American Psychological Association

Journal of Experimental Psychology: General

0096-3445/19/$12.00

2019, Vol. 148, No. 3, 475–500
http://dx.doi.org/10.1037/xge0000575

The Phonographic Language Network: Using Network Science to Investigate the Phonological and Orthographic Similarity Structure of Language

Cynthia S. Q. Siew
National University of Singapore and University of Warwick

Michael S. Vitevitch
University of Kansas

Orthographic effects in spoken word recognition and phonological effects in visual word recognition have been observed in a variety of experimental tasks, strongly suggesting that a close interrelationship exists between phonology and orthography. However, the metrics used to investigate these effects, such as consistency and neighborhood size, fail to generalize to words of various lengths or syllable structures, and do not take into account the more global similarity structure that exists between phonological and orthographic representations in the language. To address these limitations, the tools of Network Science were used to simultaneously characterize the phonological as well as orthographic similarity structure of words in English. In the phonographic network of language, links are placed between words that are both phonologically and orthographically similar to each other (e.g., words such as *pant* (/pænt/) and *punt* (/pʌnt/)). Conventional psycholinguistic experiments (auditory naming and auditory lexical decision) and an archival analysis of the English Lexicon Project (visual naming and visual lexical decision) were conducted to investigate the influence of 2 network science metrics derived from the phonographic network—phonographic degree and phonographic clustering coefficient—on spoken and visual word recognition. Results indicated a facilitatory effect of phonographic degree on visual word recognition, and a facilitatory effect of phonographic clustering coefficient on spoken word recognition. Implications of the present findings for theoretical models of spoken and visual word recognition are discussed.

Keywords: network science, phonographic language network, spoken word recognition, visual word recognition, clustering coefficient

Consider the following from T. S. Watt's (1954) poem titled "Brush up your English:"

> Beware of heard, a dreadful word
>
> that looks like beard and sounds like bird,
>
> and dead—it's said like bed, not bead.
>
> For goodness's sake, do not call it deed!"

The rest of the poem continues to warn the reader to be wary of "dreadful" English words that are not pronounced as one might expect from its spelling and not spelled in ways as one might expect from its pronunciation. Clearly, an interesting and complex interrelationship exists between the phonology and orthography of English. Years of psycholinguistic research in both visual and spoken word recognition have further demonstrated that this interrelationship influences word recognition—orthographic influences on spoken word recognition and phonological influences on visual word recognition have been observed in a variety of experimental tasks (Seidenberg & Tanenhaus, 1979; van Orden, 1987; Yates, Locker, & Simpson, 2004; Ziegler, Muneaux, & Grainger, 2003).

An overview of prior work investigating orthographic effects in spoken word recognition and phonological effects in visual word recognition is first provided. We argue that prior work ultimately fails to (a) generalize to words of various lengths or syllable structures and (b) take into account the more global similarity structure that exists between phonological and orthographic representations in the language. To address the limitations of prior work, we propose an alternative approach that uses the tools of network science to characterize, simultaneously, the phonological as well as orthographic similarity structure of words (of all lengths) in the language.

Cynthia S. Q. Siew, Department of Psychology, National University of Singapore, and Department of Psychology, University of Warwick; Michael S. Vitevitch, Department of Psychology, University of Kansas.

The experiments and analyses reported in this article partially fulfilled the requirements for the degree of Doctor of Philosophy in Psychology awarded to Cynthia S. Q. Siew. We thank members of the committee (Susan Kemper, Evangelia Chrysikou, Kelsie Forbush, Joan Sereno, and Allard Jongman) for their comments and suggestions. We also thank Joseph Denning for helping with data collection. A portion of this work was presented at 58th Annual Meeting of the Psychonomic Society in Vancouver, British Columbia, Canada.

Correspondence concerning this article should be addressed to Cynthia S. Q. Siew, Department of Psychology, University of Warwick, Humanities Building 3.49, Coventry CV4 7AL, United Kingdom. E-mail: cynsiewsq@gmail.com

Phonological Effects in Visual Word Recognition

Language researchers have long noted the close interrelationship between spoken language and its writing system (e.g., Liberman, 1992). Indeed, several years of research have showed that phonology plays an important role in reading, as demonstrated by the influence of various phonological effects on visual word recognition. For instance, the *homophone effect* refers to the finding that homophones, words that have different spellings and meanings but sound identical (e.g., "rose"/"rows" and "boar"/"bore"), are more slowly processed in the presence of homophone foils (Ferrand & Grainger, 2003; Grainger & Ferrand, 1994; van Orden, 1987). Another robust finding is the feedforward consistency effect, where words containing bodies with inconsistent pronunciations (e.g., the body "-ave" is pronounced /-æv/ as in "have" and /-eIv/ as in "wave") are more slowly processed in a number of visual word recognition tasks (Cortese & Simpson, 2000; Jared, 1997; Jared, McRae, & Seidenberg, 1990; Stone, Vanhoy, & van Orden, 1997; Ziegler, Montant, & Jacobs, 1997). Recent work has also found that the number of phonological neighbors influences visual word processing tasks, such that the processing of words with many phonological neighbors is facilitated as compared to words with few phonological neighbors (Grainger, Muneaux, Farioli, & Ziegler, 2005; Yates, 2005; Yates et al., 2004).

Orthographic Effects in Spoken Word Recognition

Seidenberg and Tanenhaus's (1979) paper was one of the first to show that orthography influences the processing of spoken words. Using a rhyme detection task, Seidenberg and Tanenhaus showed that the time taken to decide if two words rhymed was influenced by their orthographic similarity. Participants took a longer time to decide that "tie" and "rye" (orthographically dissimilar pair) rhymed as compared to "tie" and "pie" (orthographically similar pair). Subsequently, several studies have also found orthographic effects in a variety of online tasks such as naming (Ziegler, Ferrand, & Montant, 2004; Ziegler et al., 2003) and auditory lexical decision (Dich, 2011; Roux & Bonin, 2013; Ziegler & Ferrand, 1998; Ziegler et al., 2003, 2004).

One of the most robust orthographic effects found in the studies cited above is the feedback consistency effect, which refers to the finding that words containing sound-to-spelling inconsistencies (e.g., the rime /ip/ can be spelled as "-eep" as in "deep" or "-eap" as in "heap") are more slowly and less accurately responded to in lexical decision and word naming tasks (Ziegler & Ferrand, 1998; Ziegler et al., 2003, 2004). The size of the orthographic neighborhood also influences spoken language processing. Various studies have shown that when phonological neighborhood size is controlled for, the size of the orthographic neighborhood facilitates spoken word recognition—words with many orthographic neighbors are produced and recognized more quickly and accurately than words with few orthographic neighbors (Muneaux & Ziegler, 2004; Ziegler et al., 2003).

To recapitulate, effects of phonological variables have been observed in visual word recognition tasks and effects of orthographic variables have been observed in spoken word recognition tasks, pointing to the presence of a close interrelationship between phonology and orthography in language processing (see Table 1 for a summary of the previous literature). Importantly, these findings raise key questions about the nature of lexical representations that are stored within long-term memory and the cognitive processes that support lexical retrieval.

Limitations of Current Approaches

Metrics such as consistency and neighborhood size represent different ways of capturing the relationship between orthography and phonology in a language—neighborhood size is calculated based on evaluating the orthographic or phonological similarity of a target word to other words in the lexicon (Coltheart, Davelaar, Jonasson, & Besner, 1977; Luce & Pisoni, 1998), and consistency is determined by calculating how often the body of the target word is pronounced or spelled among words that also share the same body or rime (Kessler & Treiman, 1997). Nevertheless, these metrics do not capture the overall relationship between orthography and phonology in a language because the way in which these metrics have been operationalized restricts their applicability to a subset of words within the entire mental lexicon. That is, consistency measures are limited to words with a vowel-consonant structure. Phonological or orthographic neighborhood measures capture only one particular aspect of similarity in the language rather than the interrelationship between phonology and orthography. To make continued progress in our understanding of phonological and orthographic influences on language processing we argue that an alternative theoretical approach is required, one that explicitly considers how cognitive processes operate in a complex network that represents the interrelationships among orthographic and phonological information in words.

Consider the asymmetric nature of phonological neighborhood and orthographic neighborhood effects in visual and auditory tasks. Ziegler et al. (2003) found an inhibitory phonological neighborhood effect in spoken word recognition tasks, whereas Yates et al. (2004, 2005; also Grainger et al., 2005) found a facilitatory phonological neighborhood effect in visual word recognition tasks.

Table 1
Summary of Phonological Effects in Visual Word Recognition and Orthographic Effects in Spoken Word Recognition

Phonological effects in visual word recognition	Orthographic effects in spoken word recognition
• Homophone effect (Ferrand & Grainger, 2003; Grainger & Ferrand, 1994; Van Orden, 1987)	• Feedback consistency effect (Ziegler & Ferrand, 1998; Ziegler, Ferrand, & Montant, 2003, 2004)
• Feedforward consistency effect (Cortese & Simpson, 2000; Jared, 1997; Jared, McRae, & Seidenberg, 1990; Stone, Vanhoy, & Van Orden, 1997; Ziegler, Montant, & Jacobs, 1997)	• Orthographic neighborhood size (Muneaux & Ziegler, 2004; Ziegler, Muneaux, & Grainger, 2003)
• Phonological neighborhood size (Grainger, Muneaux, Farioli, & Ziegler, 2005; Yates, 2005; Yates, Locker, & Simpson, 2004)	

On the other hand, orthographic neighborhood effects appear to be facilitatory in both visual and spoken word recognition tasks (Ziegler et al., 2003). These findings hint at differences in the way that phonological and orthographic information influence lexical processing in different modalities. However, to date, it is not clear why the effects of similarity on lexical processing differ across different modalities. Furthermore, there has been little attempt to account for and integrate these findings within a single model or framework that considers both the overall phonological and orthographic structure of language.

To address these limitations we make use of the tools of network science to simultaneously represent the overall phonological and orthographic similarity structure of English words. We first provide a brief introduction to the field of network science and show how network science has contributed to our understanding of the cognitive and language sciences.

Introduction to Network Science

Network science is an emerging interdisciplinary field that uses mathematical techniques to characterize and analyze the structure of complex networks in various domains (Barabási, 2009; Watts, 2004). Examples of complex networks include friendship networks on social media websites (Lewis, Kaufman, Gonzalez, Wimmer, & Christakis, 2008), air transportation networks (Cardillo et al., 2013), and the mental lexicon (Steyvers & Tenenbaum, 2005; Vitevitch, 2008).

Networks consist of nodes that are connected to each other via links. For instance, nodes can represent individuals in a social network, or airports in an air transportation network. The links that connect individual nodes in networks represent relationships that exist between pairs of nodes. In a social network, a link could be placed between individuals who are friends with each other on a social media website such as Facebook. In an air transportation network, links represent the presence of flights between airports.

Network scientists recognize that the processes that occur within these networks are affected by the structure of the network (Strogatz, 2001). For instance, the structure of the social network affects the way in which information spreads among people, and the structure of air transportation networks affects the way air travel is rerouted when there are major airport closures. In addition to offering researchers a theoretical framework to study complex networks, network science provides a comprehensive suite of methodological techniques to derive a variety of network measures that describe structure at various levels of the network. These include metrics that describe the network's global or macrolevel structure (e.g., average path length, average clustering coefficient, overall degree distribution), the local or microlevel structure (e.g., degree, clustering coefficient of individual nodes), as well as the mesolevel structure (i.e., the level that falls between the macro- and microlevels; e.g., community structure).

The tools of network science have been used to study the structure of the mental lexicon, which consists of all the words that a person knows that are stored in long-term memory. Language researchers have used these tools to model phonological (Vitevitch, 2008), orthographic (Kello & Beltz, 2009), and semantic (Hills, Maouene, Maouene, Sheya, & Smith, 2009; Solé, Corominas-Murtra, Valverde, & Steels, 2010; Steyvers & Tenenbaum, 2005) networks of words in the mental lexicon. In these networks, each node represents a word, but the ways in which links connect the nodes differ (i.e., based on phonological, orthographic or semantic relationships among words).

Phonological Language Network

The phonological network examined in Vitevitch (2008) consisted of the phonological transcriptions of 19,340 words obtained from the 1964 *Merriam-Webster Pocket Dictionary*. The words in the *Merriam-Webster Pocket Dictionary* were used to represent the mental lexicon of an average native adult speaker of English. Although it must be noted that individual differences do exist with respect to the size and internal contents of one's mental lexicon, computational analyses of corpora reveal that there exists an overall kernel lexicon of words common to all speakers of a particular language (Ferrer-i-Cancho & Sole, 2001). Hence, the words in the phonological network could be said to be an approximation of the kernel lexicon of the English language. In this network, nodes represented phonological word forms and links represented phonological similarity between words. Two words were considered phonologically similar if the first word could be transformed to the other by either substituting, adding, or deleting one phoneme in any position (Landauer & Streeter, 1973; Luce & Pisoni, 1998). For instance, the word /kæt/ ("cat") would be connected to /æt/ ("at"), /bæt/ ("bat"), and /skæt/ ("scat").

Prior work by Vitevitch and colleagues demonstrated that various aspects of the structure of the phonological network influences spoken word recognition (Siew & Vitevitch, 2016) and production (Chan & Vitevitch, 2010), word learning (Vitevitch & Goldstein, 2014), and short- and long-term memory processes (Vitevitch, Chan, & Roodenrys, 2012). At the local or microlevel of the phonological network, the clustering coefficient, *C*, of a word, which represents the extent to which the phonological neighbors of a word are also phonological neighbors of each other, had measurable effects on a variety of psycholinguistic tasks such as perceptual identification, lexical decision, and picture naming (Chan & Vitevitch, 2009, 2010; Vitevitch et al., 2012; Vitevitch & Goldstein, 2014).

At the macrolevel of analysis, Vitevitch and Goldstein (2014) found a processing advantage for "keywords"—a set of words that, when removed, would cause the network to fracture into several smaller components—as compared to nonkeywords with comparable lexical characteristics. Another measure of the macrolevel structure of the network is assortative mixing by degree, which refers to the tendency for highly connected nodes to be connected to other highly connected nodes in the network (Newman, 2002). Vitevitch, Chan, and Goldstein (2014) analyzed instances of failed lexical retrieval by participants and found that the errors reflected the presence of high assortative mixing by degree in the phonological network. Another study of the macrostructure of the phonological network (Siew & Vitevitch, 2016) showed that words from lexical islands (small groups of words connected to each other but not to the rest of the network) were recognized more quickly than words from the giant component (the part of the network where most of the words are found).

Together, these findings suggest that, in addition to the microlevel (as exemplified by the clustering coefficient network metric), the macrolevel structure of the phonological network also

has important implications for understanding lexical processes. These findings also show that language researchers can investigate language in new ways and provide novel insights into the psychological mechanisms that support lexical processing by using the tools of network science.

Orthographic Language Network

In contrast to the work done with phonology, there has not been as much research using the tools of network science to study orthographic relationships among words. One exception is the analysis conducted by Kello and Beltz (2009), who constructed an orthographic word form network whereby links were placed between words that were substrings of other words. For instance, the word "air" would be connected to the words "fair" and "aired". However, Kello and Beltz's operationalization of orthographic similarity (i.e., placing links between words that were substrings of other words) differs significantly from the way orthographic similarity has been typically operationalized in the psycholinguistic literature, where words are considered to be orthographically similar if they differ by the substitution of a single letter (Coltheart et al., 1977).

It is also important to note that Kello and Beltz (2009) merely conducted a computational analysis of the orthographic network. To date, there has not been any behavioral or experimental work investigating how the network structure of the orthographic lexicon might influence lexical processing (but see Siew, 2018). However, research by Iyengar, Veni Madhavan, Zweig, and Natarajan (2012) suggests that the orthographic structure of language could have key implications for navigating the mental lexicon. Participants played a word-morph game where they had to find a sequence of words such that the first word could be transformed to the second word (of the same length) by changing a single letter. For example, the sequence of words to get from "try" to "pot" was "try-toy-ton-tot-pot".

Iyengar et al. (2012) found that participants were much faster at the game when they learned to make use of "landmark" words to find the sequence of words. That is, participants would repeatedly morph the start words to the landmark word, then morph to the end word. Iyengar et al. further examined these landmark words and found that these words had high closeness centrality—a network science measure indicating the inverse of the sum of distances of a node to all other nodes in the network (Borgatti, 2005). High closeness centrality words were close to many other words in the network, making them easy targets to morph words into and then to another word. Iyengar et al.'s findings suggest that the network structure of orthographic word-forms (albeit one that contained only three-letter words) has behavioral consequences as one navigates the mental lexicon. Although the word-morph game is an offline task, the results of the study by Iyengar et al. suggests that there could be similar implications for lexical retrieval.

Introduction to Multiplex Networks

Words can be phonologically, orthographically, and semantically related to each other. To date, language networks have been constructed based on a single type of relationship among words and analyzed independently of other types of language networks, thereby failing to capture the multiplexity inherent in language. To better capture the phonological and orthographic relationships that exist among words it might be better to use something called a *multiplex network*.

A *multiplex network* (also known as a *multilayer network* or simply a *multiplex*) consists of multiple layers of networks, whereby the links within each layer represent a different type of relationship among a common set of nodes (Battiston, Nicosia, & Latora, 2014). Figure 1 shows a simple multiplex. One example of a multiplex in the real world is the different kinds of relationships such as platonic, romantic, or sexual relationships that exist between people (Lewis et al., 2008). Constructing a social network based on a single type of relationship is merely a crude approximation to reality. Because multiplexity is an inherent feature of most real-world systems it is important to examine multiplex networks in areas of research such as language processing.

Introducing the Phonographic Network of Language

A phonographic multiplex was constructed with phonological and orthographic layers. The phonological layer was identical to the phonological network constructed by Vitevitch (2008). The orthographic layer consisted of the same words (nodes) in the phonological network, but the links in this layer are based on orthographic similarity. Words were connected to each other if they differed by the substitution, addition, or deletion of a single phoneme in any word position in the phonological network or a single letter in any word position in the orthographic network.

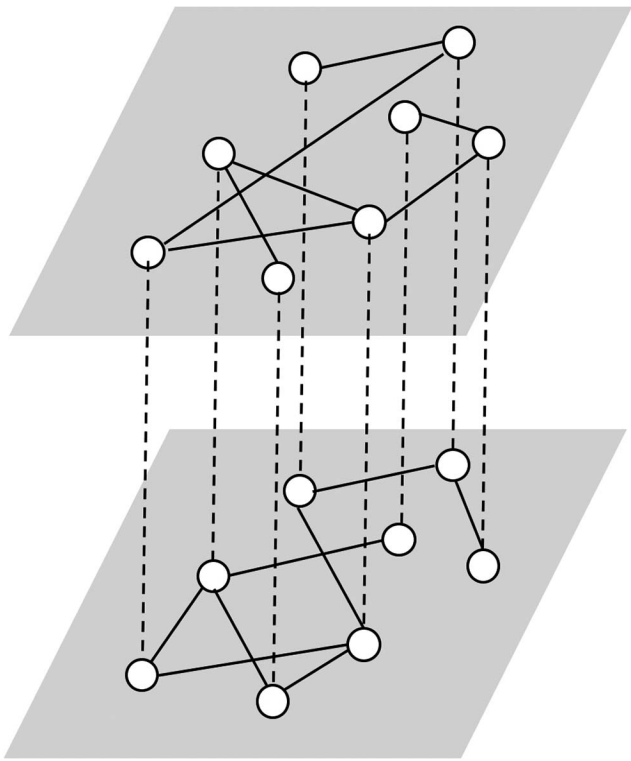

Figure 1. A two-layer multiplex. The same nodes are represented in both layers, and the links within each layer represent a different type of relationship among nodes (Battiston et al., 2014).

These definitions of phonological and orthographic similarity among words have a long history in the field of psycholinguistics (e.g., Coltheart et al., 1977; Greenberg & Jenkins, 1964).

In the phonographic multiplex, words can be (a) phonologically and orthographically related to each other, (b) only phonologically related to each other, (c) only orthographically related to each other, or (d) neither phonologically nor orthographically related to each other. Examining the overlapping areas of the phonological and orthographic layers in the phonographic multiplex (i.e., (a) words that are phonologically and orthographically related to each other) could be particularly relevant for studying the interaction of phonology and orthography in various language processes. Therefore, in this article we focus on the part of the phonographic multiplex where the phonological and orthographic links overlap, that is, the section consisting of links that are found in both phonological and orthographic layers of the multiplex. This section of the phonographic multiplex is hence known as the *phonographic network* (named after the phonological and orthographic layers of the multiplex).

The phonographic network consisted of 5,896 nodes and 11,702 links. The largest connected component of the phonographic network, also known as the giant component, consisted of 3,292 nodes (approximately 55.8% of the entire phonographic network) and 9,583 links. The remainder of the phonographic network consisted of several (~800) lexical islands, smaller connected components of the network that are not connected to the giant component. Note that not all words (~70%) from the original set of 19,340 words (from Vitevitch, 2008) were represented in the phonographic network. The words that were not represented are essentially hermits in the phonographic network, because they were not phonologically and orthographically similar to any words.

Below, the results of a computational analysis of the macro- and mesolevel structure of the largest connected component (LCC) of the phonographic network are briefly reported. Note that only the largest connected component of the network was analyzed, in line with the network science literature, as meaningful network measures cannot be computed if disconnected nodes were included (Newman, 2006; Watts & Strogatz, 1998). To provide a baseline for making comparisons of the structure of the phonographic network, a similarly sized random network was constructed by randomly placing links between nodes (Erdös & Rényi, 1960).

Macrostructure of the Phonographic Network

Small world structure. The average path length of the LCC of the phonographic network was 7.14. On average, approximately seven links had to be traversed to connect any two nodes in the network. The average path length of the random network was 4.79. Although the average path length of the phonographic network was somewhat larger than that of a comparably sized random network, the conventions used in network science would consider these values comparable (Watts & Strogatz, 1998).

The average clustering coefficient of the LCC of the phonographic network was 0.284. The average clustering coefficient of the random network was 0.002. The average clustering coefficient of the phonographic network was much larger by several orders of magnitude than that of a comparably sized random network, indicating that the neighbors of a given node in the phonographic network are more likely to be neighbors of each other, as compared to the neighbors of a given node in the random network.

According to Watts and Strogatz (1998), a small-world network has (a) an average path length comparable to the average path length of a random network, but (b) an average clustering coefficient much larger than the average clustering coefficient of a random network with the same number of nodes and edges. Several real world networks, such as the network of scientific collaborations (Newman, 2004a) and the human brain (Bullmore & Sporns, 2009), possess these two characteristics and are said to have a small-world structure. The above results suggest that, similar to the phonological (Vitevitch, 2008) and semantic (Steyvers & Tenenbaum, 2005) networks of language, the phonographic network has the features of a small-world network.

Degree distribution. A power law degree distribution is a common feature of several real-world networks (Albert & Barabási, 2002). Therefore, analyzing the degree distribution of a network can reveal additional information regarding the overall structure of the network. *Degree distribution* refers to the proportion of nodes that have a given number of links (i.e., degree). If a degree distribution resembles a normal distribution, most nodes have the average number of links per node. If a degree distribution resembles a power law, many nodes have few links (low degree) and a few nodes have many links (high degree). To be consistent with prior theoretical analyses, the degree distribution of words found in the largest component of the phonographic network, rather than of the entire network, was analyzed.

Various distributions (power law, log-normal, exponential) were fit to the degree distribution of the giant component of the phonographic network. The results indicate that the degree distribution of the giant component of the phonographic network was best fit by a log-normal distribution (Kolmogorov–Smirnov statistic = 0.0129, $p = .78$), and not by a power law (Kolmogorov–Smirnov statistic = 0.0677, $p < .001$) or exponential distribution (Kolmogorov–Smirnov statistic = 0.0281, $p < .001$). Note that nonsignificant p values indicate that the degree distribution did not significantly differ from the fitted distribution, whereas significant p values indicate that the degree distribution significantly differed from the fitted distribution.

The degree distributions of phonological networks of different languages (e.g., English, Spanish, Basque; see Arbesman, Strogatz, & Vitevitch, 2010) resembled a truncated power law whereas the degree distribution of the semantic network resembled a power law (Steyvers & Tenenbaum, 2005). In comparison, the degree distribution of the phonographic network was best fit by a log-normal distribution (which indicates that the logarithm of the variable of interest, degree, is normally distributed). Both log-normal and power law distributions are examples of heavy- or fat-tailed distributions, where higher probabilities of extreme values tend to occur (i.e., nodes with very high degree) as compared to a normal distribution.

Mesostructure of the Phonographic Network

In addition to delineating the overall topology of a network (i.e., the macrolevel), the tools of network science also permit us to investigate the mesolevel of a network that is typically exemplified by a network's community structure. *Community structure* refers

to the presence of several smaller groups of nodes within a larger network, where smaller groups form such that there are many links among nodes within a group, but fewer links exist between nodes belonging to different groups (Newman & Girvan, 2004). Communities have been commonly observed in real-world networks such as the structure of the human brain (Wu et al., 2011), the World Wide Web (Newman, 2004b), as well as the phonological network of language (Siew, 2013).

A preliminary community detection analysis was conducted on the giant component of the phonographic network and on the random network. Modularity, Q, is a measure of the density of links within communities as compared to the density of links between communities (Newman, 2006). Positive Q values that are close to the maximum value of 1.0 indicate the presence of high quality communities, where the density of links within communities is high relative to the density of links between communities (Fortunato, 2010). Using the Louvain community detection algorithm, 28 communities with $Q = 0.820$ were detected in the phonographic network. The large positive modularity value implies the presence of robust community structure in the phonographic network—which was also observed in the phonological language network (Siew, 2013). In comparison, 38 communities with a much lower Q of 0.377 were detected in the random network.

Overall, the above analysis of the phonographic network at the macro- and meso-levels reveal that several features of its overall structure are similar to those observed in other real-world networks. Similar to the phonological language network reported in Vitevitch (2008), the phonographic network possesses a small-world structure (i.e., short average path length and high average clustering coefficient), a "small" giant component, and robust community structure. The degree distribution of the phonographic network appeared to follow a log-normal distribution. Overall, this analysis suggests that the structure of the phonographic network is not merely random and may be worth exploring further. In the remainder of the article we examine the influence of two microlevel network metrics on language processing: (a) phonographic degree and (b) phonographic clustering coefficient. These two network metrics will be described in further detail below.

Phonographic Degree

Phonographic degree refers to the number of words that are both phonological and orthographic neighbors of a given word. Therefore, phonographic neighbors differ from the target word by the substitution, deletion, or addition of one phoneme and the substitution, deletion, or addition of one letter. For instance, the phonographic neighbors of "peep" /pip/ include "deep" /dip/, "keep," /kip/, and "pep" /pɛp/, among others. Note that, as shown in the case of "pep" /pɛp/, it is possible that a phonographic neighbor differs from the target word by the substitution of one phoneme and the deletion of one letter—rather than by the substitution of one phoneme and one letter, or the addition of one phoneme and one letter, and so on. As an additional example, consider the word "pant" /pænt/: Its phonographic neighbors include "punt" /pənt/ and "past" /pæst/, but not "panel" /pænL/ (phonological neighbor) and "want" /wɔnt/ (orthographic neighbor). Based on the words in the giant component of the phonographic network, the mean phonographic degree was 5.82 (SD = 4.56) with a range from 1 to 26.

In the visual word recognition literature, there is a body of research investigating the influence of phonographic neighborhood size on language processing (Adelman & Brown, 2007; Muneaux & Ziegler, 2004; Peereman & Content, 1997). The general finding is that the presence of phonographic neighbors facilitates naming of visually presented words (Adelman & Brown, 2007; Peereman & Content, 1997). Phonographic neighborhood size is the same as the phonographic degree measure in the phonographic network, which represents the number of phonographic neighbors a given word has (i.e., the links in the phonographic network). Whereas previous psycholinguistic work did not consider the interconnectivity within a word's neighborhood, the network science approach allows us to quantify the internal structure of a word's neighborhood, as demonstrated below using the clustering coefficient measure. Based on the past literature, one would predict a facilitatory effect of phonographic degree on visual word recognition.

On the other hand, to date there has not been any work studying the role of phonographic neighbors in spoken word recognition, and it is unclear if the presence of more phonographic neighbors would facilitate or inhibit recognition. The presence of more phonographic neighbors could inhibit recognition by contributing greater competition among activated neighbors (Luce & Pisoni, 1998). However, recall that these metrics are obtained from the phonographic network, which represents the part of the phonographic multiplex where the phonological and orthographic layers overlap. Based on this, one would predict that the presence of more phonographic neighbors would facilitate processing in the first layer of the multiplex (e.g., phonological) by providing more of a "boost" in activation in the second layer of the multiplex (e.g., orthographic).

Phonographic Clustering Coefficient

Clustering coefficient (C) measures the extent to which nodes tend to cluster together. That is, to what extent are neighbors of a given word also neighbors of each other. A word with high phonographic C would have phonographic neighbors that tend to also be neighbors of each other whereas a word with low phonographic C would have phonographic neighbors that do not tend to be neighbors of each other. Consider the following two words: "mold" and "pant." Both "mold" and "pant" have 14 phonographic neighbors; however, "mold" has a higher phonographic C (0.440) as compared to "pant" (phonographic $C = 0.121$). As shown in Figure 2 below, the phonographic neighbors of "mold" tend to also be phonographic neighbors of each other (greater density of links within the phonographic neighborhood), whereas the phonographic neighbors of "pant" do not tend to be phonographic neighbors of each other (lower density of links within the phonographic neighborhood). The mean phonographic C of the words in the giant component of the phonographic network was 0.284 ($SD = 0.278$) with values covering the full range of C from 0 to 1 (the convention in network science is to compute C just on nodes in the giant component).

Given previous work showing that phonological C influences lexical retrieval (Chan & Vitevitch, 2010), one might also expect that phonographic C would affect the speed and accuracy of lexical processes. Specifically, Chan and Vitevitch found an inhibitory

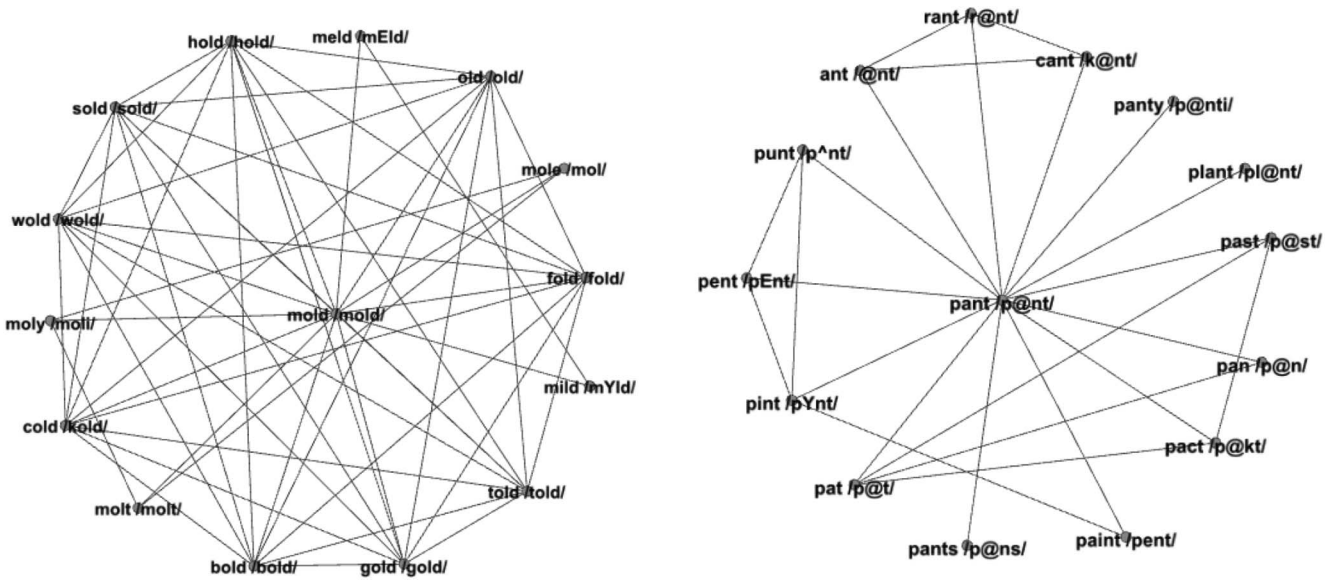

Figure 2. The phonographic neighborhood of *mold* (high phonographic *C*) is shown on the left and the phonographic neighborhood of *pant* (low phonographic *C*) is shown on the right. Both words have the same number of phonographic neighbors but differ in the level of interconnectivity within their neighborhoods. Nodes are labeled with conventional orthography and a computer readable phonological transcription.

effect of phonological *C* on various spoken word recognition tasks. One might therefore expect an inhibitory effect of phonographic *C* on visual and spoken word recognition. Nevertheless, it is important to note that phonological *C* metric used by Chan and Vitevitch was based on the (single) phonological layer of the phonographic multiplex—whereas phonographic *C* measures the internal structure of a word's phonographic neighborhood (based on both layers of the phonographic multiplex). A greater value of phonographic *C* indicates greater similarity in the phonological and orthographic neighborhood structures of a given word. Given past work on "conspiracy models" of word pronunciation, which has shown that words with more consistent neighbors tend to be more quickly named (e.g., Taraban & McClelland, 1987), one would expect that the activation dynamics that occur among similar phonological and orthographic network structures would also "conspire" and lead to the facilitation, rather than inhibition, of lexical retrieval. In summary, with respect to phonographic degree, we expect to replicate the facilitatory effect of phonographic neighborhoods observed in visual word recognition (Adelman & Brown, 2007). With respect to phonographic *C*, previous work on phonological *C* effects (Chan & Vitevitch, 2009) predicts an inhibitory effect of phonographic *C* on spoken and visual word recognition, whereas previous work on conspiracy effects in word recognition (Taraban & McClelland, 1987) predicts a facilitatory effect of phonographic *C* on spoken and visual word recognition.

Experiment 1: Auditory Naming Task

In Experiment 1, a conventional psycholinguistic task was used to examine how phonographic degree and phonographic *C* might influence spoken word recognition. In the auditory naming task, participants repeated the words they heard out loud as quickly and accurately as possible.

The traditional approach in psycholinguistics is the factorial experiment, which entails the selection of two sets of words that are closely matched on a number of variables while manipulating the variable of interest. As linguistic variables tend to be correlated with each other (e.g., words with high phonological degree also tend to occur frequently in the language; Frauenfelder, Baayen, & Hellwig, 1993), it is sometimes difficult to select stimuli that are perfectly matched on all extraneous variables. One solution is to broadly select stimuli and then use a multilevel modeling approach to statistically control for these variables during the analysis of the responses. Given the difficulty of matching stimuli in the present experiments on all extraneous variables, multilevel models will be used to analyze the data from Experiments 1 and 2.

Method

Participants. Sixty native English speakers were recruited from the introductory psychology subject pool at the University of Kansas. All participants had no previous history of speech or hearing disorders and received partial course credit for their participation. The Institution Review Board of the University of Kansas approved all studies reported in this article.

Materials. Two sets of monosyllabic English words were selected as stimuli. The first set consisted of words that varied in phonographic degree and the second set consisted of words that varied in phonographic *C*. The first set of words was selected such that a range of phonographic degree values was represented while allowing phonographic clustering coefficient to vary freely. The second set of words was selected such that a range of phonographic clustering coefficients was represented while allowing phonographic degree to vary freely. There were a total of 160 words: 80 words varying in phonographic degree and 80 words

varying in phonographic *C*. A list of the stimuli is provided in Appendix A.

The stimuli were chosen so as to capture a representative range of lexical variables, while excluding words that had an extreme value of any of the following lexical characteristics: number of phonemes, number of letters, subjective familiarity (measured on a seven-point scale; Nusbaum, Pisoni, & Davis, 1984), word frequency (log-base 10 of frequency counts from the $SUBTLEX_{US}$ corpus; Brysbaert & New, 2009), phonological degree (number of words that are phonologically similar to a given word [i.e., phonological neighborhood density; Luce & Pisoni, 1998]), phonological *C* (measures the extent to which the phonological neighbors of a word are also neighbors of each other; Chan & Vitevitch, 2009), phonological neighborhood frequency (average frequency of the phonological neighbors of a target word), two measures of phonotactic probability (positional segment probability and biphone probability were obtained from the Phonotactic Probability Calculator; Vitevitch & Luce, 2004), orthographic degree (number of words that are orthographically similar to a given word, based on the substitution, addition, or deletion of one letter in a given word), orthographic *C* (measures the extent to which the orthographic neighbors of a word are also neighbors of each other), orthographic neighborhood frequency (average frequency of the orthographic neighbors of a target word), two measures of bigram frequency (average bigram frequency counts and sum of bigram frequency counts by position were obtained from the English Lexicon Project (ELP); Balota et al., 2007). These lexical variables will be included as covariates in the multilevel model.

The key variables of interest are phonographic degree and phonographic clustering coefficient. Phonographic degree refers to the number of words that are both phonological and orthographic neighbors of a given word. Therefore, phonographic neighbors differ from the target word by the substitution, deletion, or addition of one phoneme *and* the substitution, deletion, or addition of one letter. Phonographic clustering coefficient, *C*, refers to the extent to which the phonographic neighbors of a word are also neighbors of each other. To calculate clustering coefficient, the number of links between neighbors of a target word was counted and divided by the number of possible links that could exist among the neighbors. Therefore, the clustering coefficient is the ratio of the actual number of links existing among neighbors to the number of all possible links among neighbors if every neighbor were connected. The value of the clustering coefficient ranges from 0 to 1; when $C = 1$ all neighbors of the word are neighbors of each other; when $C = 0$ no neighbors of the word are neighbors of each other.

A large set of words that varied across phonographic measures was selected for the experiments, and this included words with a phonographic degree of 2. It should be noted that the phonographic *C* value of these words was either 0 or 1 (i.e., a binary value) and is not an accurate representation of the level of interconnectivity among a word's phonographic neighbors. On the other hand, for words with more than two neighbors, clustering coefficient is a continuous variable that ranges from 0 to 1 that represents the extent to which a word's neighbors are also neighbors of each other. Indeed, one known limitation of the *C* measure is that its value can be biased by the node's degree, whereby nodes with few neighbors tend to have a larger clustering coefficient as compared to nodes with several neighbors (see Opsahl & Panzarasa, 2009; Soffer & Vázquez, 2005); although it is important to note that *C* and degree are not correlated in the phonological network of language (Vitevitch et al., 2012). To ensure that the analysis is not biased by words with phonographic *C*s that distort the level of interconnectivity among neighbors, the following items with a phonographic degree of 2 were excluded from the analysis: "balm," "cue," "crime" (phonographic $C = 1$), and "bleed," "slur," and "tomb" (phonographic $C = 0$).

A male native speaker of American English (Michael S. Vitevitch) produced the stimuli by speaking at a normal speaking rate into a high-quality microphone in an Industrial Acoustics Company sound-attenuated booth. Individual sound files for each word were edited from the digital recording with SoundEdit16 (Macromedia, Inc., San Francisco, California). The Normalization function in SoundEdit16 was used to ensure that all sound files were comparable in amplitude.

Procedure. Participants were tested individually. Each participant was seated in front of an iMac computer that was connected to a New Micros response box. PsyScope 1.2.2 was used to randomize and present the stimuli via headphones at a comfortable listening level. A response box containing a dedicated timing board provided millisecond accuracy for the recording of response times.

In each trial, the word "READY" appeared on the screen for 500 ms. Participants heard one of the randomly selected stimuli and were instructed to repeat the word as quickly and accurately as possible. Reaction times (RTs) were measured from stimulus onset to the onset of the participant's verbal response. Verbal responses were recorded for offline scoring of accuracy. The next trial began 1 s after the participant's response was made. Prior to the experimental trials, each participant received five practice trials to become familiar with the task; these trials were not included in the subsequent analyses.

Results

Accuracy was scored offline by an undergraduate research assistant. Trials containing mispronunciations of the word or responses that triggered the voice-key prematurely (e.g., coughing, "uh") were coded as errors. The first author (Cynthia S. Q. Siew) also independently scored ~10% of the data. There was a high level of agreement between the two independent scorers (Cohen's $\kappa = 0.74$; Cohen, 1960).

For the RT data, errors were first excluded, after which responses below 200 ms and above 2,000 ms were eliminated before the overall mean and *SD* of each participant's RT was calculated. Trials with latencies that were 2 *SD*s above or below each participant's mean RT were considered outliers and excluded from analysis. This resulted in ~5% of the data being removed (i.e., ~95% of the trials were included in the analysis). Trials from two items were also excluded from the analysis due to very low overall item accuracies in the naming task (i.e., outliers that were more than 3 *SD*s below the mean accuracy): "lung" (60%) and "mount" (72%).

Using the *lme4* package in R, a linear mixed effects (LME) model was used to predict RTs from the naming data and a

generalized linear mixed effects (GLM) model was used to predict accuracy from the naming data (Bates, Maechler, Bolker, & Walker, 2014). LME and GLM models are regression models where the effects of interest (i.e., phonographic degree and phonographic *C*) are included as predictors in the model, and these effects are evaluated by examining the magnitude and direction of the regression coefficients (Baayen, Davidson, & Bates, 2008; Janssen, 2012). Mixed effects models are increasingly used to analyze psycholinguistic data as the model can take into account the random effects of the participant as well as the effects of specific items used in the experiment (Baayen et al., 2008). The RT model included the following predictors: (a) random effects of participants and items, (b) fixed effects of phonographic degree and phonographic *C*. The accuracy model included the same predictors: (a) random effects of participants and items, (b) fixed effects of phonographic degree and phonographic *C*. For the RT model, additional lexical variables (i.e., subjective familiarity, word frequency, number of phonemes, phonological degree, phonological *C*, phonological neighborhood frequency, phonotactic probability, number of letters, orthographic degree, orthographic *C*, orthographic neighborhood frequency, bigram frequency) were included as covariates to control for any influences these variables may have on word recognition times. For both models all predictor variables were standardized and for the RT model all covariate variables were standardized.

Note that the inclusion of lexical variables (e.g., word frequency, familiarity) as covariates led to convergence issues in the accuracy model. Such models can fail to converge when the number of predictors in the model is high relative to the number of trials or data points (i.e., a complex or imbalanced data structure); and this is particularly true for LME or logistic models with binary responses (see Eager & Roy, 2017). Therefore, these lexical variables were not included as covariates in the accuracy model, and a simpler model that included the main variables of interest (i.e., phonographic degree and phonographic *C*) was fitted to the accuracy data instead.

Finally, we briefly note that in LME models, values for degrees of freedom and *p* values can only be obtained by approximations. Although a number of approximation techniques have been developed to estimate these values, we adopt the Satterthwaite's method as implemented in the *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2017) to obtain the degrees of freedom and *p* values for all the linear mixed effects models reported in this paper (see Tables 2–5).

For the RT model, a two-step hierarchical approach was used. Number of letters, number of phonemes, subjective familiarity, word frequency, orthographic degree, orthographic clustering coefficient, orthographic neighborhood frequency, mean bigram frequency counts and mean bigram frequency counts by position, phonological degree, phonological clustering coefficient, phonological neighborhood frequency, mean positional segment probability, and mean biphone probability were entered into the LME model in Step 1. Phonographic degree and phonographic clustering coefficient were entered into the LME model in Step 2, in addition to the previously entered variables. Partitioning the analysis into two steps was done to determine if the network measures from the phonographic network accounted for additional variance over conventional lexical variables.

RT. Table 2 presents the results of the LME model for naming RTs. The overall mean RT was 916 ms (*SD* = 175 ms). The following fixed effects were significant: phonographic *C*, familiarity, number of phonemes, number of letters, orthographic degree, orthographic *C*, orthographic neighborhood frequency, average bigram counts, and sum of bigram counts by position. Phonographic degree did not significantly predict naming RTs, standardized $\beta = -14.67$, $t = -1.72$, $p = .088$. Phonographic *C* significantly predicted naming RTs, standardized $\beta = -12.39$, $t = -2.40$, $p = .017$, such that words with higher phonographic *C* were more quickly named as compared to words with lower phonographic *C*. For each standardized unit increase in phonographic *C* (approximately 0.153), the average decrease in naming RTs was 12 ms. The likelihood ratio test indicated that the inclusion of phonographic network measures significantly improved model fit, $\chi^2 = 7.80$, $df = 2$, $p = .020$.[1]

Accuracy. Table 3 presents the results of the generalized LME model for naming accuracies. The overall mean accuracy was 98.18% (*SD* = 13.37). No fixed effects were significant. Both phonographic degree and phonographic *C* did not significantly predict naming accuracies, both *p*s $> .05$.

Discussion

The results of Experiment 1 showed that phonographic *C* predicted naming RTs. Higher phonographic *C* words were named more quickly than lower phonographic *C* words, after taking into account the variance contributed by several lexical variables known to influence language processing.

Recall that phonographic *C* refers to the extent to which the phonographic neighbors of a word are also phonographic neighbors of each other, and that phonographic neighbors are words that both phonologically and orthographically similar to a target word. In the present study, a facilitatory effect was observed for words with a higher level of interconnectivity among its phonographic neighbors. At first glance, this result appears to contradict previous work investigating the influence of the phonological clustering coefficient on spoken word recognition, which found that words with high phonological clustering coefficients were more slowly and less accurately processed. Note that Chan and Vitevitch (2009) used a factorial design, therefore the label "high" is appropriate; see also Siew, 2017 for a similar finding with the network density measure.

A simple diffusion framework was used to account for this finding. In this framework, activation spreads back and forth between the target word, its neighbors, and other words in the network (see also the computer simulation reported in Vitevitch, Ercal, & Adagarla, 2011). For words with highly interconnected neighborhoods, over time a greater amount of activation will remain within the neighborhood, instead of diffusing to the rest of

[1] Note that an examination of the variance inflation factors (VIFs) for predictors in this model (and all regression models in the article) revealed that lexical predictors were moderately correlated with each other (see Table 6). Nevertheless, VIFs of the predictors were generally within the acceptable range and the model does not suffer from severe multicollinearity issues. Table 6 shows a correlation matrix of the lexical variables included in the linear mixed effects model and the results of the analysis should be interpreted with the understanding that lexical variables tend to be correlated with each other.

Table 2
Linear Mixed Effects Model Estimates for Fixed and Random Effects for the Auditory Naming Experiment (Reaction time; Experiment 1)

Variable	Variance	β	*SD*	*SE*	*df*	*t*	*p*-value
Random effects							
Items							
Intercept	2096.00		45.78				
Participants							
Intercept	22,115.00		148.71				
Fixed effects							
Step 1							
Number of phonemes		19.89		11.15	150.49	1.78	.076
Phonological degree		−2.74		7.25	150.59	−.38	.706
Phonological *C*		−2.70		4.51	150.74	−.60	.551
Phonological neighborhood frequency		−6.79		5.24	150.59	−1.30	.197
Positional segment probability		−1.28		7.11	150.63	−.18	.858
Biphone probability		−7.78		7.35	150.64	−1.06	.292
Number of letters		16.11		7.20	150.67	2.24	.027*
Orthographic degree		12.77		5.50	150.76	2.32	.022*
Orthographic *C*		9.58		4.24	151.15	2.26	.025*
Orthographic neighborhood frequency		11.65		4.37	150.86	2.67	.009**
Average bigram counts		14.41		5.24	150.74	2.75	.007**
Sum of bigram counts by position		−15.90		6.71	150.50	−2.37	.019*
Subjective familiarity		7.58		4.37	150.76	1.73	.085
Word frequency		−.55		4.60	150.77	−.12	.905
Step 2							
Phonographic degree[a]		−14.67		8.55	151.06	−1.72	.088
Phonographic *C*[a]		−12.34		5.13	150.68	−2.40	.017*
Number of phonemes		24.21		11.01	150.50	2.20	.029*
Phonological degree		2.66		7.69	150.67	.35	.730
Phonological *C*		−1.37		4.49	150.77	−.31	.760
Phonological neighborhood frequency		−6.18		5.11	150.60	−1.21	.228
Positional segment probability		−.37		6.94	150.64	−.05	.958
Biphone probability		−8.62		7.18	150.67	−1.20	.232
Number of letters		15.22		7.05	150.68	2.16	.033*
Orthographic degree		23.81		8.48	151.04	2.81	.006**
Orthographic *C*		17.75		5.25	151.13	3.38	<.001***
Orthographic neighborhood frequency		11.36		4.30	150.88	2.64	.009**
Average bigram counts		14.53		5.11	150.76	2.85	.005**
Sum of bigram counts by position		−16.22		6.65	150.47	−2.44	.016*
Subjective familiarity		8.60		4.28	150.80	2.01	.046*
Word frequency		−1.76		4.50	150.83	−.39	.696

[a] Variables added in Step 2.
* $p < .05$. ** $p < .01$. *** $p < .001$.

the network. On the other hand, for words with less interconnected neighborhoods, over time most of the activation will be spread to the rest of the network. Based on this account, it is more difficult for words with highly interconnected neighborhoods (i.e., words with high phonological clustering coefficients) to "stand out" from its competitors as compared to words with less interconnected neighborhoods (i.e., words with low phonological clustering coefficients).

However, it is important to note that although both phonological *C* and phonographic *C* represent the amount of interconnectivity among a word's neighbors, these two measures are different in that phonological *C* represents the structure of a word's phonological neighborhood (i.e., the phonological layer in the phonographic multiplex), whereas phonographic *C* represents the structure of a word's phonological and orthographic neighborhoods (i.e., the phonological and orthographic layers of the phonographic multiplex). More specifically, phonographic *C* can be viewed as a metric that represents the internal structure of the area where the phonological and orthographic neighborhoods of words overlap. Therefore, a facilitatory effect might be expected in this case because higher phonographic *C* values indicate greater overlap in the similarity structures of the phonological and orthographic neighborhoods of words. Based on the activation diffusion framework described earlier, one might expect that for high phonographic *C* words, the diffusion process operates on both the phonological layer and the orthographic layer of the phonographic multiplex, leading to similar, overlapping patterns of activation in the parts of the multiplex where phonological and orthographic similarity overlap, which reinforce each other during processing and facilitate the recognition of the target word. It is also noteworthy that the facilitatory effects of phonographic *C* align well with the widely known notion of conspiracy effects, where overlapping, consistent phonological and orthographic information "conspire" to produce facilitative effects in word naming (e.g., Taraban & McClelland, 1987). Finally, it is important to highlight that the diffusion process might predict either inhibitory or facil-

Table 3
Generalized Linear Mixed Effects Model Estimates for Fixed and Random Effects for the Auditory Naming Experiment (Accuracy; Experiment 1)

Variable	Variance	β	*SD*	*SE*	*z*	*p*-value
Random effects						
Items						
Intercept	1.74		1.32			
Participants						
Intercept	.32		.56			
Fixed effects						
Intercept		4.74		.19	24.78	<.001***
Phonographic degree		−.01		.14	−.08	.93
Phonographic *C*		−.21		.13	−1.53	.13

*** $p < .001$.

itative effects depending on the nature of the structural representation that the process is implemented on. Specifically, greater network interconnectivity might predict facilitation when the measure is computed on a representation that represents overlapping similarity relations of phonology and orthography (i.e., the phonographic network), whereas greater network interconnectivity might predict increased competition and hence inhibition when the measure is computed on a representation that only considered one type of similarity relation between words (e.g., phonological similarity only, Chan & Vitevitch, 2010; Luce & Pisoni, 1998).

The present result is significant because it is the first to demonstrate that a network science metric—the phonographic clustering coefficient—which simultaneously represents the phonological and orthographic structure of language, influences spoken word recognition. However, it is important to note that all psycholinguistic tasks have unique task demands and the auditory naming task might reflect, to a certain extent, nonlexical processes involving repetition ability (Nozari, Kittredge, Dell, & Schwartz, 2010). To ensure that these findings are not task-specific and that they can indeed be replicated using a different psycholinguistic task, the next experiment sought to replicate the present findings using another traditional task from psycholinguistics—auditory lexical decision.

Experiment 2: Auditory Lexical Decision

The aim of Experiment 2 was to replicate the findings of Experiment 1 with another commonly used psycholinguistic task—auditory lexical decision. In this task, participants are auditorily presented with words and nonwords and have to decide if the given stimulus was a real word or not.

Method

Participants. Sixty-five native English speakers were recruited from the same population described in Experiment 1. All participants were right-handed and had no previous history of speech or hearing disorders; none took part in Experiment 1.

Materials. The word stimuli for the present experiment consisted of the same 160 words used in Experiment 1. In addition, 160 phonotactically legal nonwords were constructed by replacing a phoneme (at any position except the first and last positions) of the word stimuli with another phoneme. For instance, the nonword "brame" (/bɹem/) was created by replacing /l/ in the word "blame" (/blem/) with /ɹ/. The phonological transcriptions of the nonwords are listed in Appendix B. The nonwords were recorded by the same male speaker in a similar manner as in Experiment 1. The same method for editing and digitizing the word stimuli was used to create individual sound files for each nonword. The Normalization function in SoundEdit16 was used to ensure that all word and nonword sound files were comparable in amplitude. Stimuli durations were equivalent across both words and nonwords, $t(318) < 1$, $p = .92$.

Procedure. Participants were tested in groups no larger than three. The same equipment used in Experiment 1 was used in the present experiment, except that a response box containing a dedicated timing board was used to record response times.

In each trial, the word "READY" appeared on the screen for 500 ms. Participants heard one of the randomly selected stimuli and were instructed to decide, as quickly and accurately as possible, whether the item heard was a real English word or a nonword. If the item was a word, participants pressed the button labeled "WORD" with their right index finger. If the item was a nonword, participants pressed the button labeled "NONWORD" with their left index finger. RTs were measured from stimulus onset to the onset of the participant's button press. The next trial began 1s after the participant's response was made. Prior to the experimental trials, each participant received eight practice trials to become familiar with the task; these trials were not included in the subsequent analyses.

Results

The trimming procedure is identical to that used in Experiment 1. For the RT data, errors were first excluded, after which responses below 200 ms and above 2,000 ms were eliminated before the overall mean and *SD* of each participant's RT was calculated. Trials with latencies that were 2 *SD*s above or below each participant's mean RT were removed. This resulted in ~7% of the data being removed (i.e., ~93% of the trials were included in the analysis). Nonword trials were also excluded and not analyzed further. Trials from four items were also excluded from the analysis due to very low overall item accuracies in the lexical decision task (i.e., outliers that were more than 3 *SD*s below the mean accuracy): "clod" (23%), "balk" (32%), "plume" (38%), and "posh" (46%).

As in Experiment 1, LME and GLM models were used to predict RTs and accuracy respectively from the lexical decision data. The RT model included the following predictors: (a) random effects of participants and items, (b) fixed effects of phonographic degree and phonographic *C*. The accuracy model included the same predictors: (a) random effects of participants and items, (b) fixed effects of phonographic degree and phonographic *C*. For the RT model, additional lexical variables (i.e., subjective familiarity, word frequency, number of phonemes, phonological degree, phonological *C*, phonological neighborhood frequency, phonotactic probability, number of letters, orthographic degree, orthographic *C*, orthographic neighborhood frequency, bigram frequency) were included as covariates to control for any influences these variables may have on word recognition times. For both models all predictor

Table 4
Linear Mixed Effects Model Estimates for Fixed and Random Effects for the Auditory Lexical Decision Experiment (Reaction Time; Experiment 2)

Variable	Variance	β	*SD*	*SE*	*df*	*t*	*p*-value
Random effects							
Items							
Intercept	3127.00		55.92				
Participants							
Intercept	7138.00		84.49				
Fixed effects							
Step 1							
Number of phonemes		7.54		14.26	149.08	.53	.598
Phonological degree		−5.03		9.33	148.19	−.54	.591
Phonological *C*		−5.51		5.64	149.39	−.98	.331
Phonological neighborhood frequency		4.17		6.63	148.18	.63	.531
Positional segment probability		5.94		9.11	148.80	.65	.515
Biphone probability		−8.51		9.39	147.95	−.91	.366
Number of letters		7.52		9.18	147.75	.82	.414
Orthographic degree		8.27		7.11	147.58	1.16	.247
Orthographic *C*		9.38		5.52	147.48	1.70	.092
Orthographic neighborhood frequency		6.52		5.50	148.63	1.19	.238
Average bigram counts		15.38		6.58	147.77	2.34	.021*
Sum of bigram counts by position		−13.74		8.64	147.28	−1.59	.114
Subjective familiarity		−12.77		5.15	151.63	−2.48	.014*
Word frequency		−10.06		5.73	148.32	−1.76	.081
Step 2							
Phonographic degree[a]		−17.73		11.18	147.79	−1.59	.115
Phonographic *C*[a]		−18.41		7.07	147.50	−2.60	.010*
Number of phonemes		12.06		14.00	149.05	.86	.390
Phonological degree		.46		9.76	148.07	.05	.963
Phonological *C*		−2.64		5.68	149.35	−.47	.642
Phonological neighborhood frequency		5.14		6.45	147.91	.80	.427
Positional segment probability		7.37		8.87	148.60	.83	.407
Biphone probability		−10.00		9.14	147.77	−1.09	.276
Number of letters		6.16		8.96	147.47	.69	.493
Orthographic degree		21.80		11.17	147.47	1.95	.053
Orthographic *C*		22.33		7.22	147.09	3.09	.002**
Orthographic neighborhood frequency		5.36		5.39	148.14	.99	.322
Average bigram counts		15.61		6.40	147.48	2.44	.016*
Sum of bigram counts by position		−13.42		8.53	146.93	−1.57	.118
Subjective familiarity		−13.08		5.01	151.58	−2.61	.010*
Word frequency		−11.72		5.60	148.23	−2.09	.038*

[a] Variables added in Step 2.
* $p < .05$. ** $p < .01$.

variables were standardized and for the RT model all covariate variables were standardized.

As described previously, the inclusion of lexical variables (e.g., word frequency, familiarity) as covariates led to convergence issues in the accuracy model. Therefore, these lexical variables were not included as covariates in the accuracy model, and a simpler model that included the main variables of interest (i.e., phonographic degree and phonographic *C*) was fitted to the accuracy data instead.

For the RT model, a two-step hierarchical approach was used. Number of letters, number of phonemes, subjective familiarity, word frequency, orthographic degree, orthographic clustering coefficient, orthographic neighborhood frequency, mean bigram frequency counts and mean bigram frequency counts by position, phonological degree, phonological clustering coefficient, phonological neighborhood frequency, mean positional segment probability and mean biphone probability were entered into the LME model in Step 1. Phonographic degree and phonographic clustering coefficient were entered into the LME model in Step 2, in addition to the previously entered variables. Partitioning the analysis into two steps was done to determine if the network measures from the phonographic network accounted for additional variance over conventional lexical variables.

RT. Table 4 presents the results of the LME model for lexical decision RTs. The overall mean RT was 898 ms ($SD = 177$ ms). In Step 2, the following fixed effects were significant: phonographic *C*, frequency, familiarity, orthographic *C*, and average bigram counts. Phonographic degree did not significantly predict lexical decision RTs, standardized $\beta = -17.73$, $t = -1.59$, $p = .115$. Phonographic *C* significantly predicted lexical decision RTs, standardized $\beta = -18.41$, $t = -2.60$, $p = .010$, such that words with higher phonographic *C* were more quickly responded to as compared to words with lower phonographic *C*. For each standardized unit increase in phonographic *C* (approximately 0.153), the average decrease in naming RTs was 18 ms. The likelihood ratio test indicated that the inclusion of phonographic network

Table 5
Generalized Linear Mixed Effects Model Estimates for Fixed and Random Effects for the Auditory Lexical Decision Experiment (Accuracy; Experiment 2)

Variable	Variance	β	*SD*	*SE*	*z*	*p*-value
Random effects						
Items						
Intercept	1.52		1.23			
Participants						
Intercept	.37		.61			
Fixed effects						
Intercept		2.75		.13	20.63	<.001***
Phonographic degree		.12		.11	1.08	.28
Phonographic *C*		.07		.11	.64	.53

*** $p < .001$.

measures significantly improved model fit, $\chi^2 = 8.40$, $df = 2$, $p = .015$.

Accuracy. Table 5 presents the results of the GLM model for lexical decision accuracies. The overall mean accuracy was 88.06% (*SD* = 32.43). No fixed effects were significant. Both phonographic degree and phonographic *C* did not significantly predict lexical decision accuracies, both *p*s > .05.

Discussion

The results of Experiment 2 showed that phonographic *C* predicted lexical decision RTs, replicating the results of Experiment 1. High phonographic *C* words were recognized more quickly than low phonographic *C* words, after taking into account the variance contributed by several lexical variables known to influence language processing.

As discussed earlier, phonographic *C* represents the internal structure of the area where the phonological and orthographic neighborhoods of words overlap, such that higher phonographic *C* values indicate greater overlap in the similarity structures of the phonological and orthographic neighborhoods of words. Based on the activation diffusion framework described above, similar, overlapping patterns of activation are more likely to occur in the phonological and orthographic neighborhood structures of words with higher phonographic *C* words, as compared to words with lower phonographic *C*. These similar, overlapping patterns of activation reinforce each other during processing, and hence serve to facilitate the recognition of the target word.

In addition, it is worth noting that phonographic *C* was a significant predictor in both experiments. The effect of phonographic degree was in the same direction as phonographic *C* (higher phonographic degree was associated with faster RTs), although it is important to note that this effect was marginally significant in both experiments (*p*s = .09 and .11 in auditory naming and lexical decision respectively). Both measures of degree and *C* capture somewhat different aspects of the phonological and orthographic similarity structure of language. Phonographic degree simply represents the number of words that are both phonologically and orthographically similar to a given word, whereas phonographic *C* captures more subtle aspects of the similarity structure—namely, the internal connectivity among these phonographic neighbors. Overall, the results suggest that phonographic

Table 6
Correlation Matrix of All Predictors Included in the Linear Mixed Effect and Regression Models

Variable	Phonographic *C*	Subjective familiarity	Word frequency	Number of phonemes	Phonological degree	Phonological *C*	Phonological neighborhood frequency	Positional segment probability	Biphone probability	Number of letters	Orthographic degree	Orthographic *C*	Orthographic neighborhood frequency	Average bigram counts	Sum of bigram counts by position
Phonographic degree	−.11	.02	.07	−.29	.57	.02	−.03	−.11	−.18	−.42	.86	−.07	.12	.12	−.18
Phonographic *C*		.05	−.03	−.05	−.03	.20	.01	−.08	−.05	−.03	−.03	.54	.08	−.16	−.08
Subjective familiarity			.35	−.10	.02	−.07	.17	−.13	−.06	−.07	.04	.02	.07	−.04	−.01
Word frequency				−.15	.17	.02	.28	−.05	.01	.09	.11	−.11	.13	.25	.27
Number of phonemes					−.75	−.19	−.36	.73	.68	.58	−.37	−.07	−.24	.15	.22
Phonological degree						.11	.27	−.45	−.46	−.45	.54	−.03	.26	.09	−.08
Phonological *C*							.12	−.34	−.14	.05	.08	.23	.00	−.12	−.05
Phonological neighborhood frequency								−.22	−.02	.04	.03	−.07	.33	.24	.37
Positional segment probability									.71	.32	−.16	−.20	−.17	.27	.34
Biphone probability										.35	−.19	−.18	−.10	.40	.53
Number of letters											−.52	.02	−.12	.26	.52
Orthographic degree												−.09	.23	.11	−.15
Orthographic *C*													.07	−.27	−.25
Orthographic neighborhood frequency														.04	.08
Average bigram counts															.55

C is a better predictor of spoken word recognition performance than phonographic degree. Together the results of Experiments 1 and 2 demonstrate that the phonographic clustering coefficient, a network science metric that simultaneously represents the phonological and orthographic structure of language, influences spoken word recognition.

English Lexicon Project Analyses

The availability of item-level behavioral data for a large set of words resulting from megastudies of visual word recognition (New et al., 2006; Yap & Balota, 2009) offers another way to complement the conventional psycholinguistic approach of factorial experiments in a small-scale study (Balota, Yap, Hutchison, & Cortese, 2012). In a factorial experiment, psycholinguists typically carefully select word stimuli such that groups of words are matched on a variety of lexical characteristics while manipulating the lexical variable of interest, whereas in the megastudy approach, extraneous lexical variables can be statistically controlled for. In addition, whereas lab-based experiments with carefully controlled stimuli can answer the question of whether phonographic degree and phonographic *C* influence word recognition, the large-database approach can answer the slightly different question of how much influence phonographic degree and phonographic *C* have on word recognition performance, after taking into account the influence of other lexical variables on word recognition. The database approach also allows for replication using a larger set of stimuli. In this section a regression analysis of words in the ELP was conducted to determine if phonographic degree and phonographic clustering coefficient are significant predictors of performance of speeded visual naming and visual lexical decision for a large set of words, after taking into account the contributions of other lexical variables.

Method

Database. The English Lexicon Project is a large database that contains descriptive and behavioral data for over 40,000 words (see Balota et al., 2007 for a complete description of the database). It is available at http://elexicon.wustl.edu.

Dataset/materials. ELP behavioral data exist for 2,914 of the 3,292 (~90%) words in the giant component of the phonographic network. It is important to note that some of the words in the phonographic network do not have a "meaningful" phonographic clustering coefficient value. For instance, it is not possible to calculate the clustering coefficient for words with either zero or one phonographic neighbor(s), that is, phonographic *C* for these words is undefined. As discussed earlier, for words with more than two neighbors, clustering coefficient is a continuous variable that ranges from zero to one that represents the extent to which a word's neighbors are also neighbors of each other. However, the phonographic clustering coefficient for words with two phonographic neighbors is binary (i.e., either zero or one) and does not accurately represent the level of interconnectivity among a word's phonographic neighbors. Therefore, to ensure that the analysis was not biased by the presence of several words with an undefined phonographic *C* (i.e., words with a phonographic degree of 1), or by words with phonographic *C*s that distort the level of interconnectivity among neighbors (i.e., words with a phonographic degree of 2), words with two or fewer phonographic neighbors were excluded, resulting in a total of 2,120 words for the subsequent regression analyses.

Results

Item-level regression analyses were conducted on the mean RTs and accuracies for 2,120 words for speeded naming and visual lexical decision tasks that were obtained from the ELP. The dependent variables consisted of *z*-scored RTs and accuracy rates, averaged across participants for each word, for both speeded naming and lexical decision tasks. Each participant's raw naming and lexical decision latencies were first standardized using a z-score transformation, and the mean z-score for all participants presented with a particular word is then computed for that word (Balota et al., 2007). Although both raw and *z*-scored RTs are available in the ELP, *z*-scored RTs, instead of raw RTs, were analyzed to reduce the likelihood that a single participant may disproportionately influence the item means (Balota et al., 2007).

A two-step hierarchical approach was used. Number of letters, number of phonemes, subjective familiarity, word frequency, orthographic degree, orthographic clustering coefficient, orthographic neighborhood frequency, mean bigram frequency counts and mean bigram frequency counts by position, phonological degree, phonological clustering coefficient, phonological neighborhood frequency, mean positional segment probability and mean biphone probability were entered into the regression model in Step 1. Phonographic degree and phonographic clustering coefficient were entered into the regression model in Step 2, in addition to the previously entered variables. Partitioning the regression analysis into two steps was done to determine if the network measures from the phonographic network accounted for additional variance over conventional lexical variables.

Speeded naming.

RTs. Table 7 shows the results of the regression analysis on *z*-scored naming RTs. In Step 1, the following variables significantly predicted naming RTs: number of phonemes, phonological degree, positional segment probability, biphone probability, number of letters, orthographic *C*, average bigram counts, sum of bigram counts by position, familiarity, and frequency. Together, the variables entered at Step 1 explained 28.0% of the variance in naming RTs, accounting for a significant proportion of the variance in naming RTs, $R^2 = .280$, $F(14, 2103) = 58.27$, $p < .001$.

In Step 2, the following variables significantly predicted RTs: positional segment probability, biphone probability, number of letters, orthographic degree, average bigram counts, sum of bigram counts by position, familiarity, frequency, and phonographic degree. Phonographic degree significantly predicted visual naming RTs, standardized $\beta = -0.0763$, $t(2101) = -7.72$, $p < .001$, such that words with higher phonographic degree were more quickly named as compared to words with lower phonographic degree. For each standardized unit increase in phonographic degree (approximately 4.31), the average decrease in *z*-scored naming RTs was 0.076 *SD*s. Phonographic *C* did not significantly predict naming RTs, standardized $\beta = -0.0102$, $t(2101) = -1.37$, $p = .172$. The influence of phonographic variables accounted for an additional 1.9% of the variance, $\Delta R^2 = .019$, $F(2, 2101) = 29.85$, $p < .001$. Together, the variables entered at both steps explained 29.9% of the variance in naming RTs, accounting for a significant propor-

Table 7
Regression Results for English Lexicon Project Naming Reaction Times

Variable	β	*SE*	*t*	*p*	R^2	ΔR^2
Step 1						
Number of phonemes	−.0254	.0103	−2.46	.014*		
Phonological degree	−.0191	.00724	−2.64	.008**		
Phonological *C*	.00742	.00465	1.60	.11		
Phonological neighborhood frequency	.00356	.00529	.673	.50		
Positional segment probability	.0435	.00734	5.92	<.001***		
Biphone probability	−.0252	.00710	−3.55	<.001***		
Number of letters	.0723	.00832	8.69	<.001***		
Orthographic degree	−.00843	.00669	−1.26	.21		
Orthographic *C*	−.0132	.00453	−2.92	.004**		
Orthographic neighborhood frequency	.00734	.00525	1.40	.16		
Average bigram counts	.0235	.00501	4.70	<.001***		
Sum of bigram counts by position	−.0323	.00648	−4.99	<.001***		
Subjective familiarity	−.0516	.00451	−11.44	<.001***		
Word frequency	−.0450	.00486	−9.26	<.001***		
					.280***	
Step 2						
Number of phonemes	−.0138	.0103	−1.34	.18		
Phonological degree	−.00337	.00773	.436	.66		
Phonological *C*	.00786	.00468	1.68	.09		
Phonological neighborhood frequency	.000791	.00524	.151	.88		
Positional segment probability	.0447	.00726	6.16	<.001***		
Biphone probability	−.0227	.00701	−3.24	.001**		
Number of letters	.0758	.00822	9.22	<.001***		
Orthographic degree	.0523	.0103	5.10	<.001***		
Orthographic *C*	−.000485	.00739	−.066	.95		
Orthographic neighborhood frequency	.00799	.00518	1.54	.12		
Average bigram counts	.0260	.00495	5.25	<.001***		
Sum of bigram counts by position	−.0420	.00652	−6.44	<.001***		
Subjective familiarity	−.0501	.00445	−11.26	<.001***		
Word frequency	−.0489	.00482	−10.15	<.001***		
Phonographic degree	−.0763	.00989	−7.72	<.001***		
Phonographic *C*	−.0102	.00744	−1.37	.17		
					.299***	.019***

Note. $N = 2{,}120$. "Phonographic degree" and "Phonographic C" were * variables added in Step 2, as in Table 2 and Table 4.
* $p < .05$. ** $p < .01$. *** $p < .001$.

tion of the variance in naming RTs, $R^2 = .299$, $F(16, 2101) = 56.12$, $p < .001$.

Accuracy. Table 8 shows the results of the regression analysis on visual naming accuracies. In Step 1, the following variables significantly predicted accuracies: number of phonemes, phonological neighborhood frequency, biphone probability, orthographic *C*, orthographic neighborhood frequency, sum of bigram counts by position, familiarity, and frequency. Together, the variables entered at Step 1 explained 27.5% of the variance in naming accuracies, accounting for a significant proportion of the variance in naming accuracies, $R^2 = .275$, $F(14, 2103) = 56.86$, $p < .001$.

In Step 2, the following variables significantly predicted naming accuracies: biphone probability, orthographic degree, orthographic neighborhood frequency, sum of bigram counts by position, familiarity, frequency, and phonographic degree. Phonographic degree significantly predicted naming accuracies, standardized $\beta = 0.0108$, $t(2101) = 4.07$, $p < .001$, such that words with higher phonographic degree were more accurately named as compared to words with lower phonographic degree. For each standardized unit increase in phonographic degree (approximately 4.31), the average increase in naming accuracies was 1.09%. Phonographic *C* did not significantly predict naming accuracies, standardized $\beta = -0.0017$, $t(2101) < 1$, $p = .405$. The influence of phonographic variables accounted for an additional 0.6% of the variance, $\Delta R^2 = .006$, $F(2, 2101) = 9.22$, $p < .001$. Together, the variables entered at both steps explained 28.1% of the variance in visual naming accuracies, accounting for a significant proportion of the variance in naming accuracies, $R^2 = .281$, $F(16, 2101) = 51.29$, $p < .001$.

Visual lexical decision.

RTs. Table 9 shows the results of the regression analysis on *z*-scored lexical decision RTs. In Step 1, the following variables significantly predicted visual lexical decision RTs: number of letters, orthographic degree, orthographic *C*, familiarity, and frequency. Together, the variables entered at Step 1 explained 50.5% of the variance in lexical decision RTs, accounting for a significant proportion of the variance in lexical decision RTs, $R^2 = .505$, $F(14, 2103) = 153.50$, $p < .001$.

In Step 2, the following variables significantly predicted lexical decision RTs: number of letters, average bigram counts, familiarity, frequency, and phonographic degree. Phonographic degree significantly predicted lexical decision RTs, standardized $\beta = -0.0348$, $t(2101) = -3.01$, $p = .003$, such that words with higher phonographic degree were more quickly responded to as compared to words with lower phonographic degree. For each

Table 8
Regression Results for English Lexicon Project Naming Accuracies

Variable	β	*SE*	*t*	*p*	R^2	ΔR^2
Step 1						
Number of phonemes	.00605	.00275	2.20	.028*		
Phonological degree	.00293	.00193	1.52	.13		
Phonological *C*	.000633	.00124	.512	.61		
Phonological neighborhood frequency	−.00327	.00141	−2.32	.020*		
Positional segment probability	−.00349	.00195	−1.79	.074†		
Biphone probability	.00534	.00189	2.83	.005**		
Number of letters	.00111	.00221	.502	.62		
Orthographic degree	.00159	.00178	.892	.37		
Orthographic *C*	.00243	.00120	2.02	.043*		
Orthographic neighborhood frequency	.00290	.00140	2.08	.038*		
Average bigram counts	−.00130	.00133	−.973	.33		
Sum of bigram counts by position	−.00516	.00172	−3.00	.003**		
Subjective familiarity	.0274	.00120	22.86	<.001***		
Word frequency	.00353	.00129	2.73	.006**		
					.275***	
Step 2						
Number of phonemes	.00432	.00277	1.56	.12		
Phonological degree	−.000402	.00208	−.194	.85		
Phonological *C*	.000977	.00126	.777	.44		
Phonological neighborhood frequency	−.00274	.00141	−1.95	.051†		
Positional segment probability	−.00349	.00195	−1.79	.073†		
Biphone probability	.00490	.00188	2.60	.009**		
Number of letters	.000615	.00221	.279	.78		
Orthographic degree	−.00679	.00276	−2.46	.014*		
Orthographic *C*	.00289	.00198	1.46	.15		
Orthographic neighborhood frequency	.00282	.00139	2.03	.043*		
Average bigram counts	−.00169	.00133	−1.27	.20		
Sum of bigram counts by position	−.00376	.00175	−2.15	.032*		
Subjective familiarity	.0272	.00120	22.78	<.001***		
Word frequency	.00411	.00129	3.17	.002**		
Phonographic degree	.0108	.00265	4.07	<.001***		
Phonographic *C*	−.00167	.00200	−.83	.40		
					.281***	.006***

Note. $N = 2{,}120$. "Phonographic degree" and "Phonographic C" were * variables added in Step 2, as in Table 2 and Table 4.
† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

standardized unit increase in phonographic degree (approximately 4.31), the average decrease in *z*-scored lexical decision RTs was 0.035 *SD*s. Phonographic *C* did not significantly predict lexical decision RTs, standardized $\beta = 0.0010$, $t(2101) < 1$, $p = .909$. The influence of phonographic variables accounted for an additional 0.3% of the variance, $\Delta R^2 = .003$, $F(2, 2101) = 4.66$, $p = .010$. Together, the variables entered at both steps explained 50.8% of the variance in lexical decision RTs, accounting for a significant proportion of the variance in lexical decision RTs, $R^2 = .508$, $F(16, 2101) = 135.40$, $p < .001$.

Accuracy. Table 10 shows the results of the regression analysis on visual lexical decision accuracies. In Step 1, the following variables significantly predicted lexical decision accuracies: phonological neighborhood frequency, number of letters, orthographic degree, sum of bigram counts by position, familiarity, and frequency. Together, the variables entered at Step 1 explained 65.0% of the variance in lexical decision accuracies, accounting for a significant proportion of the variance in lexical decision accuracies, $R^2 = .650$, $F(14, 2103) = 279.20$, $p < .001$.

In Step 2, the following variables significantly predicted lexical decision accuracies: phonological degree, phonological neighborhood frequency, number of letters, familiarity, frequency, and phonographic degree. Phonographic degree significantly predicted visual lexical decision accuracies, standardized $\beta = 0.0176$, $t(2101) = 3.05$, $p = .002$, such that words with higher phonographic degree were more accurately responded to as compared to words with lower phonographic degree. For each standardized unit increase in phonographic degree (approximately 4.31), the average increase in lexical decision accuracies was 1.76%. Phonographic *C* did not significantly predict lexical decision accuracies, standardized $\beta = 0.0017$, $t(2101) < 1$, $p = .692$. The influence of phonographic variables accounted for an additional 0.2% of the variance, $\Delta R^2 = .002$, $F(2, 2101) = 4.66$, $p = .01$. Together, the variables entered at both steps explained 65.2% of the variance in lexical decision accuracies, accounting for a significant proportion of the variance in lexical decision accuracies, $R^2 = .652$, $F(16, 2101) = 245.70$, $p < .001$.

Discussion

The results of the ELP analyses showed that phonographic degree, but not phonographic *C*, predicted naming and lexical decision RTs in visual word recognition. Words with higher phonographic degree were named and recognized more quickly than

Table 9
Regression Results for English Lexicon Project Visual Lexical Decision Reaction Times

Variable	β	*SE*	*t*	*p*	R^2	ΔR^2
Step 1						
Number of phonemes	.00948	.00119	.795	.43		
Phonological degree	.000773	.00836	.092	.93		
Phonological *C*	.00787	.00537	1.47	.14		
Phonological neighborhood frequency	.00906	.00611	1.48	.14		
Positional segment probability	−.0100	.00848	−1.185	.24		
Biphone probability	.00173	.00820	.210	.83		
Number of letters	−.0206	.00961	−2.15	.032*		
Orthographic degree	−.0242	.00772	−3.14	.002**		
Orthographic *C*	−.0119	.00523	−2.27	.023*		
Orthographic neighborhood frequency	−.00159	.00606	−.261	.79		
Average bigram counts	.0113	.00579	1.96	.050†		
Sum of bigram counts by position	.00240	.00748	.321	.75		
Subjective familiarity	−.133	.00521	−25.45	<.001***		
Word frequency	−.121	.00561	−21.49	<.001***		
					.505***	
Step 2						
Number of phonemes	.0149	.0120	1.24	.22		
Phonological degree	.0113	.00903	1.25	.21		
Phonological *C*	.00733	.00548	1.34	.18		
Phonological neighborhood frequency	.00756	.00613	1.23	.22		
Positional segment probability	−.00979	.00848	−1.15	.25		
Biphone probability	.00302	.00820	.369	.71		
Number of letters	−.0190	.00961	−1.98	.048*		
Orthographic degree	.00307	.0120	.255	.80		
Orthographic *C*	−.00999	.00863	−1.16	.25		
Orthographic neighborhood frequency	−.00130	.00605	−.215	.83		
Average bigram counts	.0125	.00579	2.17	.030*		
Sum of bigram counts by position	−.00206	.00761	−.271	.79		
Subjective familiarity	−.132	.00520	−25.35	<.001***		
Word frequency	−.122	.00563	−21.72	<.001***		
Phonographic degree	−.0348	.0116	−3.01	.003**		
Phonographic *C*	.000994	.00870	.114	.91		
					.508***	.003**

Note. $N = 2{,}120$. "Phonographic degree" and "Phonographic C" were * variables added in Step 2, as in Table 2 and Table 4.
† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

words with lower phonographic degree, after taking into account the variance contributed by several lexical variables known to influence language processing. These analyses replicated previous work showing that the presence of phonographic neighbors facilitates naming of visually presented words (Adelman & Brown, 2007; Peereman & Content, 1997) and extends that previous research to demonstrate similar effects in the visual lexical-decision task.

Overall, the results from the ELP analyses and psycholinguistic tasks generally show that greater phonological and orthographic similarity facilitates word recognition in both visual and auditory modalities. Furthermore, it appears that phonographic degree influences visual word recognition but not spoken word recognition, whereas phonographic *C* influences spoken word recognition but not visual word recognition. It may simply be the case that one network measure is capturing more of the variance in one modality than another—perhaps reflecting differences in the way phonographic similarity is processed in different modalities. In visual word recognition, a measure that simply represents the number of phonographic neighbors such as phonographic degree may be the better predictor, whereas a subtler metric such as phonographic *C* that captures the interconnectivity among those phonographic neighbors may be the better predictor in spoken word recognition. Unlike visually presented words, auditory signals unfold over time, which may allow for more time for activation to spread, not just from the target word to its neighbors, but also among its neighbors such that the internal structure of the phonographic neighborhood plays a role in lexical retrieval. Within the visual modality, however, the size of the phonographic neighborhood may take precedence over its internal structure because the initial activation of a target word's phonographic neighbors may already be sufficient to "nudge" the visual word recognition system over the threshold for recognition.

General Discussion

To recapitulate, the main findings were that phonographic degree significantly influenced visual word recognition and not spoken word recognition, whereas phonographic *C* significantly influenced spoken word recognition and not visual word recognition. Specifically, the presence of more phonographic neighbors (i.e., degree) facilitated word recognition in the visual modality, and greater interconnectivity within the phonographic neighborhood (i.e., *C*) facilitated word recognition in the auditory modality.

Table 10
Regression Results for English Lexicon Project Visual Lexical Decision Accuracies

Variable	β	*SE*	*t*	*p*	R^2	ΔR^2
Step 1						
Number of phonemes	.000508	.00597	.085	.93		
Phonological degree	−.00514	.00418	−1.23	.22		
Phonological *C*	−.0000376	.00268	−.014	.99		
Phonological neighborhood frequency	−.00806	.00306	−2.64	.008**		
Positional segment probability	−.000768	.00424	−.181	.86		
Biphone probability	.00329	.00410	.802	.42		
Number of letters	.0243	.00481	5.06	<.001***		
Orthographic degree	.0156	.00386	4.05	<.001***		
Orthographic *C*	.00295	.00261	1.13	.26		
Orthographic neighborhood frequency	.00145	.00303	.478	.63		
Average bigram counts	.000382	.00289	.132	.90		
Sum of bigram counts by position	−.00751	.00374	−2.01	.045*		
Subjective familiarity	.130	.00260	49.93	<.001***		
Word frequency	.0294	.00281	10.46	<.001***		
					.650***	
Step 2						
Number of phonemes	−.00218	.00602	−.363	.72		
Phonological degree	−.0104	.00452	−2.30	.022*		
Phonological *C*	−.0000581	.00274	−.021	.98		
Phonological neighborhood frequency	−.00739	.00306	−2.41	.016*		
Positional segment probability	−.00102	.00424	−.241	.81		
Biphone probability	.00270	.00410	.658	.51		
Number of letters	.0235	.00481	4.89	<.001***		
Orthographic degree	.00165	.00600	.274	.78		
Orthographic *C*	.000269	.00432	.062	.95		
Orthographic neighborhood frequency	.00130	.00303	.430	.67		
Average bigram counts	−.000200	.00290	−.069	.95		
Sum of bigram counts by position	−.00528	.00381	−1.39	.17		
Subjective familiarity	.130	.00260	49.84	<.001***		
Word frequency	.0303	.00282	10.74	<.001***		
Phonographic degree	.0176	.00578	3.05	.002**		
Phonographic *C*	.00172	.00435	.396	.69		
					.652***	.002*

Note. $N = 2{,}120$. "Phonographic degree" and "Phonographic C" were * variables added in Step 2, as in Table 2 and Table 4.
* $p < .05$. ** $p < .01$. *** $p < .001$.

The finding of a significant effect of phonographic degree on visual word recognition is consistent with previous literature (Adelman & Brown, 2007), although it should be noted that in their analyses Adelman and Brown used a more limited definition of phonographic neighbors by only considering words that differed by the substitution of one phoneme and letter. In the present study, the phonographic neighbors included words that differed from the target word by the substitution, addition, or deletion of either one phoneme or one letter. On the other hand, a significant influence of phonographic *C* on spoken word recognition represents a novel finding, as the influence of phonographic neighbors has never been previously examined in the spoken modality.

These results indicate that the phonographic relationships among words play an important role in both spoken and visual word recognition. Recall that the phonographic network represented the section of the phonographic multiplex where phonological and orthographic links overlapped. Therefore, phonographic degree and phonographic *C* represent the extent to which the similarity structure in both layers of the individual layers "mirror" each other, such that they reinforce the similarity structure in both layers of the phonographic multiplex.

The key takeaway from these experiments and analyses is that the presence of phonographic links, which represent both phonological and orthographic similarity relationships among words, as well as the structure of these links, facilitates spoken and visual word recognition, even after taking into account the influence of (a) conventional measures of orthographic and phonological similarity (i.e., phonological and orthographic degree or neighborhood density), and (b) "single-layer" network measures of orthographic and phonological similarity (i.e., phonological and orthographic clustering coefficient). The results demonstrate how simultaneously representing the phonological and orthographic similarity of words within a phonographic multiplex can lead to a more nuanced understanding of how similarity influences spoken and visual word recognition.

Similarity Effects in Spoken and Visual Word Recognition

An intriguing aspect of the present findings is that phonographic degree facilitated visual word recognition but not spoken word recognition, and phonographic *C* facilitated spoken word recogni-

tion but not visual word recognition. This divergence may reflect differences in the way that written and spoken words are processed. A long-standing question within psycholinguistics is whether similarity among phonological and orthographic word forms facilitates or hinders word recognition. Across various measures of similarity (e.g., degree/neighborhood density, Levenshtein distance, clustering coefficient), the results from the literature indicate that greater similarity among orthographic representations facilitates visual word recognition (Andrews, 1997; Siew, 2018; Yarkoni, Balota, & Yap, 2008) and greater similarity among phonological representations inhibits spoken word recognition (Chan & Vitevitch, 2009; Goh, Suárez, Yap, & Tan, 2009; Luce & Pisoni, 1998; Suárez, Tan, Yap, & Goh, 2011). On the other hand, the pattern of results is less clear in studies that investigated the influence of phonological similarity on visual word recognition (Grainger et al., 2005; Yates et al., 2004; Yates, 2013) and the influence of orthographic similarity on spoken word recognition (Muneaux & Ziegler, 2004; Ziegler et al., 2003).

The contradictory effects of neighborhood density in visual and spoken word recognition have been examined in a computational study conducted by Chen and Mirman (2012). In this paper, Chen and Mirman simulated the dynamics of interactive activation and competition in a simple model and showed that having more (orthographic) neighbors facilitated visual word recognition due to weakly activated neighbors leading to an overall facilitatory effect, whereas having more (phonological) neighbors inhibited spoken word recognition due to strongly activated neighbors leading to an overall inhibitory effect. Their model could provide one possible mechanistic account of the finding that phonographic neighbors facilitated visual word recognition. Specifically, words with more phonographic neighbors could lead to the activation of several weakly active neighbors that could result in a net facilitative effect, resulting in facilitatory effect of phonographic degree in visual word recognition.

On the other hand, Chen and Mirman (2012) suggested that the inhibitory effects of *C* on spoken word recognition were due to "particular patterns of neighbor clustering lead(ing) the neighbors to enhance one another's activation" (p. 12), resulting in net inhibitory effects of *C* as reported by Chan and Vitevitch (2009). Based on this explanation, it is unclear how this model could be extended to account for the facilitatory effects of phonographic *C* on spoken word recognition.

Before proceeding, it is important to note that the present article differs from Chen and Mirman's simulations as well as previous studies in a fundamental way. Adopting a multiplex approach from network science to simultaneously represent a language's phonological and orthographic similarity structure represents a novel theoretical conceptualization that departs from conventional ways of investigating phonological and orthographic similarity effects on word recognition. Such an approach can provide network metrics (a few of which are investigated in this work) that simultaneously represented the phonological and orthographic similarity structure of words in the language. However, in previous work, phonology and orthography were treated as separate influences to be manipulated or controlled for, making it difficult to assess the seemingly contradictory effects of similarity on word recognition. For instance, Chen and Mirman (2012) examined the effects of phonological neighborhood density on spoken word recognition and orthographic neighborhood density on visual word recognition, but not vice versa (i.e., orthographic neighborhood effects on spoken word recognition and phonological neighborhood effects on visual word recognition) as their model did not explicitly consider how the overlap or interaction between phonological and orthographic similarity structures affected word recognition.

Consider the finding that phonological degree has an inhibitory effect in spoken word recognition (e.g., Goh et al., 2009) but a facilitatory effect in visual word recognition (e.g., Grainger et al., 2005), whereas phonological clustering coefficient has an inhibitory effect in both spoken (Chan & Vitevitch, 2010) and visual word recognition (Yates, 2013). After closely controlling for the effect of orthographic degree, Grainger et al. found that phonological degree facilitated visual word processing, such that the processing of words with many phonological neighbors is facilitated as compared to words with few phonological neighbors across various tasks. Grainger and colleagues argue that greater consistency between phonology and orthography contributed to this facilitative effect. However, in Yates (2013), it is not clear if orthographic similarity (i.e., orthographic clustering coefficient) among the word stimuli was explicitly controlled for. This raises questions about the inhibitory effect of phonological clustering coefficient in visual word recognition reported by Yates (2013). Overall, conceptualizing phonology and orthography as separate effects does not necessarily appear to be the most productive way of addressing the question of whether similarity among phonological and orthographic word forms facilitates or hinders word recognition.

In contrast, the network metrics investigated in this article, degree and clustering coefficient, represent microlevel network measures of both phonological and orthographic similarity, with each metric representing different structural aspects of the phonographic neighborhood of a particular target word. Recall that degree simply refers to the number of phonographic neighbors, whereas clustering coefficient refers to the extent to which phonographic neighbors are also phonographic neighbors of each other.

One possible explanation for the observed finding that phonographic degree facilitated visual word recognition (but not spoken word recognition) and phonographic *C* facilitated spoken word recognition (but not visual word recognition) is that similarity effects depend on, and reflect, differences in the nature of "bottom-up" auditory and visual information. Visual information is instantaneous and immediately available, whereas acoustic information unfolds over time and is more ambiguous. Because of the nature of auditory information, there is more time for activation to spread among a target word's neighbors, allowing for subtle similarity effects as measured by *C* to "emerge", and diminishing the influence of partially activated neighbors (as captured by degree) in spoken word recognition.

Consider the following study by Seidenberg, Waters, Barnes, and Tanenhaus (1984), who found that words with irregular pronunciations were named more slowly than words with regular pronunciations, but this was only true for low frequency words and not high frequency words (Andrews, 1982). High frequency irregular words (e.g., "have") were named just as quickly as frequency-matched regular words. According to Seidenberg (1985), in visual word recognition, orthographic and phonological information are activated at different latencies within a single interactive process, with phonology lagging behind orthography. As it takes a longer

time to recognize low frequency words, it allows more time for phonological information (i.e., irregular pronunciations) to be activated and hence influence naming latencies.

The general argument from the above study is that when processing is difficult or effortful in some way (such as low frequency words), it permits more time for additional influences (such as phonology) to come into play. A variant of this argument could be applied to explain the present set of findings: Processing spoken words, which are more ambiguous than written words due to the nature of auditory input, may permit more time for more subtle similarity effects such as *C* to develop and subsequently affect recognition. In other words, ambiguity in the bottom up signal may lead to greater sensitivity to nuances in the similarity space. Although these explanations are somewhat speculative in nature, they could be tested and investigated in future experimental work and computer simulations (e.g., Chen & Mirman, 2012; Vitevitch et al., 2011).

Theoretical Implications for Models of Word Recognition

Several well-established theories have been put forth to account for multiple aspects of visual and spoken word recognition. The leading models of visual word recognition can be broadly classified into two groups: dual route models, which posit the presence of two distinct, independent pathways in visual word recognition (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler's, 2001 dual route cascaded model) and parallel, distributed models, which consist of orthographic units, phonological units, and a set of hidden units that interface between the orthographic and phonological units (e.g., Seidenberg & McClelland's, 1989 parallel distributed processing [PDP] model). The cohort model (Marslen-Wilson, 1987), TRACE (McClelland & Elman, 1986), Shortlist B (Norris & McQueen, 2008), and neighborhood activation model (Luce & Pisoni, 1998) represent the prominent models of spoken word recognition.

For the purposes of the present discussion, we will focus on the Seidenberg and McClelland (1989) PDP model of visual word recognition and pronunciation as an example, and consider how it may or may not be able to account for the phonographic effects observed in the present studies.

The PDP model consists of orthographic units, phonological units, and a set of hidden units that interface between the orthographic and phonological units. One key feature of distributed models is the ability of the model to learn—and thereby approximate the language acquisition process in children—by modifying connection weights between units via a back-propagation learning algorithm during training (Seidenberg & McClelland, 1989). In a connectionist model, the relative influence of orthography and phonology on lexical retrieval depends on the extent to which orthographic and phonological codes overlap (Harm & Seidenberg, 2004). A greater amount of overlap in orthography and phonology would be expected to speed up processing and lead to faster access to a word's meaning (Harm & Seidenberg, 2004); this is consistent with the present finding that phonographic similarity generally facilitates processing in both visual and spoken word recognition. However, without explicitly considering the linguistic structure of language, it is not entirely clear how more subtle effects of the phonographic similarity structure (i.e., the level of interconnectivity among similar words, as exemplified by the phonographic *C* metric) would be implemented in the model.

It is important to emphasize that although the architecture of the PDP model may seem to resemble a network of sorts (i.e., units connected to each other; see Figure 3), it differs considerably from the language network generated via the network science approach. In the PDP model, all units are connected to all other units, with connection weights that update after training. The model is distributed, such that phonological or orthographic codes are represented by a pattern of activation distributed over primitive orthographic, phonological, and hidden units (Seidenberg & McClelland, 1989). In contrast, the network science approach explicitly models the overall similarity structure of language. Nodes represent lexical forms and unweighted links are placed between similar word forms as defined by a straightforward operationalization of similarity (substitution, addition, deletion of one phoneme or letter; Landauer & Streeter, 1973; see Figure 4).

The network science approach, despite being based on simple assumptions, reveals a complex language network structure whereby a simple diffusion of activation mechanism can be implemented to account for behavioral findings such as the clustering coefficient effect (Vitevitch et al., 2011). Simulations conducted by Chan and Vitevitch (2009) using jTRACE, the computational implementation of the TRACE model of speech perception (Strauss, Harris, & Magnuson, 2007), were unable to account for the clustering coefficient effect. In contrast, a model that implemented a simple diffusion of activation process on a network structure (Vitevitch et al., 2011) was able to simulate the inhibitory *C* effect in spoken word recognition. Furthermore, even though sublexical units (such as letters or phonemes) are not explicitly represented as individual nodes or entities within the language network, one could potentially account for both lexical and sublexical effects by examining the language network at differing levels of the network (micro-, meso-, macro-). For instance, Siew (2013) speculated that (sublexical) phonotactic probability effects could emerge at the mesolevel (community) structure of the network and (lexical) neighborhood density effects could arise at the microlevel of the network.

As the above discussion demonstrates, the network science approach differs from contemporary approaches in psycholinguis-

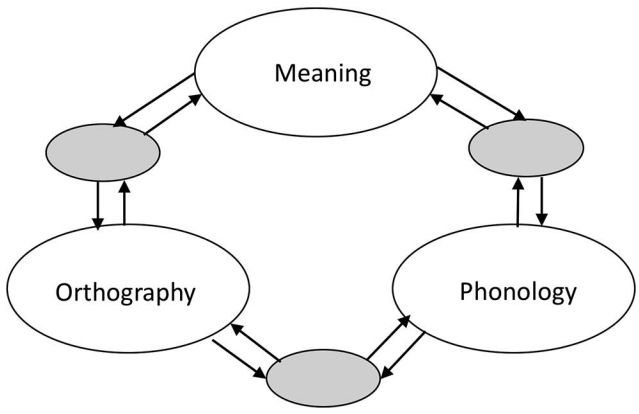

Figure 3. The general framework of the Parallel Distributed Processing model (Seidenberg & McClelland, 1989). Each ellipse represents a set of primitive units, with grey ellipses representing hidden units.

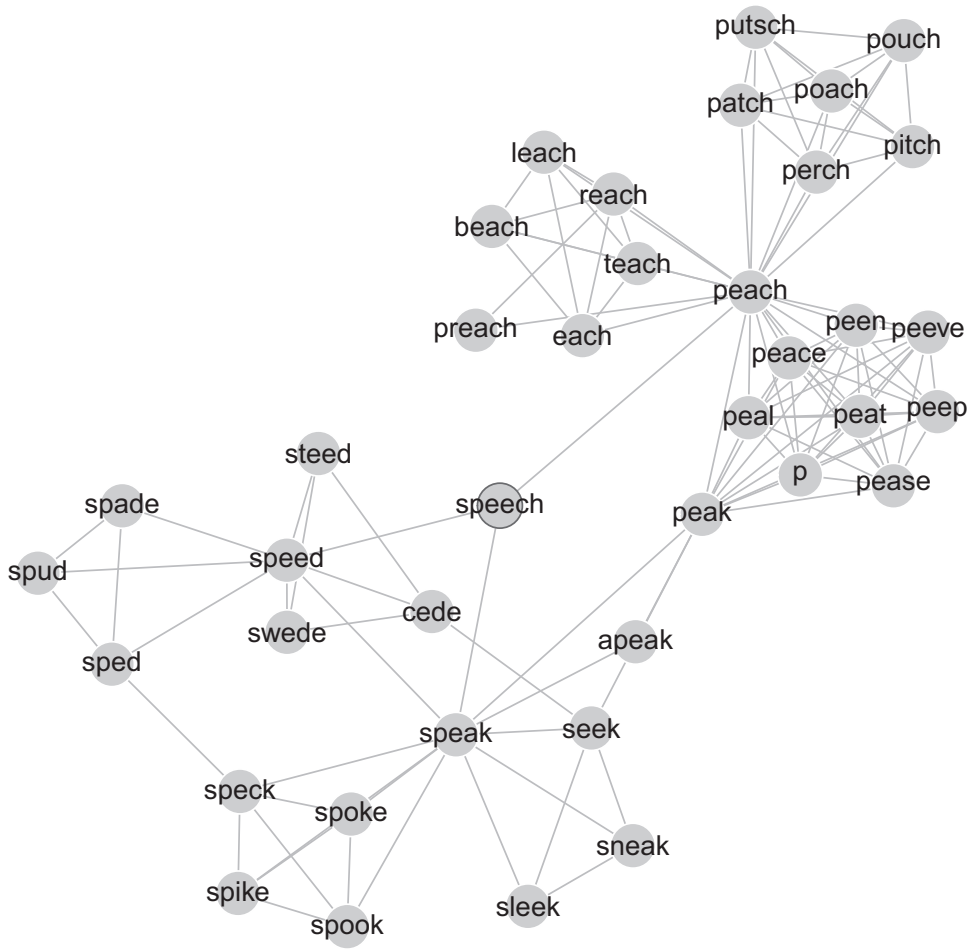

Figure 4. A section of the phonological network of language showing the word "speech", its phonological neighbors, and the phonological neighbors of its phonological neighbors. Links are placed between words that differ by the substitution, addition, or deletion of one phoneme.

tics. The former approach emphasizes the role of the underlying language structure on lexical retrieval whereas the latter focuses on investigating the cognitive processes that underlie lexical retrieval. Similarly, the measures investigated in this paper, phonographic degree and phonographic *C*, are derived directly from the phonographic language network and represent the structural characteristics of words in that network. In addition, the network science framework provides language researchers with the tools to measure and quantify various structural aspects of the phonographic network that was introduced in this article. For instance, one can examine the global structural properties of words in the phonographic network by using the closeness centrality network measure (see Goldstein & Vitevitch, 2017, who examined the influence of phonological closeness centrality on spoken word recognition), or examine the mesolevel structure (i.e., communities or clusters) of words in the phonographic network (see Siew, 2013, for an example of how community detection techniques were used to analyze the phonological language network). Adopting a network science framework would not only lead to the development of new network metrics that can help move the field toward a more integrated understanding of phonological effects on visual word recognition and orthographic effects on spoken word recognition, but also provide additional empirical findings that can further constrain and test current models of lexical retrieval in both the spoken and written modalities.

Without explicitly considering how the overall phonological and orthographic similarity structure of language affects lexical retrieval, it is unclear how current models of word recognition would be able to account for the present findings. In addition, models of spoken word recognition (e.g., Cohort, TRACE, Shortlist, neighborhood activation models) do not consider the role of orthographic information on speech processing and would not predict any orthographic effects in spoken word recognition in the first place. Unlike the above models of lexical retrieval, which tend to focus on the process of lexical retrieval, the network science approach emphasizes the importance of explicitly considering the structure of the system in which lexical processes operate on and provides a mathematical framework for quantifying and measuring the structure of the lexicon. Indeed, a simple process like a random walk implemented on a structured semantic network has been able to produce retrieval patterns from memory that resemble the retrieval patterns produced by a more complex and strategic algo-

rithm for optimal foraging (Abbott, Austerweil, & Griffiths, 2015), as well as the well-known meaning-frequency law (Ferrer-i-Cancho & Vitevitch, 2018) and word frequency law discovered by Zipf (Allegrini, Grigolini, & Palatella, 2004). Through the present article, we hope to shift, or at least tilt, the overwhelming focus of these models on processing to one that simultaneously considers how process and structure work in tandem to produce behavioral phenomena.

Finally, it is important to acknowledge that there is still much to done. In the present article, the analyses were conducted on a subset of shorter, more frequent words in the English language and it is crucial for future empirical work to investigate how the structure of the phonographic multiplex at various levels (macro-, meso-, micro-) influence processing and begin examining other areas of the multiplex beyond the phonographic network that would include longer, less frequent, multisyllabic words that are not found in the giant component of the phonographic network. Future computational simulations will also need to be conducted to formalize and test models to provide computational support for the present findings of an effect of phonographic degree on visual word recognition and an effect of phonographic clustering coefficient on spoken word recognition. Nevertheless, the present findings suggest that models of word recognition will need to explicitly consider how the cognitive processes that underlie lexical retrieval operate within a complex language structure as represented by the phonographic multiplex.

References

Abbott, J. T., Austerweil, J. L., & Griffiths, T. L. (2015). Random walks on semantic networks can resemble optimal foraging. *Psychological Review, 122,* 558–569. http://dx.doi.org/10.1037/a0038693

Adelman, J. S., & Brown, G. D. (2007). Phonographic neighbors, not orthographic neighbors, determine word naming latencies. *Psychonomic Bulletin & Review, 14,* 455–459. http://dx.doi.org/10.3758/BF03194088

Albert, R., & Barabási, A. L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics, 74,* 47.

Allegrini, P., Grigolini, P., & Palatella, L. (2004). Intermittency and scale-free networks: A dynamical model for human language complexity. *Chaos, Solitons, and Fractals, 20,* 95–105. http://dx.doi.org/10.1016/S0960-0779(03)00432-6

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory & Cognition, 10,* 565–575. http://dx.doi.org/10.3758/BF03202439

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review, 4,* 439–461.

Arbesman, S., Strogatz, S. H., & Vitevitch, M. S. (2010). The structure of phonological networks across multiple languages. *International Journal of Bifurcation and Chaos, 20,* 679–685.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., . . . Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39,* 445–459. http://dx.doi.org/10.3758/BF03193014

Balota, D. A., Yap, M. J., Hutchison, K. A., & Cortese, M. J. (2012). 5 megastudies. In J. S. Adelman (Ed.), *Visual word recognition: Vol. 1. Models and methods, orthography and phonology* (pp. 90–115). New York, NY: Psychology Press.

Barabási, A.-L. (2009). Scale-free networks: A decade and beyond. *Science, 325,* 412–413. http://dx.doi.org/10.1126/science.1173299

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4 (R package version 1.0–6) [Computer Software]. Retrieved from https://cran.r-project.org/web/packages/lme4/index.html

Battiston, F., Nicosia, V., & Latora, V. (2014). Structural measures for multiplex networks. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics, 89,* 032804. http://dx.doi.org/10.1103/PhysRevE.89.032804

Borgatti, S. P. (2005). Centrality and network flow. *Social Networks, 27,* 55–71. http://dx.doi.org/10.1016/j.socnet.2004.11.008

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, Instruments & Computers, 41,* 977–990. http://dx.doi.org/10.3758/BRM.41.4.977

Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience, 10,* 186.

Cardillo, A., Zanin, M., Gómez-Gardeñes, J., Romance, M., García del Amo, A. J., & Boccaletti, S. (2013). Modeling the multi-layer nature of the European Air Transport Network: Resilience and passengers re-scheduling under random failures. *The European Physical Journal Special Topics, 215,* 23–33. http://dx.doi.org/10.1140/epjst/e2013-01712-8

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

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

Chen, Q., & Mirman, D. (2012). Competition and cooperation among similar representations: Toward a unified account of facilitative and inhibitory effects of lexical neighbors. *Psychological Review, 119,* 417–430. http://dx.doi.org/10.1037/a0027175

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37–46.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535–555). Hillsdale, NJ: Erlbaum.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256. http://dx.doi.org/10.1037/0033-295X.108.1.204

Cortese, M. J., & Simpson, G. B. (2000). Regularity effects in word naming: What are they? *Memory & Cognition, 28,* 1269–1276. http://dx.doi.org/10.3758/BF03211827

Dich, N. (2011). Individual differences in the size of orthographic effects in spoken word recognition: The role of listeners' orthographic skills. *Applied Psycholinguistics, 32,* 169–186. http://dx.doi.org/10.1017/S0142716410000330

Eager, C., & Roy, J. (2017). *Mixed effects models are sometimes terrible*. Retrieved from https://arxiv.org/abs/1701.04858

Erdös, P., & Rényi, A. (1960). On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences, 5,* 17–60.

Ferrand, L., & Grainger, J. (2003). Homophone interference effects in visual word recognition. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 56,* 403–419. http://dx.doi.org/10.1080/02724980244000422

Ferrer-i-Cancho, R., & Sole, R. V. (2001). The small world of human language. *Proceedings of the Royal Society of London B, Biological Sciences, 268,* 2261–2265.

Ferrer-i-Cancho, R., & Vitevitch, M. S. (2018). The origins of Zipf's meaning-frequency law. *Journal of the American Society for Information Science and Technology, 69,* 1369–1379.

Fortunato, S. (2010). Community detection in graphs. *Physics Reports, 486*(3–5), 75–174.

Frauenfelder, U. H., Baayen, R. H., & Hellwig, F. M. (1993). Neighborhood density and frequency across languages and modalities. *Journal of Memory and Language, 32,* 781–804. http://dx.doi.org/10.1006/jmla.1993.1039

Goh, W. D., Suárez, L., Yap, M. J., & Tan, S. H. (2009). Distributional analyses in auditory lexical decision: Neighborhood density and word-frequency effects. *Psychonomic Bulletin & Review, 16,* 882–887.

Goldstein, R., & Vitevitch, M. S. (2017). The influence of closeness centrality on lexical processing. *Frontiers in Psychology, 8,* 1683.

Grainger, J., & Ferrand, L. (1994). Phonology and orthography in visual word recognition: Effects of masked homophone primes. *Journal of Memory and Language, 33,* 218–233. http://dx.doi.org/10.1006/jmla.1994.1011

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

Greenberg, J. H., & Jenkins, J. J. (1964). Studies in the psychological correlates of the sound system of American English. *Word, 20,* 157–177. http://dx.doi.org/10.1080/00437956.1964.11659816

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111,* 662–720. http://dx.doi.org/10.1037/0033-295X.111.3.662

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009). Longitudinal analysis of early semantic networks: Preferential attachment or preferential acquisition? *Psychological Science, 20,* 729–739. http://dx.doi.org/10.1111/j.1467-9280.2009.02365.x

Iyengar, S. R., Veni Madhavan, C. E., Zweig, K. A., & Natarajan, A. (2012). Understanding human navigation using network analysis. *Topics in Cognitive Science, 4,* 121–134. http://dx.doi.org/10.1111/j.1756-8765.2011.01178.x

Janssen, D. P. (2012). Twice random, once mixed: Applying mixed models to simultaneously analyze random effects of language and participants. *Behavior Research Methods, 44,* 232–247.

Jared, D. (1997). Spelling–sound consistency affects the naming of high-frequency words. *Journal of Memory and Language, 36,* 505–529. http://dx.doi.org/10.1006/jmla.1997.2496

Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language, 29,* 687–715. http://dx.doi.org/10.1016/0749-596X(90)90044-Z

Kello, C. T., & Beltz, B. C. (2009). Scale-free networks in phonological and orthographic wordform lexicons. In I. Chitoran, C. Coupe, E. Marsico, & F. Pellegrino (Eds.), *Approaches to phonological complexity* (pp. 171–190). Berlin, Germany: Mouton de Gruyter.

Kessler, B., & Treiman, R. (1997). Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language, 37,* 295–311. http://dx.doi.org/10.1006/jmla.1997.2522

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*. Advance online publication. http://dx.doi.org/10.18637/jss.v082.i13

Landauer, T. K., & Streeter, L. A. (1973). Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. *Journal of Verbal Learning & Verbal Behavior, 12,* 119–131. http://dx.doi.org/10.1016/S0022-5371(73)80001-5

Lewis, K., Kaufman, J., Gonzalez, M., Wimmer, A., & Christakis, N. (2008). Tastes, ties, and time: A new social network dataset using Facebook.com. *Social Networks, 30,* 330–342. http://dx.doi.org/10.1016/j.socnet.2008.07.002

Liberman, A. M. (1992). The relation of speech to reading and writing. *Advances in Psychology, 94,* 167–178. http://dx.doi.org/10.1016/S0166-4115(08)62794-6

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

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25*(1–2), 71–102. http://dx.doi.org/10.1016/0010-0277(87)90005-9

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18,* 1–86. http://dx.doi.org/10.1016/0010-0285(86)90015-0

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

New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review, 13,* 45–52. http://dx.doi.org/10.3758/BF03193811

Newman, M. E. J. (2002). Assortative mixing in networks. *Physical Review Letters, 89,* 208701. http://dx.doi.org/10.1103/PhysRevLett.89.208701

Newman, M. E. (2004a). Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences, 101*(Suppl. 1), 5200–5205.

Newman, M. E. (2004b). Fast algorithm for detecting community structure in networks. *Physical Review E, 69,* 066133.

Newman, M. E. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences, 103,* 8577–8582.

Newman, M. E., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E, 69,* 026113.

Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review, 115,* 357–395. http://dx.doi.org/10.1037/0033-295X.115.2.357

Nozari, N., Kittredge, A. K., Dell, G. S., & Schwartz, M. F. (2010). Naming and repetition in aphasia: Steps, routes, and frequency effects. *Journal of Memory and Language, 63,* 541–559.

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words. *Research on Speech Perception Progress Report, 10,* 357–376.

Opsahl, T., & Panzarasa, P. (2009). Clustering in weighted networks. *Social Networks, 31,* 155–163. http://dx.doi.org/10.1016/j.socnet.2009.02.002

Peereman, R., & Content, A. (1997). Orthographic and phonological neighborhoods in naming: Not all neighbors are equally influential in orthographic space. *Journal of Memory and Language, 37,* 382–410. http://dx.doi.org/10.1006/jmla.1997.2516

Roux, S., & Bonin, P. (2013). "With a little help from my friends": Orthographic influences in spoken word recognition. *L'Année Psychologique, 113,* 35–48. http://dx.doi.org/10.4074/S0003503313001024

Seidenberg, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition, 19,* 1–30. http://dx.doi.org/10.1016/0010-0277(85)90029-0

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523–568. http://dx.doi.org/10.1037/0033-295X.96.4.523

Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 546–554. http://dx.doi.org/10.1037/0278-7393.5.6.546

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning & Verbal Behavior, 23,* 383–404. http://dx.doi.org/10.1016/S0022-5371(84)90270-6

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

Siew, C. S. Q. (2017). The influence of 2-hop network density on spoken word recognition. *Psychonomic Bulletin & Review, 24,* 496–502. http://dx.doi.org/10.3758/s13423-016-1103-9

Siew, C. S. Q. (2018). The orthographic similarity structure of English words: Insights from network science. *Applied Network Science, 3,* 13. http://dx.doi.org/10.1007/s41109-018-0068-1

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

Soffer, S. N., & Vázquez, A. (2005). Network clustering coefficient without degree-correlation biases. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics, 71*(5, Part 2), 057101. http://dx.doi.org/10.1103/PhysRevE.71.057101

Solé, R. V., Corominas-Murtra, B., Valverde, S., & Steels, L. (2010). Language networks: Their structure, function, and evolution. *Complexity, 15,* 20–26. http://dx.doi.org/10.1002/cplx.20305

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

Stone, G. O., Vanhoy, M., & van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language, 36,* 337–359. http://dx.doi.org/10.1006/jmla.1996.2487

Strauss, T. J., Harris, H. D., & Magnuson, J. S. (2007). jTRACE: A reimplementation and extension of the TRACE model of speech perception and spoken word recognition. *Behavior Research Methods, 39,* 19–30. http://dx.doi.org/10.3758/BF03192840

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

Suárez, L., Tan, S. H., Yap, M. J., & Goh, W. D. (2011). Observing neighborhood effects without neighbors. *Psychonomic Bulletin & Review, 18,* 605–611.

Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory and Language, 26,* 608–631.

van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition, 15,* 181–198. http://dx.doi.org/10.3758/BF03197716

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

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

Vitevitch, M. S., Chan, K. Y., & Roodenrys, S. (2012). Complex network structure influences processing in long-term and short-term memory. *Journal of Memory and Language, 67,* 30–44.

Vitevitch, M. S., Ercal, G., & Adagarla, B. (2011). Simulating retrieval from a highly clustered network: Implications for spoken word recognition. *Frontiers in Psychology, 2,* 369. http://dx.doi.org/10.3389/fpsyg.2011.00369

Vitevitch, M. S., & Goldstein, R. (2014). Keywords in the mental lexicon. *Journal of Memory and Language, 73,* 131–147. http://dx.doi.org/10.1016/j.jml.2014.03.005

Vitevitch, M. S., & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments & Computers, 36,* 481–487. http://dx.doi.org/10.3758/BF03195594

Watts, D. J. (2004). The "new" science of networks. *Annual Review of Sociology, 30,* 243–270. http://dx.doi.org/10.1146/annurev.soc.30.020404.104342

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393,* 440–442.

Watt, T. S. (1954, June 21). Brush up your English. *Manchester Guardian*. Retrieved from https://linguistlist.org/issues/13/13-3353.html

Wu, K., Taki, Y., Sato, K., Sassa, Y., Inoue, K., Goto, R., . . . Fukuda, H. (2011). The overlapping community structure of structural brain network in young healthy individuals. *PLoS ONE, 6,* e19608.

Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language, 60,* 502–529. http://dx.doi.org/10.1016/j.jml.2009.02.001

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review, 15,* 971–979. http://dx.doi.org/10.3758/PBR.15.5.971

Yates, M. (2005). Phonological neighbors speed visual word processing: Evidence from multiple tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1385–1397. http://dx.doi.org/10.1037/0278-7393.31.6.1385

Yates, M. (2013). How the clustering of phonological neighbors affects visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 1649–1656. http://dx.doi.org/10.1037/a0032422

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

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review, 5,* 683–689. http://dx.doi.org/10.3758/BF03208845

Ziegler, J. C., Ferrand, L., & Montant, M. (2004). Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. *Memory & Cognition, 32,* 732–741. http://dx.doi.org/10.3758/BF03195863

Ziegler, J. C., Montant, M., & Jacobs, A. M. (1997). The feedback consistency effect in lexical decision and naming. *Journal of Memory and Language, 37,* 533–554. http://dx.doi.org/10.1006/jmla.1997.2525

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

Appendix A

List of Word Stimuli Used in Experiments 1 and 2

Phonographic degree words		Phonographic *C* words	
High degree	Low degree	High *C*	Low *C*
bloat	balm	balk	bleed
brace	blame	bland	blob
brew	blow	bleak	bread
chart	brood	blown	brink
clip	brook	boss	broom
deck	charm	bride	cave
draft	chase	clot	clod
drag	clean	count	drive
drew	cleat	crime	drove
drip	clove	dream	duct
flake	clump	drown	flat
flew	cue	duke	hack
flick	doom	dwell	hawk
float	dorm	flip	hive
flush	dread	flop	hoop
gripe	drum	haze	lab
gum	fled	hook	lobe
gust	food	husk	mile
hulk	grab	loaf	moat
hurl	halt	lunch	mount
loud	hurt	lung	nose
moist	limb	mean	pluck
mood	plea	paid	plum
pleat	porch	posh	plume
scoop	range	pulp	raft
shame	roof	pump	ramp
shock	scab	rack	ripe
slid	scan	reef	rose
slum	scorn	rope	run
slump	sheep	save	side
spool	short	skin	slap
spunk	shout	slam	slur
stall	smack	snip	snap
steep	spear	spice	space
stub	stark	stock	swell
swim	start	tab	tomb
swing	swap	tile	tool
swoop	sweep	tote	trail
trust	wand	trace	tread
weep	yarn	truce	wheat

(*Appendices continue*)

Appendix B

List of Nonword Stimuli Used in Experiment 2

Nonword foils for phonographic degree words		Nonword foils for phonographic *C* words	
bles	bɪm	blɪk	blaɪd
blet	blɛ	blɝn	bɹɑb
bɹɛ	blʊk	blund	bɹɔɪm
tʃeɹt	bɹem	blaɪd	bɹʌd
dɔk	bɹɪd	bɔɪk	bɹʌŋk
dɹɛft	tʃɪs	bos	dɑkt
dɹɪg	tʃoɹm	dek	dɹɑv
dɹop	dɔɪm	dɹom	dɹɛv
dɹaʊ	dɹem	dɹaɪn	flet
floʃ	dɹʌd	dwɑl	hik
flʊk	dʊɹm	flup	hok
flʌt	flud	fɹɑp	hov
flɪk	fʊd	hɑsk	hʊp
fɹu	glæb	hɑz	hwɑt
gæm	hɛt	hek	kɛv
gɑst	hilt	klʊt	kɹɑd
gɹʌp	klɔɪmp	kɔɪnt	lɛb
hɔɪl	klut	kɹaʊm	lʊb
holk	klʊv	lɑŋ	mont
klup	kɹin	lef	mɔɪt
lʌd	kjɑ	lʊntʃ	maʊl
mɜ˞d	lom	mɔn	nɪz
mʊst	piɹtʃ	pɑmp	plɛm
plut	plʊ	peʃ	pɹum
ʃem	ɹæf	pilp	pɹʌk
skɔɪp	ɹinʤ	pɔɪd	ɹemp
slimp	ʃɑɹt	ɹof	ɹɪn
slom	skɪb	ɹuk	ɹɪz
slaʊd	skʊn	ɹaʊp	ɹoft
ʃok	sloɹn	skɔɪn	ɹʊp
spæl	smɛk	slɑk	sles
spɪŋk	ʃɔɪt	slem	slɔɪ
spʌb	spaʊɹ	snop	snɜ˞p
stop	steɹt	sov	stɑp
stʊl	stiɹk	spos	stɛl
swɔŋ	ʃup	tɛt	saʊd
swɜ˞p	swɜ˞p	tɪb	tɛb
swum	swʌp	tɹɑs	tɹɛl
tɹɔɪst	wɔɪnd	tɹɪs	tɹʌd
wæp	joɹn	tɜ˞l	tʌl

Received October 20, 2017
Revision received December 3, 2018
Accepted January 4, 2019 ■