



## Using complex networks to understand the mental lexicon

Michael S. Vitevitch\*, Rutherford Goldstein, Cynthia S.Q. Siew,  
Nichol Castro

Department of Psychology, University of Kansas

\*mvitev@ku.edu

### Abstract

Network science is an emerging discipline drawing from sociology, computer science, physics and a number of other fields to examine complex systems in economical, biological, social, and technological domains. To examine these complex systems, nodes are used to represent individual entities, and links are used to represent relationships between entities, forming a web-like structure, or network, of the entire system. The structure that emerges in these complex networks influences the dynamics of that system. We provide a short review of how this mathematical approach has been used to examine the structure found in the phonological lexicon, and of how subsequent psycholinguistic investigations demonstrate that several of the structural characteristics of the phonological network influence various language-related processes, including word retrieval during the recognition and production of spoken words, recovery from instances of failed lexical retrieval, and the acquisition of word-forms. This approach allows researchers to examine the lexicon at the micro-, meso-, and macro-levels, holding much promise for increasing our understanding of language-related processes and representations.

**Keywords:** mental lexicon; network science; word recognition; word production.

### 1. Introduction to Network Science

Network science is an emerging discipline that draws on techniques developed in sociology, computer science, physics, and a number of other fields to examine complex systems in economical, biological, social, and technological domains (Barabási 2009). *Nodes* (or vertices) are used to represent individual entities, and *connections* (or edges) are used to represent relationships between entities, forming a web-like structure, or *network*, of the entire system. In the psychological and brain sciences, networks have been used to re-

conceptualize the diagnostic criteria of psychological disorders (Borsboom and Cramer 2013), with nodes representing symptoms and connections linking symptoms that co-occur, and to better understand connectivity of various regions in the brain (Sporns 2010).

In the language sciences, the structure of the mental lexicon has been examined in a variety of ways, including networks where the links between words represent semantic relationships (Griffiths et al. 2007; Hills et al. 2009; Steyvers and Tenenbaum 2005), and networks where the links between words represent phonological relationships (Vitevitch 2008; Arbesman et al. 2010). (A useful bibliography of research using networks to study various aspects of language can be found at: [http://www.lsi.upc.edu/~rferrericacho/linguistic\\_and\\_cognitive\\_networks.html](http://www.lsi.upc.edu/~rferrericacho/linguistic_and_cognitive_networks.html).)

The structure of a network can be measured in a number of ways at the micro-level, the meso-level, and the macro-level. At the micro-level one can analyze individual agents in the network. Using macro-level measures one can describe the overall structure of the network. Finally, at the meso-level one can describe the network at scales between the micro- and macro-levels. An important tenet of network science is that the structure of a network influences the dynamics that take place in that network (Watts and Strogatz 1998). The remainder of this review will summarize some of the analyses and psycholinguistic experiments that have examined the structure found among the phonological word-forms in the mental lexicon (Vitevitch 2008), and how the structure of that lexical network influences language-related processes.

## 2. The structure of the mental lexicon

Vitevitch (2008) applied the tools of network science to the mental lexicon by creating a network with approximately 20,000 English words as nodes. A connection was placed between nodes if the addition, deletion, or substitution of a phoneme turned one word into the other. Other logical and linguistically motivated ways to define phonological similarity exist, including connecting a longer word to shorter words that are contained in that word (e.g., *catalog* would connect to *cat*, *a*, *at*, *log*, etc.), connecting words that share morphemes (e.g., *dog* would connect to *dogs*, *dogged*, *doggedly*, etc.), or connecting words that contain phonetically related phonemes (e.g., *bull* would connect to *veer*; see Goldinger et al. 1992). We used the “one-phoneme met-

ric” because of its simplicity and fairly widespread use as a way to operationalize “phonological similarity” (e.g., Luce and Pisoni 1998). Despite the simplicity of this metric, it seems to adequately capture what language users implicitly use to decide if two words are phonologically similar (see Luce and Large 2001; Otake and Cutler 2013). Figure 1 shows a small portion of the network that is created by using the one-phoneme metric to connect phonologically similar word-forms (see Hills et al. 2009, and Steyvers and Tenenbaum 2005 for networks based on semantic rather than phonological relationships among words).

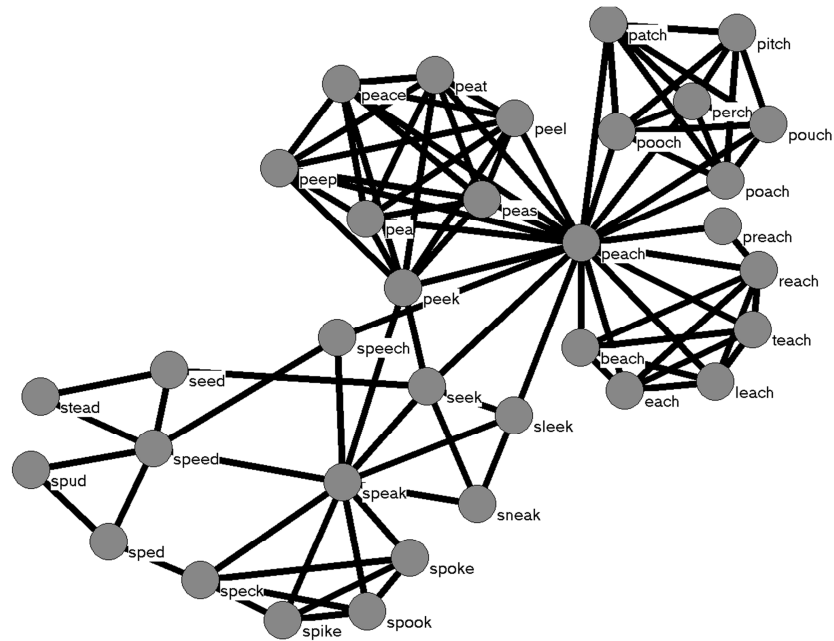


Figure 1. A sample of words from the phonological network analyzed in Vitevitch (2008). The word *speech* and its phonological neighbors (i.e., words that differ by the addition, deletion or substitution of a phoneme) are shown. The phonological neighbors of those neighbors (i.e., the 2-hop neighborhood of *speech*) are also shown.

Network analyses of the English phonological network revealed several noteworthy characteristics about the macro-structure of the mental lexicon. Vitevitch (2008) found that the phonological network had: (1) a large, highly

interconnected component, as well as many islands (words that were related to each other – such as *faction*, *fiction*, and *fission* – but not to other words in the large component) and many hermits, or words with no neighbors (known as isolates in the network science literature); the largest component exhibited (2) small-world characteristics (“short” average path length and, relative to a random graph, a high clustering coefficient; Watts and Strogatz 1998), (3) assortative mixing by degree (which means a word with many neighbors tends to have neighbors that also have many neighbors; Newman 2002), and (4) a degree distribution that deviated from a power-law. In the phonological network examined in Vitevitch (2008), the degree distribution indicates how many nodes have  $x$  phonological neighbors. Vitevitch (2008) found that there were many nodes in the network with few phonological neighbors, and only a few words with many phonological neighbors. In many real-world networks the degree distribution can be best-fit by a power-law (Albert and Barabási 2002) – that is, when the axes of the degree distribution are plotted in a logarithmic scale, a straight line is observed – but that was not the case for the phonological network examined in Vitevitch (2008).

Arbesman et al. (2010) found similar features in phonological networks of Spanish, Mandarin, Hawaiian, and Basque, and elaborated on the significance of these characteristics. For example, the giant component (i.e., the largest group of nodes that are connected to each other) of the phonological networks contained, in some cases, less than 50% of the nodes; for comparison, social networks have been observed with giant components that contain over 90% of the nodes. Arbesman et al. (2010) also noted that assortative mixing by degree is found in social networks. Recall that assortative mixing by degree means a word with many neighbors tends to have neighbors that also have many neighbors (in the context of a social network, a person with many friends has as friends people who also have many friends). The typical values for assortative mixing by degree in social networks range from .1–.3, whereas the phonological networks examined by Arbesman et al. (2010) were as high as .7.

Finally, most of the languages examined by Arbesman et al. exhibited degree distributions fit by truncated power-laws, which resembles two straight lines joined together when the axes of the degree distribution are plotted in a logarithmic scale (but the degree distribution for Mandarin was better fit by an exponential function). The degree distribution is a convenient way to statistically characterize the topology of large networks, and is essentially a frequency distribution that summarizes the number of nodes in a net-

work that have  $x$  number of connections. Networks with degree distributions that follow a power-law are known as scale-free networks. Scale-free networks have attracted attention because of certain structural and dynamic properties, such as remaining relatively intact in the face of random failures in the system, but becoming vulnerable when attacks are targeted at well-connected nodes (Albert and Barabási 2002; Albert et al. 2000). That is, if a system like the power grid that supplies electricity to a nation has a random transformer malfunction, the entire country will not experience a power outage (i.e., the system remains intact). However, if terrorists were to sabotage several of the most connected electrical transformers in the system, it is likely that most of the nation would experience a power outage (i.e., a vulnerability in the system is observed). This pattern of stability/vulnerability is typical of scale-free systems (i.e., those with degree distributions that follow a power-law), but systems with degree distributions that are best-fit by other functions exhibit other characteristics of stability/vulnerability; see work by Amaral et al. (2000) for the implications on the dynamic properties of networks with degree distributions that deviate from a power-law in certain ways.

Observing similar features in the networks of several languages (Arbesman et al. 2010) is somewhat surprising given the variety in the linguistic properties of the languages examined, and the various “families” from which the languages were sampled. For example, English is from the Germanic branch of Indo-European languages, whereas Spanish is from the Romance branch of Indo-European languages. Mandarin is not only a Sino-Tibetan language, but it further differs from English, Spanish, Hawaiian and Basque in that it uses tones to convey word meanings (*n.b.*, tone was not represented in the phonological network, however). Finally, Hawaiian is an Austronesian language with a phoneme inventory that is smaller than the inventories found in English, Spanish, Mandarin, and Basque. Basque is, of course, notable for being a linguistic isolate, or not known to be related to any other language.

Observing similar network features across a variety of languages may indicate that such networks capture important and relevant information about the similarity relationships that exist among phonological word-forms in the mental lexicon. Given the fundamental assumption of network science that the structure of a network influences processing, it is important to examine how the network features described above may influence various language processes, which is the focus of the next section.

### 3. The influence of the structure on processing

Previous studies that have examined the influence of several network measures on language-related processing will be summarized in this section. The research summarized in this section is a somewhat biased sample (focusing on work coming mostly from our lab), and is by no means a complete sampling of such work. Our goal is simply to give the reader a taste of what the tools of network science can offer language scientists. The network measures that will be considered include: degree, clustering coefficient, assortative mixing by degree, path length, and nodes that occupy “key” positions in the network (Borgatti 2006). Table 1 provides a brief definition of these network metrics.

Table 1. Brief descriptions of the network measures discussed in the text.

Network measure	Scale	Definition
Degree	Micro-level	The number of connections incident to a node.
Clustering Coefficient	Micro-level	The extent to which the neighbors of a given word are also neighbors of each other in a phonological network like that in Vitevitch (2008).
Path Length	Meso-level	The number of connections that one must traverse to get from one node to another node in a network.
Community	Meso-level	Highly interconnected sub-groups of nodes found in a network.
Assortative mixing by degree	Macro-level	Highly connected nodes tend to connect to other highly connected nodes, and less-connected nodes tend to connect to nodes that are also less-connected.
Keyplayers/keywords	Macro-level	Nodes whose removal results in the network fracturing into several smaller parts.

### 3.1. Micro-level: Degree

Degree refers to the number of connections incident to a given node. In the present review we discuss degree as a micro-level measure (i.e., a measure of individual nodes), but creating a frequency distribution of nodes varying in degree in a network (known as a degree distribution) can provide a macro-level measure of the overall network.

In the phonological network of Vitevitch (2008), degree corresponds to the number of word-forms that sound similar to a given word. In the field of psycholinguistics, the number of words that sound similar to a given word is better known as *phonological neighborhood density*. A word with few similar sounding words is said to have a sparse phonological neighborhood (or low degree), whereas a word with many similar sounding words is said to have a dense phonological neighborhood (or high degree).

The discussion of the micro-level measure of network structure known as degree is meant to highlight an interesting point of convergence between conventional psycholinguistics and network science, not to suggest that network science discovered something that is already well-known in psycholinguistics. Indeed, a number of studies have demonstrated that degree/phonological neighborhood density influences a variety of language-related processes.

For example, in studies of spoken word recognition, Luce and Pisoni (1998) found that English words with sparse phonological neighborhoods (low degree) are responded to more quickly and accurately than words with dense phonological neighborhoods (high degree). (See Vitevitch and Rodríguez 2005 for a different result in Spanish.) In studies of spoken word production, Vitevitch (2002) found that English words with dense phonological neighborhoods (high degree) are named more quickly and accurately than words with sparse phonological neighborhoods (low degree). (See Vitevitch and Stamer 2006 for a different result in Spanish.)

In the case of learning new words in English, corpus analyses (Charles-Luce and Luce 1990) and experiments using nonsense words that conform to the phonotactics of English (Storkel 2004) both suggest that it is “easier” to learn new words if they sound similar to many known words in the lexicon. That is, easily learned novel words have dense phonological neighborhoods or high degree. Finally, Roodenrys et al. (2002) showed that reintegration – that is, the use of representations in long-term memory to reconstruct representations in short-term memory – was more effective for words with dense

phonological neighborhoods (high degree) than for words with sparse phonological neighborhoods (low degree). The conceptual convergence of phonological neighborhoods/degree suggested to researchers in our lab that it might be worthwhile to examine how other network science measures of the structure of the mental lexicon influence various language-related processes.

### 3.2 Micro-level: Clustering Coefficient

In the context of a phonological network like that in Vitevitch (2008), (local) clustering coefficient, a micro-level network metric, measures the extent to which the neighbors of a given word are also neighbors of each other. Note that the mean value of the local clustering coefficient of each node in the network can be computed, providing a macro-level measure of the network. (This macro-level measure is one of the measures – the other measure being average path-length – used to classify a real-world network as having small-world characteristics instead of being either a regular or a random graph.) When clustering coefficient is low, few of the neighbors of a given word are neighbors of each other. When clustering coefficient is high, many neighbors of a given word are also neighbors with each other.

More precisely, local clustering coefficient,  $C$ , is computed:

$$C_i = \frac{2|\{e_{jk}\}|}{k_i(k_i - 1)}$$

$e_{jk}$  refers to the presence of a connection (or edge) between two neighbors ( $j$  and  $k$ ) of node  $i$ ,  $|\dots|$  is used to indicate cardinality, or the number of elements in the set (not absolute value), and  $k_i$  refers to the degree (i.e., neighborhood density) of node  $i$ . By convention, a node with degree of 0 or 1 (which results in division by 0 – an undefined value) is assigned a clustering coefficient value of 0. Thus, the (local) clustering coefficient is the proportion of connections that exist among the neighbors of a given node divided by the number of connections that could exist among the neighbors of a given node.

Although degree/neighborhood density and clustering coefficient,  $C$ , are conceptually similar, it is important to note that they are, by definition, distinct concepts. In psycholinguistic experiments examining the role of  $C$  on processing, neighborhood density/degree was controlled in the stimuli. For



example, in studies of spoken word recognition Chan and Vitevitch (2009) found that words with low  $C$  were responded to more quickly than words with high  $C$ . They suggested that words with low  $C$  have neighbors that tend to be related to other words elsewhere in the network. This relationship results in activation being broadly dispersed to the rest of the network, allowing the target word to “stand out” from its closest competitors. In contrast, words with high  $C$  have neighbors that tend to be related to other neighbors of the target word, resulting in activation being trapped in a more restricted area of the network, and making it difficult to distinguish the target word from the neighbors.

The influence of  $C$  on word recognition has been replicated in a study of visual word recognition (Yates 2013). And, importantly, Vitevitch et al. (2011) observed independent effects of degree/neighborhood density and of clustering coefficient in their computer simulation of word recognition processes, which employed activation diffusing through a network representation of the lexicon.

Chan and Vitevitch (2010) observed a similar influence of  $C$  on processing in the case of speech production. That is, words with low  $C$  were named more quickly and accurately than words with high  $C$ . A similar explanation based on the dispersion of activation for words with low  $C$  versus the trapping of activation for words with high  $C$  accounted for the speech production findings.

Vitevitch et al. (2012) further explored the dispersion of activation for words with low  $C$  versus the trapping of activation for words with high  $C$  in the context of long-term and short-term memory. In one experiment they used the phonological false memory paradigm (Sommers and Lewis 1999; see Roediger and McDermott 1995 for semantic false memories) to explore long-term memory processes. In the phonological false memory paradigm words such as *at*, *scat*, *mat*, *cut*, and *cap* are presented to participants for study. In subsequent recognition tasks, participants are likely to indicate that the word *cat* appeared in the study list, when in fact it did not (Sommers and Lewis 1999).

In the phonological false memory experiment in Vitevitch et al. (2012) the phonological neighbors of words with high  $C$  and with low  $C$  were presented to participants. For words with high  $C$ , it was predicted that activation would tend to re-circulate amongst the neighbors, whereas for low  $C$  words activation would tend to disperse to the unstudied word with low  $C$ , producing more false memories for words with low  $C$  than for words with high  $C$ .

The results of the phonological false memory experiment (and of a recognition memory task) confirmed those predictions.

Vitevitch et al. (2012) also examined the process of reintegration, in which representations in long-term memory are used to reconstruct decaying representations in short-term memory. In a serial recall task, in which participants heard a list of 6 words that were either high or low in *C*, they found higher recall for lists comprised of high *C* words than for lists comprised of low *C* words. In contrast to the previous studies, the re-circulation of activation amongst words with high *C* was now beneficial for processing by facilitating the reconstruction of decayed representations in short-term memory, whereas the dispersion of activation associated with words with low *C* provided little support to the decayed representations in short-term memory.

### 3.3. Meso-level: Paths through the lexical network

A path in a network refers to the (number of) connections that one must traverse to get from one node to another node in a network. In the present review we categorize the path (or the length of the path) in a network as a meso-level measure that describes the network between the micro-level (which describes individual nodes) and the macro-level (which describes the overall structure of the entire network). However, like degree and clustering coefficient, when a mean value is computed path-length can also be categorized as a macro-level measure.

Work by Iyengar et al. (2012) in a network of orthographic word-forms suggests that participants can use their knowledge of the paths between words to navigate from one word to another disparate word. Iyengar et al. used a game called word-morph, in which participants were given a word, and asked to form a disparate word by changing one letter at a time to form intermediate words. For example, when asked to “morph” the word *bay* into the word *egg* participants might have changed *bay* into *bad-bid-aid-add-ado-ago-ego* and finally into *egg*. Once participants in this task identified certain “landmark” words in the lexicon the task of navigating from one word to another became trivial, enabling the participants to solve subsequent word-morph puzzles very quickly. The time it took to find a solution dropped from 10-18 minutes in the first 10 games, to about 2 minutes after playing 15 games, to about 30 seconds after playing 28 games. Thus, knowledge of the

paths between words enabled participants to quickly navigate the lexical network.

Similarly, in an examination of the phonological network, Vitevitch et al. (in press) found evidence for paths existing between words during the misperception of spoken words. In a naturally occurring slip of the ear, a listener hears a word, but does not perceive the word that was uttered. Instead the listener perceives a word that is phonologically similar to the uttered word. As a laboratory analogue of what happens when one experiences a slip of the ear, Vitevitch et al. presented participants with an English word over a pair of headphones and asked participants to respond with the first word that came to mind that sounded like the word they heard (see Luce and Large 2001). In the phonological associate task – a variant of the well-known semantic associate task – participants were allowed to define for themselves what “sounded like” meant, thereby allowing the researchers to explore the parameters that listeners may use (implicitly) to define phonological similarity.

Two results from the phonological associate task used in Vitevitch et al. (in press) are relevant to the present discussion of paths in the lexical network. First, about 80% of the responses were 1 link away from the stimulus word. That is, the response differed from the stimulus word by a single phoneme. Importantly, the location of the phoneme change was distributed across positions in the word, suggesting that participants did not adopt the naïve strategy of simply producing a word that rhymed with the stimulus word.

Second, responses that differed by more than one phoneme tended to be connected to the stimulus word by a path of real words in the lexicon. The existence of lexical intermediaries raises some concerns about measures of word-form similarity that ignore such items, such as the Orthographic Levenshtein Distance-20 (OLD-20; Yarkoni et al. 2008), and the Phonological Levenshtein Distance-20 (PLD-20; Suárez et al. 2011). In OLD-20/PLD-20, Levenshtein distance is computed between a target word and all other words in the lexicon. OLD-20/PLD-20 is then the mean edit distance (i.e., the number of letter or phoneme changes) in the 20 closest neighbors. The computations of OLD-20/PLD-20 do not consider whether real-word intermediaries exist or not; the measure only considers the number of letter/phoneme changes (respectively) that are required to turn one word into another. However, the work of Iyengar et al. (2012) and of Vitevitch et al. (in press) suggest that even distant lexical neighbors tend to be connected to a word via a path of real words, raising questions about the psychological validity of met-

rics such as OLD-20 and PLD-20 that do not consider the influence of these lexical intermediaries.

### 3.4. Meso-level: Community structure

Recall that the meso-level describes structures that lie between the micro-level (which describes individual nodes) and the macro-level (which describes the overall structure of the entire network). At the meso-level researchers often look for evidence of community structure. That is, within the whole network, do enclaves or sub-groups of nodes exist, such that the nodes within a group (or community) are more highly interconnected with each other than they are to nodes in other groups?

Communities in a larger network have been observed in the human brain, in social networks, and in the World Wide Web (Porter et al. 2009). Siew (2013) examined the giant component of the phonological network described in Vitevitch (2008), and found 17 communities using a common community detection technique developed by Girvan and Newman (2002). The communities ranged in size from 31 to 697 words. Interestingly, the larger communities tended to consist of shorter, more frequent words of high degree and earlier age of acquisition, whereas smaller communities tended to consist of longer, less frequent words of low degree and later age of acquisition. This pattern was similar to the pattern found in the phonological network overall – the giant component contains shorter more frequent words with high degree, the lexical islands tend to be slightly longer and lower in frequency and degree, and the lexical hermits tend to be longer still, with even lower word frequency (and degree = 0) – suggesting a remarkable self-similarity in the phonological network across scales; that is, certain features of the phonological network appear when you look at the whole system, or at a smaller part the system.

The importance of community structure can be determined by examining the constituents of the various communities to see what the members of a community have in common, and how the constituents of one community differ from the constituents of other communities. Siew (2013) observed that different communities tended to contain certain sequences of phonological segments. For example, in one community the most common sequences of phonological segments were /ŋk/, /ŋj/, and /rɪ/, and were found in words like *brink*, *drink* and *wrinkle*. Siew suggested that knowledge of phonotactic in-

formation may emerge from lexical representations rather than from separate representations of phonological segments and sequences of segments as has been proposed in certain models of spoken word recognition (e.g., *Shortlist*; Norris, 1994). That is, processing focused on the micro-level of the network (i.e., on independent lexical representations) might result in “lexical effects” such as the commonly observed neighborhood density effects or word frequency effects, whereas processing focused on the meso-level of the network might result in phonotactic probability effects being observed (e.g., Vitevitch and Luce 2005; Vitevitch 2003). This novel insight provided by the analysis for community structure provides an alternative (and, perhaps, more parsimonious) account of neighborhood density and phonotactic probability effects observed in the literature (Vitevitch and Luce 2005; Vitevitch 2003).

Siew (2013) further suggested that the community structure of the giant component might play a role in spoken word recognition by limiting the spread of activation in the network, thereby constraining the number of possible lexical competitors. Recall that Vitevitch (2008) observed a small-world structure in the giant component of the phonological network. Further recall that one of the defining features of small-world networks is that (on average) a very short path exists between any two nodes in the network, implying that activation spreading in a network will very quickly activate the entire network. The community structure in the phonological network observed by Siew (2013) may keep activation localized within a community and slow its progress to other communities, thereby reducing the possibility of rampant activation of the mental lexicon.

Although the speculations of Siew (2013) remain to be tested empirically, the ideas put forward are novel and provocative. Furthermore, the ideas put forward by Siew show the innovative predictions that can be generated when the tools of network science are employed to examine psycholinguistic processes and representations.

### **3.5. Macro-level: Assortative mixing by degree**

Assortative mixing by degree is a macro-level metric that refers to the tendency in a network for a highly connected node to be connected to other highly connected nodes (Newman 2002), and for less-connected nodes to connect to nodes that are also less-connected. In other words, when looking

at the network overall, there is a positive correlation between the degree of a node and the degree of its neighbor.

Assortative mixing by degree is often observed in social networks (Newman 2002). Note that a negative correlation between the degree of a node and the degree of its neighbors is also possible, and is known as disassortative mixing by degree. Disassortative mixing by degree is often found in networks representing technological systems, like the World Wide Web (Newman 2002). Networks with a correlation of zero between the degree of a node and the degree of its neighbors are also possible, and also have been observed (Newman 2002).

Mathematical simulations by Newman (2002) found that network connectivity (i.e., the existence of paths between pairs of nodes) was easier to disrupt (by a factor of five to ten) by removing nodes with high-degree in networks with disassortative mixing than in networks with assortative mixing by degree. In other words, networks with assortative mixing by degree are better able to maintain processing pathways than networks with disassortative mixing by degree in the face of targeted attacks to the system. In the context of a social network representing a criminal organization a “highly connected” node (representing a criminal in that organization) could be removed from the system by arresting that person, thereby disrupting the flow of information and illicit goods through the criminal organization.

Recall that Arbesman et al. (2010) observed values of assortative mixing by degree in the phonological networks of several languages that were as high as .7 (compared to values ranging from .1–.3 in social networks). They also examined how connected the network remained (indicating how well the network could continue to transmit information, etc.) in response to the removal of nodes in the network of English words. In the context of a lexical network a node might be “removed” (at least temporarily) due to the tip-of-the-tongue state, in which the phonological representation of a word fails to be retrieved. Arbesman et al. observed that the language network remained relatively well connected when nodes were removed at random and when the highly connected nodes were targeted for removal. This pattern of network resiliency differs from that typically seen in other networks, which tend to remain well-connected in the face of random attacks on the network, but quickly become disconnected when highly connected nodes are targeted for removal (e.g., Albert et al. 2002).

Given the high levels of assortative mixing by degree, and the ability of the networks to remain intact despite random and targeted removal of nodes

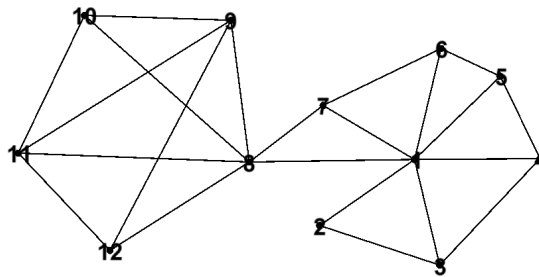


Figure 2. Although node 1 is connected to the most nodes in the network (nodes 2–8) it is not a “key player” in the network, because its removal does not disconnect the network. However, removal of node 8 (a “key player”) results in the network fracturing into two smaller components that are disconnected from each other.

Adapted from Borgatti (2006).

that was observed in the language networks examined by Arbesman et al. (2010), Vitevitch et al. (2014) examined instances of lexical retrieval failures in a computer simulation (jTRACE; Strauss et al. 2007), real lexical retrieval failures from a slips-of-the-ear corpus (Bond 1999), and elicited lexical retrieval failures from three psycholinguistic experiments for evidence of assortative mixing by degree. In all cases the same result was observed: a positive correlation was found between the degree/neighborhood density of the stimulus word and the degree/neighborhood density of the phonological neighbors that were (erroneously) retrieved when lexical retrieval failed. These results suggest that failed lexical retrieval, a less-commonly examined aspect of language processing, is influenced by the macro-level structure of the lexical network known as assortative mixing by degree. These findings highlight how the network approach can be used to examine less-commonly asked questions in psycholinguistics, and how the account provided for the observed results (i.e., the structure of the lexical network influences processing) can also account for a wide-range of other findings.

### 3.6. Macro-level: Key players in a network/keywords in the lexicon

Key players are “important” nodes in a network. In Figure 2, node 1 appears to be very “important” because it connects to many other nodes in the net-

work. What makes key players so “important” is that the removal of such nodes results in the network fracturing into several smaller components (or connected groups of nodes that are not connected to other groups of nodes). Removing node 1 (and its connections) from the network does not “fracture” the network (or in network science terminology, partition the network). Information (or disease, etc.) can still get from node 2 to nodes 8–12; the pathway from node 2 to nodes 8–12 is longer now, but a path still exists between those nodes.

Now consider what happens if node 8 (and its connections) is removed from the network. We are now left with two smaller components (one comprised of nodes 1–7, and the other comprised of nodes 9–12) that cannot communicate with each other because there is no path that connects them, resulting in a much more significant disruption to the system than what occurs when node 1 is removed.

Goldstein and Vitevitch (2014) found a set of such nodes in the phonological network, and selected another set of words (i.e., foils) from the lexical network that were comparable to the “keywords” on a number of standard psycholinguistic characteristics (word length, word frequency, familiarity, neighborhood density/degree, phonotactic probability, etc.) and network characteristics (clustering coefficient, closeness centrality, etc.). Examples of the keywords include: *fish*, *misty*, *pocket*, *spring* and *tense*; examples of the foils include: *firm*, *mystic*, *party*, *squirt*, and *test*. Goldstein and Vitevitch (2014) presented the keywords and foils as stimuli in three conventional tasks commonly used in psycholinguistics: perceptual identification task, auditory naming task, and auditory lexical decision task. Across tasks the results showed that the keywords were responded to more quickly and accurately than the words that were comparable to the keywords in their psycholinguistic and network characteristics.

The results of Goldstein and Vitevitch (2014) – indeed, of all the studies summarized above – suggest that the characteristics of individual words that language scientists typically examine – word frequency, word length, neighborhood density, phonotactic probability – are not the only things that influence lexical processing. Importantly, the relationships that exist among words in the lexicon also matter. The structure that emerges in the lexical network from the similarity relations among words appears to influence a number of language-related processes, including word retrieval during the recognition and production of spoken words, recovery from instances of



failed lexical retrieval, and the acquisition of word-forms (see Goldstein and Vitevitch 2014).

#### 4. Summary

One way to measure the structure of complex systems such as the mental lexicon at multiple scales – micro-, meso- and macro-level – is with the tools provided by network science. Important discoveries in the domains of biology, technology, and social interaction have been made using the tools of this approach (for a brief review see Albert and Barabási 2002; Brandes et al. 2013). The studies summarized in this brief review demonstrate that network science can lead to new discoveries in the psychological and linguistic sciences as well. Although we did not discuss this issue in the present review, network science also provides a framework for examining changes in a system on multiple time-scales, such as the development of language in an individual (Hills et al. 2009), diachronic changes in a language over longer time scales, and even changes on a longer, evolutionary time-scale (Solé et al. 2010).

Network science holds much promise for various areas of psychology and the language sciences, and can be used to address a wide variety of problems and research questions. It is important to note, however, that network science is not a panacea. Researchers who desire to use the theoretical framework and analytic tools of network science should think carefully about how well entities and relationships among those entities in a given domain map onto nodes and connections in a network representation (see Valente 2012 for a similar point regarding network interventions).

#### References

- Albert, R. and A.L. Barabási. 2002. “Statistical mechanics of complex networks”. *Review of Modern Physics* 74. 47–97.
- Albert, R., H. Jeong and A.L. Barabási. 2000. “Error and attack tolerance of complex networks”. *Nature* 406. 378–382.
- Arbesman, S., S.H. Strogatz and M.S. Vitevitch. 2010. “The structure of phonological networks across multiple languages”. *International Journal of Bifurcation and Chaos* 20. 679–685.

- Barabási, A.L. 2009. "Scale-free networks: A decade and beyond". *Science* 325. 412–413.
- Bond, Z.S. 1999. *Slips of the ear: Errors in the perception of casual conversation*. New York: Academic Press.
- Borgatti, S.P. 2006. "Identifying sets of key players in a network". *Computational, Mathematical and Organizational Theory* 12. 21–34.
- Borsboom, D. and A.O.J. Cramer. 2013. "Network analysis: An integrative approach to the structure of psychopathology". *Annual Review of Clinical Psychology* 9. 91–121.
- Brandes, U., G. Robins, A. McCranie and S. Wasserman. 2013. "Editorial: What is network science?" *Network Science* 1. 1–15.
- Chan, K.Y. and M.S. Vitevitch. 2009. "The influence of the phonological neighborhood clustering-coefficient on spoken word recognition". *Journal of Experimental Psychology: Human Perception and Performance* 35. 1934–1949.
- Chan, K.Y. and M.S. Vitevitch. 2010. "Network structure influences speech production". *Cognitive Science* 34. 685–697.
- Charles-Luce, J. and P.A. Luce. 1990. "Similarity neighborhoods of words in young children's lexicons". *Journal of Child Language* 17. 205–215.
- Girvan, M. and M.E.J. Newman. 2002. "Community structure in social and biological networks". *Proceedings of the National Academy of Sciences* 99. 7821–7826.
- Goldinger, S.D., P.A. Luce, D.B. Pisoni and J.K. Marcario. 1992. "Form-based priming in spoken word recognition: The roles of competition and bias". *Journal of Experimental Psychology: Learning, Memory and Cognition* 18. 1211–1238.
- Goldstein, R. and M.S. Vitevitch. 2014. "The influence of clustering coefficient on word-learning: How groups of similar sounding words facilitate acquisition". *Frontiers in Psychology* 5. 1307.
- Griffiths, T.L., M. Steyvers and A. Firl. 2007. "Google and the mind: Predicting fluency with PageRank". *Psychological Science* 18. 1069–1076.
- Hills, T.T., M. Maouene, J. Maouene, A. Sheya and L. Smith. 2009. "Longitudinal analysis of early semantic networks: Preferential attachment or preferential acquisition?" *Psychological Science* 20. 729–739.
- Iyengar, S.R.S., C.E.V. Madhavan, K.A. Zweig and A. Natarajan. 2012. "Understanding human navigation using network analysis". *Topics in Cognitive Science* 4. 121–134.
- Luce, P.A. and N.R. Large. 2001. "Phonotactics, density and entropy in spoken word recognition". *Language and Cognitive Processes* 16. 565–581.
- Luce, P.A. and D.B. Pisoni. 1998. "Recognizing spoken words: The neighborhood activation model". *Ear and Hearing* 19. 1–36.
- Newman, M.E.J. 2001. "The structure of scientific collaboration networks". *Proceedings of the National Academy of Sciences* 98. 404–409.
- Newman, M.E.J. 2002. "Assortative mixing in networks". *Physical Review Letters*, 89. 20889701.
- Norris, D. 1994. "Shortlist: A connectionist model of continuous speech recognition". *Cognition* 52. 189–234.
- Otake, T. and A. Cutler. 2013. "Lexical selection in action: Evidence from spontaneous punning". *Language and Speech* 56. 555–573.

- Porter, M.A., J. P. Onnela and P.J. Mucha. 2009. "Communities in networks". *Notices of the American Mathematical Society* 56. 1082–1166.
- Roediger, H. L. and K.B. McDermott. 1995. "Creating false memories: Remembering words not presented in lists". *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21. 803–814.
- Roodenrys, S., C. Hulme, A. Lethbridge, M. Hinton and L.M. Nimmo. 2002. "Word-frequency and phonological-neighborhood effects on verbal short-term memory". *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28. 1019–1034.
- Siew, C.S.Q. 2013. "Community structure in the phonological network". *Frontiers in Psychology* 4. 553.
- Solé, R.V., B. Corominas-Murtra, S. Valverde and L. Steels. 2010. "Language networks: Their structure, function and evolution". *Complexity* 15. 20–26.
- Sommers, M.S. and B.P. Lewis. 1999. "Who really lives next door: Creating false memories with phonological neighbors". *Journal of Memory and Language* 40. 83–108.
- Sporns, O. 2010. *Networks of the brain*. MIT Press.
- Steyvers, M. and J. Tenenbaum. 2005. "The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth". *Cognitive Science* 29. 41–78.
- Storkel, H.L. 2004. "Do children acquire dense neighborhoods? An investigation of similarity neighborhoods in lexical acquisition". *Applied Psycholinguistics* 25. 201–221.
- Strauss, T.J., H.D. Harris and J.S. Magnuson. 2007. "jTRACE: A reimplementation and extension of the TRACE model of speech perception and spoken word recognition". *Behavior Research Methods* 39. 19–30.
- Suárez, L., S.H. Tan, M.J. Yap and W.D. Goh. 2011. "Observing neighborhood effects without neighbors". *Psychonomic Bulletin and Review* 18. 605–611.
- Valente, T.W. 2012. "Network interventions". *Science* 337. 49–53.
- Vitevitch, M.S. 2002. "The influence of phonological similarity neighborhoods on speech production". *Journal of Experimental Psychology: Learning, Memory and Cognition* 28. 735–747.
- Vitevitch, M.S. 2003. "The influence of sublexical and lexical representations on the processing of spoken words in English". *Clinical Linguistics and Phonetics* 17. 487–499.
- Vitevitch, M.S. 2008. "What can graph theory tell us about word learning and lexical retrieval?" *Journal of Speech Language Hearing Research* 51. 408–422.
- Vitevitch, M.S., K.Y. Chan and R. Goldstein. 2014. "Insights into failed lexical retrieval from network science". *Cognitive Psychology* 68. 1–32.
- Vitevitch, M.S., G. Ercal and B. Adagarla. 2011. "Simulating retrieval from a highly clustered network: Implications for spoken word recognition". *Frontiers in Psychology* 2. 369.
- Vitevitch, M.S., R. Goldstein and E. Johnson. In press. "Path-length and the misperception of speech: Insights from Network Science and Psycholinguistics". In: Mehler, A., P. Blanchard, B. Job and S. Banish (eds.), *Towards a theoretical*

- framework for analyzing complex linguistic networks. (Understanding Complex Systems series.)* New York: Springer.
- Vitevitch, M.S. and P.A. Luce. 2005. "Increases in phonotactic probability facilitate spoken nonword repetition". *Journal of Memory and Language* 52. 193–204.
- Vitevitch, M.S. and E. Rodríguez. 2005. "Neighborhood density effects in spoken word recognition in Spanish". *Journal of Multilingual Communication Disorders* 3. 64–73.
- Vitevitch, M.S. and M.K. Stamer. 2006. "The curious case of competition in Spanish speech production". *Language and Cognitive Processes* 21. 760–770.
- Watts, D.J. and S.H. Strogatz. 1998. "Collective dynamics of 'small-world' networks". *Nature* 393. 409–410.
- Yarkoni, T., D. Balota and M.J. Yap. 2008. "Moving beyond Coltheart's N: A new measure of orthographic similarity". *Psychonomic Bulletin and Review* 15. 971–979.
- Yates, M. 2013. "How the clustering of phonological neighbors affects visual word recognition". *Journal of Experimental Psychology: Learning, Memory, and Cognition* 39. 1649–1656.

**Address correspondence to**

Michael S. Vitevitch  
Department of Psychology  
1415 Jayhawk Blvd.  
Fraser Hall, Room 426  
Lawrence, KS 66045-7556  
USA  
vitevitch@ukans.edu