View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by KU ScholarWorks

NIH Public Access
Author Manuscript
Cogn Psychol. Author manuscript; available in PMC 2015 February 01.

Published in final edited form as:
Cogn Psychol. 2014 February ; 68: 1–32. doi:10.1016/j.cogpsych.2013.10.002.

Insights into failed lexical retrieval from network science

Michael S. Vitevitch[*], **Kit Ying Chan**, and **Rutherford Goldstein**
Department of Psychology, University of Kansas, United States

Abstract

Previous network analyses of the phonological lexicon (Vitevitch, 2008) observed a web-like structure that exhibited assortative mixing by degree: words with dense phonological neighborhoods tend to have as neighbors words that also have dense phonological neighborhoods, and words with sparse phonological neighborhoods tend to have as neighbors words that also have sparse phonological neighborhoods. Given the role that assortative mixing by degree plays in network resilience, we examined instances of real and simulated lexical retrieval failures in computer simulations, analysis of a slips-of-the-ear corpus, and three psycholinguistic experiments for evidence of this network characteristic in human behavior. The results of the various analyses support the hypothesis that the structure of words in the mental lexicon influences lexical processing. The implications of network science for current models of spoken word recognition, language processing, and cognitive psychology more generally are discussed.

Keywords

Network science; Spoken word recognition; Mental lexicon

1. Introduction

Network science draws on work from mathematics, sociology, computer science, physics and a number of other fields that examine complex systems using *nodes* (or vertices) to represent individual entities, and *connections* (or edges) to represent relationships between entities, to form a web-like structure, or *network*, of the entire system. This approach has been used to examine complex systems in economical, biological, social, and technological domains (Barabási, 2009). More relevant to the cognitive sciences, this approach has also increased our understanding of connectivity in the brain (Sporns, 2010), the categorization of psychological disorders (Cramer, Waldorp, van der Maas, & Borsboom, 2010), and the cognitive processes and representations involved in human navigation (Iyengar, Madhavan, Zweig, & Natarajan, 2012), semantic memory (Griffiths, Steyvers, & Firl, 2007; Hills, Maouene, Maouene, Sheya, & Smith, 2009; Steyvers & Tenenbaum, 2005), and human collective behavior (Goldstone, Roberts, & Gureckis, 2008).

Cognitive science has appealed to networks in the past, including artificial neural networks (Rosenblatt, 1958), networks of semantic memory (Quillian, 1967), and network-like models of language (e.g., linguistic nections: Lamb, 1970; Node Structure Theory: MacKay, 1992). What separates network science from these previous network approaches is that network science is equal parts theory and equal parts methodology: "…networks offer both a theoretical framework for understanding the world and a methodology for using this

© 2013 Elsevier Inc. All rights reserved.

[*]Corresponding author. Address: Spoken Language Laboratory, Department of Psychology, 1415 Jayhawk Blvd., University of Kansas, Lawrence, KS 66045, United States. mvitevit@ku.edu (M.S. Vitevitch).

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

framework to collect data, test hypotheses, and draw conclusions" (Neal, 2013, p. 5). Regarding methodology, network science offers a wide array of statistical and computational tools to analyze individual agents in a complex system (often referred to as the *microlevel*), characteristics of the over-all structure of a system (often referred to as the *macro-level*), as well as various levels in between (often referred to as the *meso-level*). See Appendix A for definitions of various network measures that can be used to assess the different levels of a system.

One could argue, without oversimplifying the case, that psycholinguistic research has predominantly engaged in micro-level analyses to identify specific lexical characteristics—such as word frequency, age of acquisition, word length, phonotactic probability, etc.—that affect the speed and accuracy with which a word can be retrieved from the lexicon. Indeed, Cutler (1981) provided a long list of lexical characteristics that were known to affect the retrieval of a word; this list has only grown since then (e.g., Vitevitch, 2002a; Vitevitch, 2007). Although much has been learned about the micro-structure of words in the mental lexicon, what was not known then—and is still missing in mainstream approaches—is an understanding of how the macro-structure and meso-structure of the lexicon influences lexical processing. Stated more brusquely: studying the individual pieces does not help us understand how they all fit together, nor how the entire system works as a whole (for a similar critique of reductionism in other fields see Barabási, 2012).

To better understand the complex cognitive system known as the mental lexicon, we used the mathematical tools of network science to examine how the mental lexicon might be structured at the macro-level, and how the macro-structure of the lexicon might influence lexical processing. To be sure, psycholinguists have hinted at the influence that the structure of the lexicon might have on lexical processing. Consider this statement by Forster (1978, p. 3): "[a] structured information-retrieval system permits speakers to recognize words in their language effortlessly and easily" as well as the verbal model of lexical retrieval that he proposed.

Also consider this statement by Luce and Pisoni (1998, p. 1): "…similarity relations among the sound patterns of spoken words represent one of the earliest stages at which the structural organization of the lexicon comes into play." What was missing, however, were the right tools to measure the structure of the lexicon. Statements by Barabási suggest that the tools of network science are ideally suited to examine the micro-, meso- and macro-structure of the mental lexicon, as well as other complex systems:

> All systems perceived to be complex, from the cell to the Internet and from social to economic systems, consist of an extraordinarily large number of components that interact via intricate networks. To be sure, we were aware of these networks before. Yet, only recently have we acquired the data and tools to probe their topology, helping us realize that the underlying connectivity has such a strong impact on a system's behavior that no approach to complex systems can succeed unless it exploits the network topology. (Barabási, 2009, p. 413)

Vitevitch (2008) applied the tools of network science to the mental lexicon by creating a network with approximately 20,000 English words as nodes, and connections between words that were phonologically similar (using the one-phoneme metric used in Luce and Pisoni (1998)). Fig. 1 shows a small portion of this network (see Hills et al. (2009) and Steyvers and Tenenbaum (2005) for lexical networks based on semantic rather than phonological relationships among words).

Network analyses of the English phonological network revealed several noteworthy characteristics about the macro-structure of the mental lexicon. Vitevitch (2008) found that

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

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 & Strogatz, 1998), (3) assortative mixing by degree (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.

Arbesman, Strogatz, and Vitevitch (2010b) found the same constellation of structural features in phonological networks of Spanish, Mandarin, Hawaiian, and Basque, and elaborated on the significance of these characteristics. For example, the giant component of the phonological networks contained, in some cases, less than 50% of the nodes; networks observed in other domains often have giant components that contain 80–90% of the nodes. Arbesman et al. (2010b) also noted that assortative mixing by degree is found in networks in other domains. However, typical values for assortative mixing by degree in social networks range from .1–.3, whereas the phonological networks examined by Arbesman et al. were as high as .7.

Finally, most of the languages examined by Arbesman et al. exhibited degree distributions fit by truncated power-laws (but the degree distribution for Mandarin was better fit by an exponential function). 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 vulnerability when attacks are targeted at well-connected nodes (Albert & Barabási, 2002; Albert, Jeong, & Barabási, 2000). See work by Amaral, Scala, Barthélémy, and Stanley (2000) for the implications on the dynamic properties of networks with degree distributions that deviate from a power-law in certain ways.

It is important to note that Kello and Beltz (2009) demonstrated that certain characteristics they observed in the networks of several real languages (English, Dutch, German, Russian, and Spanish) were unlikely to arise simply because connections between nodes were based on substring relations (like the one-phoneme metric used to connect nodes in the networks examined in Vitevitch (2008) and Arbesman, Strogatz, and Vitevitch (2010a, 2010b). Kello and Beltz used methods like those used by Mandelbrot (1953) and Miller (1957) to create variable-length random letter strings, and found that the degree distribution of a network comprised of such overlapping items differed radically from the degree distribution of the networks containing words from real languages.

Similarly, Gruenenfelder and Pisoni (2009) created three pseudolexicons of randomly generated "words." Only the pseudolexicon that contained items that closely resembled real-words in natural languages—words that varied in length, the proportions of words at a given length that matched those proportions found in English, phonotactic-like constraints on the ordering of consonants and vowels in the words—produced a network structure that began to resemble the network structure observed in natural languages. Thus, networks of word-forms from natural languages capture important information at the micro- meso- and macro-levels about the structural relations among those word-forms above and beyond what might be expected from a network formed from nodes that simply share overlapping features.

One of the fundamental assumptions of network science is that the structure of a network influences the dynamics of that system (Watts & Strogatz, 1998). A certain process might operate very efficiently on a network with a certain structure. However, in a network with the same number of nodes and same number of connections—but with those nodes

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

connected in a slightly different way—the same process might now be woefully inefficient. Given the fundamental assumption that the structure of a network influences the dynamics of that system, then the structure among phonological word-forms in the mental lexicon observed by Vitevitch (2008) should influence certain language- related processes. Indeed, psycholinguistic evidence demonstrates that three network features at the micro-level of the network—degree, clustering coefficient, and closeness centrality of individual words—influence a number of language-related processes.

Degree refers to the number of connections incident to a given node. In the context of a phonological network like that of Vitevitch (2008), degree corresponds to the number of word-forms that sound similar to a given word. Many psycholinguistic studies have shown that degree—better known in the psycholinguistic literature as *phonological neighborhood density*—influences spoken word recognition (Luce & Pisoni, 1998), spoken word production (Vitevitch, 2002b), word-learning (Charles-Luce & Luce, 1990; Storkel, 2004), and phonological short-term memory (Roodenrys, Hulme, Lethbridge, Hinton, & Nimmo, 2002). Our discussion of degree is not meant to suggest that our use of network science to examine the mental lexicon led us to discover something new (*i.e*., the already well-known effects of phonological neighborhood density on processing). Rather, we discuss degree and neighborhood density to show an interesting point of convergence between conventional psycholinguistics and network science. This conceptual convergence inspired us to examine how other network science measures of the structure of the mental lexicon might influence various language-related processes.

One of those other network science measures is the *clustering coefficient*, which—in the context of a phonological network like that in Vitevitch (2008)—measures the extent to which the neighbors of a given node are also neighbors of each other. We initially explored the influence of the clustering coefficient on processing because it provides a measure of the "internal structure" of a phonological neighborhood. When clustering coefficient is low, few of the neighbors of a target node are neighbors of each other. When clustering coefficient is high, many neighbors of a target word are also neighbors with each other. Given the well-known and widely replicated effects of degree/neighborhood density, we reasoned that we would likely be able to observe influences of the internal structure of a phonological neighborhood on processing. Indeed, the results of several studies—using a variety of conventional psycholinguistic and memory tasks, as well as computer simulations—demonstrated that clustering coefficient influences language-related processes like spoken word recognition, word production, retrieval from long-term memory, and redintegration in short-term memory (Chan & Vitevitch, 2009, 2010; Vitevitch, Chan, & Roodenrys, 2012; Vitevitch, Ercal, & Adagarla, 2011).

Although degree/neighborhood density and clustering coefficient are conceptually similar, it is important to note that they are, by definition, distinct concepts. Furthermore, Vitevitch et al. (2012) demonstrated that degree and clustering coefficient are not correlated in the phonological network of English. Moreover, Vitevitch et al. (2011) observed in a computer simulation independent effects of degree/neighborhood density and of clustering coefficient on the diffusion of activation in their network representation of the lexicon.

Using 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, Iyengar et al. (2012) demonstrated the influence of another network science measure—*closeness centrality*—on a language-related process. Closeness centrality is one type of centrality measure (others being degree centrality, betweenness centrality, reach centrality, and eigenvector centrality) that attempts to capture in some way which nodes are "important" in the network. Closeness centrality assesses how far away other nodes in the network are from a given node.[1] A node with high

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

closeness centrality is very close to many other nodes in the network, requiring that, on average, only a few links be traversed to reach another node. A node with low closeness centrality is, on average, far away from other nodes in the network, requiring the traversal of many links to reach another node.

Iyengar et al. (2012) found that words with high closeness centrality in a network of the orthographic lexicon allowed participants in the word-morph game to quickly transform one word into another. For example, 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*. Similarly, when asked to "morph" the word *ass* into the word *ear* participants might have changed *ass* into *ask–ark–arm–aim–**aid**–bid–bad–bar* and finally into *ear*. Once participants in this task identified certain "landmark" words in the lexicon—words that had high closeness centrality, like the word ***aid*** in the examples above—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 min in the first 10 games, to about 2 min after playing 15 games, to about 30 s after playing 28 games, because participants would "morph" the start-word (e.g., *bay* or *ass*) into one of the landmark words that were high in closeness centrality (e.g., *aid*), then morph the landmark-word into the desired end-word (e.g., *egg* or *ear*). Although this task is a contrived word-game rather than a conventional psycholinguistic task that assesses on-line lexical processing, the results of Iyengar et al. (2012) nevertheless illustrate further how the tools of network science can be used to provide insights about cognitive processes and representations.

Our review of previous studies that examined how degree, clustering coefficient, and closeness centrality influence various language-related processes clearly shows that the tools of network science can be used to provide novel insights about cognitive processes and representations. Note, however, that these metrics assess the micro-level of a system, providing information about individual nodes rather than the system as a whole. Therefore, one could argue that these applications of network science to psycholinguistics are also guilty of the charge we leveled against mainstream psycholinguistic research: the myopic focus of mainstream psycholinguistics on the trees (i.e., individual characteristics of words) has prevented mainstream psycholinguistics from seeing the forest (i.e., the mental lexicon as a system). To illustrate further how the tools of network science can be used to provide novel insights about cognitive processes and representations (and to acquit ourselves of this crime of perspective), we examined in the present study how the macro-level measure from network science known as *assortative mixing by degree* might influence language-related processing.

To define *assortative mixing by degree* we will consider each component of this term in turn. *Mixing* describes a preference for how nodes in a network tend to connect to each other. This preference can be based on a variety of characteristics. For example in a social network, mixing may occur based on age, gender, race, etc. *Assortative* mixing (*a.k.a.* homophily) means that "like goes with like." Again using the example of a social network,

[1]Closeness centrality is defined as the inverse of the sum of distances from that node to all other reachable nodes in the network. More precisely:

$$Closeness(i) = \frac{1}{\sum_j d_{ij}}$$

where *i* is the node of interest, *j* is another node in the network, and d_{ij} is the shortest distance between these two nodes.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

people of similar age tend to connect to each other. There are, of course, instances in which younger people are friends with older people, but overall there is a general tendency in the network for people of similar ages to be connected. Assortative mixing can be contrasted with disassortative mixing, which means that dissimilar entities will tend to be connected. For example, in a network of a heterosexual dating website, males and females would be connected, but not males and males, nor females and females. It is also possible that no mixing preferences are exhibited in a network.

In addition to gender or age, degree, or the number of edges incident on a vertex, is another way in which nodes in a network may exhibit a preference for mixing. Therefore, assortative mixing by degree refers to the tendency in a network for a highly connected node to be connected to other highly connected nodes (Newman, 2002). 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 neighbors. 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).

The relatively high values of assortative mixing by degree in the phonological networks examined by Arbesman et al. (2010b)—values as high as .7 compared to values ranging from .1–.3 in social networks— demand additional investigation. One possibility is that the values observed by Arbesman et al. are just a statistical quirk or a mathematical curiosity. Indeed, there are many interesting relationships and correlations that have been observed among words or in language more generally. Consider Menzerath's law, which states that the larger a particular unit is, the smaller its constituents will be, such that a long word will be comprised of small or simple syllables (such as CVs or Vs), whereas a short word will be comprised of a large or more complex syllable (such as CCCVCCC as in the monosyllabic word, *strengths*). The relationship between unit and constituent size known as Menzerath's law is not only observed in language, but has also been observed in music and genomes (Ferrer-i-Cancho, Forns, Hernandez-Fernandez, Bel-Enguix, & Baixeries, 2012).

Consider also Martin's Law (the relationship between the number of definitions of a word and the generality of those definitions), as well as the relationship between phoneme inventory and entropy, and the relationship between polysemy and word length. Finally, consider the numerous observations by Zipf (1935), including the correlation between word-length and word frequency (long words tend to occur less often in the language; see also Baayen, 1991, 2001, 2010), and the observation that high frequency words tend to have many phonological neighbors (e.g., Frauenfelder, Baayen, Hellwig, & Schreuder, 1993; Landauer & Streeter, 1973). Given the plethora of relationships observed among words and in language more generally (and sometimes even in other domains), the observation of assortative mixing by degree in Arbesman et al. (2010b) could simply be another statistical curiosity.

We grant that it is interesting to discover and document the presence of such relationships in and across languages. We further grant that much has been learned by studying how the variables that contribute to those relationships influence processing. Consider all of the work investigating the influence of word frequency on various language and memory processes (dating back at least to Lester, 1922), studies of word-length on processing (e.g., Vitevitch, Stamer, & Sereno, 2008), all of the work on orthographic and phonological neighborhood density on processing (e.g., Laxon, Coltheart, & Keating, 1988; Pisoni, Nusbaum, Luce, &

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Slowiaczek, 1985), and the work on neighborhood frequency on processing (e.g., Grainger, O'Regan, Jacobs, & Segui, 1989). What appears to be lacking are studies that demonstrate that relationships between variables—such as the observation that high frequency words tend to have many phonological neighbors (e.g., Frauenfelder et al., 1993; Landauer & Streeter, 1973)—directly influence processing. From the perspective of cognitive psychology, it is important to demonstrate that statistical relationships between variables are more than interesting mathematical quirks. That is, one must demonstrate that such statistical relationships influence *cognitive processing* in some way. In the work that follows, we will examine how the macro-level measure of a network known as assortative mixing by degree might influence certain aspects of language related processing.

Note that there have been many studies on Menzerath's law, Martin's law, and other relationships among words in the language, such as the general relationships observed about word frequency (e.g., Baayen, 1991, 2001, 2010; Zipf, 1935), but most of the previous studies of these statistical relationships attempted to determine the origin of the global pattern observed in the language. To be clear, the goal of the present work is *not* to determine the origin of assortative mixing by degree in the phonological lexicon, or to propose a model that could generate such a macro-level pattern in the language (for such work see the stochastic model described in Baayen (1991)). Instead, we take the observations of Arbesman et al. (2010b) as a given: assortative mixing by degree exists in the mental lexicon. The goal of the present research is to determine if this statistical relationship observed at the macro-level of the lexicon influences cognitive processing in some way.

Furthermore, given the work of Keller (2005) and others, we caution against the practice of "inverse inference," that is, inferring from an observed pattern in the data back to the model that might have generated it. Keller (2005) criticized the once-common practice in network science of observing a power-law degree distribution (*N.B.* Zipfian distributions of word frequency follow a power law) and then inferring that the network was generated by a particular mechanism (*i.e.*, growth with preferential attachment in the case of scale-free networks). To illustrate the problems of inversely inferring the mechanism that produced the observed distribution, Keller discussed the work of Herbert Simon and a number of others that shows that such power-law distributions can be produced by a large number of algorithms/mechanisms. Given that there are often many ways to produce a particular pattern in the data, and often no way to discern which of those mechanisms is actually responsible for the observed data, the field of network science has essentially abandoned the practice of inferring the generating mechanism when a scale-free degree distribution is observed.

Again, regardless of how assortative mixing by degree came to be in the lexicon, the goal of the present research is to determine if this statistical relationship observed at the macro-level of the lexicon influences cognitive processing in some way. We begin by considering how the pattern of mixing influences processing in other domains. Mathematical simulations suggest that the overall pattern of mixing exhibited in a network has implications for the ability of the system to maintain processing in the face of damage to the network, a concept somewhat reminiscent of graceful degradation in cognitive science. 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.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

More directly related to language, Arbesman et al. (2010b) examined network resilience in response to simulated attacks on nodes in a network of English words. Recall that English and a variety of other languages exhibited relatively high values of assortative mixing by degree (.5–.8 in the language networks examined by Arbesman et al., whereas .1–.3 is typically observed in social networks). Arbesman et al. observed similar and high levels of resilience in connectivity in the language network when either a random attack or an attack targeting highly connected nodes was carried out. This pattern of network resiliency differs from that typically seen in other networks, which tend to be resilient to random attacks on the network, but eventually succumb to failure when highly connected nodes are targeted for removal (e.g., Albert et al., 2000; Newman, 2002). Given the high levels of assortative mixing by degree, and high levels of resiliency to both random and targeted attacks observed in the language networks examined by Arbesman et al. (2010b), it is not unreasonable to suggest that the assortative mixing by degree found in the mental lexicon may contribute to the resilience of some language- related processes.

While it is unclear what the equivalent of a targeted attack on a phonological network might be in real-life, or how a word (*i.e*., node) could be *permanently* removed from the network, there are instances, as in the tip-of-the-tongue phenomenon (e.g., Brown & McNeill, 1966) and in certain forms of aphasia, in which words are temporarily “unavailable” during lexical retrieval. Given the influence that assortative mixing by degree has in maintaining network integrity, and the large values of assortative mixing by degree observed in several languages, we reasoned that the influence that assortative mixing by degree might have on lexical processing might be more easily observed when lexical retrieval *failed*.

Examining the influence of assortative mixing by degree during failed lexical retrieval complements previous studies that examined the influence of other network characteristics on processing in a number of ways. First, previous studies focused on micro-level measures of the network (e.g., degree, clustering coefficient, closeness centrality of individual nodes), whereas the present study focuses on the macro-level characteristic of mixing. Second, previous studies focused predominantly on quickly retrieving (or navigating to) a desired lexical item, whereas in the present study we focus on instances of failed lexical retrieval (a topic that is, as described below, relatively less examined in psycholinguistics). Finally, previous studies and network analyses have, in many cases, simply reported statistical measures of the network or the language in general. In the present case, we, like a few notable exceptions in the previous literature, are examining how an observed structure in the mental lexicon might influence cognitive processing (see Borge-Holthoefer and Arenas (2010) for the importance of tying measures of language networks to cognitive processing).

Investigations of various types of lexical retrieval failures in the area of speech *production*—such as slips of the tongue, malapropisms, and tip-of-the-tongue experiences—played an important role in increasing our understanding of the process of speech production. Arguably without the pioneering work of Fromkin (1971), Fay and Cutler (1977), Brown and McNeill (1966) and others, including work examining on-line processing with reaction time based tasks (Levelt, Roelofs, & Meyer, 1999), our understanding of the process of speech production would have been significantly impeded. Although investigations of speech production errors continue to play a crucial role in increasing our understanding of speech production, comparatively less work has investigated errors during speech *perception*. Like speech production errors, perceptual errors of speech—such as mondegreens and slips of the ear—can be examined scientifically rather than anecdotally (see the pioneering work summarized in Bond (1999)). As in the case of models of speech production, perceptual speech errors have the potential to inform models of speech perception and spoken word recognition.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Although models of spoken word recognition have existed for several decades, and some models have undergone significant revisions in that time, none of the widely-accepted models of spoken word recognition have been used to predict what will happen when lexical retrieval fails (*NAM*: Luce & Pisoni, 1998; *PARSYN*: Luce, Goldinger, Auer, & Vitevitch, 2000; *Shortlist*: Norris, 1994; Norris & McQueen, 2008; *Cohort*: Gaskell & Marslen-Wilson, 1997; Marslen-Wilson, 1987; *TRACE:* McClelland & Elman, 1986). Given the basic assumptions of these models—multiple wordforms that resemble the acoustic–phonetic input are activated and then compete with each other for recognition—how the system recovered from failed lexical retrieval might have appeared so obvious as to not require comment: one of the other partially-activated competitors will be retrieved if the desired target word cannot, for some reason, be retrieved.

Alternatively, the design characteristics of many models of spoken word recognition are such that the desired target word wins the competition for recognition (e.g., *Shortlist*: Norris, 1994; Norris & McQueen, 2008; *Cohort*: Gaskell & Marslen-Wilson, 1997; Marslen-Wilson, 1987; *TRACE:* McClelland & Elman, 1986). That is, models of spoken word recognition aren't designed to make errors.[2] A similar state of affairs is found in models of spoken word production where one class of models addresses chronometric aspects of speech production, but don't make any errors or predictions about speech errors (e.g., Levelt, Roelofs, & Meyer, 1999), and another class of models addresses speech production errors, but don't make predictions regarding chronometric aspects of speech production (e.g., Dell, 1986, 1988, etc.).

Regardless of the reason for the silence on this matter, the lack of attention to recognition errors is unfortunate, because, as Norman (1981, p. 13) reminds us: "By examining errors, we are forced to demonstrate that our theoretical ideas can have some relevance to real behavior." To demonstrate that the theoretical idea we put forward—the macro-level structure of the mental lexicon influences language-related processes—has some relevance to real behavior, we examined (simulated and real) instances in which lexical retrieval failed for a "cognitive footprint" (Waller & Zimbelman, 2003), or indirect evidence, of assortative mixing by degree in the mental lexicon. If the macro-level structure of the lexicon observed by Vitevitch (2008)—a positive correlation between the degree of a given word and the degree of its neighbors—has consequences for human behavior, then, in the context of failed lexical retrieval, we would expect to observe a positive correlation between the degree of the target word and the degree of the "erroneously" retrieved word. In the present study we used computer simulations, a corpus analysis of actual speech perception errors, and several laboratory-based tasks that captured certain relevant aspects of failed lexical retrieval to look for a "cognitive footprint" of assortative mixing by degree.

The focus of the present study on demonstrating that the macro-structure of the phonological lexicon, as measured by assortative mixing by degree, influences cognitive processing also contrasts with many previous studies that observed various statistical relationships among words and in language more generally. Consider the correlation that exists between the frequency with which a given word occurs, and the number of phonological neighbors that the given word has: high frequency words tend to have many phonological neighbors, whereas low frequency words tend to have few phonological neighbors (e.g., Frauenfelder et al., 1993; Landauer & Streeter, 1973). This relationship is found between two different lexical characteristics in the same word, whereas we examined a relationship between the same lexical characteristic (i.e., degree/neighborhood density) in two different words (albeit the words were phonological neighbors of each other). Less trivial, the present study

[2] Note that the neighborhood activation model (*NAM*; Luce & Pisoni, 1998) produces as outputs probabilities of numerous word-forms being retrieved. Therefore, it could be used to account for failed lexical retrieval in a straightforward manner.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

examined how this relationship might influence cognitive processing. There have been few (if any) studies that have demonstrated an influence of global patterns observed in language—such as the various relationships observed by Zipf (1935)—on cognitive processing, making the present study a significant contribution to the field of cognitive psychology rather than simply a quantitative observation about the constituents of language.

2. Simulation 1 using jTRACE

Lewandowsky (1993) described several rewards (as well as some hazards) associated with using computer simulations to explore new ideas in cognitive psychology. Among the benefits of using computer simulations is that they provide a low-cost way to examine novel predictions about human behavior. As a preliminary, low-cost test of the hypothesis that assortative mixing by degree influences some aspect of lexical processing we conducted a simulation using jTRACE (Strauss, Harris, & Magnuson, 2007), an updated implementation of the TRACE model of spoken word recognition (McClelland & Elman, 1986). jTRACE/TRACE shares many features with (but does have important differences from) other widely-accepted models of spoken word recognition (e.g., Shortlist, NAM, cohort, etc.).

Because the design characteristics of jTRACE make actual retrieval errors rare, if not impossible, we "simulated" a failure in lexical retrieval by presenting the model with a target item, and took the next most active lexical item in the set of competitors as the retrieval error. We recognize that taking the next-most active competitor in jTRACE to be the item that would be erroneously retrieved if lexical retrieval failed may be a less than ideal method of simulating failed lexical retrieval in the model. However, in the absence of a computational model of speech perception errors—the word recognition equivalent of models of speech production errors like those of Dell (1986, 1988)—we found this method to be acceptable as a preliminary test of our hypothesis. The outcome of this preliminary test could help us determine if further investigation of psycholinguistic behavior in human participants might be warranted.

Note that our decision to use jTRACE should not be seen as an endorsement of the TRACE model (McClelland & Elman, 1986). Indeed, our previous work demonstrated that jTRACE/TRACE (and Shortlist) could not account for the influence of clustering coefficient on spoken word recognition (Chan & Vitevitch, 2009). Rather, our decision to use this computer model in our simulation was based solely on practical concerns: it was the only computer model of spoken word recognition that was available in a ready-to-use format, with an easy-to-use interface, and it would enable us to test our hypothesis in a low-cost setting. Therefore, we wish to set aside a number of theoretical issues that the psycholinguistics literature has (hotly) debated over the years—the existence of feedback connections between levels of representation, the required number of levels of representation, the nature of those representations, etc.—and instead focus on the advantages of using computational models to explore novel hypotheses, as suggested by Lewandowsky (1993).

If the macro-level structure of the phonological network observed by Vitevitch (2008; see also Arbesman et al., 2010b) influences lexical processing, then we should see evidence of this influence in the output of a model of spoken word recognition that captures certain important aspects of human behavior. Specifically, we should observe a positive correlation between the degree/neighborhood density of the target word and the degree/neighborhood density of the erroneously retrieved word (*i.e.*, the next most active lexical item in jTRACE).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

2.1. Method

2.1.1. Stimuli—We used the 28 words with high clustering coefficient and the 28 words with low clustering coefficient used in the jTRACE simulation in Chan and Vitevitch (2009). These items consisted of monosyllabic words with 3 phonemes found in initial_lexicon in jTRACE. All of the words were found in initial_lexicon (one of the lexicons available for use in jTRACE) and were also comparable in neighborhood density, as reported in Chan and Vitevitch (2009). Given the previous failure of jTRACE to produce processing differences in these words in an examination of clustering coefficient (Chan & Vitevitch, 2009), we reasoned that these words would provide the model with a significant challenge in the present simulation investigating the influence of assortative mixing by degree on processing. If effects of assortative mixing by degree are observed on processing in the present simulation, then our confidence that similar results might be observed in human behavior is increased.

2.1.2. Procedure—The default parameter settings of the jTRACE computer model (Strauss et al., 2007) were used, including use of initial_lexicon as the lexicon. Importantly, the network structure of the words in initial_ lexicon exhibits the same characteristics as those observed in Vitevitch (2008) for a larger sample of English words; this is not surprising because initial_lexicon was designed to reflect key characteristics of the English language. As in Chan and Vitevitch (2009), we allowed the target words to reach maximal activation. In the present simulation we then identified the next-most active item from the set of lexical competitors. We assumed that the second-most active item would be the item retrieved from the lexicon if lexical retrieval of the target word failed. The target words and the next most active items (i.e., the "perceptual error") are listed in Appendix B.

2.2. Results and discussion

If the macro-structure of the mental lexicon (as assessed by the network science measure of assortative mixing by degree) influences lexical processing, then we should see evidence of this influence in the output of jTRACE. Specifically, we should observe a positive correlation between the degree/neighborhood density of the target word and the degree/neighborhood density of the erroneously retrieved word (*i.e.*, the next most active lexical item in jTRACE). As described by Newman (2002, p. 1), "…the Pearson correlation coefficient of the degrees at either ends of an edge…." is the statistic used to assess mixing by degree in the network science literature. We found a significant, positive relationship ($r(56) = +.46$, $p < .001$) between the degree of the target items (*mean* = 5.28 neighbors) and the degree of the next most active lexical item in jTRACE (*mean* = 4.68 neighbors).

Note that some of the words ($n = 10$) that were the next most active lexical item in jTRACE differed from the target word by more than a single phoneme. Given the relatively small number of words in initial_lexicon compared to the lexicon of the average human language user, it is not surprising that such words are included in the set of lexical competitors (*i.e.*, these words are the closest matches in the initial_lexicon, but not necessarily the closest matches in a real lexicon). Such words would not be directly connected to the target word in the network model constructed in Vitevitch (2008). Therefore, to make our test of the assortative mixing by degree hypothesis comparable to the network analysis of Vitevitch (2008) we again correlated the degree of the target item to the degree of the next most active lexical item in jTRACE, but restricted the analysis to next-most-active items that differed from the target word by the addition, deletion, or substitution of a single phoneme (i.e., there is a link between these words in the phonological network in Vitevitch, 2008). We again found a significant, positive relationship ($r(46) = +.44$, $p < .01$) between the degree of the target items (*mean* = 5.63 neighbors) and the degree of the next most active lexical item in jTRACE (*mean* = 5.34 neighbors).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Thus, the output of a computer simulation that captures certain important aspects of human psycholinguistic behavior does indeed show a "cognitive footprint" for the network feature of assortative mixing by degree. If the network feature of assortative mixing by degree observed in a network analysis by Vitevitch (2008) was simply a spurious mathematical phenomenon, or resulted only from the way in which the network was constructed, and did not have any consequences for psycholinguistic behavior, then there should be no evidence of this relationship among word-forms in the output of a computer simulation of psycholinguistic behavior. Observing evidence for assortative mixing by degree when lexical retrieval "fails" suggests that the structure of the phonological lexicon—as observed with tools from network science—influences lexical processing.

3. Simulation 2 using jTRACE and lexicons with different types of mixing by degree

Another advantage of computer simulations is that they can be used in "computational experiments" to explore questions that are difficult—for ethical or practical reasons—to examine in the real world (Plunkett & Elman, 1997; see also Vitevitch & Storkel, 2012). In the present simulation we further explored how the structure of the lexicon might influence (failed) lexical retrieval by manipulating mixing by degree in the lexicon. That is, we created a lexicon that exhibited assortative mixing by degree, a lexicon that exhibited disassortative mixing by degree (where words with high degree tend to connect to words with low degree, and vice versa; a negative correlation in degree), and a lexicon that exhibited no mixing by degree (zero correlation in degree, indicating that it is equally likely that a word with high degree connects to a word with low degree or to a word with high degree, and vice versa).

It would be difficult to carry out this manipulation in a controlled laboratory setting with human language-users using a conventional psycholinguistic task because assortative mixing by degree appears to be a general characteristic of phonological word-forms in real languages (Arbesman et al., 2010b). That means that an artificial vocabulary would need to be constructed with the characteristics that we wish to examine, and participants would need to be trained on this new artificial vocabulary before we could test for the influence of these characteristics on processing; a difficult and time-consuming endeavor, indeed.

We again assumed that the most active lexical competitor in jTRACE would be retrieved if lexical retrieval of the target item failed. If the macro-level structure of the lexicon indeed influences processing, then we expect to find a cognitive footprint of this influence in the output of the computational model. That is, the correlation between the degree of the target word and the degree of the next most active lexical item in the set of competitors should match the direction of the correlation between the degree of the target word and the degree of the neighbors in each lexicon that we constructed. Specifically, we should see a positive correlation in the output of the model for the lexicon that exhibits assortative mixing by degree, a negative correlation in the output of the model for the lexicon that exhibits disassortative mixing by degree, and zero correlation in the output of the model for the lexicon that does not exhibit any mixing by degree.

3.1. Method

3.1.1. Stimuli—Our focus in the present simulations was on the macro-level structure of the lexicon, so we simply used strings of phonemes to create items that—for the sake of convenience—we refer to as words (*N.B.* any resemblance of these specially created words to real words in English is strictly coincidental). All of the words in the specially created lexicons were three phonemes long, and used the 14 phoneme symbols available in jTRACE

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

(^, a, b, d, g, i, l, k, p, r, s, S, t, u). A word was considered a neighbor of another word if a single phoneme could be substituted into the word to form another word that appeared in the lexicon. (Additions and deletions of phonemes were not included in the assessment of phonological neighbors in order to keep all of the words the same length.)

Three separate lexicons were created that varied in the type of mixing by degree. The lexicon that exhibited assortative mixing by degree contained 56 words, the lexicon that exhibited disassortative mixing by degree contained 56 words, and the lexicon that exhibited no mixing by degree contained 48 words. The small (and unequal) number of words in each lexicon is a result of the constraints of the number of phonemes in the model, from maintaining the same word-length, and creating neighbors (by substituting one phoneme in the target word to form a new word) that would exhibit the desired type of mixing by degree. See Appendix C for the words that constituted the three different lexicons. Correlational analyses of the degree of each word and the degree of the neighbors of each word confirmed that the three lexicons exhibited the desired type of mixing by degree. In the case of the lexicon that exhibited assortative mixing by degree, $r = +1.0$ ($p < .0001$). For the lexicon that exhibited disassortative mixing by degree, $r = -1.0$ ($p < .0001$), and for the lexicon that exhibited no mixing by degree, $r = 0$ ($p = 1$).

Although the relationship between the degree of the target and neighbors (i.e., mixing by degree) differed among the three lexicons, the three lexicons contained words that were comparable in degree/neighborhood density ($F(2,157) = 2.01$, $p = .14$). In the case of the lexicon that exhibited assortative mixing by degree, the mean number of neighbors per word was 1.7 (*sd* = .46). For the lexicon that exhibited disassortative mixing by degree, the mean number of neighbors per word was 1.5 (*sd* = .87). For the lexicon that exhibited no mixing by degree, the mean number of neighbors per word was 1.5 (*sd* = .51).

We recognize that the small number of words and the few neighbors that each word has in the current lexicons might be considered a “simplified” lexicon, especially when compared to the size of the lexicon used in Simulation 1 or to the lexicon of a typical human. However, it is important to keep in mind that “[m]odels are not intended to capture fully the processes they attempt to elucidate. Rather, they are explorations of ideas about the nature of cognitive processes. In these explorations, simplification is essential—through simplification, the implications of the central ideas become more transparent” (McClelland, 2009, p. 11). The “simplified” lexicons used in the present simulation provided us with the opportunity to explore the idea of how the macro-level structure of a network—as measured by mixing by degree—might influence lexical processing.

3.1.2. Procedure—The default parameter settings of the jTRACE computer model (Strauss et al., 2007) were used in the present simulations, with the exception of using the specially created lexicons that varied in mixing by degree. Each word in each lexicon was presented as input. After 180 time slices the identity of the second-most active item in the set of lexical competitors was determined. We assumed that the second-most active item would be the item retrieved from the lexicon if lexical retrieval failed.

3.2. Results and discussion

As in Simulation 1, we again correlated the degree of the target word and the degree of the nextmost active word in the set of competitors to determine if there was a “cognitive footprint” in the output of the model indicative of the type of mixing by degree exhibited by the lexicon. If the macro-level structure of the lexicon influences (failed) lexical retrieval, then the correlation between the degree of the target word and the degree of the next most active word in the output of the model should be positive for the lexicon that exhibits

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

assortative mixing by degree, negative for the lexicon that exhibits disassortative mixing by degree, and zero for the lexicon that does not exhibit mixing by degree.

For the lexicon that exhibited assortative mixing by degree, the correlation between the degree of the target word and the degree of the next most active lexical item in the set of competitors was $r(54) = +.91$ ($p < .0001$). For the lexicon that exhibited disassortative mixing by degree, the correlation between the degree of the target word and the degree of the next most active lexical item in the set of competitors was $r(54) = -.75$ ($p < .0001$). Note that in three instances in the disassortative lexicon (indicated in the appendix) the target word itself was the second-most active item. In those instances, we used the (incorrect) item with the highest activation as the item that would be retrieved if lexical retrieval failed. For the lexicon that did not exhibit mixing by degree, the correlation between the degree of the target word and the degree of the next most active lexical item in the set of competitors was $r(46) = 0.0$ ($p = 1$).

The results of Simulation 1 and 2 indicate that the type of mixing by degree exhibited by the phonological lexicon influences certain aspects of lexical processing, suggesting that the structure of the mental lexicon influences lexical processing. Because the computer model used in these simulations, jTRACE, captures certain relevant aspects of spoken word recognition behavior in humans we are encouraged to look for a cognitive footprint of assortative mixing by degree in human behavior. If the network feature of assortative mixing by degree is not simply a mathematical artifact of the network analysis performed in Vitevitch (2008), and the type of mixing observed in the English phonological lexicon influences processing, then we should be able to observe evidence of assortative mixing by degree when lexical retrieval "fails" in humans. We first analyzed a corpus of speech perception errors, that is, actual, attested failures in lexical retrieval, for a cognitive footprint of assortative mixing by degree in human behavior. We then used several psycholinguistic tasks to create situations in the laboratory that captured certain relevant aspects of "failed" lexical retrieval to look further for a cognitive footprint of assortative mixing by degree in human behavior.

4. Slip of the ear analysis

Although analyses of naturally occurring speech errors laid the foundation for understanding the process of speech production (Fromkin, 1971), our understanding of the process of word recognition has been guided primarily by evidence obtained from experiments employing reaction-time methodologies (e.g., priming, lexical decision, and naming tasks) rather than from naturally occurring errors. Speech perception errors do exist, and are perhaps better known as *slips of the ear*. A slip of the ear is a misperception of a correctly produced utterance. Slips of the ear should not be confused with slips of the tongue (*i.e.*, speech *production* errors), where the speaker intends to say one thing but erroneously produces something else. In slips of the ear, the utterance is produced correctly (*i.e.*, as intended), but the perceiver "hears" something else. An example (from Bond, 1999) of a slip of the ear is when a speaker says, "Stir this!" but the listener perceives the utterance as "Store this!"

As described in Bond (1999), the misperceptions that typically occur in slips of the ear involve misperceiving a vowel or consonant in a word (resulting in the erroneous perception of another word), the addition or deletion of a syllable in a word, or the mis-segmentation of words in a phrase (resulting in the perception of more or fewer words than was produced). Although there are likely to be many factors that contribute to slips of the ear, certain errors have not been observed, like hearing *cat* instead of *dog*, suggesting that semantic information does not strongly contribute to the initial stages of speech perception or to

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

misperceptions; this contrasts with speech *production* errors where such substitutions have been observed (Fromkin, 1971).

Analyses of slips of the ear often provide evidence that is consistent with and complements the findings obtained from laboratory-based experiments (see the work of Bond (1999), Felty, Buchwald, Gruenenfelder, and Pisoni (2013), and others). For example, Vitevitch (2002b) found that the target word in 88 slips of the ear (obtained from the materials in the appendix of Bond (1999)) had denser phonological neighborhoods and higher neighborhood frequency than words comparable in length and word class that were randomly selected from the lexicon. This result—in accord with predictions derived from the neighborhood activation model and tested with several conventional psycholinguistic tasks in the laboratory (Luce & Pisoni, 1998)—suggests that words with dense neighborhoods and high neighborhood frequency are more difficult to perceive than words with sparse neighborhoods and low neighborhood frequency (hence their disproportionate appearance in slips of the ear corpora). In the present analysis we examined in a different way the same 88 slips of the ear examined in Vitevitch (2002b) for a cognitive footprint of assortative mixing by degree in human performance.

4.1. Methods

The items examined in Vitevitch (2002b) came from the Bond (1999) corpus and consisted of misperceptions made by adults that were classified by Bond as misperceptions of vowels or of consonants. This led to the exclusion of “complex errors” and parsing errors that resulted in “extensive mismatch between utterance and perception.” Also excluded were proper nouns, foreign words or phrases, acronyms or auditory spellings, and domain-specific technical terms.

4.2. Results and discussion

In the present analysis we examined the neighborhood density (i.e., degree) of the target words (*mean* = 12.3 neighbors, *sd* = 7.8) and the neighborhood density of the erroneously perceived words (*mean* = 13.1 neighbors, *sd* = 8.2). When the neighborhood density/degree of the 88 spoken words was correlated with the neighborhood density/degree of the corresponding misperceived words, a significant, positive correlation was observed ($r(88) = +.68$, $p < .0001$). This result suggests that naturally occurring instances of failed lexical retrieval contain a cognitive footprint of the network structure known as assortative mixing by degree that was observed in a network analysis of the English phonological lexicon (Vitevitch, 2008), indicating further that the structure of the mental lexicon influences certain language-related processes.

Observing that assortative mixing by degree influences human behavior is neither a trivial result, nor a foregone outcome given the observation of assortative mixing by degree in the phonological network examined in Vitevitch (2008). One need only look at the other studies reported in Vitevitch (2002b)—where the slips of the ear used in the present analysis were initially analyzed (see also Bond, 1999)—for evidence that human behavior doesn’t always conform to the predictions of psychological theory or to statistical distributions in the language. In addition to the findings regarding neighborhood density and neighborhood frequency summarized above in Section 4, Vitevitch (2002b) found that the slip of the ear tokens were higher in frequency of occurrence than words in general. This observation was unexpected given the predictions of models of spoken word recognition, which indicate that words with low frequency of occurrence are difficult to perceive, and would therefore lead to the prediction that slips of the ear should be low in frequency of occurrence rather than high as was observed in Vitevitch (2002b).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Furthermore, Zipf (1935) observed that there are many words in the language that are relatively low in frequency of occurrence, and few words in the language that are relatively high in frequency of occurrence. The statistical distribution of the number of words with a given word-frequency in the language would further lead to the prediction that, based on the principles of statistical sampling, it is most likely that slips of the ear would be words that are low in frequency of occurrence, instead of high in frequency of occurrence as was observed in Vitevitch (2002b). We note that the perceptual identification experiment and lexical decision experiment reported in Vitevitch (2002b) provided evidence that the counter-intuitive finding regarding word-frequency observed in the same study might have been due to the more rapid production of high frequency words compared to low frequency words that typically occurs in naturalistic settings (see also Wright, 1979). We further note that the counter-intuitive finding of word frequency in slips of the ear serves as an example that human behavior does not always conform to psychological theory or to statistical distributions in the language, making the present observation of an influence of assortative mixing by degree on slips of the ear a significant finding that provides important evidence that the structure of the mental lexicon influences certain language-related processes.

The combination of naturalistic observation and laboratory-based methods in Vitevitch (2002b) required to resolve the counter-intuitive finding of word frequency in slips of the ear emphasized to us the importance of using multiple methods to examine a research question. Like all research methodologies, naturalistic observations such as naturally occurring speech errors have certain strengths and weaknesses (see discussion in Bond (1999)). Rather than rely solely on this source of human behavioral data to examine how assortative mixing by degree might influence lexical retrieval processes, we conducted three psycholinguistic experiments. Given the role that assortative mixing by degree plays in network resilience, we used psycholinguistic tasks that captured certain relevant aspects of “failed” lexical retrieval. Like all laboratory-based tasks, the psycholinguistic tasks that we used are somewhat contrived, and performance in each of them is likely to be influenced by several factors. Nevertheless, these tasks may allow us to glimpse a cognitive footprint of the network structure of the mental lexicon. While each of these tasks individually provides limited information, together the results of these tasks offer a more complete view of how the structure of the mental lexicon might influence processing.

5. Experiment 1: Perceptual identification task

In the perceptual identification task, participants hear a word mixed with noise and type or say the word that was heard. Pisoni (1996) provided a concise review of the perceptual identification task and its use. He noted that differences in signal-to-noise ratios, speaking rate, response format (open- versus closed-set), the types of noise used (e.g., white, pink, envelope-shaped, etc.), and a number of other factors influence the responses made in this task. In addition, responses in this task are thought to reflect both bottom-up processing of the acoustic–phonetic input, as well as top-down processing strategies that may be employed in this “off-line” task (meaning that a response deadline is not typically imposed on participants). We acknowledge these concerns regarding the perceptual identification task, but also recognize certain benefits of this task that make it well-suited to our present need as a proxy of failed lexical retrieval. Indeed, we are not the only ones to recognize that the perceptual identification task captures certain relevant aspects of failed lexical retrieval; Felty et al. (2013) used this task to make a corpus of speech perception errors collected under controlled laboratory conditions.

In the version of the perceptual identification task used in the present experiment, participants heard real English words embedded in white noise at a signal-to-noise ratio that would lead to a reasonable number of erroneous responses (S/N = +5). Although overall

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

accuracy of responses is typically assessed in this task, we instead analyzed the incorrect responses, which, by definition, are instances of failed lexical retrieval. Lower signal-to-noise ratios would increase the number of erroneous responses, but the increased noise would also likely discourage the participants in the task, and perhaps lead to idiosyncratic and task-specific strategies that would influence task performance. Higher signalto- noise ratios would likely lead to performance that approached "ceiling," leaving us a small number of errors to examine. Therefore, we used a signal-to-noise ratio that would provide us with a reasonable number of "genuine" erroneous responses.

Recall that *degree* in network science terminology is equivalent to *phonological neighborhood density* in psycholinguistics. It is well-known that phonological neighborhood density is correlated with a number of other factors, like word frequency, phonotactic probability, etc., that also influence spoken word recognition (Frauenfelder et al., 1993; Landauer & Streeter, 1973). Also recall the over-representation of words with high neighborhood density observed by Vitevitch (2002b) in a corpus of speech errors. To take advantage of the control afforded by laboratory conditions to minimize the potential influence of other factors known to influence spoken word recognition, and to examine a broader range of words from the mental lexicon, we selected an equal number of words that had relatively high neighborhood density and that had relatively low neighborhood density (Luce & Pisoni, 1998), but were similar in terms of several other lexical characteristics (see Section 5.1.2 for details). We then correlated the neighborhood density of the *erroneous responses* to the neighborhood density of the target word that was embedded in noise. If the macro-level characteristic of the mental lexicon known as assortative mixing by degree influences processing, then we should observe in the perceptual identification task a correlation between the neighborhood density of the erroneous response and the neighborhood density of the target word.

5.1. Method

5.1.1. Participants—Twelve native English-speaking students enrolled at the University of Kansas gave their written consent to participate in the present experiment. None of the participants reported a history of speech or hearing disorders, nor participated in the other experiments reported here.

5.1.2. Materials—One hundred English monosyllabic words containing three phonemes in a consonant–vowel–consonant syllable structure were used as stimuli in this experiment (and are listed in Appendix D). A male native speaker of American English (the first author) produced all of the stimuli by speaking at a normal speaking rate and loudness in an IAC sound attenuated booth into a high-quality microphone. Isolated words were recorded digitally at a sampling rate of 44.1 kHz. The pronunciation of each word was verified for correctness. Each stimulus word was edited into an individual sound file using SoundEdit 16 (Macromedia, Inc.). The amplitude of the individual sound files was increased to their maximum without distorting the sound or changing the pitch of the words by using the Normalization function in SoundEdit 16. The same program was used to degrade the stimuli by adding white noise equal in duration to the sound file. The white noise was 5 dB less in amplitude than the mean amplitude of the sound files. Thus, the resulting stimuli were presented at a +5 dB signal to noise ratio (S/N).

Neighborhood density refers to the number of words that sound similar to the stimulus word based on the addition, deletion or substitution of a single phoneme in that word (Luce & Pisoni, 1998). A word like *cat*, which has many neighbors (e.g., *at, bat, mat, rat, scat, pat, sat, vat, cab, cad, calf, cash, cap, can, cot, kit, cut, coat*), is said to have a *dense phonological neighborhood*, whereas a word, like *dog*, that has few neighbors (e.g., *dig,*

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

dug, dot, fog) is said to have a *sparse phonological neighborhood* (*N.B.*, each word has additional neighbors, but only a few were listed for illustrative purposes). Half of the stimuli had dense phonological neighborhoods (*mean* = 27.7 neighbors, *sd* = 1.6), and the remaining stimuli had sparse phonological neighborhoods (*mean* = 14.9 neighbors, *sd* = 1.5; $F(1,98) = 1648.62$, $p < .0001$). Although the stimuli differed in neighborhood density, they were comparable on a number of other characteristics as described below.

Subjective familiarity: Subjective familiarity was measured on a seven-point scale (Nusbaum, Pisoni, & Davis, 1984). Words with dense neighborhoods had a mean familiarity value of 6.87 (*sd* = .22) and words with sparse neighborhoods had a mean familiarity value of 6.82 (*sd* = .28, $F(1,98) = 1.50$, $p = .22$). The mean familiarity value for the words in the two groups indicates that all of the words were highly familiar.

Frequency of occurrence in the language: Average log word frequency ($\log_{10}$ of the raw values from Kučera and Francis (1967)) was 1.03 (*sd* = .58) for the words with dense neighborhoods, and 1.00 (*sd* = .58) for the words with sparse neighborhoods ($F(1,98) = .08$, $p = .77$).

Neighborhood frequency: Neighborhood frequency is the mean word frequency of the neighbors of the target word. Words with dense neighborhoods had a mean log neighborhood frequency value of 2.03 (*sd* = .24), and words with sparse neighborhoods had a mean log neighborhood frequency value of 1.94 (*sd* = .25; $F(1,98) = 2.99$, $p = .09$).

Phonotactic probability: The phonotactic probability was measured by how often a certain segment occurs in a certain position in a word (*positional segment frequency*) and by the segment-to-segment co-occurrence probability (*biphone frequency*; as in Vitevitch and Luce (2005)). The mean positional segment frequency for words with dense neighborhoods was .147 (*sd* = .02) and .140 (*sd* = .02, $F(1,98) = 2.11$, $p = .15$), respectively. The mean biphone frequency for words with dense neighborhoods was .007 (*sd* = .003) and for words with sparse neighborhoods was .007 (*sd* = .003, $F(1,98) = .009$, $p = .93$). These values were obtained from the web-based calculator described in Vitevitch and Luce (2004).

5.1.3. Procedure—Participants were tested individually. Each participant was seated in front of an iMac computer running PsyScope 1.2.2 (Cohen, MacWhinney, Flatt, & Provost, 1993), which controlled the presentation of stimuli and the collection of responses. In each trial, the word "READY" appeared on the computer screen for 500 ms. Participants then heard one of the randomly selected stimulus words imbedded in white noise through a set of Beyerdynamic DT 100 headphones at a comfortable listening level. Each stimulus was presented only once. The participants were instructed to use the computer keyboard to enter their response (or their best guess) for each word they heard over the headphones. They were instructed to type "?" if they were absolutely unable to identify the word. The participants could use as much time as they needed to respond. Participants were able to see their responses on the computer screen when they were typing and could make corrections to their responses before they hit the RETURN key, which initiated the next trial. The experiment lasted about 15 min. Prior to the experiment, each participant received five practice trials to become familiar with the task. These practice trials were not included in the data analyses.

5.2. Results and discussion

A response was scored as correct if the phonological transcription of the response matched the phonological transcription of the stimulus. Misspelled words and typographical errors in the responses were scored as correct responses in certain conditions: (1) adjacent letters in

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

the word were transposed, (2) the omission of a letter in a word was scored as a correct response only if the response did not form another English word, or (3) the addition of a single letter in the word was scored as a correct response if the letter was within one key of the target letter on the keyboard. Of the 1200 responses, 328 were correct (27%) and 14 responses (1%) received the response of "absolutely unable to identify" (i.e., "?"); these responses could not, of course, be analyzed further.

Turning to the 858 incorrect responses (72% of the 1200 responses), 398 (46% of the incorrect responses) were phonological neighbors of the target words based on the one phoneme metric used in Luce and Pisoni (1998), and 460 (54% of the incorrect responses) were not phonological neighbors of the target words. Given the way in which assortative mixing by degree was calculated in the network analysis of Vitevitch (2008), we further analyzed only the incorrect responses that were phonological neighbors of the target words (based on the metric used in Luce & Pisoni, 1998). We recognize that the remaining incorrect responses might be useful for developing an account of all types of perceptual errors, for developing an algorithm that would predict the likelihood of a misperception occurring, etc. Such questions are interesting, but they are not the focus of the present investigation and do not directly address the question at hand: can we find in human language behavior a cognitive footprint of assortative mixing by degree? Only the incorrect responses that are phonological neighbors of the target words allow us to directly address that question. Furthermore, the same restriction was used in the analysis of actual speech perception errors and in the jTRACE simulations. Using the same restriction in each analysis enables us to equitably compare the results across research methods.

To look for a trace of assortative mixing by degree in the responses in the perceptual identification task we correlated the neighborhood density of the target word ($n = 87$, *mean* = 21.06, $sd = 6.61$) to the mean neighborhood density of the incorrect responses to that target ($n = 87$, *mean* = 22.36, $sd = 6.22$; note that all of the participants correctly identified 13 words reducing n in the above analyses to 87 instead of 100). We observed a positive and statistically significant correlation between degree/neighborhood density of the targets and the mean of the incorrect responses that were neighbors of the targets ($r(85) = +.22$, $p < .01$).

The magnitude of the correlation in the present experiment is smaller than that observed in the network analysis reported in Vitevitch (2008), and in the analysis of slips of the ear in Section 4.2 above. This should not be surprising given that we selected an equal number of words with dense phonological neighborhoods (i.e., high degree) and with sparse phonological neighborhoods (i.e., low degree) to use as stimuli in the perceptual identification experiment. Recall that the slips of the ear analysis showed that misperceptions typically occur for words with dense phonological neighborhoods (i.e., high degree), so inclusion of items with sparse phonological neighborhoods/low degree may have reduced the overall magnitude of the present correlation.

Furthermore, the perceptual identification task captures certain important aspects of what happens when lexical retrieval fails, but it is unlikely that this (or any) laboratory-based task provides a perfect analog of cognitive processes used in the real world, contributing further to the reduced magnitude of the present correlation. This, too, should not be surprising, because "…the laboratory setting is not necessarily representative of the social world within which people act. As a consequence, we cannot use laboratory findings to estimate the likelihood that a certain class of responses will occur in naturalistic situations…. Experiments are not conducted to yield such an estimate." (Berkowitz & Donnerstein, 1982, p. 255). Rather, the present experiment (like most experiments) was conducted to demonstrate that changes in variable X *can* lead to changes in variable Y.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

The result observed in the present experiment indicates that listeners who incorrectly perceived a neighbor of a target word with a dense phonological neighborhood (mixed with noise) tended to respond with a word that also had a dense phonological neighborhood. Concomitantly, listeners who incorrectly perceived a neighbor of a target word with a sparse phonological neighborhood (mixed with noise) tended to respond with a word that also had a sparse phonological neighborhood. This result provides evidence of a cognitive footprint of assortative mixing by degree observed in a network analysis of the mental lexicon by Vitevitch (2008), and also observed in the present report in a computer simulation and an analysis of a speech error corpus. This result further suggests that the structure of the mental lexicon influences cognitive processing, and that the tools of network science can be used to provide novel insights about those cognitive processes and representations.

6. Experiment 2: Similar sounding word task

The perceptual identification task used in Experiment 1 captured certain relevant aspects of failed lexical retrieval (see also Felty et al., 2013), allowing us to examine whether the macro-structure observed among word-forms in the phonological lexicon of English (*i.e.*, assortative mixing by degree) might influence in some way human language processing. However, we also recognize that the noise used in the perceptual identification task might differentially mask different phonemes, and might induce task-specific strategies in addition to the strategies typically used during spoken word recognition. Therefore, in the present experiment we used a task that did not employ degraded stimuli. In Experiment 2 of Luce and Large (2001), participants heard a nonword, and were asked to say the first real English word that came to mind that sounded like that nonword. For example, a participant might have heard/fin/ and responded with the real word *feet*. We reasoned that a modified version of the task—using as stimuli real English words instead of nonwords—might capture certain relevant aspects of failed lexical retrieval and allow us to see an influence of assortative mixing by degree on processing. We recognize that asking participants to respond with a word other than the one they heard might seem unnatural, but we believe this task is no more unnatural than other tasks used in psycholinguistics (e.g., semantic associate task, Nelson, McEvoy, & Schreiber, 1998).

Participants in the present experiment were presented with the same English words used in Experiment 1—this time without noise—and were asked to respond with the first English word that came to mind that sounded like the word they heard. Similar to the jTRACE simulation, the word that participants produced in this task was viewed as the word that would be retrieved when lexical retrieval failed. As in Experiment 1, we correlated the neighborhood density/degree of the target word (i.*e*., what was presented) to the mean neighborhood density/degree of responses that were phonological neighbors.

6.1. Method

6.1.1. Participants—Fourteen native English speaking students enrolled at the University of Kansas gave their written consent to participate in the present experiment. None of the participants reported a history of speech or hearing disorders.

6.1.2. Materials—The same stimuli used in Experiment 1 were used in the present experiment. The only exception being that the words were not mixed with white noise in the present experiment.

6.1.3. Procedure—The same equipment and procedure used in Experiment 1 were used in the present experiment, with the following exception. Instead of being asked to identify the degraded stimulus presented to them over a set of headphones, participants were asked to type in the first English word that came to mind that “sounded like” the word (presented

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

without noise) that they heard over a set of headphones. In addition, the option to indicate that they were "unable to identify the word" (i.e., "?") was not provided to participants.

6.2. Results and discussion

Misspelled words and typographical errors in the responses were corrected according to the following criteria: (1) transposition of adjacent letters in the word was corrected, and (2) the addition of a single letter in the word was removed if the letter was within one key of the target letter on the keyboard. Of the 1400 responses that were made, nonwords (2.8%), proper nouns or slang words (2.2%), responses that were identical to the stimulus (0.7%), words that differed from the target by more than 1 phoneme (15.1%), or that were semantically rather than phonologically related (4.7%) were not included in the analysis. The remaining 1043 (74.5%)[3] responses were phonological neighbors of the target words and analyzed further.

As in Experiment 1, we correlated the degree/neighborhood density of the target words (n = 99, *mean* = 21.39, *sd* = 6.61) to the mean degree/neighborhood density of the responses that were phonological neighbors of the target word (n = 99, *mean* = 21.87, *sd* = 4.63; note that none of the participants provided a response that was a neighbor to the word *doll*, reducing n in the above analyses to 99 instead of 100). We observed a positive and statistically significant correlation between neighborhood density/degree of the targets and responses ($r(97) = +.38$, $p < .01$). This result indicates that listeners who heard a word with a dense phonological neighborhood tended to respond with a similar sounding word that also had a dense phonological neighborhood. Concomitantly, listeners who heard a word with a sparse phonological neighborhood tended to respond with a similar sounding word that also had a sparse phonological neighborhood.

The result observed in the present experiment replicates the results observed in the perceptual identification task in Experiment 1, and in the slip of the ear analysis. Observing the same result across experimental tasks (perceptual identification and the similar word task), research methodologies (network analysis, psycholinguistic experiments, naturalistic observation, and computer simulation), and in different samples of words is reassuring. However, in Experiment 3 we wished to further examine the influence of assortative mixing by degree in human language behavior by using a less conventional task from psycholinguistics that again captured certain aspects of failed lexical retrieval.

7. Experiment 3: Verbal transformation illusion

To further examine how assortative mixing by degree might influence human language processing we used the *verbal transformation illusion* first described by Warren and Gregory (1958), and used more recently by Shoaf and Pitt (2002). In the verbal transformation illusion a listener hears a continuously repeating word (or phrase) that they perceive to change. As described in Warren (1968, p. 261), "[t]hese "verbal transformations" (VT's) range from a word that rhymes with the actual stimulus to extreme phonetic distortions."

One account of the verbal transformation illusion suggests that repeated stimulation "satiates" the representational unit corresponding to the repeated word (MacKay, Wulf, Yin,

[3] This proportion is comparable to the proportion of phonological neighbors obtained as responses in Experiment 2 of Luce and Large (2001). They found that about 70% of the real word responses to their nonword stimuli differed from the nonword by the *substitution* of a single phoneme in the onset, vowel, or final position. Furthermore, Luce and Large found that 27% of the changes took place in the initial consonant of the nonword, 10% took place in the vowel, and 34% took place in the final consonant. The distribution of responses across word positions in the present experiment was also similar to that observed by Luce and Large, suggesting that participants in the present experiment did not adopt the naïve strategy of simply producing words that rhymed with the stimulus.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

& Abrams, 1993; see also Warren, 1968)—analogous to the way a single neuron habituates with repeated stimulation—such that increasingly greater amounts of stimulation are required to activate that representational unit. In the presence of a constant level of stimulation (*i.e*., the continuous repetition of the word), it becomes more difficult to retrieve the representational unit corresponding to the repeated word. With the representational unit corresponding to the repeated word "temporarily removed" from the competition occurring among similar sounding word-forms, another representational unit (that sounds similar to the repeated word) will be retrieved instead of the representational unit corresponding to the repeated word. Although the verbal transformation illusion might be considered a somewhat unconventional psycholinguistic task, it creates in the controlled conditions of the laboratory a situation that captures certain aspects of failed lexical retrieval, and an opportunity for us to find a cognitive footprint of assortative mixing by degree.

Participants in the present experiment heard 350 repetitions of a word (with a 150 ms inter-stimulus- interval), and were asked to report the "new" word they "heard" whenever their percept changed. Listeners took part in four such trials (2 words had dense neighborhoods and 2 words had sparse neighborhoods). As in the analysis of the speech error corpus and the previous experiments, we correlated the neighborhood density of the word that was presented to the mean neighborhood density of the reported percepts that were phonological neighbors.

7.1. Methods

7.1.1. Participants—Forty-eight undergraduates were recruited from Introductory Psychology classes at the University of Kansas. They received partial credit towards a course requirement for participating. All participants reported no history of speech or hearing disorders, and all were native speakers of English.

7.1.2. Materials—The stimuli consisted of eight monosyllabic English words. Four words (*pail, loan, jail, peel*) had dense phonological neighborhoods (*mean* = 33.75; *sd* = .5) and four words (*poise, lodge, guess, peg*) had sparse phonological neighborhoods (*mean* = 9.0; *sd* = 2.71). These eight items were split into two lists of four words such that each participant received two words with high density and two words with low density. The words with high density and the words with low density that each participant received started with the same phoneme (e.g. *lodge* and *loan*). The words that participants heard were presented in a counterbalanced order such that all possible orders of the four stimuli were presented, resulting in a total of 48 different orderings. Each participant randomly received one of the different orderings.

The words chosen were matched on familiarity, log frequency, sum of phones, sum of biphones, and log neighborhood frequency. No significant differences were found between the dense and sparse words on these variables, *familiarity* ($F(1,6) = 2.49$, $p = .17$), *log frequency* ($F(1,6) = .04$, $p = .86$), *sum of the phones* ($F(1,6) = 1.46$, $p = .27$), *sum of the biphones* ($F(1,6) = .104$, $p = .76$), *log neighborhood frequency* ($F(1,6) = .12$, $p = .74$), nor *stimulus duration* of the recorded words ($F(1,6) = 1.96$, $p = .21$). See Table 1 for the mean and standard deviation values of these variables.

A male native speaker of American English (the third author) produced the words by speaking at a normal rate and volume into a high-quality microphone in an Industrial Acoustics Company sound attenuated booth. Individual words were recorded digitally at a sampling rate of 44.1 kHz. The pronunciation of each word was verified for correctness. Each stimulus word was edited with SoundEdit 16 (Macromedia, Inc.) into an individual sound file.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

7.1.3. Procedure—The same equipment used in the previous experiments was used in the present experiment. Participants in the present experiment were tested individually. They were instructed that they would hear a word repeated over a set of headphones, and that the word they heard might occasionally change to another word or to a nonsense word. The participants were informed that the changes could be subtle or very noticeable and that they were to report (out loud into a microphone placed in front of the seated participant) any and all changes, words and nonsense words alike. The participants were assured that there were no correct or incorrect responses, and that if they did not hear any changes, they were to say nothing.

After receiving verbal instructions, each participant was seated in front of an iMac computer running PsyScope 1.2.2 (Cohen et al., 1993). Participants were given a short practice session in which the word *ball* was repeated 50 times. After the practice session, the experiment proper began. Each of the four words was presented for 350 repetitions, with a 150 ms of silence between each repetition. The entire session lasted approximately 30 min. Participant's responses were recorded digitally for later analysis and verification.

7.2. Results and discussion

All of the responses were phonologically transcribed. Of the 513 responses, we further analyzed 143 real word neighbors (27.88%). We did not analyze further the 282 responses that were nonwords (54.97%),[4] the 72 real words that differed from the target word by 2 phonemes (14.04%), or the 16 real words that differed from the target word by more than 2 phonemes (3.12%). For the 143 responses that were phonological neighbors of the target words we correlated the degree/neighborhood density of the repeated words (*mean* = 21.37, *sd* = 13.35) to the mean degree/neighborhood density of the percepts in the verbal transformation illusion (*mean* = 13.84, *sd* = 10.70), and observed a positive, statistically significant correlation between neighborhood density/degree of the targets and responses ($r(141) = +.66$, $p < .0001$).

This result indicates that listeners who repeatedly heard a word with a dense phonological neighborhood tended to transform that word into a similar sounding word that also had a dense phonological neighborhood. Concomitantly, listeners who repeatedly heard a word with a sparse phonological neighborhood tended to transform that word into a similar sounding word that also had a sparse phonological neighborhood. This result is consistent with the hypothesis that assortative mixing by degree— first observed in a network analysis of the mental lexicon by Vitevitch (2008) and again in the results of the other studies reported here—might influence some aspect of language processing.

8. General discussion

In the present study we used computer simulations, naturalistic observation (*i.e.*, a slips of the ear corpus), and several psycholinguistic experiments to look for a cognitive footprint in human language behavior of the macro-level measure of network structure known as assortative mixing by degree. Given the role that assortative mixing by degree plays in network resilience (as demonstrated in network simulations), we focused our investigations on instances of real and simulated "failed" lexical retrieval. Although various methodologies were employed in the present study, different words were used as stimuli, and there was variability in the number of responses that contributed to each analysis, the same result was

[4] The proportion of nonword responses observed in the present study is comparable to the proportion of nonword responses observed in the experiments reported in Shoaf and Pitt (2002). As they note, nonword responses should not be surprising given that our instructions to the participants (like the instructions they gave to their participants) explicitly stated that words as well as nonwords were possible percepts.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

observed: a positive correlation was found between the degree/neighborhood density of a target word and the degree/neighborhood density of the phonological neighbors that were retrieved when lexical retrieval failed. This result suggests that human language behavior does indeed seem to be influenced by the macro-level structure of the lexical network, as measured by the network science metric known as assortative mixing by degree.

To be clear, our exploration of failed lexical retrieval in the present study was not attempting to provide an account of all types of perceptual errors, to develop an algorithm that would predict the likelihood of a misperception occurring, or to predict a list of words that would most likely be retrieved when lexical retrieval fails. Such questions are interesting, but they are not the focus of the present investigation. The results of the present studies suggest that the network perspective may be useful in addressing those questions—and perhaps for optimizing error recovery procedures in automatic speech recognition in computers (Scharenborg, ten Bosch, Boves, & Norris, 2003)—but we acknowledge that there may be other factors that lead to perceptual errors than simply phonological ones.

Importantly, the present findings suggest that the macro-level measure of assortative mixing by degree is not simply a spurious, mathematical phenomenon observed in a network analysis of phonological word-forms (Vitevitch, 2008). Rather, the present results suggest that the macro-level structure of the mental lexicon has consequences for psycholinguistic processing. Finding psycholinguistic evidence that is consistent with assortative mixing by degree observed in a network analysis demonstrates that the tools of network science may be useful in studying the structure of cognitive representations in memory, and in examining how that structure influences various cognitive processes. Borge-Holthoefer and Arenas (2010) argued that measures of cognitive performance serve to strengthen network analyses of language. Similarly, we suggest that the tools of network science have much to offer studies of language processing and cognitive psychology more generally.

To examine the mental lexicon with the tools of network science, one must, of course, make several assumptions. First, one must assume that the lexicon is composed of distinct representations of words that can be accessed with a well-defined search algorithm. As Turvey and Moreno (2006) point out, this is a common assumption that is made in many views of the mental lexicon (*N.B.*, spreading activation is a common method used to model search processes), so it should not be controversial that our previous network analyses and current empirical work appeal to the same assumption.

Assuming that the lexicon is composed of distinct representations of words raises some questions about what exactly a word is. For example, should a compound word like *pushcart* (a cart that one pushes) be stored as one whole word or as its two constituents? What about opaque compound words, like *buttercup* (a flower, not a cup that contains a dairy product)? How are idioms (phrases that figuratively express a concept, such as “kick the bucket” meaning “to die”) stored: as individual words or a single phrase? What is a word in a polysynthetic language, where several morphemes may be put together to form a complex word that functions like an entire sentence? (See Arbesman et al. (2010a) for a discussion of some issues related to morphology from the network perspective; see also Celata, Calderone, & Montermini, 2011 for another approach to the morphology–phonology interface.) Recent evidence also shows that language users in analytic languages like English are sensitive to the frequency of occurrence of sentence-length phrases, such as “I don’t know why,” with a high frequency of occurrence, and “I want to sit,” with a low frequency of occurrence (Arnon & Snider, 2010; *cf.*, Baayen, Hendrix, & Ramscar, 2013). If sentences are stored in the lexicon as well, how might one represent and examine a sentence using the tools of network science (for one solution see Ferrer i Cancho, Mehler, Abramov, & Díaz-Guilera, 2007)?

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

We also assume that these distinct lexical representations are abstract, free of any indexical properties that convey information about the age, gender, and so on of the speaker (Abercrombie, 1967). This assumption differs from the view of the mental lexicon being composed of very detailed exemplar representations that do contain such information (e.g., Goldinger, 1998 Johnson, 1997). We do not, however, see the use of the tools of network science as being incompatible with certain exemplar models, such as the exemplar view proposed by McLennan and Luce (2005). In the McLennan and Luce (2005) view, the lexicon contains both abstract and specific, exemplar-like representations, with abstract representations accessed rapidly, and exemplar-like representations accessed more slowly. Although the present work does not directly address exemplars per se, we see the present work as providing insight into how those abstract lexical representations in the McLennan and Luce (2005) view might be organized in memory (and how that structure might influence processing).

If one assumes the 'abstract *and* exemplar representation' framework proposed by McLennan and Luce (2005)—as we have in other work from our lab (e.g., Vitevitch & Donoso, 2011)—then some aspects of language processing, such as the comprehension of idioms, and sentence-frequency effects (e.g., Arnon & Snider, 2010), could emerge from the exemplar-based lexicon rather than the abstract portion of the lexicon that our present work bears on. The exemplar-based lexicon might also offer an elegant account of various other word-frequency effects as well.[5]

Given that one must make some of the same assumptions as mainstream psycholinguistics to use the tools of network science to examine the mental lexicon, one may wonder what is to be gained by using the tools of network science to examine the mental lexicon. We believe that mainstream psycholinguistics has focused investigation primarily on the characteristics of individual words (e.g., word frequency, word length, phonotactic probability, etc.), leading to an ever-increasing list of variables that affect the retrieval of a word (e.g., Cutler, 1981; Vitevitch, 2002a; Vitevitch, 2007). This attention to the characteristics of individual words has limited the development of tools to explore the mental lexicon as a system. It is perhaps the lack of such tools that keeps researchers from exploring certain topics, such as failed lexical retrieval, a less-explored topic in the current psycholinguistic paradigm.

Network science provides the quantitative tools necessary to examine the mental lexicon and other cognitive systems at the micro-, meso-, and macro-level. For example, the work of Chan and Vitevitch (2009, 2010); see also Vitevitch et al. (2011, 2012) examined the structure of the mental lexicon at the micro-level, and demonstrated that differences among words in the network science measure known as the clustering coefficient demonstrably influenced various language-related processes.

Siew (2013) examined the structure of the mental lexicon at the meso-level and observed that the larger network of phonological word-forms could be broken-up into several smaller

[5] Models of spoken word recognition have used a variety of approaches to represent differences in word-frequency including differential resting activation levels of word-detectors, differential thresholds in word-detectors, the order in which the representations are searched, and having the frequency of occurrence of a word act as a bias at a later decision-stage in processing. These approaches assume that information regarding the frequency of occurrence of a word is encoded at the level of the word-form, which may not be the case. Instead, word-frequency might more appropriately reflect the relationship between word-forms and semantic information (Baayen, 2010), or be represented as part of semantic information (e.g., Bates et al., 2003). If word-frequency is indeed represented in either the connections between word-forms and semantic information or solely in semantic representations, then we need not concern ourselves with how to represent word-frequency in a network of phonological word-forms. A more provocative alternative is that the network model that we proposed in Chan and Vitevitch (2009) and simulated in Vitevitch et al. (2011)—given the relationship between word frequency and degree/neighborhood density (e.g., Frauenfelder et al., 1993; Landauer & Streeter, 1973)—already represents information related to word-frequency as degree of a node, and no additional mechanisms or accounts of word frequency are required.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

communities. More in-depth analyses of the communities revealed that each community contained sequences of phonemes, such as /ŋk/, /ɪŋ/ and /ɹɪ/, that occurred frequently in that community and could be combined to form various words found in the community, such as *brink*, *drink* and *wrinkle*. The presence of these common phoneme sequences in the phonological communities raises the intriguing possibility that a separate level of representation—such as the layer of phonological segments found in models of spoken word recognition like *TRACE* (McClelland & Elman, 1986) and *Shortlist* (Norris, 1994; Norris & McQueen, 2008)—may not be required to account for the widely observed effects of phonotactic probability on language processing (e.g., Vitevitch & Luce, 2005). Instead, such effects could arise from the information already found in the lexical representations (i.e., word-forms) in the network. If that information is processed with a lower resolution, then "lexical" effects, such as competitive neighborhood density effects, may be observed. However, if that information is processed with a higher resolution, then "sub-lexical" effects, such as facilitative phonotactic probability effects, may be observed; no additional level of representation may be required (*cf*., Vitevitch & Luce, 1999).

In the present case, we demonstrated that a macro-level measure of network structure—assortative mixing by degree—also influences certain language-related processes. Although many previous studies have used sophisticated and elegant mathematical methods to find statistical relationships at the macro-level of a language—such as the work by Zipf (1935), Landauer and Streeter (1973), Baayen (1991, 2001) and Frauenfelder et al. (1993)—few, if any, of these previous studies demonstrated that these statistical relationships had behavioral consequences in psycholinguistics tasks. Therefore, the present study, employing the computational tools of network science *and* the empirical methods of psycholinguistics, represents a significant advance over previous analyses of the statistical properties of language.

Network science brings to cognitive psychology the unprecedented ability to measure several levels of a cognitive system within a common framework. This common framework, however, should not be confused with other general frameworks that have also appealed to "networks," such as connectionism (*a.k.a*. artificial neural networks). The connectionist framework allowed researchers to examine *processing* in a wide range of domains studied by cognitive psychologists. However, artificial neural networks, especially parallel distributed processing (PDP) networks, had little to say about the nature of cognitive representations, and even less to say about the structure or organization of cognitive representations. Rather, proponents of the PDP approach went so far as to argue that "… information is not stored as such, but instead is reconstructed in response to probes…" (McClelland & Rogers, 2003, p. 312). That is, cognitive representations do not exist; they simply emerge from the processes of the connectionist network. The network science approach as used in the present study suggests instead that representations are stored in memory, and, more important, the way in which those representations are organized in memory influences cognitive processing.

Network science offers a unique and powerful set of tools to measure the structure of complex systems. 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 & Barabási, 2002; see also Boccaletti et al. (2006)). As demonstrated in the present study, adopting this approach in the cognitive sciences can lead to new discoveries (e.g., how the structure of the mental lexicon influences certain aspects of language processing), and to an expansion in the range of questions investigated by cognitive scientists (i.e., how speech perception errors can increase our understanding of the spoken word recognition system). Several other areas in the domain of cognitive psychology have benefitted from adopting the analytic tools of network science (Goldstone et al., 2008; Griffiths et al., 2007; Hills et al.,

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

2009; Iyengar et al., 2012; Sporns, 2010; Steyvers & Tenenbaum, 2005). We urge more language and cognitive scientists to consider how network science might lead to new questions and novel insights. At the same time we urge language and cognitive scientists to keep in mind that networks are not panaceas, and they should be used judiciously and appropriately (see Butts (2009)).

Acknowledgments

This research was supported in part by grants from the National Institutes of Health to the University of Kansas through the Mental Retardation and Developmental Disabilities Research Center (National Institute of Child Health and Human Development P30 HD002528), and the Center for Biobehavioral Neurosciences in Communication Disorders (NIDCD P30 DC005803). R.G. was supported in part by NIDCD T32–DC00052 Training Researchers in Language Impairments. We thank two anonymous reviewers and R. Harald Baayen for their comments and suggestions.

References

Abercrombie, D. Elements of general phonetics. Chicago, IL: Aldine; 1967.

Albert R, Barabási AL. Statistical mechanics of complex networks. Review of Modern Physics. 2002; 74:47–97.

Albert R, Jeong H, Barabási AL. Error and attack tolerance of complex networks. Nature. 2000; 406:378–382. [PubMed: 10935628]

Amaral LAN, Scala A, Barthélémy M, Stanley HE. Classes of small-world networks. Proceedings of the National Academy of Sciences. 2000; 97:11149–11152.

Arbesman S, Strogatz SH, Vitevitch MS. Comparative analysis of networks of phonologically similar words in English and Spanish. Entropy. 2010a; 12:327–337.

Arbesman S, Strogatz SH, Vitevitch MS. The structure of phonological networks across multiple languages. International Journal of Bifurcation and Chaos. 2010b; 20:679–685.

Arnon I, Snider N. More than words: Frequency effects for multi-word phrases. Journal of Memory & Language. 2010; 62:67–82.

Baayen, RH. A stochastic process for word frequency distributions. Proceedings of the 29th annual meeting of the association for computational linguistics; 1991.

Baayen, RH. Word frequency distributions. Dordrecht: Kluwer; 2001.

Baayen RH. Demythologizing the word frequency effect: A discriminative learning perspective. The Mental Lexicon. 2010; 5:436–461.

Baayen RH, Hendrix P, Ramscar M. Sidestepping the combinatorial explosion: An explanation of *n*-gram frequency effects based on naïve discriminative learning. Language & Speech. 2013; 56:329–347.

Barabási AL. Scale-free networks: A decade and beyond. Science. 2009; 325:412–413. [PubMed: 19628854]

Barabási AL. The network takeover. Nature Physics. 2012; 8:14–16.

Barabási AL, Albert R. Emergence of scaling in random networks. Science. 1999; 286:509–512. [PubMed: 10521342]

Bates E, D'Amico S, Jacobsen T, Szekely A, Andonova E, Devescovi A, et al. Timed picture naming in seven languages. Psychonomic Bulletin & Review. 2003; 10:344–380. [PubMed: 12921412]

Berkowitz L, Donnerstein E. External validity is more than skin deep: Some answers to criticisms of laboratory experiments. American Psychologist. 1982; 37:245–257.

Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D. Complex networks: Structure and dynamics. Physics Reports. 2006; 424:175–308.

Bond, ZS. Slips of the ear: Errors in the perception of casual conversation. New York: Academic Press; 1999.

Borge-Holthoefer J, Arenas A. Semantic networks: Structure and dynamics. Entropy. 2010; 12:1264–1302.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Brown R, McNeill D. The “tip of the tongue” phenomenon. Journal of Verbal Learning and Verbal Behavior. 1966; 5:325–337.

Butts CT. Revisiting the foundations of network analysis. Science. 2009; 325:414–416. [PubMed: 19628855]

Celata C, Calderone B, Montermini F. Enriched sublexical representations to access morphological structures. A psycho-computational account. Traitement Automatique des Langues. 2011; 52:123–149.

Chan KY, Vitevitch MS. The influence of the phonological neighborhood clustering-coefficient on spoken word recognition. Journal of Experimental Psychology: Human Perception & Performance. 2009; 35:1934–1949. [PubMed: 19968444]

Chan KY, Vitevitch MS. Network structure influences speech production. Cognitive Science. 2010; 34:685–697. [PubMed: 21564230]

Charles-Luce J, Luce PA. Similarity neighborhoods of words in young children’s lexicons. Journal of Child Language. 1990; 17:205–215. [PubMed: 2312642]

Cohen J, MacWhinney B, Flatt M, Provost J. PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. Behavior Research Methods, Instruments, and Computers. 1993; 25(2):257–271.

Cramer AOJ, Waldorp LJ, van der Maas H, Borsboom D. Comorbidity: A network perspective. Behavioral and Brain Sciences. 2010; 33:137–193. [PubMed: 20584369]

Cutler A. Making up materials is a confounded nuisance, or: Will we be able to run any psycholinguistic experiments at all in 1990? Cognition. 1981; 10:65–70. [PubMed: 7198562]

Dell GS. A spreading-activation theory of retrieval in sentence production. Psychological Review. 1986; 93:283–321. [PubMed: 3749399]

Dell GS. The retrieval of phonological forms in production: Tests of predictions from a connectionist model. Journal of Memory and Language. 1988; 27:124–142.

Fay D, Cutler A. Malapropisms and the structure of the mental lexicon. Linguistic Inquiry. 1977; 8:505–520.

Felty RA, Buchwald A, Gruenenfelder TM, Pisoni DB. Misperceptions of spoken words: Data from a random sample of American English words. Journal of the Acoustical Society of America. 2013; 134:572–585. [PubMed: 23862832]

Ferrer i Cancho, R.; Mehler, A.; Abramov, O.; Díaz-Guilera, A. Correlations in the organization of large-scale syntactic dependency networks. Proceedings of graph-based methods for natural language processing (TextGraphs-2) the annual conference of the North American chapter of the Association for Computational Linguistics (NAACL-HLT 2007); Rochester, New York. 2007. p. 65-72.

Ferrer-i-Cancho R, Forns N, Hernandez-Fernandez A, Bel-Enguix G, Baixeries J. The challenges of statistical patterns of language: The case of Menzerath’s Law in genomes. Complexity. 2012; 18:11–17.

Forster, KI. Accessing the mental lexicon. In: Walker, E., editor. Explorations in the biology of language. Montgomery, VT: Bradford; 1978.

Frauenfelder UH, Baayen RH, Hellwig FM, Schreuder R. Neighborhood density and frequency across languages and modalities. Journal of Memory & Language. 1993; 32:781–804.

Fromkin VA. The non-anomalous nature of anomalous utterances. Language. 1971; 47:27–52.

Gaskell MG, Marslen-Wilson WD. Integrating form and meaning: A distributed model of speech perception. Language and Cognitive Processes. 1997; 12:613–656.

Girvan M, Newman MEJ. Community structure in social and biological networks. Proceedings of the National Academy of Sciences. 2002; 99:7821–7826.

Goldinger SD. Echoes of echoes? An episodic theory of lexical access. Psychological Review. 1998; 105:251–279. [PubMed: 9577239]

Goldstone RL, Roberts ME, Gureckis TM. Emergent processes in group behavior. Current Directions in Psychological Science. 2008; 17:10–15.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Grainger J, O'Regan JK, Jacobs AM, Segui J. On the role of competing word units in visual word recognition: The neighborhood frequency effect. Perception & Psychophysics. 1989; 45:189–195. [PubMed: 2710616]

Griffiths TL, Steyvers M, Firl A. Google and the mind: Predicting fluency with PageRank. Psychological Science. 2007; 18:1069–1076. [PubMed: 18031414]

Gruenenfelder TM, Pisoni DB. The lexical restructuring hypothesis and graph theoretic analyses of networks based on random lexicons. Journal of Speech, Language, and Hearing Research. 2009; 52:596–609.

Hills TT, Maouene M, Maouene J, Sheya A, Smith L. Longitudinal analysis of early semantic networks: Preferential attachment or preferential acquisition? Psychological Science. 2009; 20:729–739. [PubMed: 19470123]

Iyengar SRS, Madhavan CEV, Zweig KA, Natarajan A. Understanding human navigation using network analysis. Topics in Cognitive Science. 2012; 4:121–134. [PubMed: 22253185]

Johnson, K. Speech perception without speaker normalization: An exemplar model. In: Mullennix; Johnson, editors. Talker variability in speech processing. San Diego: Academic Press; 1997. p. 145-165.

Keller EF. Revisiting "scale-free" networks. BioEssays. 2005; 27:1060–1068. [PubMed: 16163729]

Kello, CT.; Beltz, BC. Scale-free networks in phonological and orthographic wordform lexicons. In: Chitoran, I.; Coupé, C.; Marsico, E.; Pellegrino, F., editors. Approaches to phonological complexity. Mouton de Gruyter; 2009.

Kučera, H.; Francis, WN. Computational analysis of present day American English. Providence: Brown University Press; 1967.

Lamb, S. Linguistic and cognitive networks. In: Garvin, P., editor. Cognition: A multiple view. New York: Spartan Books; 1970. p. 195-222.

Landauer TK, Streeter LA. Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. Journal of Verbal Learning and Verbal Behavior. 1973; 12:119–131.

Laxon VJ, Coltheart V, Keating C. Children find friendly words friendly too: Words with many orthographic neighbours are easier to read and spell. British Journal of Educational Psychology. 1988; 58:103–119.

Lester JA. A study of high school spelling material II. Journal of Educational Psychology. 1922; 13:152–159.

Levelt WJM, Roelofs A, Meyer AS. A theory of lexical access in speech production. Behavioral & Brain Sciences. 1999; 22:1–75. [PubMed: 11301520]

Lewandowsky S. The rewards and hazards of computer experiments. Psychological Science. 1993; 4:236–243.

Luce PA, Goldinger SD, Auer ET Jr, Vitevitch MS. Phonetic priming, neighborhood activation, and PARSYN. Perception and Psychophysics. 2000; 62:615–625. [PubMed: 10909252]

Luce PA, Large NR. Phonotactics, density and entropy in spoken word recognition. Language and Cognitive Processes. 2001; 16:565–581.

Luce PA, Pisoni DB. Recognizing spoken words: The neighborhood activation model. Ear and Hearing. 1998; 19:1–36. [PubMed: 9504270]

MacKay, DG. Errors, ambiguity, and awareness in language perception and production. In: Baars, BJ., editor. Experimental slips and human error: Exploring the architecture of volition. New York: Plenum Press; 1992. p. 39-69.

MacKay DG, Wulf G, Yin C, Abrams L. Relations between word perception and production: New theory and data on the verbal transformation effect. Journal of Memory & Language. 1993; 32:624–646.

Mandelbrot, B. An informational theory of the statistical structure of language. In: Jackson, W., editor. Communication theory. London: Bettersworths; 1953.

Marslen-Wilson WD. Functional parallelism in spoken word-recognition. Cognition Special issue: Spoken word recognition. 1987; 25:71–102.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

McClelland JL. The place of modeling in cognitive science. Topics in Cognitive Science. 2009; 1:11–38.

McClelland JL, Elman JL. The TRACE mo4del of speech perception. Cognitive Psychology. 1986; 18:1–86. [PubMed: 3753912]

McClelland JL, Rogers TT. The parallel distributed processing approach to semantic cognition. Nature Reviews Neuroscience. 2003; 4:310–322.

McLennan CT, Luce PA. Examining the time course of indexical specificity effects in spoken word recognition. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2005; 31:306–321.

Miller G. Some effects of intermittent silence. American Journal of Psychology. 1957; 52:311–314. [PubMed: 13424784]

Neal, ZP. The Connected City: How networks are shaping the modern metropolis. New York: Routledge; 2013.

Nelson, DL.; McEvoy, CL.; Schreiber, TA. The University of South Florida word association, rhyme, and word fragment norms. 1998. <http://www.usf.edu/FreeAssociation/>

Newman MEJ. Assortative mixing in networks. Physical Review Letters. 2002; 89:20889701.

Norman DA. Categorization of action slips. Psychological Review. 1981; 88:1–15.

Norris D. Shortlist: A connectionist model of continuous speech recognition. Cognition. 1994; 52:189–234.

Norris D, McQueen JM. Shortlist B: A Bayesian model of continuous speech recognition. Psychological Review. 2008; 115:357–395. [PubMed: 18426294]

Nusbaum HC, Pisoni DB, Davis CK. Sizing up the Hoosier Mental Lexicon: Measuring the familiarity of 20,000 words. Research on Speech Perception Progress Report. 1984; 10:357–376.

Pisoni DB. Word identification in noise. Language & Cognitive Processes. 1996; 11:681–688. [PubMed: 21687807]

Pisoni DB, Nusbaum HC, Luce PA, Slowiaczek LM. Speech perception, word recognition and the structure of the lexicon. Speech Communication. 1985; 4:75–95. [PubMed: 23226910]

Plunkett, K.; Elman, JL. Exercises in rethinking innateness: A handbook for connectionist experiments. MIT Press; 1997.

Porter MA, Onnela JP, Mucha PJ. Communities in networks. Notices of the American Mathematical Society. 2009; 56:1082–1166.

Quillian R. Word concepts: A theory and simulation of some basic semantic capabilities. Behavioral Science. 1967; 12:410–430. [PubMed: 6059773]

Roodenrys S, Hulme C, Lethbridge A, Hinton M, Nimmo LM. Word-frequency and phonological-neighborhood effects on verbal short-term memory. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2002; 28:1019–1034.

Rosenblatt F. The perceptron: A probabilistic model for information storage and organization in the brain. Psychological Review. 1958; 65:386–408. [PubMed: 13602029]

Scharenborg O, ten Bosch L, Boves L, Norris D. Bridging automatic speech recognition and psycholinguistics: Extending Shortlist to an end-to-end model of human speech recognition. Journal of the Acoustical Society of America. 2003; 114:3032–3035. [PubMed: 14714783]

Shoaf LC, Pitt MA. Does node stability underlie the verbal transformation effect? A test of node structure theory. Perception & Psychophysics. 2002; 64:795–803. [PubMed: 12201338]

Siew CSQ. Community structure in the phonological network. Frontiers in Psychology. 2013; 4:533. http://dx.doi.org/10.3389/fpsyg.2013.00553. [PubMed: 23966971]

Sporns, O. Networks of the brain. MIT Press; 2010.

Steyvers M, Tenenbaum J. The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. Cognitive Science. 2005; 29:41–78. [PubMed: 21702767]

Storkel HL. Do children acquire dense neighborhoods? An investigation of similarity neighborhoods in lexical acquisition. Applied Psycholinguistics. 2004; 25:201–221.

Strauss TJ, Harris HD, Magnuson JS. JTRACE: A reimplementation and extension of the TRACE model of speech perception and spoken word recognition. Behavior Research Methods. 2007; 39:19–30. [PubMed: 17552468]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Turvey MT, Moreno M. Physical metaphors for the mental lexicon. The Mental Lexicon. 2006; 1:7–33.

Vitevitch MS. Influence of onset density on spoken-word recognition. Journal of Experimental Psychology: Human Perception and Performance. 2002a; 28:270–278. [PubMed: 11999854]

Vitevitch MS. Naturalistic and experimental analyses of word frequency and neighborhood density effects in slips of the ear. Language and Speech. 2002b; 45:407–434. [PubMed: 12866911]

Vitevitch MS. The spread of the phonological neighborhood influences spoken word recognition. Memory & Cognition. 2007; 35:166–175. [PubMed: 17533890]

Vitevitch MS. What can graph theory tell us about word learning and lexical retrieval? Journal of Speech Language Hearing Research. 2008; 51:408–422.

Vitevitch MS, Donoso A. Processing of indexical information requires time: Evidence from change deafness. Quarterly Journal of Experimental Psychology. 2011; 64:1484–1493.

Vitevitch MS, Chan KY, Roodenrys S. Complex network structure influences processing in long-term and shortterm memory. Journal of Memory & Language. 2012; 67:30–44. [PubMed: 22745522]

Vitevitch MS, Ercal G, Adagarla B. Simulating retrieval from a highly clustered network: Implications for spoken word recognition. Frontiers in Psychology. 2011; 2:369. http://dx.doi.org/10.3389/fpsyg.2011.00369. [PubMed: 22174705]

Vitevitch MS, Luce PA. Probabilistic phonotactics and neighborhood activation in spoken word recognition. Journal of Memory & Language. 1999; 40:374–408.

Vitevitch MS, Luce PA. A web-based interface to calculate phonotactic probability for words and nonwords in English. Behavior Research Methods, Instruments, and Computers. 2004; 36:481–487.

Vitevitch MS, Luce PA. Increases in phonotactic probability facilitate spoken nonword repetition. Journal of Memory & Language. 2005; 52:193–204.

Vitevitch MS, Stamer MK, Sereno JA. Word length and lexical competition: Longer is the same as shorter. Language & Speech. 2008; 51:361–383. [PubMed: 19348156]

Vitevitch, MS.; Storkel, HL. Examining the acquisition of phonological word forms with computational experiments. Language & Speech. 2012. http://dx.doi.org/10.1177/0023830912460513 (Published online before print October 23, 2012 1177/0023830912460513 (Published online before print October 23, 2012)

Waller WS, Zimbelman MF. A cognitive footprint in archival data: Generalizing the dilution effect from laboratory to field settings. Organizational Behavior and Human Decision Processes. 2003; 91:254–268.

Warren RM. Verbal transformation effect and auditory perceptual mechanisms. Psychological Bulletin. 1968; 70:261–270. [PubMed: 5722568]

Warren RM, Gregory RL. An auditory analogue of the visual reversible figure. American Journal of Psychology. 1958; 71:612–613. [PubMed: 13571475]

Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. Nature. 1998; 393:409–410. [PubMed: 9623993]

Wright CE. Duration differences between rare and common words and their implications for the interpretation of word frequency effects. Memory & Cognition. 1979; 7:411–419. [PubMed: 542114]

Zipf, GK. The psychobiology of language. Boston: Houghton Mifflin; 1935.

Appendix A. Definitions of several network science measures used to assess different levels of a system

A.1. Micro-level

Measures used to examine the behavior of individual nodes.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

A.1.1. Degree

The degree of a node, k_i, is the number of connections (or edges) that node *i* has to other nodes. In the phonological network examined in Vitevitch (2008), degree is synonymous with the psycholinguistic term "neighborhood density" (Luce & Pisoni, 1998).

A.1.2. Clustering coefficient

The extent to which the neighbors of a given node are also neighbors of each other. More precisely, it is defined as:

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

e_{jk} refers to the presence of a connection (or edge) between two neighbors (*j* and *k*) of node *i*, |…| 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* (Watts & Strogatz, 1998). Thus, the (local) clustering coefficient is the number of connections that actually exist among the neighbors of a given node divided by the number of connections that could possibly exist among the neighbors of a given node. The clustering coefficient has a range from 0 to 1. When $C = 0$, none of the neighbors of a target node are neighbors of each other. When $C = 1$, every neighbor of a node is also a neighbor of all of the other neighbors of a node. Chan and Vitevitch (2009) showed in a variety of psycholinguistic tasks that *C* influenced the recognition of spoken words such that words with low *C* were responded to more quickly and accurately than words with high *C*. Similarly, Chan and Vitevitch (2010) showed in a variety of psycholinguistic tasks that *C* influenced the production of spoken words such that words with low *C* were produced more quickly and accurately than words with high *C*.

A.1.3. Centrality measures

These measures attempt to determine the relative importance of a node in a network, and include: degree centrality, closeness centrality, betweenness centrality, and eigenvector centrality. *Degree centrality* is simply the number of connections (or edges) that a node has to other nodes. Thus a node that is connected to many other nodes in the network (i.e., it has high degree centrality) is thought to be more "important" than a node that is connected to few other nodes in the network (i.e., it has low degree centrality). *Closeness centrality* is defined as the inverse of the sum of distances from that node to all other reachable nodes in the network. More precisely:

$$Closeness(i) = \frac{1}{\sum_j d_{ij}}$$

where *i* is the node of interest, *j* is another node in the network, and d_{ij} is the shortest distance between these two nodes. See the work of Iyengar et al. (2012) for an example of how closeness centrality influences language-related processing. *Betweenness centrality* measures the number of times a node appears in the shortest path between two other nodes. A node is said to be high in betweenness centrality if it lies along many shortest paths. That is, that node lies "between" many other nodes, and therefore could control the flow of information, etc. *Eigenvector centrality* is a measure of node importance that is based on the connections to a given node. The PageRank algorithm used in the Google search engine uses a variant of eigenvector centrality to rank order web pages, such that a page is highly ranked if many highly ranked pages link to it.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

A.2. Meso-level

Measures used to examine the interaction of nodes at short distances. These measures allow you to generalize beyond the micro-level measures of individual nodes to groups of nodes that share some characteristic. The most common assessment at the meso-level is *community detection*, or finding groups of nodes such that the connections within the group are denser than the connections between groups (Girvan & Newman, 2002; Porter, Onnela, & Mucha, 2009). Siew (2013) examined the community structure of the phonological network from Vitevitch (2008).

A.3. Macro-level

General measures of network structure. If the network contains several components (*i.e.*, it is disconnected), the convention is to compute the average values described below in the largest (or giant) component.

A.3.1. Average degree <k>

The degree of a node, k_i, is the number of other nodes to which node *i* is connected. The average degree is the average value of *k* over all nodes in a network.

A.3.2. Degree distribution

The probability, $P(k)$, that a randomly chosen node has degree *k* for every possible *k*. The degree distribution is a convenient way to statistically characterize the topology of large networks. Scale-free networks exhibit a degree distribution that approximate a power-law, $P(k) \sim k^{-\gamma}$ where the exponent, γ, is usually between 2 and 3.

A.3.3. Average shortest path length (l)

The shortest path length is the length (in number of connections) of the shortest path between pairs of nodes (*i* and *j*). The average shortest path length is the average of the shortest path length over all pairs of nodes in the network. Note that average shortest path length is typically computed only in the giant component of a disconnected network because the distance between disconnected nodes is infinite. This is one of the two criteria (the other being *average clustering coefficient*) used to determine whether a network exhibits the characteristics of a *small-world network*.

A.3.4. Average clustering coefficient

The average clustering coefficient is the average of the clustering coefficient over all nodes in the network. This is one of the two criteria (the other being *average shortest path length*) used to determine whether a network exhibits the characteristics of a *small-world network*.

A.3.5. Small-world network

For a network to exhibit the characteristics of a *small-world network* it must have: (1) an *average shortest path length* that is comparable to the average shortest path length of a random network with the same number of nodes and the same average degree, and (2) an *average clustering coefficient* that is much greater than the average clustering coefficient of a random network with the same number of nodes and the same average degree (Watts & Strogatz, 1998).

A.3.6. Scale-free network

Networks with degree distributions that follow a power-law are known as scale-free networks (see also degree distribution). Scale-free networks were introduced in Barabási and

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Albert (1999), and were shown to arise from the network growing (i.e., new nodes are added to the system) such that new nodes tended to connect to nodes that were already well connected in the network (i.e., preferential attachment). The presence of a power-law in the degree distribution was thought to be an indication that the system arose though self-organization. However, as Keller (2005) and others have shown, there are a number of mechanisms other than growth via preferential attachment that can produce a power-law degree distribution.

Appendix B. Target words and "perceptual errors" from Simulation 1 with jTRACE

Target words	Perceptual errors
lig	lig^l*
tub	tu
r^S	r^b
pul	par*
p^t	pat
rut	rupi*
s^k	s^ksid*
did	dip
dip	did
lid	lig
tul	tu
ark	ar
bar	bark
dat	dart
Sut	Sat
S^t	Sat
tar	targ^t*
sis	si
lat	lak
sil	si
sik	si
sid	si
sit	si
blu	blak*
dru	drap*
d^k	d^l
gat	gad
par	part
lup	lus
r^b	r^S
tri	trit
b^s	b^t
k^p	k^p^l*

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Target words	Perceptual errors
gad	gat
rab	rad
ril	rili
rad	rab
kru	kru^l*
gru	grup
l^k	l^ki
lak	lat
dil	dip
rid	ril
tru	trup
but	bust
pap	pap^
bit	bi
kap	kapi
Sit	Si
Sat	Sap
pat	pap
bab	babi
klu	kru
dal	dart*
kip	ki
lip	lid

Note: The 10 perceptual errors marked with an asterisk (and their corresponding target word) were excluded from the analysis that used a more stringent scoring criterion.

Appendix C. The three lexicons varying in mixing by degree used in Simulation 2 with jTRACE

Assortative mixing by degree	Disassortative mixing by degree	No mixing by degree
ggg	^^^	^^^
ggp	^^a	^^b
agg	^b^	a^^
agp	d^^	a^b
iii	aaa	ba^
iir	aad	bb^
bii	aba	bbt
bir	gaa	tbb
kkk	bbb	tba
kkg	b^b	taa
^kk	bbg	btb

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Assortative mixing by degree	Disassortative mixing by degree	No mixing by degree
^kg	ibb	bta
lll	ddd	ddd
llb	bdd	ddg
ill	dkd*	idd
ilb	ddS	idg
ppp	ggg	gid
ppa	ggl	ggd
lpp	grg†	ggu
lpa	pgg	ugg
rrr	iii	ugi
rrp	^ii	uii
krr	iia	gug
krp	ipi	gui
sss	kkk	kkk
ssr	akk	kkp
^ss	kki	lkk
^sr	kuk	lkp
SSS	lll	plk
SSk	^ll	ppk
iSS	lla	ppt
iSk	lkl	tpp
ttt	ppp	tpl
Stt	^pp	tll
tts	ppd	ptp
Sts	pdp	ptl
uuu	rrr	rrr
uut	^rr	rrS
^uu	rbr	srr
^ut	rrd	srS
^ti	sss	Ssr
^tp	^ss	SSr
aiu	sbs	SSu
ait	ssg	uSS
bSu	SSS	uSs
bSp	^SS	uss
dis	SaS	SuS
dil	SSg	Sus
gba	ttt	
gbu	^tt	
idp	tbt	
idt	ttg	
kul	uuu	

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Assortative mixing by degree	Disassortative mixing by degree	No mixing by degree
kub	^uu	
lrd	uau[§]	
lrs	uud	

Notes:
[*] ddd was most activated instead.
[†] ggg was most activated instead.
[§] uuu was most activated instead.

Appendix D. The stimuli used in Experiments 1 and 2

Dense neighborhoods		Sparse neighborhoods	
lame	lace	dash	myth
raid	buck	dose	coach
reap	lease	firm	wedge
rose	lice	dive	beg
bead	rag	surge	badge
lash	tail	sham	niece
load	lock	wrath	term
loot	bag	null	folk
ride	roam	mop	numb
nail	wine	lobe	noose
luck	bite	niche	gang
kneel	gore	path	page
lag	tire	cuff	thick
rake	dial	balm	dig
tall	lamb	poach	beam
reek	peak	perch	dish
rope	seed	fish	match
rash	dot	retch	math
tuck	code	doll	foam
shore	dine	gum	bib
root	rid	leg	bath
rum	bun	curse	siege
note	suit	fig	rung
dumb	pad	tab	shed
goat	dare	deaf	nab

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

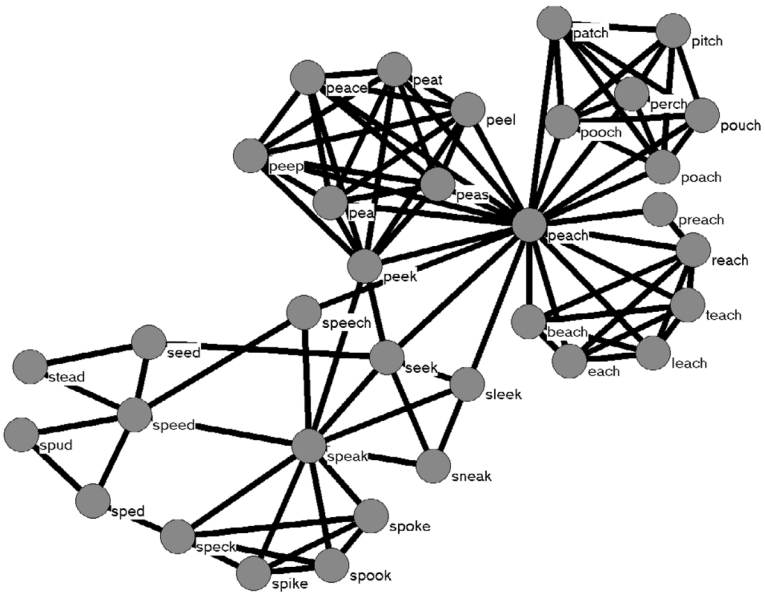

Fig. 1.

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.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 1

Means (and standard deviations) of dense and sparse words in Experiment 3.

	Dense	Sparse
Familiarity	7 (0)	6.71 (.37)
Log frequency	1.18 (.67)	1.10 (.52)
Sum of phones	.17 (.03)	.14 (.04)
Sum of biphones	.005 (.0006)	.004 (.003)
Log neighborhood freq.	2.07 (.02)	2.13 (.39)
Stimulus duration	560.9 (72.05)	640.38 (87.66)

Note: Durations are measured in milliseconds.