999), indicating that the connectivity of the MPN hinges on a small set of words/nonwords. In contrast, the PAN and 1PN did *not* exhibit a scale-free structure. These results suggest that some dimensions of “phonological similarity” may result in networks that are more resilient and less likely to fracture when the system is perturbed in some way (e.g., changes in the S/N ratio of words presented in noise, or placing participants in the phonological associate task under a cognitive load).

Next, consider mixing by degree. The MPN exhibited disassortative mixing by degree. This mixing pattern indicates that high-degree nodes tended to be connected to low-degree nodes, and vice versa. Upon closer inspection, over 77% of the nodes in the MPN had only one connection, and many of these nodes were nonword or inaccurate real word responses. Furthermore, the nodes with the most connections in the network tended to be cue words eliciting many misperceptions. These cue words tended to be long and of low word frequency, like *quotidian* and *herbivorous*. This finding suggests that only a small set of cue words that are unique in the language elicit misperceptions. Those misperceptions are somewhat idiosyncratic (possibly due to individual differences in vocabulary or hearing acuity) and are unlikely to be responses to multiple cue words.

In contrast, the PAN and IPN exhibited assortative mixing by degree. This mixing pattern indicates that high-degree nodes tended to be connected to other high-degree nodes, whereas low-degree nodes tended to be connected to other low-degree nodes. In the 1PN, connections only exist when all but one phoneme between two words is maintained. Although there is more flexibility in the amount of overlap, the phonological associate task, used to construct the PAN, also requires participants to maintain some amount of phonological overlap between the cue and their response. Because of the preference for explicit phonological overlap between nodes in the 1PN and PAN, we should consider shared phonological sequences. It is well-known that there is a limited repertoire of phonemes for a given language and sequences of those phonemes are also limited in number due to the phonotactic rules of that language. Because of this, phonological sequences are going to be shared among words and, in particular, high probability sequences are likely going to be shared among many words (Vitevitch et al., 1999). Thus, since the PAN and 1PN have preference for shared phonological overlap, it is not surprising to see assortative mixing by degree.

At this point, we recognize that there are limitations in all of the ways we have defined phonological similarity in the present study, but it is striking that the one-phoneme metric appears to capture some of the “explicit” dimensions of phonological similarity, identified through comparison with the PAN, and some of the “implicit” dimensions of phonological similarity, identified through comparison with the misperception network. Furthermore, the PAN and MPN appear to capture different dimensions of phonological similarity, despite both relying on behavioral tasks. In the next section, we further examine how the different dimensions of phonological similarity as represented in the different network structures may account for lexical processing using behavioral data from two megastudies.

4 Testing network degree using behavioral data

One of the central tenets of network science is that the structure of a network influences processing on that network (Watts & Strogatz, 1998). In the present analyses, we examined how the different dimensions of phonological similarity as represented in the different phonological networks we created might influence lexical processing. That is, do the three different network structures predict behavioral data in the same way? In this section, we tested the network measure of degree (referred to as neighborhood density in the psycholinguistic literature), given its long history of research in psycholinguistics (Vitevitch & Luce, 2016).

In spoken word recognition of English, research has shown that words with many phonological neighbors (i.e., high degree, or a dense phonological neighborhood) are responded to more slowly and less accurately than words with few phonological neighbors (i.e., low degree, or a sparse phonological neighborhood; e.g., Luce & Pisoni, 1998; Vitevitch et al., 1999; Vitevitch & Luce, 1998, 1999; Ziegler et al., 2003). In most of the previous studies, phonological neighbors were defined with the one-phoneme metric. To assess how the definition of neighborhood density/degree in our networks (PAN, 1PN, MPN) influenced processing, we compared those measures of the degree to behavioral data obtained for the same words in an auditory lexical decision task from the Auditory English Lexicon Project (Goh et al., 2020).

In visual word recognition (i.e., reading), research examining the influence of phonological neighborhood density on processing has found mixed results. For example, some studies have found that words with many phonological neighbors (i.e., high degree, or a dense phonological neighborhood) are responded to more quickly and accurately than words with few phonological neighbors (i.e., low degree, or a sparse phonological neighborhood; e.g., Yates et al., 2004, 2008), and others have found interactions of phonological and orthographic neighborhoods (e.g., Grainger et al., 2005). To assess how the definition of neighborhood density/degree in our networks (PAN, 1PN, and MPN) influenced processing, we compared those measures of degree to behavioral data obtained for the same words in a visual lexical decision task from the English Lexicon Project (ELP; Balota et al., 2007).

4.1 Method

4.1.1 Lexical decision data. We obtained data from the Auditory ELP (AELP; Goh et al., 2020), an openly available database containing auditory lexical decision reaction times. AELP data (for responses to the American English speakers) were extracted on 14 February 2020. We also obtained data from the ELP (Balota et al., 2007), an openly available database containing visual lexical decision reaction times. ELP data were extracted on 2 January 2020.

We limited our data extractions to those words that overlapped across the PAN, 1PN, and MPN, and with the ELP database ($n = 4,656$ words). We obtained several variables for the words from the ELP database: SUBTLEX word frequency, age of acquisition, orthographic neighborhood size (given its study in previous lexical decision studies), number of letters (used in the ELP analysis), number of phonemes (used in the AELP analysis), and the network measure of degree (or neighborhood density) based on each of the three networks (PAN, 1PN, and MPN). Because the 1PN is an undirected, unweighted network, we chose to obtain the undirected, unweighted degree of words in all three networks for consistency. Furthermore, note that there was missingness for word frequency and age of acquisition, leading to exclusion of 101 words, reducing the visual lexical decision dataset to 4,555 words. Finally, although all of the 4,555 words were found in the ELP database, only 3,036 of those words were found in the AELP database, further reducing the number of words we analyzed for auditory lexical decision.

For our dependent measure of lexical decision reaction time, we extracted the z-scored variable provided in the databases. This z-scored reaction time variable takes into consideration within-subject variability in responding (Balota et al., 2007). In addition, we followed a commonly used procedure in analyzing lexical decision data, considering only correct responses and excluding responses if the reaction time was <200 or >3000 ms.

Our primary predictors of interest were the three measures of degree: PAN, 1PN, and MPN degrees. For the visual lexical decision analysis, the PAN degree ranged from 1 to 52 ($M=10.42$, $SD=7.47$), the 1PN degree ranged from 1 to 53 ($M=10.66$, $SD=10.86$), and the MPN degree ranged from 0 to 35 ($M=3.41$, $SD=4.77$). For the auditory lexical decision analysis, the PAN degree ranged from 1 to 52 ($M=12.42$, $SD=7.32$), the 1PN degree ranged from 1 to 52 ($M=11.05$, $SD=10.95$), and the MPN degree ranged from 0 to 26 ($M=3.72$, $SD=4.71$). All degree measures were z-scored.

As control predictors, we also extracted from the respective databases the reported values of SUBTLEX word frequency, age of acquisition, number of letters (only for the visual lexical decision reaction time model since stimuli are presented visually), number of phonemes (only for the auditory lexical decision reaction time model since stimuli are presented aurally), and orthographic neighborhood size. All control predictors were z-scored. While a host of other variables could have been included as control predictors, we focused on these commonly studied and influential variables to lexical decision in this exploratory analysis. Included in Appendix 1 is a correlation matrix of all predictors.

4.1.2 Statistical analysis. We conducted crossed-effects mixed modeling to assess the influence of degree on predicting lexical decision reaction using R (R Core Team, 2020) and the "lme4" package (Bates et al., 2015). We conducted two sets of model-building procedures: one predicting auditory lexical decision reaction time using data from the AELP and one predicting visual lexical decision reaction time using data from the ELP. We used χ^2 tests to compare models in our model building procedure, starting with a model of only psycholinguistic control variables, then models adding a single degree measure at a time, and finally a model including all three measures of degree simultaneously.

In addition to reporting the standard output for our final models, we include the variance inflation factor (VIF) for each tested variable (Gareth et al., 2013). The VIF indicates how much a variable correlates with the other variables in the model. A rule of thumb guideline provided by Gareth et al. (2013) indicates that variables with a VIF between 5 and 10 should be considered for removal and a VIF greater than 10 must be removed due to high multicollinearity.

4.2 Results

Table 4 provides the model comparisons for the visual lexical decision and auditory lexical decision data. For visual lexical decision, all models including only one of the phonological neighborhood degrees significantly improved model fit compared to the baseline model with only the psycholinguistic control predictors (all ps $<.01$). In addition, the full model including all of the phonological degree measures simultaneously also improved model fit as compared to the baseline model, despite correction for inclusion of more predictors, χ^2 $(3)=121.45$, $p<.0001$. For auditory lexical decision, of the models including only one phonological degree, only the model including the 1PN phonological degree measure significantly improved model fit as compared to the baseline model, χ^2 $(1)=84.47$, $p<.0001$. In addition, the full model including all of the phonological degree measures simultaneously also improved model fit as compared to the baseline model, despite correction for inclusion of more predictors, χ^2 $(3)=99.25$, $p<.0001$.

Table 5 provides the full model for each lexical decision task, which includes the psycholinguistic control predictors and all three phonological network degree measures. For the visual lexical decision task, all phonological network degree measures were significant. For PAN degree, words

Table 4. Model Building Comparisons.

		AIC	BIC	LogLik	Deviance	Comparison	Chisq	df	*p*
Visual lexical decision									
1	WF+AoA+Len+ON	28,8705	28,8784	−144,345	288,689				
2	WF+AoA+Len+ON+PAN_D	288,639	288,707	−144,310	288,621	1 vs. 2	68.57	1	<.0001
3	WF+AoA+Len+ON+MPN_D	288,967	288,785	−144,339	288,679	1 vs. 3	10.40	1	.0012
4	WF+AoA+Len+ON+1PN_D	288,696	288,785	−144,339	288,678	1 vs. 4	10.80	1	.0010
5	WF+AoA+Len+ON+PAN_D+MPN_D+1PN_D	288,590	288,697	−144,284	288,568	1 vs. 5	121.45	3	< .0001
Auditory lexical decision									
1	WF+AoA+Len+ON	446,017	446,098	−223,000	446,001				
2	WF+AoA+Len+ON+PAN_D	446,018	446,110	−223,000	446,000	1 vs. 2	.11	1	.735
3	WF+AoA+Len+ON+MPN_D	446,019	446,110	−223,000	446,001	1 vs. 3	.05	1	.821
4	WF+AoA+Len+ON+1PN_D	445,934	446,025	−222,958	445,916	1 vs. 4	84.47	1	< .0001
5	WF+AoA+Len+ON+PAN_D+MPN_D+1PN_D	445,923	446,035	−222,951	445,901	1 vs. 5	99.25	3	< .0001

Note. WF: word frequency; AoA: age of acquisition; Len: word length (in letters for visual lexical decision, in phonemes for auditory lexical decision); ON: orthographic neighborhood; PAN_D: phonological association network degree; MPN_D: misperception network degree; 1PN_D: one-phoneme metric network degree.

Table 5. Model Outputs Predicting Auditory and Visual Lexical Decision Reaction Times From Control and Network Degree Predictors.

Visual lexical decision					
	β	*SE*	*t*	*p*	VIF
Controls predictors					
(Intercept)	−.402	.004	−87.011	<.0001	
Word frequency	−.111	.004	−26.135	<.0001	1.990
Age of acquisition	.086	.004	21.252	<.0001	1.800
Number of phonemes	.035	.004	8.557	<.0001	1.916
Orthographic neighborhood size	.002	.004	.435	.663	2.593
Degree predictors					
PAN degree	−.418	.004	−10.078	<.0001	1.934
1PN degree	.319	.005	6.237	<.0001	2.951
MPN degree	.011	.003	3.727	.0001	1.030
Auditory lexical decision					
	β	*SE*	*t*	*p*	
Controls predictors					
(Intercept)	−.393	.118	−33.121	<.0001	
Word frequency	−.068	.008	−7.890	<.0001	1.850
Age of acquisition	.052	.008	6.024	<.0001	1.859
Number of letters	.060	.008	6.843	<.0001	1.965
Orthographic neighborhood size	.007	.010	.727	.466	2.701
Degree predictors					
PAN degree	−.035	.009	−3.845	.0001	2.181
1PN degree	.124	.012	10.025	<.0001	3.856
MPN degree	.0001	.006	.026	.978	1.006

Note. 1PN: one-phoneme metric network; MPN: misperception network; PAN: phonological association network; VIF: variance inflation factor.

with higher degree were responded to more quickly than words with lower degree, $\beta=-.418$, $t=-10.078$, $p<.0001$. For 1PN degree, words with higher degree were responded to more slowly than words with lower degree, $\beta=.319$, $t=6.237$, $p<.0001$. For MPN degree, words with higher degree were also responded to more slowly than words with lower degree, $\beta=.11$, $t=3.727$, $p=.0001$. All VIFs were below 5, which represents a low correlation of that variable with the other variables in the model (Gareth et al., 2013).

For the auditory lexical decision task, only the PAN and 1PN network degree measures were significant. For PAN degree, words with higher degree were responded to more quickly than words with lower degree, $\beta=-.035$, $t=-3.845$, $p=.0001$. For the 1PN degree, words with higher degree were responded to more slowly than words with lower degree, $\beta=.124$, $t=10.025$, $p<.0001$. All VIFs were below 5, which represents a low correlation of that variable with the other variables in the model (Gareth et al., 2013).

4.3 Discussion

Our results for 1PN degree are consistent with previous studies of spoken word recognition in English: words with many one-phoneme metric neighbors are responded to more slowly than words with few one-phoneme metric neighbors. A small, but statistically significant effect was also found for 1PN degree on visual lexical decision reaction time: words with many one-phoneme metric neighbors were responded to more slowly than words with few one-phoneme metric neighbors. This result contrasts with the previous findings where words with many phonological neighbors were responded to more quickly and accurately than words with few phonological neighbors (Yates et al., 2004, 2008). The different findings for visual lexical decision in the present analysis compared to a previous study may, in part, be influenced by other characteristics of the words that were analyzed in the present case compared to the set of words used in previous studies (see Grainger et al., 2005).

We also found that PAN degree had a significant influence on both auditory and visual lexical decision reaction times. In both cases, words with many phonological neighbors were responded to more quickly and accurately than words with few phonological neighbors. This finding may seem to contradict the influence found for 1PN degree, but it is important to recall the demands of the lexical decision task. To complete the task, a participant must only decide whether a stimulus is a word (or not). Lexical access to a specific word is not necessarily required for the participant to respond. That is, a participant only needs enough "evidence" that the stimulus is a word to make a decision. Given this, one can hypothesize that perhaps the influence of PAN degree reflects this process of having sufficient evidence to make a quick decision. That is, target words with many phonological associate neighbors have more converging evidence from a greater variety of words (e.g., neighbors can differ by more than one phoneme, neighbors could share morphology), leading to quick reaction times for words with many neighbors, compared to target words with few phonological associate neighbors.

Finally, we show only a small, but significant effect of MPN degree in visual word recognition, but no influence of MPN degree in auditory lexical decision. We recognize the hazards associated with interpreting null results, but we consider the possibility that the dimension of phonological similarity captured by misperceptions may be more acoustic- and perception-based than the other metrics evaluated here. As such, influences of this dimension of phonological similarity may only be observed in tasks that are more perceptual in nature, such as the auditory repetition task (where a participant only needs access to the phonological system but not necessarily the lexicon as they repeat out loud the stimulus that is presented) rather than a lexical decision task which taps into more cognitive/lexical levels of processing.

Regarding the influence of MPN degree on visual word recognition, reaction times were slower for target words with many misperception neighbors than for target words with few misperception neighbors. In this case, the acoustic dimension of phonological similarity captured by this definition of neighbor may reflect varying amounts of confusion in simply perceiving the stimulus. Given that phonological processing influences visual word recognition quite early on (Braun et al., 2009), words with many misperception neighbors will have increased perceptual confusion and uncertainty about whether the stimulus is a real word or a nonword and, therefore, longer reaction times in the visual lexical decision task. For words with few misperception neighbors, the confusion and uncertainty are only about whether X or Y was heard, but given that both X and Y are real words, a rapid decision (i.e., shorter reaction times) can be made in the visual lexical decision task.

The present results demonstrate that the structure of the different phonological networks indeed affected lexical processing in some way (Watts & Strogatz, 1998). Interestingly, while the 1PN and PAN networks had similar overall structures and degrees of a word in the 1PN and PAN were positively correlated, those network structures had different influences on lexical processing in the auditory lexical decision task and visual lexical decision tasks, such that PAN degree had a facilitatory effect (i.e., faster reaction times) and 1PN degree had an inhibitory effect (i.e., slower reaction times). In addition, the MPN structure only had a significant inhibitory effect in the visual lexical decision task (no effect of MPN was found in the auditory lexical decision task). We believe the different effects on lexical decision reaction times reflect the fact that the different phonological networks capture different dimensions of phonological similarity (e.g., perceptual, cognitive, etc.).

5 General discussion

The mathematical tools of network science have been increasingly used in the Cognitive Sciences to shed new light on a number of traditional questions about cognition and to spark novel research in new areas of investigation that were previously not possible without these tools (Castro & Siew, 2020; Vitevitch, 2019). The present study further highlights how network analyses can inform us about different dimensions of phonological similarity, how those different dimensions result in differences in the way lexical representations are organized in memory, and the influence that such structures have on processing.

We compared the structure of three phonological networks that captured different dimensions of "phonological similarity." We analyzed an "explicit" phonological network derived from phonological associate data (see Supplementary Material), an "implicit" phonological network derived from the identification of words heard in noise data (Felty et al., 2013), and a computational phonological network derived from the one-phoneme metric (Vitevitch, 2008). A coarse topographical comparison of the networks using an information-theoretic technique indicated that the PAN and MPN were more similar to each other than to the 1PN, largely driven by their large giant components and significant modularity in community structure. However, upon closer examination of a variety of network measures and analyses, the 1PN and PAN networks were more similar to each other than they were to the MPN. Specifically, the 1PN and PAN networks exhibit small-world structure and assortative mixing by degree did not exhibit a scale-free structure and had positive correlations between the degree and clustering coefficient of a node in each network.

Several recent reports have questioned what a phonological network based on the one-phoneme metric is actually representing (cf., Brown et al., 2018; Vitevitch & Mullin, 2021). It has been suggested that some structural features of networks may be artifacts of the one-phoneme metric (e.g., small-world structure and clustering coefficient), while other structural features are actually due to cognitive, organizing principles (e.g., assortative mixing by degree; Stella & Brede, 2016a; Turnbull & Peperkamp, 2017). The present results, however, challenge the assertion that the structural features of phonological networks are artifacts of the one-phoneme metric. Given that the network derived from the one-phoneme metric shared many topological features with a phonological network formed from participant-provided phonological associates and with a network based on misperceptions of words heard in noise, it is unclear how those same network features could be due simply to the one-phoneme metric. Furthermore, the beta weight of the 1PN network in the auditory lexical decision task attests to the ability of this metric to account for the largest amount of variance in the lexical decision data compared to each of the other two ways of defining phonological similarity. Therefore, we instead posit that there may be different dimensions of phonological similarity and that the one-phoneme metric may simply be an imperfect way to capture some but not all aspects of those different dimensions.

Furthermore, we used data from independent psycholinguistic megastudies to ascertain whether different ways of defining phonological similarity in a network influence language processing. We considered the impact of a commonly studied measure, degree/neighborhood density, on auditory and visual lexical decision reaction times. We found that 1PN and PAN degrees had opposing effects in both lexical decision tasks; 1PN degree had an inhibitory effect (i.e., slower reaction times as the degree increased), while PAN degree had a facilitatory effect (i.e., faster reaction times as degree increased). In contrast, MPN degree only had a significant, inhibitory effect on the visual lexical decision task.

Taken together, our results indicate that the PAN, and to a lesser extent, the MPN capture some aspects of the one-phoneme metric while providing novel dimensions of phonological similarity. Prior consideration of the number of phonemes different between cues and responses in both the phonological association and misperception datasets indicates that responses tend to be one or two phonemes different (see Supplementary Material and Felty et al., 2013). Yet, we find key differences between the structure and the impact of those structures on processing between the PAN, MPN, and 1PN degrees.

We acknowledge that there are limitations in the present study. First, the methods of network construction (e.g., what defines a node or an edge) influence the resulting structure. The MPN network included nonword nodes. Although these nodes would not appear in the 1PN or PAN, we retained them because they provide valuable information about how phonological similarity would be defined in this implicit dimension of similarity. Furthermore, the PAN and MPN require participant-generated responses which may not fully encompass all words in a person's lexicon, resulting in a potentially skewed, observed network. Second, the networks do not share identical nodes, which may impact the overall structure of each network and limit our ability to directly compare words across networks (e.g., as seen in the reduced set of words that could be analyzed with the megastudy data). However, our results highlight the importance of considering multiple phonological similarity methods to fully understand the different dimensions of similarity (e.g., acoustic, perceptual, cognitive, etc.) that may be represented in the mind. These results also highlight that further work is needed to better control these factors. In what follows, we consider the pros and cons of using a PAN and an MPN and conclude by proposing a novel approach to assessing multiple dimensions of phonological similarity and their interactions simultaneously.

5.1 Considerations for using phonological association and misperception networks

First, the present results suggest that phonological associate data may capture cognitive and linguistic dimensions of phonological similarity. There are several benefits of using an association-based network, including the ability to consider the weighting and directionality of cue-response pairs and the ability to construct person-specific networks. For example, individualized semantic and phonological networks may be particularly important for developing diagnostic protocols and language interventions (Borge-Holthoefer et al., 2011; Neergaard et al., 2019; Vitevitch & Castro, 2015; Zemla & Austerweil, 2018). Given the important role that assortative mixing by a degree may have in network resilience, it might be useful to measure how this association metric changes in an individual prior to experiencing aphasia (i.e., an acquired language impairment that primarily disrupts the accessibility of words in memory; Mirman & Britt, 2014) or semantic dementia (i.e., where the concomitant language impairment occurs primarily due to the degradation of word representations from memory), and during the course of the condition. Other network measures, such as various centrality measures, might be used to target certain sets of words for therapy to partially restore communicative abilities in individuals affected by aphasia (Castro et al., 2020; Castro & Stella, 2019). Individualized language networks may indeed prove to be a new and exciting frontier for research on language networks (Neergaard et al., 2019; Zemla et al., 2016), but critically require participant-provided data.

Drawbacks to utilizing association-based networks, however, include the significant amount of data required to obtain stable estimates of weighting and directionality. In the study reported in the Supplementary Material that was used to generate the PAN network we analyzed, there was a significant proportion of cue-response pairs that were generated by only one participant (76.8%), and there was only one wave of data collection, limiting the ability to fully consider certain network measures, like in- versus out-degree. The present analyses provide a starting point for considering phonological similarity using associations generated by participants. However, given that the structure of the PAN had similar structural features (e.g., small-world structure, degree distribution, assortative mixing by degree) to networks constructed using a much simpler computational metric of phonological similarity (i.e., the one-phoneme metric) that has been well-established in the psycholinguistic literature, the additional time, effort, and expense required to generate such a database of associates (e.g., De Deyne et al., 2019) does not make for a favorable cost–benefit ratio.

Second, the present results also suggest that misperception data may capture acoustic/phonetic/perceptual dimensions of phonological similarity. As with phonological association data, misperception data provides some information on weighting and directionality that is not provided in the 1PN network (i.e., the network constructed using the one-phoneme metric). Misperception data may also provide important insights into certain speech-language disorders, such as hearing loss of various types (e.g., Kaiser et al., 2003), and captures information about confusability. However, like the association-based networks, misperception data also suffer from the need of a significant amount of data to obtain stable estimates of weighting and directionality, making for an unfavorable cost–benefit ratio.

5.2 A new method to capture dimensions of phonological similarity

One solution for advancing research on language processing is to continue to use the widely used one-phoneme metric. It has well-known shortcomings but captures some aspects of the different dimensions of phonological similarity. Alternatively, yet another increasingly more complex metric could be developed to address some of those shortcomings.

Rather than continue to use an imperfect metric or develop a more complex but still imperfect metric that loses some potentially useful information, we instead advocate for future research to use the tools of network science to combine the PAN and the MPN into a multi-layered network (Battiston et al., 2017; De Domenico et al., 2013). A *multilayered network* (sometimes called a network of networks) has two or more layers of nodes, in this case words, that are connected in one layer based on one type of relationship (e.g., phonological association) and are connected in another layer based on a different type of relationship (e.g., misperceptions). Such an approach would not only retain both types of information and both dimensions of phonological similarity but could allow us to add a layer (or several layers) of words to capture other aspects of phonological similarity, as well as semantic relationships between words to better account for how phonological and semantic information interact during various language processes.

This direction is ambitious, but some progress has already been made in connecting phonological and semantic networks to increase our understanding of language acquisition (Stella & Brede, 2016b; Stella et al., 2017) and word retrieval (Castro et al., 2020; Castro & Stella, 2019). In these works, a *multiplex* network (a specific type of multilayered network where the nodes are identical across layers) is composed of either three or four layers, with some layers representing different aspects of semantic relationships between words (e.g., free association, synonyms) and another layer representing phonological similarity between words (i.e., one-phoneme metric). Critically, Castro et al. (2020) provide evidence supporting a hypothesis that semantic and phonological layers may contribute separately and interactively during word retrieval, highlighting the importance of considering multiple word-word similarity relations simultaneously to better account for the structure and processes of the mental lexicon. Other computational and neural network models also represent multiple linguistic systems simultaneously, including Dell's (1986) Interactive Activation model, Nadeau's (2012) Parallel Distributed Processing model, and Baayen et al.'s (2019) Linear Discriminative Lexicon model.

What the network science approach provides beyond adding more linguistic domains is the possibility to bridge the abstract representations in the mind to the physical structures in the brain, much like abstract social networks emerge via various social media platforms through the various layers of physical telecommunication networks such as fiber optic cables (Vitevitch, 2019). That is, layers in the network could capture the structure of physical aspects of the brain (e.g., groups of neurons or regions in the brain) and bridge them to abstract aspects of the mental lexicon (e.g., word–word similarity relations) to account for the dynamical processes that occur within and between layers. Such an approach to connecting social networks and brain networks has already been demonstrated (Baek et al., 2021; Falk & Bassett, 2017; Weaverdyck & Parkinson, 2018), providing a framework for this approach to connect mental lexicon networks and brain networks. Although additional evidence is needed to test the viability of the multilayered network approach, this direction could significantly move the Psychological and Neurosciences forward (and closer to each other) and increase our understanding of various psychological processes.

Author note

Portions of this work fulfilled part of the dissertation requirements for Nichol Castro.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Nichol Castro https://orcid.org/0000-0001-8691-7968

Supplemental material

Supplemental material for this article is available online.

References

Albert, R., Jeong, H., & Barabasi, A. (2000). Error and attack tolerance of complex networks. *Nature*, *406*(6794), 378–382. https://doi.org/10.1038/35019019

Arbesman, S., Strogatz, S., & Vitevitch, M. (2010). The structure of phonological networks across multiple languages. *International Journal of Bifurcation and Chaos*, *20*(03), 679–685. https://doi.org/10.1142/S021812741002596X

Baayen, R., Chuang, Y., Shafaei-Bajestan, E., & Blevins, J. (2019). The discriminative lexicon: A unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de)composition but in linear discriminative learning. *Complexity*, *2019*, 4895891. https://doi.org/10.1155/2019/4895891

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology*, *18*(4), 362–365. https://doi.org/10.1080/14640746608400055

Baek, E. C., Porter, M. A., & Parkinson, C. (2021). Social network analysis for social neuroscientists. *Social Cognitive and Affective Neuroscience*, *16*(8), 883–901. https://doi.org/10.1093/SCAN/NSAA069

Bagrow, J. P., & Bollt, E. M. (2019). An information-theoretic, all-scales approach to comparing networks. *Applied Network Science*, *4*(1), 45. https://doi.org/10.1007/s41109-019-0156-x

Bagrow, J. P., Bollt, E. M., Skufca, J. D., & Ben-Avraham, D. (2008). Portraits of complex networks. *EPL (Europhysics Letters)*, *81*(6), 68004.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*(3), 445–459. https://doi.org/10.3758/BF03193014

Barabasi, A., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*(5439), 509–512. https://doi.org/10.0.4.102/science.286.5439.509

Baronchelli, A., Ferrer-i-Cancho, R., Pastor-Satorras, R., Chater, N., & Christiansen, M. H. (2013). Networks in cognitive science. *Trends in Cognitive Sciences*, *17*(7), 348–360. https://doi.org/10.1016/J.TICS.2013.04.010

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Battiston, F., Nicosia, V., & Latora, V. (2017). The new challenges of multiplex networks: Measures and models. *The European Physical Journal Special Topics*, *226*(3), 401–416. https://doi.org/10.1140/epjst/e2016-60274-8

Beckage, N. M., & Colunga, E. (2015). Language networks as models of cognition: Understanding cognition through language. In A. Mehler, A. Lucking, S. Banisch, P. Blanchard, & B. Job (Eds.), *Towards a theoretical framework for analyzing complex linguistic networks* (pp. 3–28). Springer. https://doi.org/10.1007/978-3-662-47238-5_1

Blondel, V., Guillaume, J., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *10*, P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Borge-Holthoefer, J., Moreno, Y., & Arenas, A. (2011). Modeling abnormal priming in Alzheimer's patients with a free association network. *PLOS ONE*, *6*(8), Article e22651. https://doi.org/10.1371/journal.pone.0022651

Braun, M., Hutzler, F., Ziegler, J. C., Dambacher, M., & Jacobs, A. M. (2009). Pseudohomophone effects provide evidence of early lexico-phonological processing in visual word recognition. *Human Brain Mapping*, *30*(7), 1977–1989.

Broido, A., & Clauset, A. (2019). Scale-free networks are rare. *Nature Communications*, *10*(1), 1017. https://doi.org/10.1038/s41467-019-08746-5

Brown, K., Allopenna, P., Hunt, W., Steiner, R., Saltzman, E., Mcrae, K., & Magnuson, J. (2018). Universal features in phonological neighbor networks. *Entropy*, *2*(7), 526. https://doi.org/10.3390/e20070526

Castro, N., & Siew, C. (2020). Contributions of modern network science to the cognitive sciences: Revisiting research spirals of representation and process. *Proceedings of the Royal Society A*, *476*(2238), 20190825. https://doi.org/10.31234/osf.io/gkmb8

Castro, N., & Stella, M. (2019). The multiplex structure of the mental lexicon influences picture naming in people with aphasia. *Journal of Complex Networks*, *7*(6), 913–931. https://doi.org/10.31234/osf.io/eqvmg

Castro, N., Stella, M., & Siew, C. (2020). Quantifying the interplay of semantics and phonology during failures of word retrieval by people with aphasia using a multiplex lexical network. *Cognitive Science*, *44*(9), Article e12881. https://doi.org/10.1111/cogs.12881

Chan, K., & Vitevitch, M. (2009). The influence of the phonological neighborhood clustering coefficient on spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1934–1949. https://doi.org/10.1037/a0016902

Chan, K., & Vitevitch, M. (2010). Network structure influences speech production. *Cognitive Science*, *34*(4), 685–697. https://doi.org/10.1111/j.1551-6709.2010.01100.x

Clauset, A., Newman, M., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E*, *70*(6), 066111. https://doi.org/10.1103/PhysRevE.70.066111

Connine, C., Blasko, D., & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language*, *32*(2), 193–210. https://doi.org/10.1006/jmla.1993.1011

Conrad, R., & Hull, A. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology*, *55*(4), 429–432. https://doi.org/10.1111/j.2044-8295.1964.tb00928.x

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, *1695*(5), 1–9. http://igraph.org

Cutler, A., Sebastian-Galles, N., Soler-Vilageliu, O., & Van Ooijen, B. (2000). Constraints of vowels and consonants on lexical selection: Cross-linguistic comparisons. *Memory & Cognition*, *28*(5), 746–755. https://doi.org/10.3758/BF03198409

De Cara, B., & Goswami, U. (2002). Similarity relations among spoken words: The special status of rimes in English. *Behavior Research Methods, Instruments, & Computers*, *34*(3), 416–422. https://doi.org/10.3758/bf03195470

De Deyne, S., Navarro, D., Perfors, A., Brysbaert, M., & Storms, G. (2019). The "small world of words" English word association norms for over 12,000 cue words. *Behavior Research Methods*, *51*(3), 987–1006. https://doi.org/10.3758/s13428-018-1115-7

De Domenico, M., Solé-Ribalta, A., Cozzo, E., Kivelä, M., Moreno, Y., Porter, M., Gómez, S., & Arenas, A. (2013). Mathematical formulation of multilayer networks. *Physical Review X*, *3*(4), 041022. https://doi.org/10.1103/PhysRevX.3.041022

Dell, G. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Falk, E. B., & Bassett, D. S. (2017). Brain and social networks: Fundamental building blocks of human experience. *Trends in Cognitive Sciences*, *21*(9), 674–690. https://doi.org/10.1016/J.TICS.2017.06.009

Felty, R., Buchwald, A., Gruenenfelder, T., & Pisoni, D. (2013). Misperceptions of spoken words: Data from a random sample of American English words. *The Journal of the Acoustical Society of America*, *134*(1), 572–585. https://doi.org/10.1121/1.4809540

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, *486*(3–5), 75–174. https://doi.org/10.1016/j.physrep.2009.11.002

Fricke, M., Baese-Berk, M., & Goldrick, M. (2016). Dimensions of similarity in the mental lexicon. *Language, Cognition and Neuroscience*, *31*(5), 639–645. https://doi.org/10.1080/23273798.2015.1130234

Gareth, J., Daniela, W., Trevor, H., & Robert, T. (2013). *An introduction to statistical learning: With applications in R*. Springer.

Goh, W., & Pisoni, D. (2003). Effects of lexical competition on immediate memory span for spoken words. *Quarterly Journal of Experimental Psychology Section A*, *56*(6), 929–954. https://doi.org/10.1080/02724980244000710

Goh, W., Yap, M., & Chee, Q. (2020). The Auditory English Lexicon Project: A multi-talker, multi-region psycholinguistic database of 10,170 spoken words and nonwords. *Behavior Research Methods*, *52*, 2202–2231. https://doi.org/10.3758/s13428-020-01352-0

Goldstein, R., & Vitevitch, M. (2014). The influence of clustering coefficient on word-learning: How groups of similar sounding words facilitate acquisition. *Frontiers in Psychology*, *5*, Article 1307. https://doi.org/10.3389/fpsyg.2014.01307

Goldstein, R., & Vitevitch, M. (2017). The influence of closeness centrality on lexical processing. *Frontiers in Psychology*, *8*, Article 1683. https://doi.org/10.3389/fpsyg.2017.01683

Gollan, T., & Acenas, L. (2004). What Is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and Tagalog-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(1), 246–269. https://doi.org/10.1037/0278-7393.30.1.246

Gordon, J. (2002). Phonological neighborhood effects in aphasic speech errors: Spontaneous and structured contexts. *Brain and Language*, *82*(2), 113–145. https://doi.org/10.1016/S0093-934X(02)00001-9

Grainger, J., Muneaux, M., Farioli, F., & Ziegler, J. C. (2005). Effects of phonological and orthographic neighbourhood density interact in visual word recognition. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *58*(6), 981–998. https://doi.org/10.1080/02724980443000386

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

Gruenenfelder, T. M., & Pisoni, D. B. (2009). The lexical restructuring hypothesis and graph theoretic analyses of networks based on random lexicons. *Journal of Speech, Language, and Hearing Research*, *52*(3), 596–609. https://doi.org/10.1044/1092-4388(2009/08-0004)

Hahn, U., & Bailey, T. (2005). What makes words sound similar? *Cognition*, *97*(3), 227–267. https://doi.org/10.1016/j.cognition.2004.09.006

Humphries, M. D., & Gurney, K. (2008). Network 'small-world-ness': A quantitative method for determining canonical network equivalence. *PLOS ONE*, *3*(4), Article e0002051. https://doi.org/10.1371/journal.pone.0002051

Kaiser, A., Kirk, K., Lachs, L., & Pisoni, D. (2003). Talker and lexical effects on audiovisual word recognition by adults with cochlear implants. *Journal of Speech, Language, and Hearing Research*, *46*(2), 390–404. https://doi.org/10.1044/1092-4388(2003/032)

Kleinberg, J. (2000). Navigation in a small world. *Nature*, *406*(6798), 845.

Knoop, C., Blohm, S., Kraxenberger, M., & Menninghaus, W. (2021). How perfect are imperfect rhymes? Effects of phonological similarity and verse context on rhyme perception. *Psychology of Aesthetics, Creativity, and the Arts*, *15*(3), 560–572. https://doi.org/10.1037/aca0000277

Landauer, T., & Streeter, L. (1973). Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. *Journal of Memory and Language*, *12*(2), 119–131. https://doi.org/10.1016/S0022-5371(73)80001-5

Lindell, A., & Lum, J. (2008). Priming vs. rhyming: Orthographic and phonological representations in the left and right hemispheres. *Brain and Cognition*, *68*(2), 193–203. https://doi.org/10.1016/j.bandc.2008.04.005

Luce, P. (1986). *Neighborhoods of words in the mental lexicon: Research on speech perception* (Technical Report No. 6). https://files.eric.ed.gov/fulltext/ED353610.pdf

Luce, P., Goldinger, S., Auer, E., & Vitevitch, M. (2000). Phonetic priming, neighborhood activation, and PARSYN. *Perception and Psychophysics*, *62*(3), 615–625. https://doi.org/10.3758/BF03212113

Luce, P., & Large, N. (2001). Phonotactics, density, and entropy in spoken word recognition. *Language and Cognitive Processes*, *16*(5–6), 565–581. https://doi.org/10.1080/01690960143000137

Luce, P., & Pisoni, D. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*(1), 1–36. https://doi.org/10.1097/00003446-199802000-00001

Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 576–585. https://doi.org/10.1037/0096-1523.15.3.576

Milgram, S. (1967). The small-world problem. *Psychology Today*, *2*, 60–67.

Miller, G. (1995). WordNet: A lexical database for English. *Communications of the ACM*, *38*(11), 39–41. https://doi.org/10.1145/219717.219748

Miller, G., & Nicely, P. (1955). An analysis of perceptual confusions among some English consonants. *The Journal of the Acoustical Society of America*, *27*(2), 338–352. https://doi.org/10.1121/1.1907526

Mirman, D., & Britt, A. (2014). What we talk about when we talk about access deficits. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634), 20120388. https://doi.org/10.1098/rstb.2012.0388

Muneaux, M., & Ziegler, J. (2004). Locus of orthographic effects in spoken word recognition: Novel insights from the neighbour generation task. *Language and Cognition*, *19*(5), 641–660. https://doi.org/10.1080/01690960444000052

Munson, B., & Solomon, N. (2004). The effect of phonological neighborhood density on vowel articulation. *Journal of Speech, Language, and Hearing Research*, *47*(5), 1048–1058. https://doi.org/10.1044/1092-4388(2004/078)

Nadeau, S. (2012). *The neural architecture of grammar*. MIT Press.

Neergaard, K., & Huang, C. (2019). Constructing the Mandarin phonological network: Novel syllable inventory used to identify schematic segmentation. *Complexity*, *2019*, 6979830. https://doi.org/10.1155/2019/6979830

Neergaard, K., Luo, J., & Huang, C. (2019). Phonological network fluency identifies phonological restructuring through mental search. *Scientific Reports*, *9*(1), 15984. https://doi.org/10.1038/s41598-019-52433-w

Nelson, D. L., McEvoy, C. L., & Dennis, S. (2000). What is free association and what does it measure? *Memory & Cognition*, *28*(6), 887–899. https://doi.org/10.3758/BF03209337

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402–407. https://doi.org/10.3758/BF03195588

Newman, M. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(2), 404–409. https://doi.org/10.1073/pnas.98.2.404

Newman, M. (2003a). Mixing patterns in networks. *Physical Review E*, *67*(2), 026126. https://doi.org/10.1103/PhysRevE.67.026126

Newman, M. (2003b). The structure and function of complex networks. *SIAM Review*, *45*(2), 167–256. https://doi.org/10.1137/S003614450342480

Newman, M., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, *69*(2), 026113. https://doi.org/10.1103/PhysRevE.69.026113

Pisoni, D., Nusbaum, H., Luce, P., & Slowiaczek, L. (1985). Speech perception, word recognition and the structure of the lexicon. *Speech Communication*, *4*(1–3), 75–95. https://doi.org/10.1016/0167-6393(85)90037-8

R Core Team (2020). R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Ravasz, E., & Barabási, A-L. (2003). Hierarchical organization in complex networks. *Physical Review E*, *67*, 026112. https://doi.org/10.1103/PhysRevE.67.026112

Roget, P. (1911). *Roget's thesaurus of English words and phrases*. Thomas Y. Crowell. http://www.gutenberg.org/ebooks/10681

Scarborough, R. (2012). Lexical similarity and speech production: Neighborhoods for nonwords. *Lingua*, *122*(2), 164–176. https://doi.org/10.1016/j.lingua.2011.06.006

Scarborough, R., & Zellou, G. (2013). Clarity in communication: "Clear" speech authenticity and lexical neighborhood density effects in speech production and perception. *The Journal of the Acoustical Society of America*, *134*(5), 3793–3807. https://doi.org/10.1121/1.4824120

Schweppe, J., Grice, M., & Rummer, R. (2011). What models of verbal working memory can learn from phonological theory: Decomposing the phonological similarity effect. *Journal of Memory and Language*, *64*(3), 256–269. https://doi.org/10.1016/j.jml.2010.11.006

Siew, C. (2013). Community structure in the phonological network. *Frontiers in Psychology*, *4*, Article 553. https://doi.org/10.3389/fpsyg.2013.00553

Siew, C., & Vitevitch, M. (2016). Spoken word recognition and serial recall of words from components in the phonological network. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(3), 394–410. https://doi.org/10.1037/xlm0000139

Siew, C., Wulff, D., Beckage, N., & Kenett, Y. (2019). Cognitive network science: A review of research on cognition through the lens of network representations, processes, and dynamics. *Complexity*, *2019*, 2108423. https://doi.org/10.1155/2019/2108423

Sommers, M. (1996). The structural organization of the mental lexicon and its contribution to age-related declines in spoken-word recognition. *Psychology and Aging*, *11*(2), 333–341. https://doi.org/10.1037/0882-7974.11.2.333

Stamer, M., & Vitevitch, M. (2012). Phonological similarity influences word learning in adults learning Spanish as a foreign language. *Bilingualism: Language and Cognition*, *15*(3), 490–502. https://10.0.3.249/S1366728911000216

Stella, M., Beckage, N., & Brede, M. (2017). Multiplex lexical networks reveal patterns in early word acquisition in children. *Scientific Reports*, *7*, 46730. https://doi.org/10.1038/srep46730

Stella, M., & Brede, M. (2016a). Investigating the phonetic organisation of the English language via phonological networks, percolation and Markov models. In S. Battiston, F. De Pellegrini, G. Caldarelli, & E. Merelli (Eds.), *Proceedings of ECCS 2014* (pp. 219–229). Springer. https://doi.org/10.1007/978-3-319-29228-1_19

Stella, M., & Brede, M. (2016b). Mental lexicon growth modelling reveals the multiplexity of the English language. In H. Cherifi, B. Goncalves, R. Menezes, & R. Sinatra (Eds.), *Complex networks VII: Studies in computational intelligence* (Vol. *644*, pp. 267–279). Springer. https://doi.org/10.1007/978-3-319-30569-1_20

Stemberger, J. P. (1992). The reliability and replicability of naturalistic speech error data. In B. J. Baars (Ed.), *Experimental slips and human error* (pp. 195–215). Springer. https://doi.org/10.1007/978-1-4899-1164-3_8

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, *29*(1), 41–78. https://doi.org/10.1207/s15516709cog2901_3

Storkel, H. (2004). Do children acquire dense neighborhoods? An investigation of similarity neighborhoods in lexical acquisition. *Applied Psycholinguistics*, *25*, 201–221. https://doi.org/10.1017/S0142716404001109

Strand, J. F. (2014). Phi-square Lexical Competition Database (Phi-Lex): An online tool for quantifying auditory and visual lexical competition. *Behavior Research Methods*, *46*(1), 148–158. https://doi.org/10.3758/s13428-013-0356-8

Strogatz, S. H. (2001). Exploring complex networks. *Nature*, *410*(6825), 268–276. https://doi.org/10.1038/35065725

Suárez, L., Tan, S. H., Yap, M. J., & Goh, W. D. (2011). Observing neighborhood effects without neighbors. *Psychonomic Bulletin and Review*, *18*(3), 605–611. https://doi.org/10.3758/s13423-011-0078-9

Turnbull, R., & Peperkamp, S. (2017). What governs a language's lexicon? Determining the organizing principles of phonological neighbourhood networks. In H. Cherifi, S. Gaito, W. Quattrociocchi, & A. Sala (Eds.), *International workshop on complex networks and their applications* (Vol. *693*, pp. 83–94). Springer. https://doi.org/10.1007/978-3-319-50901-3_7

Van Ooijen, B. (1996). Vowel mutability and lexical selection in English: Evidence from a word reconstruction task. *Memory & Cognition*, *24*(5), 573–583. https://doi.org/10.3758/BF03201084

Vitevitch, M. (1997). The neighborhood characteristics of malapropisms. *Language and Speech*, *40*(3), 211–228. https://doi.org/10.1177/002383099704000301

Vitevitch, M. (2002). Influence of onset density on spoken-word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(2), 270–278. https://doi.org/10.1037/0096-1523.28.2.270

Vitevitch, M. (2007). The spread of the phonological neighborhood influences spoken word recognition. *Memory & Cognition*, *35*(1), 166–175. https://doi.org/10.3758/BF03195952

Vitevitch, M. (2008). What can graph theory tell us about word learning and lexical retrieval? *Journal of Speech, Language, and Hearing Research*, *51*, 408–422. https://doi.org/10.1044/1092-4388(2008/030)

Vitevitch, M. (2012). What do foreign neighbors say about the mental lexicon? *Bilingualism: Language and Cognition*, *15*(1), 167–172. https://doi.org/10.1017/S1366728911000149

Vitevitch, M. (2019). Can network science connect mind, brain, and behavior?. In M. Vitevitch (Ed.), *Network science in cognitive psychology* (1st ed., pp. 184–197). Taylor & Francis.

Vitevitch, M., & Castro, N. (2015). Using network science in the language sciences and clinic. *International Journal of Speech-Language Pathology*, *17*(1), 13–25. https://doi.org/10.3109/17549507.2014.987819

Vitevitch, M., Chan, K., & Goldstein, R. (2014). Insights into failed lexical retrieval from network science. *Cognitive Psychology*, *68*, 1–32. https://doi.org/10.1016/j.cogpsych.2013.10.002

Vitevitch, M., Chan, K., & Roodenrys, S. (2012). Complex network structure influences processing in long-term and short-term memory. *Journal of Memory and Language*, *67*(1), 30–44. https://doi.org/10.1016/j.jml.2012.02.008

Vitevitch, M., Goldstein, R., & Johnson, E. (2016). Path-length and the misperception of speech: Insights from network science and psycholinguistics. In A. Mehler, A. Lucking, S. Banisch, O. Blanchard, & B. Job (Eds.), *Towards a theoretical framework for analyzing complex linguistic networks* (pp. 29–45). Springer. https://doi.org/10.1007/978-3-662-47238-5_2

Vitevitch, M., & Luce, P. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science*, *9*(4), 325–329. https://doi.org/10.1111/1467-9280.00064

Vitevitch, M., & Luce, P. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, *40*(3), 374–408. https://doi.org/10.1006/jmla.1998.2618

Vitevitch, M., & Luce, P. (2016). Phonological neighborhood effects in spoken word perception and production. *Annual Review of Linguistics*, *2*, 75–94. https://doi.org/10.1146/annurev-linguistics-030514-124832

Vitevitch, M., Luce, P., Pisoni, D., & Auer, E. (1999). Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and Language*, *68*(1–2), 306–311. https://doi.org/10.1006/brln.1999.2116

Vitevitch, M., & Mullin, G. (2021). What do cognitive networks do? Simulations of spoken word recognition using the cognitive network science approach. *Brain Sciences*, *11*(12), 1628. https://doi.org/10.3390/BRAINSCI11121628

Vitevitch, M., Niehorster-Cook, L., & Niehorster-Cook, S. (2021). Exploring how phonotactic knowledge can be represented in cognitive networks. *Big Data and Cognitive Computing*, *5*(4), 47. https://doi.org/10.3390/bdcc5040047

Vitevitch, M., & Stamer, M. (2006). The curious case of competition in Spanish speech production. *Language and Cognitive Processes*, *21*(6), 760–770. https://doi.org/10.1080/01690960500287196

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature*, *393*(6684), 440–442. https://doi.org/10.1038/30918

Weaverdyck, M. E., & Parkinson, C. (2018). The neural representation of social networks. *Current Opinion in Psychology*, *24*, 58–66. https://doi.org/10.1016/J.COPSYC.2018.05.009

Wieling, M., Nerbonne, J., Bloem, J., Gooskens, C., Heeringa, W., & Baayen, R. (2014). A cognitively grounded measure of pronunciation distance. *PLOS ONE*, *9*(1), Article e75734. https://doi.org/10.1371/journal.pone.0075734

Yates, M., Friend, J., & Ploetz, D. (2008). The effect of phonological neighborhood density on eye movements during reading. *Cognition*, *107*(2), 685–692. https://doi.org/10.1016/j.cognition.2007.07.020

Yates, M., Locker, L., & Simpson, G. (2004). The influence of phonological neighborhood on visual word perception. *Psychonomic Bulletin & Review*, *11*(3), 452–457. https://doi.org/10.3758/BF03196594

Zemla, J., & Austerweil, J. (2018). Estimating semantic networks of groups and individuals from fluency data. *Computational Brain & Behavior*, *1*(1), 36–58. https://doi.org/10.1007/s42113-018-0003-7

Zemla, J., Kenett, Y., Jun, K., & Austerweil, J. (2016). U-INVITE: Estimating individual semantic networks from fluency data. In A. Papafragou, D. Grodner, D. Mirman, & J. Trueswell (Eds.), *Proceedings of the 38th Annual Meeting of the Cognitive Science Society* (pp. 1907–1912). Cognitive Science Society.

Ziegler, J., Muneaux, M., & Grainger, J. (2003). Neighborhood effects in auditory word recognition: Phonological competition and orthographic facilitation. *Journal of Memory and Language*, *48*(4), 779–793. https://doi.org/10.1016/S0749-596X(03)00006-8

Appendix 1. Correlation Matrix of Model Predictors.

	MPN_D	PAN_D	1PN_D	WF	AOA	NPhon	NLet
MPN_D							
PAN_D	.024						
1PN_D	.044*	.674***					
WF	.018	.475***	.339***				
AOA	.022	−.448***	−.359***	−.642***			
NPhon	−.053**	−.470***	−.679***	−.331***	.392***		
NLet	−.055**	−.472***	−.648***	−.318***	.370***	.855***	
OND	.022	.601***	.785***	.286***	−.329***	−.561***	−.646***

Note. 1PN_D: one-phoneme metric network degree; AOA: age of acquisition; MPN_D: misperception network degree; NLet: number of letters; NPhon: number of phonemes; OND: orthographic neighborhood density; PAN_D: phonological association network degree; WF: word frequency.
$* < .05$, $** < .01$, $*** < .001$.