An Analysis of Semantic and Phonological Associations Using Network Science

By
© 2017

Nichol Castro
M.A., University of Kansas, 2014

Submitted to the graduate degree program in the Department of Psychology and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Chair: Dr. Michael S. Vitevitch

Dr. Susan Kemper

Dr. Paul Atchley

Dr. Tamara Baker

Dr. Jonathan S. Brumberg

Date Defended: August 24, 2017

The dissertation committee for Nichol Castro certifies that this is the approved version of the following dissertation:

An Analysis of Semantic and Phonological Associations Using Network Science

Chair: Dr. Michael S. Vitevitch

Date Approved: August 24, 2017

Abstract

Semantic and phonological systems interact during word processing. However, the current approaches to studying these systems tend to examine them as separate entities with a focus on processes that occur in those systems. An alternative approach is to examine the underlying representations of these systems with the use of the computational and mathematical tools of Network Science. The analysis of language networks, where nodes represent words and edges represent relationships, have shown that network structure influences language processes. The present study analyzes a novel phonological network using collected phonological association data. 1,018 participants provided up to three phonological associates to a cue word. The cue and response words were used as nodes in the phonological association network, and edges were placed between cue and response pairs. The resulting phonological association network structure exhibited several characteristics, like small-world structure and assortative mixing by degree that were similar to the well-studied one-phoneme difference phonological network, but the phonological association network was also different in structure from the well-studied one-phoneme difference phonological network. In addition, three age-related phonological association networks were examined that represented young adulthood, early middle adulthood, and late middle adulthood. However, there was little phonological network structure change across these age-related networks. Lastly, cutting-edge research in Network Science that uses multiplex networks was employed to examine the semantic and phonological systems simultaneously. This multiplex consisted of two layers: semantic associations and phonological associations. Cue and response words were used as nodes and edges were placed between cue and response pairs in their respective layers. The two layers are distinctly different in their network structure as they represent different aspects of the mental lexicon. However, there was

overlap between layers, or instances where a pair of words was connected in both the semantic and phonological layers. Regression analyses were conducted to further assess the influence of single-layer and multiplex network structure on behavioral performance. Specifically, the reaction time for visual lexical decision and naming were predicted using semantic degree, phonological degree, aggregated multiplex degree, multidegree, and the interaction between semantic and phonological degree. The results of a model building procedure indicated that all of the degree measures were needed in the regression analysis model, providing evidence that multiplex structure and the interaction between layers is important to word processing. In sum, the findings from this study provide evidence that phonological associations can be used to construct a representation of the phonological system, that phonological network structure does not significantly change with increasing age, and that the multiplex structure is important to language processing.

Acknowledgements

This dissertation is dedicated to my sister, Victoria. You will always be the light guiding my way, the strength in my weak moments, and the joy radiating in my laughter.

The completion of this dissertation would not have been possible without an amazing community of mentors, family, and friends. First, I would like to thank my advisor, Dr. Michael Vitevitch, for his unending support and encouragement. I am forever grateful for the opportunity to work with you these past years. You have made me a better researcher and academic, as well as a better person.

I would also like to thank my committee members, Drs. Susan Kemper, Paul Atchley, Tamara Baker, and Jonathan Brumberg. Thank you for your advice and questions, which have shaped this work and will continue to shape my future work. Your support through this process has made me even more confident in my ideas, and I look forward to continuing in this research endeavor.

I also want to thank Dr. Lori James, my undergraduate advisor, who was the initial inspiration and motivator to my academic career. Thank you Lori for continuing to be an amazing mentor as I jumped head first into graduate school. You continue to be a role model for me as I embark on the next stages of my career.

Rounding out my academic community of support are several graduate students and faculty in the Departments of Psychology, Speech-Language-Hearing, and Linguistics (too many to list here). I have always taken on an interdisciplinary approach to research and education, and have thoroughly enjoyed being welcomed by so many people in different departments. I am especially grateful to my cohort mate, Cynthia Siew, who was my sounding board for crazy ideas, an assurance when things were tough, and the person to celebrate with for all our successes. Graduate school would not have been as manageable without you.

Lastly, I want to thank my family and friends both near and far. At times, the toll of graduate school could be exhausting and isolating. However, you have never wavered in your praise, love, and support, making this journey so much more bearable. I know that I have made you all proud, and I promise to continue doing so. An especially big thank you to my parents, Larry and Michelle, and my sister, Samantha.

List of Figures

List of Tables

List of Appendices

Table of Contents

Chapter 1: Using Network Science to Understand Language Processing

Complex systems exist in all aspects of our lives from the way we travel and how we surf the Internet to the way people communicate with others or a body communicates with itself (Newman, 2008). By defining two parameters, nodes and edges, these complex systems can be examined as a network, where nodes represent an entity (e.g., people) and edges represent relationships (e.g., friendship). With the tools of Network Science, we can describe the underlying structure of the complex system and make inferences about how processes occur given that structure. For instance, in an analysis of a social network of friends, the tools of Network Science can be used to assess which people in the network are the most connected and can spread information very quickly to many others. By being able to model the structure and processes of a complex system, Network Science has emerged as a useful tool in a variety of disciplines, including cognitive psychology. In particular, language networks that represent aspects of the mental lexicon (or the place in memory where all the words a person knows are stored) have been examined as a way to better understand how words are represented, organized, and used.

A network of the mental lexicon is constructed with nodes representing words and edges connecting words that are related. Relatedness could be defined in several ways, including meaning (i.e., semantic relationships) or sound (i.e., phonological relationships). The connections found in "semantic networks" have been defined in a number of ways, including connecting words that are associates (e.g., De Deyne, Navarro, & Storms, 2013; Hills, Maouene, Maouene, Sheya, & Smith, 2009; Morais, Olsson, & Schooler, 2013; Nelson, McEvoy, & Dennis, 2000; Steyvers & Tenenbaum, 2005), connecting words if they share features (e.g., Hills, et al., 2009), connecting words that are synonyms or antonyms of each other (e.g., Motter, Moura, Lai, & Dasgupta, 2002; Ravasz & Barabási, 2003; Steyvers & Tenenbaum, 2005), or connecting words if they co-occur in usage (e.g., Ferrer i Cancho & Solé, 2001; Lund & Burgess, 1996). In a "phonological network", on the other hand, the overlap of strings of phonology determine relatedness between words. For example, Vitevitch (2008) connects word with high phonological overlap, where words are connected that differ by only one phoneme either through addition, deletion, or substitution (Luce & Pisoni, 1998). Other phonological networks have been examined that measure lesser degrees of overlap, where one phonological string is a subset of another (e.g., Kello & Beltz, 2009).

Importantly, with Network Science, we can examine not just the individual properties of words as is commonly done in the traditional psycholinguistic approach, but also the relationships that exist among them. The structure that emerges from these connections in a language network will have important implications for how processing occurs. Consider this scenario: Two networks are created with the same number of nodes and the same number of edges. The only difference between these two networks lies in the structure that emerges from how those edges connect the nodes. In one network, the edges are placed randomly, while in the other network, the edges are placed according to a defined relationship. The structure of the edges in the latter network may allow for more efficient processing than in the former network, highlighting the importance of how edges are defined in the network and the structure that emerges. Through structural examination of networks, researchers can determine which network best models the mental lexicon by testing derived predictions with behavioral experiments. Therefore, continuing to model and understand the structure of the mental lexicon will provide new insight on word retrieval processes that cannot be done with the standard psycholinguistic approach alone, which typically considers only processes or only representations.

Single-Layer Networks of the Mental Lexicon

As stated previously, understanding the structure of a network will provide insight into how processing will occur in that network. Specifically, a single-layer network is one in which there is only a single defined type of edge placed between nodes. There are many standard network measures commonly obtained when examining single-layer networks (see Appendix A and below for descriptions of these measures), and there are measures that assess three different levels of the network structure. Micro-level measures examine individual nodes in the network and the nodes immediately connected to that individual (i.e., “neighbors”) Macro-level measures examine the whole network and general tendencies of that network. And, in between the micro- and macro-levels, the meso-level measures focus on sub-sets or communities of nodes. By examining the network structure at these different levels, researchers can consider how the structure of the mental lexicon might influence processing during word retrieval beyond some of the more traditional psycholinguistic measures that focus only on the characteristics of individual words (e.g., word frequency and word length).

Micro-level analysis. One measure that has received much consideration is *degree*, or the number of immediate connections of a particular node. In psycholinguistic research, degree in a phonological network has also been termed *phonological neighborhood density* (Luce & Pisoni, 1998). However, I will use the term *degree* in the remainder of this paper. A node with high degree is connected to many similar words, whereas a node with low degree is connected to few similar words.

Research has found that degree of a node can influence the ease and speed with which the associated word is recalled or produced. For example, individuals produce more speech errors and tip-of-the-tongue states for words with low degree and are slower to produce low degree words than high degree words (Harley & Bown, 1998; Vitevitch, 1997; 2002; Vitevitch & Sommers, 2003). In contrast, words with high semantic associate set size (equivalent to degree) are slower and less accurately recalled (Nelson, Bennett, Gee, Schreiber, & McKinney, 1993; Nelson & McEvoy, 1979; Schreiber & Nelson, 1998). The difference in effect for degree in these two examples may be due to 1) the system of examination (phonological vs semantic), and/or 2) the task itself (a production task vs a recall task).

A second micro-level measure is *clustering coefficient*, which has been shown to also influence word retrieval processes. The clustering coefficient assesses the extent to which neighbors of a node are also connected to each other (Watts & Strogatz, 1998). A node with high clustering coefficient has many connections amongst its neighbors, whereas a node with low clustering coefficient has few connections amongst its neighbors.

In the phonological network, Chan & Vitevitch (2010) found that participants produced words with high clustering coefficient more slowly and less accurately than words with low clustering coefficient. Having a more interconnected local neighborhood was more detrimental for word production processes than having a less interconnected local neighborhood. In the semantic association network, Nelson and colleagues found that participants recalled more words with higher interconnectivity (equivalent to clustering coefficient) amongst associates than words with lower interconnectivity amongst associates (Nelson, et al., 1993). Similar to the findings of degree, the differing effects of clustering coefficient may be due to differences in the system being examined and/or the task.

It is important to note that although degree and clustering coefficient are both micro-level measures, they describe different aspects of the micro-level structure and can have different effects on processing (see Figure 1). For example, in the phonological network, low degree is more detrimental for word production, whereas high clustering coefficient is more detrimental. In other words, having few neighbors that sound similar can disrupt word production processes, but high interconnectivity amongst the neighbors (regardless of how many neighbors) can also disrupt word production processes. Figure 1 shows an example from Chan & Vitevitch (2010) of two words (*badge* and *log*) that have the same degree and would be considered as having high degree. However, despite both having the same high degree, the word *badge* would be more difficult to produce than the word *log* due to their differences in clustering coefficient. Taken together, these findings highlight the importance of looking at multiple types of network measures even at the same level of analysis to gain the most complete picture of how structure influences processing.

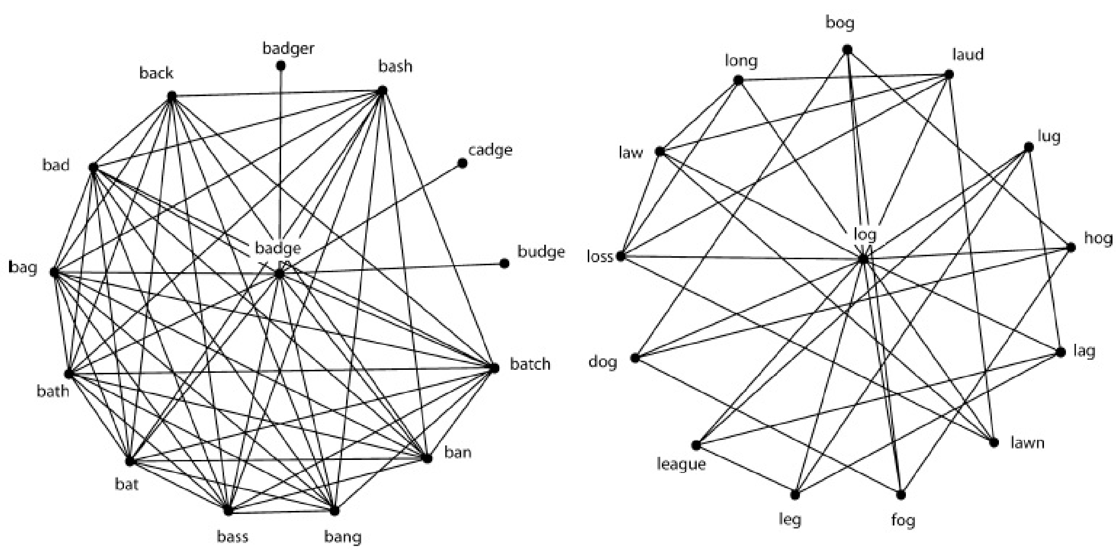

Figure 1. Example words from the phonological network with the same degree but different clustering coefficients. Reprinted from Chan & Vitevitch (2010). The words *badge* and *log* have degree of 13, but *badge* has a clustering coefficient of 0.58 and *log* has a clustering coefficient of 0.28. Despite having the same degree, *badge* and *log* would still have different rates of success in production (e.g., speed and accuracy).

Meso-level analysis. At the meso-level, the unit of analysis is *communities* of nodes within the larger network. Nodes within a community are densely connected to one another, with few connections between communities (Newman & Girvan, 2004; Ravasz & Barabási, 2003). This community structure is thought to occur through the natural division of a larger network into smaller groups that share features. Indeed, Siew (2013) found in the phonological network that words within a given community shared similar phonological segments and lexical characteristics compared to words in other communities. Additionally, Ravasz & Barabási (2003) suggested that the semantic network also has a community structure, where the communities share meaning. Furthermore, they suggest that "important" highly connected nodes, or hubs, serve as a bridge between communities allowing for the formation of a larger, robust network (Ravasz & Barabási, 2003).

The division of a large network into smaller communities may allow for more efficient processing. For example, in the semantic network, the presence of communities could make the initial search process for a target word more efficient by reducing the required search space from the entire network to just the community that the word resides within. This hypothesis could explain the semantic interference and facilitation effects in picture-word interference tasks. That is, associatively related distractors that provide facilitation (e.g., carrot and rabbit; Sailor, et al., 2009) would be located within the same semantic community. But, categorically related distractors that interfere (e.g., chipmunk and rabbit; Damian, Vigliocco, & Levelt, 2001; Hantsch, Jescheniak, & Schriefers, 2005; Rahman & Aristei, 2010) would be members of separate communities competing for activation.

In addition, the typically found facilitation effects for phonology can also be explained by the community structure of the phonological network. Recall that members of a given phonological community share phonological segments. Communities may provide an indirect way to send priming to needed phonology during word retrieval. For example, during a TOT elicitation task, priming of the target word's phonology has been shown to reduce the frequency of TOTs if presented before the word retrieval task and has been shown to increase TOT resolution if presented after the word retrieval task (James & Burke, 2000). In these studies, only partial phonological information, from one phoneme to one syllable, is presented with each prime. These primes likely reside in the same phonological communities as their target word. Therefore, community structure could facilitate retrieval of phonology and reduce word retrieval failures, and lends well for testing with behavioral studies and simulations.

Macro-level of analysis. At the macro-level, we consider the whole network structure. Many of these measures are the average of all the node's micro-level measures (e.g., average clustering coefficient), but additional measures of node location, path length, mixing patterns, and network description (e.g., small-world and scale-free structure) are also used to describe the overall network structure.

All nodes in the network are located in one of three places: the giant component, an island (or smaller component), or as an isolated hermit (Vitevitch, 2008). The giant component is the largest grouping of nodes in the network that are all connected in some way. Islands are separate, smaller components (i.e., fewer nodes than the giant component), where nodes in an island are all connected to each other. Lastly, hermits are nodes that have no connection with any other node in the network; in other words, they are isolates.

A comparison of giant component size of semantic and phonological networks provides interesting insight into the overall connectedness of these networks. In particular, the giant component of a semantic network has been shown to be quite large (e.g., about 96% of all nodes in the network; Steyvers & Tenenbaum, 2005), whereas the giant component of a phonological network is much smaller (e.g., about 34% of all nodes in the network; Vitevitch, 2008). The way in which similarity is defined in semantics versus phonology lends to these stark differences in giant component size. For example, in an association network, where edges are placed between cue and response words, it is much more difficult to get an island and hermit word due to the nature of the association task.

Taking note of not just the size of these components, but also the way the structure of these components influences processing, is important. Little research has explicitly examined how location of nodes influences word processing. However, it is often noted that words located in the giant component tend to be of shorter length, higher word frequency, and earlier age of acquisition than words located in islands or as hermits (Siew, 2013). The traditional psycholinguistic hypothesis would be that words in the giant component should be easier to retrieve and produce given their item-level characteristics. However, a study by Vitevitch and Castro (2015) examining archival picture naming data of healthy older adults and individuals with aphasia, highlight the importance of looking closer at the influence of location on processing. This initial examination showed that words located *outside* of the giant component were easier to name than words located inside of the giant component for both healthy older adults and individuals with aphasia (Vitevitch & Castro, 2015). Further research is needed to test these effects in young adults as a test for the influence of age, as well as using a continuous variable of component size (rather than inside versus outside of the giant component).

A second way to assess the macro-level structure of the network is to determine "distance" measures, like *average shortest path length*. Path length is the number of connections that must be traversed to get from one node in the network to another node in the network (Watts

& Strogatz, 1998). The average shortest path length is computed by taking the average of all the shortest path lengths of all pairs of nodes in the network. Having short average path length suggests that traversing across even a large network can be done very easily by taking "shortcuts."

Both semantic and phonological networks have short average path lengths. In the semantic networks explored by Steyvers and Tenenbaun (2005), the average path length was 3 with a maximum path length of 5, whereas in the phonological network of Vitevitch (2008) the average path length was approximately 6. The smaller average path length of the semantic network may be due to more interconnectivity amongst nodes as compared to the phonological network, possibly due to the constraints of phonology. For example, there are only so many phonemes in the English language and only a set number of ways to combine those phonemes to create English words (i.e., phonotactic constraints).

Some work has been done with semantic networks to assess the influence of distance on word processing through the examination of "near" and "far" neighbors of a target word. Given a target word (e.g., bottle), "near" neighbors (e.g., jar) would be more similar in meaning than "far" neighbors (e.g., skillet). In a picture-word interference paradigm, naming latencies of a target word were slower when presented with a semantically "near" neighbor than a semantically "far" neighbor (Vieth, McMahon, & de Zubicaray, 2014; but see Hutson & Damian (2014) for no effect). Additionally, in a blocked naming task, blocks that contained items from two "near" categories (e.g., body parts and clothing) were named slower than items from two "far" categories (e.g., body parts and vehicles; Vigliocco, Vinson, Damian, & Levelt, 2002). This finding is consistent with the previous hypothesis regarding communities of semantic categories, in that words within a community would facilitate processing and words in different communities would interfere in processing. The addition of distance suggested here would be that closer communities would have more interference than farther communities. Important to note is that these current measures of distance (i.e., "near" vs "far") are based on subjective ratings, rather than defined Network Science measures. By using a Network Science measure, like path length, we can quantify exactly how many connections lie between one node and another in the network, thereby providing a more precise definition of "near" and "far."

Mixing patterns refer to the way in which nodes tend to connect. For example, in a social network, people tend to be friends if they have the same gender, race, or age. Pertinent to language networks, mixing patterns among words are also found. In semantic networks, mixing has been found for a variety of measures including part of speech, valence, dominance, arousal, and concreteness (Van Rensbergen, Storms, & De Deyne, 2015). Participants tend to produce words that are similar on these characteristics as the cue word presented to them and highlight different ways in which "meaning similarity" can be subjectively defined in the network.

Network properties of a node, for example degree, can also be used to describe mixing patterns, and have been found in phonological networks. Two examples of mixing by degree is assortative mixing by degree and disassortative mixing by degree. Assortative mixing by degree is the notion that nodes with high degree tend to connect to other nodes with high degree, whereas disassortative mixing by degree is the notion that nodes with high degree tend to connect to other nodes with low degree (Newman, 2002).

The phonological network of Vitevitch (2008) has assortative mixing by degree, and has been found to influence word retrieval. Specifically, Vitevitch, Chan, and Goldstein (2014) found that participants are more likely to respond with a word of the same degree in a variety of psycholinguistic tasks (e.g., hear a high degree word and respond with a similar sounding high degree word), and may be useful in assessing the definition of "phonological similarity."

Mixing by degree also has important implications for the resiliency of a network; in other words, how connected does the network remain with the removal of nodes. For example, Newman (2002) examined how the targeted removal of nodes by degree (e.g., removing nodes with the highest degree first) affects network structure as compared to the random removal of nodes. Indeed, Newman (2002) found that a network with assortative mixing by degree is more resilient to a targeted attack than a network with disassortative mixing by degree. One explanation for this finding is that assortatively mixed networks have a more highly interconnected giant component with many redundant pathways of connections, and removing one of the high degree nodes will have little impact on processing. However, in the dissassortatively mixed network, the connections of high degree nodes are more likely to be diffused across the network, and their loss will be more detrimental to the network (Newman, 2002). This has also been seen in the relatively consistent average path length of a phonological network after the targeted removal of high degree nodes (Arbesman, Strogatz, & Vitevitch, 2010).

Given that both semantic and phonological networks exhibit assortative mixing by degree, language networks would be hypothesized as being resilient to damage. However, there are changes to language processing with diseases, and even normal, healthy aging. Therefore, it is important to further study the way in which these measures can be used to assess changes in network structure over time and test different models of "damage." For instance, it may not be the case that random or targeted removal of nodes occurs with age or disease (because this would be equivalent to a word or concept being removed from the lexicon), but rather a weakening of connections between nodes (Borge-Holthoefer, Moreno, & Arenas, 2011; Steyvers & Tenenbaum, 2005). One way to assess how these different types of damage to the network manifest in actual behavioral changes would be to conduct computer simulations. For example, we could simulate a network where nodes are removed and compare that to a network where edges are weakened to see which approach better accounts for real data of patients with dementia or aphasia.

Lastly, networks can be classified in different ways based on their macro-level structure, like small-world and scale-free structure. Small-world structure is the notion that despite being large in size (i.e., many nodes), the network is easy to traverse. This is the commonly understood notion of "six degrees of separation" discussed in social psychology (Milgram, 1967), whereby there are, on average, six people between you and any other person in the world. A network is said to have a small-world structure when average path length is approximately equivalent to, but average clustering coefficient is much greater than a comparably-sized random network (Watts & Strogatz, 1998).

On the other hand, scale-free structure is the notion that few nodes have many connections, and many nodes have few connections. Those nodes with many connections are sometimes called "hubs", which have been found to be critical in mechanisms of network growth, network resiliency, and the spread of processing across a network (Albert, Jeong, & Barabasi, 2000; Newman, 2008). A network is said to have a scale-free structure when the degree distribution of the network follows a power-law, which contrasts with the degree distribution of a comparably-sized random network (i.e., same number of nodes and edges, but one where edges are placed randomly) that follows a Poisson distribution (Newman, 2008).

Small-world structure and scale-free structure have been assessed for networks of the mental lexicon. Both semantic networks (Morais, et al., 2013; Motter, et al., 2002) and phonological networks (Vitevitch, 2008) have been defined as having small-world structure, which allows for an efficient and rapid search of the network (Kleinberg, 2000; Vitevitch, 2008; Watts & Strogatz, 1998). Interestingly, though, language networks are more mixed on scale-free structure. Some semantic networks have been shown to exhibit a scale-free structure with degree distributions following a power-law (Steyvers & Tenenbaum, 2005). However, in other semantic networks, the degree distribution is better fit by a logarithmic scale with an exponential cut off (Morais, et al., 2013). And furthermore, the phonological network of Vitevitch (2008) best fits an exponential distribution. For language networks, it seems more plausible to not exhibit scale-free structure. Morais, et al. (2013) argue that there is a limit to the storage and processing of information. Specifically, having many words with few connections would be ideal so as not to have an overly connected network that slows processing, and also having a boundary for the maximum number of connections a node can have is important for capacity limits.

In sum, Network Science measures can be utilized to examine word retrieval in ways that cannot be done by using traditional psycholinguistic approaches, namely through the consideration of how structure influences processing. Another advantage of the Network Science approach is that it can model multiple layers of information simultaneously. Current research on networks of the mental lexicon tend to focus on only one layer of information (i.e., just examination of semantic relationships or just examination of phonological relationships). However, emerging work in Network Science is examining multiplex networks, which are networks that contain two (or more) different types of relationships. Thus, it is possible to create a network that includes representations for both semantic and phonological relationships at the same time, and will provide a more comprehensive model for testing.

Multiplex Network of the Mental Lexicon

As with most traditional psycholinguistic research, examination of networks of the mental lexicon has only focused on one type of relationship at a time. These exclusive analyses may be due to the way that Linguistics categorizes the field (e.g., phonology and semantics are different sub-disciplines) to the “modules” found in most models of speech perception and speech production, or simply for simplicity’s sake. Indeed, Strogatz (2001) highlights that although there are many useful avenues of investigation with networks, different disciplines will suppress some aspects of networks to focus on others. This includes ignoring other potentially relevant “layers” of information (i.e., examination of only one type of edge between nodes) to fully understand what is happening in only one layer.

However, the possibility exists to include multiple types of edges in a network to examine multiple layers simultaneously (see Kivela, et al., 2014). A multiplex network is a specific kind of multilayer network in which all layers share the same nodes (Figure 2). Given that most models of speech perception and production have a notion of a semantic “module” and a phonological “module”, it is important that both are examined simultaneously, where one layer represents semantic relationships and another layer represents phonological relationships. Indeed, one multiplex has been examined thus far that includes aspects of semantic, syntactic, and phonological relationships amongst words (Stella, Beckage, & Brede, 2017), and has been shown to be a better predictor of word acquisition in children than using single-layer networks alone. Therefore, the use of a *multiplex* network will enable a more inclusive examination of the structure of both the semantic and phonological relationships among words, and allow for testing of how this complex structure influences a variety of language processing.

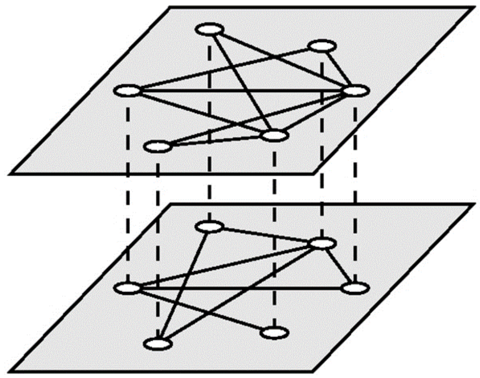

Figure 2. A depiction of a two-layer multiplex network. Reprinted from Gómez, et al. (2013). Each layer is represented by a plane and each plane consists of the same nodes as indicated by the dashed lines. However, edges within a layer can be different as indicated by the solid lines.

Multiplex analysis. There are different ways to visualize and analyze a multiplex network. For example, the multiplex can be constructed to include two types of edges (intra- and inter- layer edges), like done in other types of multilayer networks. Intra-layer edges are those edges placed between nodes within a given layer (e.g., the solid lines in Figure 2), whereas inter-layer edges are edges placed between nodes across layers (e.g., the dashed lines in Figure 2). An alternative method for visualizing and analyzing a multiplex is to reduce the layers into one network by using colored edges (see Figure 3), where a different color is used for the edges of each layer. In the unique case of multiplex networks, inter-layer edges only represent one-to-one mappings of words, and are not often included in network analysis. From this edge-colored multiplex, similar structural measures to that examined with single-layer networks can be obtained and analyzed.

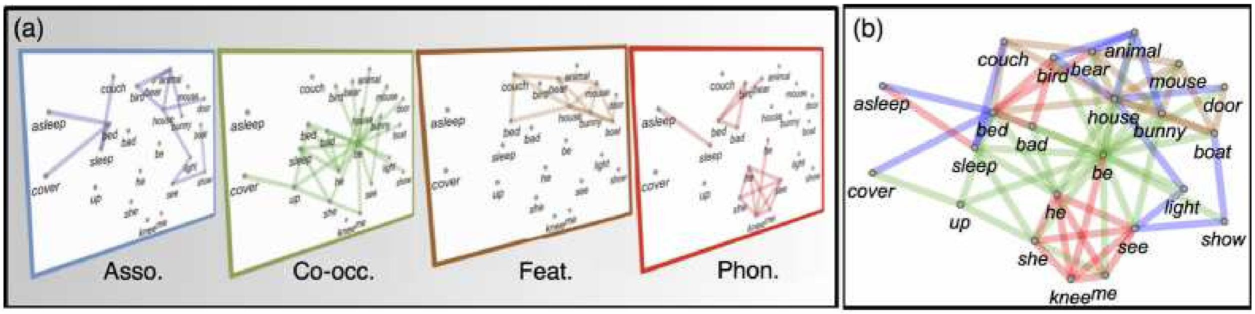

Figure 3. An edge-colored multiplex of the Mental Lexicon. Reprinted from Stella, Beckage, & Brede, (2017). Panel A depicts four single-layer networks: Association, Co-occurrences, features, and phonological similarity. Panel B represents the multiplex network with four edge-colored layers.

First, degree is still a useful measure in a multiplex network and represents the number of neighbors for a given node. Degree can be defined within each intra-layer and calculated as done previously with single-layer networks (Kivela, et al., 2014), allowing for a comparison of degree for a particular node across layers. For example, the semantic degree of a particular node can be compared to its phonological degree. With a multiplex, though, we can also determine *multidegree* (Bianconi, 2013; Kivela, et al., 2014). If a pair of nodes are connected in multiple layers, then a *multilink* can be placed between these nodes. For example, if two nodes are connected in both the semantic layer and the phonological layer, then a multilink would be placed between the node pair. *Multidegree* then is the number of multilinks of a given node. This measure provides some idea of the amount of overlap between the layers of a multiplex. For example, *rat-cat* share both semantic and phonological edges, and would have a multilink, as contrasted to pairs of nodes that are only connected within one layer (e.g., *dog-cat* in the semantic layer and *mat-cat* in the phonological layer).

Clustering coefficient can also be defined in a multiplex, but suffers from a complexity issue of deciding whether to consider edges in one layer or multiple layers (Kivela, et. al., 2014).

Recall that in the single-layer network measure, clustering coefficient is a measure of how connected the neighbors of a node are to one another. Intra-layer clustering coefficients can be calculated from the multiplex and would be identical to the clustering coefficients obtained when examining that layer's single-layer network.

To examine multiplex clustering coefficient that takes into consideration the edges in multiple layers, Cozzo, et al. (2013) suggest an examination of "3-cycles," in which the node of interest and two connected neighbors form a closed triangle, regardless of what layer those edges reside in. This form of clustering coefficient considers how many closed triangles can be created for a particular node. Each closed triangle is formed by taking a total of 3 steps starting and finishing at the node of interest (Figure 4). For example, panel A in Figure 4 shows the standard single-layer closed triangle, where the node of interest forms a closed triangle of edges with two neighbors in the same layer. Panels B-D in Figure 4 show ways in which a closed triangle can be achieved with two layers. Importantly, the use of closed triangles to measure clustering coefficient of a multiplex can help provide insight on how processing moves between layers of a multiplex.

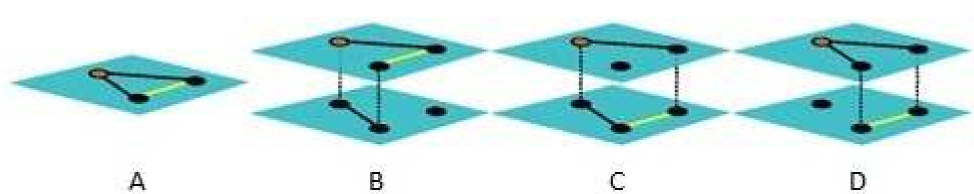

Figure 4. Depiction of closed triangles in single-layer and multi-layer networks. Adapted from Cozzo, et al. (2015). Closed triangles are formed by making 3 steps starting at the orange node, or the node of interest. The solid black lines represent edges between nodes. The yellow lines also represent edges between nodes and indicate the second step in each triangle. Dashed lines indicate identical nodes across layers. Depiction A shows a closed triangle in a single-layer network. Panels B-D show different ways a closed triangle can be completed when considering nodes in two layers.

Careful note must be made regarding transitivity, or how processes move, within the multiplex. Traversing across a multiplex is more complicated than traversing just a single-layer network. Kivela, et al. (2014) state that there are two important theoretical questions that must be answered: 1) does moving from one layer to another layer count as a “step” in the process, and 2) are intra-layer “steps” equivalent across layers. These questions need careful theoretical consideration. Regarding the first question, in the current models of speech perception and production, moving between a semantic “module” and a phonological “module” is assumed to incur some kind of “cost.” Therefore, the traversal between semantic and phonological layers in the network should be considered as a “step.” For example, in tip-of-the-tongue (TOT) states, the inability to successfully activate all necessary phonological nodes leads to a disruption in word retrieval (Burke, MacKay, Worthley, & Wade, 1991). Activating the semantic information is done successfully, but that activation does not spread successfully in the phonological system. Moving activation from the semantic to the phonological system resulted in a “cost” that impacted successful word retrieval.

Regarding the second question, a step within one layer is likely equivalent to a step in another layer. Specifically, moving from one word to another in the semantic system would have the same “cost” as moving from one word to another in the phonological system. Simulations within a multiplex network representing semantic and phonological relationships may provide additional insights into the costs associated with these intra- and inter- layer steps.

Shortest path lengths can also be determined from a multiplex with the same cautions regarding transitivity. The freedom to move between layers may shorten average path length overall. But, it is important to note that theoretically this may not be applicable in all language processing contexts. The extent to which processing travels between semantic and phonological systems, for example, is a contested issue in psycholinguistics. For example, it is the case that processing travels top-down from the semantic system to the phonological system during word production, but the amount of interaction back and forth between the semantic and phonological systems is thought to be minimal, if it is possible at all. However, such switching may occur when a particular search strategy is used, for example in a language- or word-game task.

In addition to these network measures that are similar between single-layer and multiplex networks, there are additional measures that can be calculated to determine the amount of overlap between layers in a multiplex. For example, the *degree of multiplexity* assesses the ratio of node pairs with multilinks to the total number of all node pairs in the network (Kivela, et al., 2014). If there is high overlap between two layers in the multiplex, then the degree of multiplexity will be close to 1. However, if there is little overlap between the layers in the multiplex then the degree of multiplexity will be close to 0.

In sum, there are several measures that can be used to investigate the structure of a multiplex, and these measures are complex due to accounting for multiple layers. It is important to note that the usage of multiplex networks and measures is still an emerging area within the larger discipline of Network Science. Using a well-studied domain, like language processing, it would be possible for network scientists to further develop these multiplex measures. However, the purpose of this paper is to begin the initial construction of a multiplex for the mental lexicon, one that includes semantic and phonological relationships. By using some of the established measures and a visual exploration of the multiplex, I will be able to see how the structure of these two layers are different and how they overlap.

Given that words hold both semantic and phonological information, it would be appropriate to have a multiplex that includes the same nodes (i.e., words) in both layers and represents both aspects of the words. In addition, for this multiplex, it is important to also consider having similar operational definitions for edges in each layer. This will help ease interpretation of any effects derived from this multiplex. Semantic association data has been made available to help achieve this goal. Chapter 2 describes the collection and analysis of phonological association data, which is then used to create a phonological association network in Chapter 3. In addition, to comparing different types of phonological network structures in Chapter 3, a comparison of age-related phonological networks will be done in Chapter 4. Then, Chapter 5 consists of an analysis of the multiplex structure that includes both semantic and phonological associations. Finally, a discussion of the current findings and their limitations, as well as future directions, is given in Chapter 6.

Chapter 2: Phonological Association Task

Introduction

One common way to define edges in semantic networks is to use association norms (Nelson, et al., 2000). That is, participants are given a cue word and are asked to respond with the first word that comes to mind. Nodes are the cue and response items, and edges are placed between cue-response pairs (e.g., De Deyne & Storms, 2008; Morais, et al., 2013; Steyvers & Tenenbaum, 2005). These semantic association networks are created from participant-driven data, rather than corpora or printed materials (e.g., dictionary or thesaurus).

On the other hand, the commonly studied phonological network of Vitevitch (2008) uses the one-phoneme metric to define edges. Nodes are words from the Merriam Webster Pocket Dictionary, and edges are placed between words that differ by one phoneme (through addition, substitution, or deletion; Luce & Pisoni, 1998). This phonological network, although shown to explain several psycholinguistic findings for word recognition and production (Vitevitch, Goldstein, Siew, & Castro, 2014; Vitevitch & Luce, 2016), is not derived from participant data, and to some may seem like an arbitrary measure of phonological similarity (e.g., why only a one phoneme difference, rather than two or more?).

Conducting the phonological association task will provide two benefits. First, similar operational definitions of edges in the semantic and phonological layers of the multiplex analyzed in a later chapter will allow for an easier interpretation of findings, particularly when comparing words that are semantically related to words that are phonologically related. Second, a comparison of the phonological association network to the well-studied phonological network of Vitevitch (2008) will help determine if different operational definitions of "phonological similarity" have significant influences on the overall structure of the network.

Method

Participants. A total of 1,051 participants completed the phonological association task. However, data from only 1,018 participants is described here. Participants were dropped according to data pre-processing steps detailed later. Table 1 provides detailed demographic information of the 1,018 participants. All participants (37.6% male) were native English speakers from the United States. Participants ranged in age from 18 to 99 years ($M = 42$, $SD = 16$). Education level of participants ranged from high school to doctorate. Participants were recruited from Amazon Mechanical Turk (87.3%) and from the University of Kansas SONA-Systems pool of undergraduate psychology students. Amazon Mechanical Turk participants received monetary compensation, whereas SONA participants received partial course credit. It should be noted that there were no participants over the age of 25 recruited from the SONA-Systems pool. Therefore, all of the adults in middle to late adulthood participating in this study were sampled using Amazon Mechanical Turk. Given the level of computer literacy required to use Amazon Mechanical Turk it is not very likely that the middle to late adulthood participants had significant cognitive deficits.

Table 1.

Number of Participants by Age and Education Level.

	Education Level					
Age (years)	High School	Some College	Bachelor's Degree	Master's Degree	Doctorate Degree	TOTAL
18 to 24	8	146	14	1	0	169
25 to 34	24	73	124	16	2	239
35 to 44	18	48	63	25	6	160
45 to 54	21	39	41	17	9	127
55 to 64	28	84	79	32	9	232
65+	11	26	30	17	7	91
TOTAL	110	416	351	108	33	1018

Determination for the appropriate sample size of this study was made based on the sample sizes and cue-response parameters of semantic association tasks. Sample size does vary among studies, for example, from 300 (Nelson, et al., 2000) to more than 70,000 (De Deyne, et al., 2013) participants. Additionally, these studies vary in the number of cue words each participant receives and the number of responses participants are expected to provide. For example, Nelson, et al., (2000) presented 60 cue words and requested 2 responses, whereas De Deyne, et al. (2013) presented on average 18 cue words (ranged from 7 to 30) and requested 3 responses. Importantly, presenting fewer cue words and requesting fewer responses will require more participants. In the present case 60 cue words were presented and 3 responses were requested as a way to compensate for the relatively small sample size that serves as a starting norm dataset for phonological associations.

Materials. In order to create the multiplex in a later chapter, it is important to have cue words that are used to obtain both the semantic and phonological association responses. A semantic association dataset is already available with permission from S. De Deyne. The same cue words from the semantic association dataset were used as the cue words for the phonological association task in this study. The original set of cue words from S. De Deyne had 10,050 words. Items that were more than one word (e.g., apple juice), proper nouns (e.g., America), represented in different spellings (e.g., labour vs labor; the American version was maintained), or inappropriate (e.g., taboo words) were removed from the list, leaving 9,371 words.

Qualtrics was used to administer the phonological association task and data was analyzed using R version 3.4.0 (R Core Team, 2017). An informed consent statement was given to participants, and they indicated consent before the study began. Demographic questions capturing age, education, and whether they were a native English speaker were shown. Instructions for participation were then displayed, followed by the phonological association task. During the phonological association task, instructions remained on the screen followed by a cue word and text box for the input of responses. Cue words that were homographs (n=154 words) were presented with a sample sentence using the intended pronunciation (see Appendix B). Intended pronunciation was determined by the semantic association data collected by S. De Deyne.

Procedure. After providing consent and answering basic demographic questions, participants were presented with the following instructions: *Your task is to provide up to three words that **SOUND** similar to the word provided. Type those responses that <u>immediately come to mind</u>. Do not spend too long on any one item. If you do not know the word provided or do not have any responses that immediately come to mind, please type **"DK"**. Note that it is acceptable to type only one or two words, but no more than three. Please use commas (,) to separate your responses.*

After reading the instructions, participants then moved on to the association task. A cue word was presented on the screen and a text box was available for participants to provide their responses. In addition, the instructions remained on the screen for reference if needed. Each participant received a total of 60 random cue words, and each cue word was responded to by at least 6 participants (max = 8).

Results

Cleaning of Data. Before performing any analyses on the network, preprocessing was completed like that done with the semantic association data (De Deyne, et al., 2013). First, data from participants who responded with 65% or more "Don't Know" responses were removed. De Deyne, et al. (2013) had a cutoff of 50% or more "Don't Know" responses; however, suppressing semantic associates and providing phonological associates is a harder task, hence the increase in allowable "Don't Know" responses. Also, participants who responded with semantic associates were also excluded. Responses from 15 random cue words for each participant were examined to determine if semantic associates were provided that contained no phonological overlap (e.g., CLIMATE-WEATHER). If those responses were primarily semantic associates, then their data were excluded. These criteria helped to ensure that participants completed the phonological association task according to task instructions with effort, and resulted in the loss of 33 participants' data.

Next, all responses were examined to ensure they were real words by comparing responses to the commonly used word corpora of SUBTLEX-US (Brysbaert & New, 2009), Kučera & Francis (1967), and CELEX (Baayen, Piepenbrock, & van Rijn, 1993). Spelling errors were corrected if it was clear what the intended word was (e.g., recieve). Any remaining words not found in the word corpora were examined further. In some cases, the word was a real word as found in the dictionary, but had an uncommonly used morphological form (e.g., anointer). These words were not listed in the corpora, but were retained. The remaining words not found in the word corpora and determined not to be a real word by checking in the online Merriam-Webster dictionary (www.merriam-webster.com) were removed. In addition, two-word responses, proper nouns, and inappropriate responses (including those words that were removed from the original cue list) were also excluded. This resulted in 8,575 responses (~10% of the data) being excluded from analysis.

After removal of data as indicated above and only examining actual word responses (i.e., removal of "Don't Know" responses), there was a total of 77,451 cue and response pairs generated by participants. Of these responses, there were 56,747 unique cue and response pairs. Table 2 provides the number of participants providing each unique cue-response pair. It can be seen that a large number of unique cue-response pairs were made by only one participant (76.8%). What is often done in association datasets is to remove those responses that are not frequently made as a way to ensure that responses are reflective of "most" people. For example, the minimum cut-off would be to remove those responses that are not generated by at least two people (Nelson, et al., 2000). However, a number of appropriate responses would likely be discarded in the current study should this minimum cut-off be used. Rather than following removal cut-offs as done in previous work, weights on the edges will provide an alternative way to maintain the data but acknowledge the frequency of responses. The weighting of edges will be included in the phonological association network analyzed in the next chapter.

Table 2.

Count and Frequency of the Number of Participants Generating the Same Response for a Particular Cue Word.

Number of Participants with Same Response to a Particular Cue	Number of Unique Cue and Response Pairs	Percentage of Unique Cue and Response Pairs
1	43,602	76.8%
2	8,223	14.5%
3	3,087	5.4%
4	1,206	2.1%
5	476	0.8%
6	134	0.2%
7	18	<0.1%
8	1	<0.1%

Description of Words. There was a total of 20,575 unique words in the phonological association dataset. This set of words consisted of 9,329 of the original cue words (9,298 were responded to as cues and 7,669 were provided as responses) and another 11,246 new words.

Two standard psycholinguistic measures were calculated for the phonological association dataset: length and word frequency. Words varied in length as measured by the number of phonemes from 1 to 15 (M = 5.8, SD = 2.0). Word frequency was determined by extracting from standard corpora the log of word frequency, since it is known that word frequency is highly skewed. First, the log of word frequency was taken from SUBTLEX-US (Brysbaert & New, 2009), with 557 words not found in SUBTLEX. Kučera and Francis (1967) was used to obtain the log of word frequency for another 174 words, and CELEX (Baayen, et al., 1993) was used to obtain the log of word frequency for another 302 words. The remaining 81 words were verified as real words in the dictionary and given a value of 0 for the log of word frequency. SUBTLEX-US was chosen as the starting word corpus because it has been shown to be more reliable in predicting participant performance on standard psycholinguistic tasks (Brysbaert & New, 2009). In the phonological association dataset, the log of word frequency varied from 0 to 6.33 (*M* = 1.88, *SD* = 0.92).

An additional 40 cue words had no responses to them and were not provided as responses. These words ranged in the number of phonemes from 3 to 10 (*M* = 6.5, *SD* = 1.6) and ranged in the log of word frequency from 0 to 3.26 (*M* = 1.98, *SD* = 0.76).

Age differences were also examined given that word finding problems increase with age and may impact performance during the association task. Three age groups were examined that resulted in a fairly even distribution of the sample: young, early middle, and late middle adults. The young adult group ranged in age from 18-34 years (*M* = XX, *SD* = XX). The early middle adult group ranged in age from 35-54 years *M* = XX, *SD* = XX). And, the late middle adult group ranged in age from 55 to 99 years (*M* = XX, *SD* = XX). Although there were some participants over the age of 75 that could represent an older adult category, their number is small. Thus, the sample is more representative of an early middle and a late middle adulthood range.

First, the time to complete the task was examined. On average, the task took 24.64 min (*SD* = 16.38). Young adults took on average 22.26 min (*SD* = 17.90), early middle adults took 24.55 min (*SD* = 12.38), and late middle adults took 27.72 min (*SD* = 17.03). This finding suggests that the task took longer as age increased, and could be reflective of increased word finding difficulties (but see below for an alternative explanation).

The proportion of responses by response number was also examined. Table 3 provides the proportion of responses for each age category and response number. Young adults provided 28,467 responses, early middle adults provided 22,638 responses, and late middle adults provided 26,346 responses. The proportion of first responses appears to decrease with age, while the proportion of second and third responses appears to increase with age. This may indicate that older adults are more likely to provide multiple responses than younger adults, and may contribute to the increased amount of time spent on the task.

Table 3.

Proportion of Responses for Each Age Category and Response Number.

Age Category	First Responses	Second Responses	Third Responses
Young Adults	55.8%	29.3%	14.9%
Early Middle Adults	53.0%	30.8%	16.2%
Late Middle Adults	51.3%	31.3%	17.4%

Cue and Response Pairs. Of interest to this study is the amount of phonological overlap between cue-response pairs. The number of phonemes different between each cue-response pair was calculated following the one-phoneme difference metric of Luce and Pisoni (1998), where phoneme changes include addition, substitution, and deletion. Cue-response pairs differed between 0 (e.g., be → bee) and 11 phonemes (e.g., especially → unfortunately). Using the same three age groups, the number of phonemes different for cue-response pairs was examined. Young adults ranged in number of phonemes different from 0 to 11, with a mean of 2.38 phonemes. Early middle adults ranged in the number of phonemes different from 0 to 11, with a mean of 2.36 phonemes. Late middle adults ranged in the number of phonemes different from 0 to 10, with a mean of 2.32 phonemes. Therefore, age did not impact the range or average number of phonemes different. Table 4 lists the count and frequency of phoneme differences for cue-response pairs, regardless of age.

Table 4.

Count and Frequency for the Number of Phonemes Different in Cue and Response Pairs.

Number of Phonemes Different	Count	% of All Cue-Response Pairs	Cumulative % of All Cue-Response Pairs
0	655	0.8%	0.08%
1	27566	35.6%	36.4%
2	22234	28.7%	65.1%
3	11608	15.0%	80.1%
4	7460	9.6%	89.8%
5	4299	5.5%	95.3%
6	2178	2.8%	98.1%
7	938	1.2%	99.3%
8	372	0.5%	99.8%
9	107	0.1%	99.9%
10	30	<0.1%	99.9%
11	4	<0.1%	100.0%

Even though over 60% of cue-response pairs generated were different by only one or two phonemes, there was a large range of phoneme differences, sparking further interest in the cue-response pairs. Several additional analyses were conducted to better understand the relationship between cue-response pairs provided by participants.

First, recall that participants could provide up to three responses to a given cue word. There has been debate over the utility of this particular protocol of allowing multiple responses versus one response. De Deyne, et al. (2013) argue that multiple responses provide richer data that captures a larger portion of the mental lexicon, including weaker edges between words that might not be captured if only collecting one response (i.e., strong edges only between a pair of words). However, Nelson, et al. (2000) argue that although weak edges are added, the data becomes less reliable in capturing similarity between words. For example, participants may be making new responses in relation to their own earlier responses, rather than the cue word.

To address this concern, the number of phonemes different between cue and response pairs for each response number was examined, as well as the number of phonemes different between first and second responses, second and third responses, and first and third responses for a cue word. If participants are making later responses in relation to their earlier responses, and not the cue word, it would be expected that the number of phonemes different between cue and response will increase as response number increases. A one-way ANOVA examining the mean number of phonemes different between a cue and each response number was statistically significant, $F(2, 22401) = 15.75$, $p < .0001$. Tukey's HSD showed that the mean number of phonemes different was higher for cue-second responses ($M = 2.83$, $SD = 1.52$) than cue-first responses ($M = 2.70$, $SD = 1.40$) and cue-third responses ($M = 2.75$, $SD = 1.63$), ps < .01, with no significant difference between cue-first responses and cue-third responses. This finding suggests that second responses were furthest from the cue word in terms of phoneme overlap.

Additional tests were conducted to further examine the number of phonemes different between responses of a cue word. If participants are making responses in relation to earlier responses, rather than the cue word, the response to response phoneme difference should be smaller than the cue to response phoneme difference. In other words, later responses should be more phonologically different from the cue word than to an earlier response. When examining second responses made to cues, a t-test showed that the number of phonemes different between cue-second response pairs ($M = 2.83$, $SD = 1.52$) and first-second response pairs ($M = 2.84$, $SD = 1.61$) were not different, $t(15359) = 0.40$, $p = 0.69$. However, when examining third responses made to cues, a one-way ANOVA was statistically significant, $F(2, 16454) = 9.86$, $p < .0001$. The number of phonemes different between cue-third response pairs ($M = 2.75$, $SD = 1.63$) and first-third response pairs ($M = 2.70$, $SD = 1.66$) were larger than second-third response pairs (M

= 2.61, *SD* = 1.65), *p*s < .05, with no significant difference between cue-third response pairs and first-third response pairs.

These findings taken together suggest that second responses were likely made in relation to the cue word presented, but third responses may have been made in response to earlier responses. However, it is also interesting that even though third responses may have been influenced by previous responses, they were also more likely to be closer in phonology to the cue word than second responses.

In addition to the large range in the number of phonemes different between cue and response pairs, the small proportion of responses that are one phoneme different (36.4%) is also surprising. Previous studies eliciting phonological similarity associations from participants have found high rates of one phoneme differences between cue-response pairs. For example, Luce & Large (2001) found 71% of responses to nonwords to be one phoneme different, and Vitevitch, et al. (2014) found 74.5% of responses to real words to be one phoneme different. In addition, Vitevitch, Goldstein, & Johnson (2016) report proportions of responses at each difference in number of phonemes, with 84.2% of the responses being one phoneme different from the cue. However, these previous studies have only used cue words that are 3 phonemes in length, whereas the present data used cue words that ranged from 1 to 14 phonemes in length.

To assess whether the present findings compare to previous findings regarding the number of phonemes different between cue-response pairs, only cue words that are three phonemes in length and their responses were analyzed. There was a total of 15,697 cue-response pairs where the cue word was 3 phonemes in length. Responses in this subset of the data ranged from 0 to 9 phonemes (*M* = 1.34, *SD* = 0.67). Table 5 provides the proportion of this subset of data for each number of phonemes different, along with the reported results from Vitevitch, et al.

(2016). Interestingly, the findings for this subset of the data resemble the previous findings. In this data, 69.5% of the cue-response pairs had a one phoneme difference between them, with 23.5% of the cue-response pairs having a two-phoneme difference between them.

Table 5.

Count and Frequency for the Number of Phonemes Different between Cue and Response Pairs where Cue Length = 3.

Number of Phonemes Different	Number of Cue-Response Pairs	Frequency of Cue-Response Pairs	Vitevitch, Goldstein, & Johnson (2016)
0	287	1.8%	
1	10914	69.5%	84.2%
2	3695	23.5%	13.5%
3	627	4.0%	2.1%
4	124	0.8%	
5	32	0.2%	
6	10	< 0.1%	.07%
7	5	< 0.1%	
8	1	< 0.1%	.07%
9	2	< 0.1%	

Given that these findings for cue words with a length of three phonemes are consistent with previous findings, it was of interest to further understand how the length of cue words impacts the number of phonemes different between cue and response pairs. The large range in number of phonemes different in the present phonological association data may be influenced by cue word length, given that cue words were as long as 14 phonemes. Specifically, as the number of phonemes increase in a cue word, the more phonemes that must be held constant to maintain phonological overlap. This decreases the number of possible responses that a participant might be able to provide for a cue word. Indeed, a Pearson's correlation shows that as the number of phonemes in the cue word increases, the difference in the number of phonemes between cue and response also increases, *r* = 0.68. The correlation is plotted in Figure 5 with the best fitting line of $y = 0.5416x – 0.3485$ and $R^2 = .47$. In sum, these results provide evidence that the number of phonemes different between cue and response pairs may be driven by factors like cue length and response number.

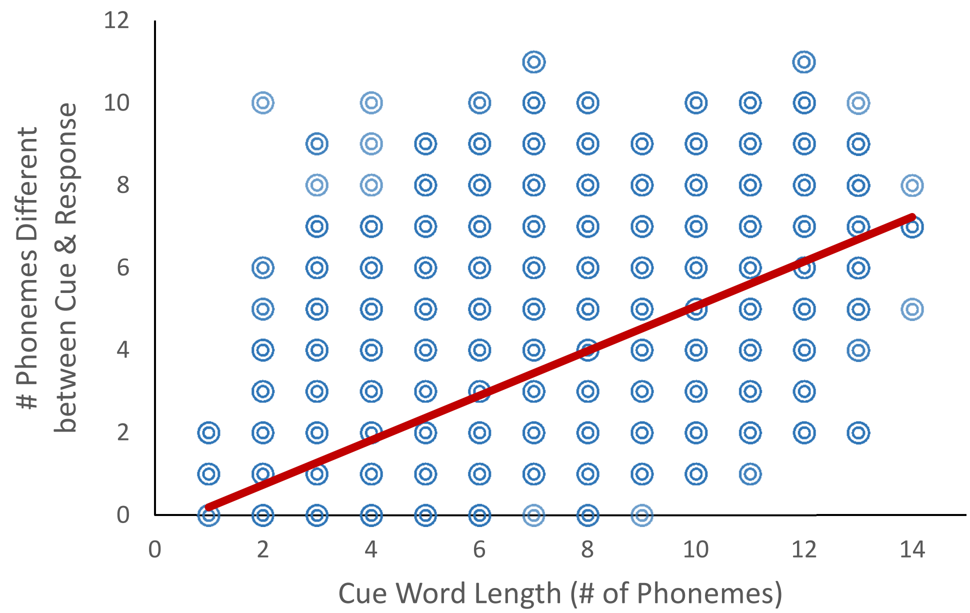

Figure 5. Scatterplot of cue word length and number of phonemes different between cue and response. The Pearson's correlation r = .68. The best fitting line is given in red, with darker blues representing a larger proportion of responses.

Lastly, an analysis was done that considers the nature of the association task itself in producing similarity associations. Indeed, an association task is hypothesized to capture semantic relationships when individuals are asked to provide the first word that come to mind. In the present phonological association task, participants are still required to provide responses that immediately come to mind, but must also determine whether those responses sound similar to the cue word. In this case, it may be possible that participants are still having semantic associates come to mind from which phonological associates are filtered out, rather than the intention that only words that sound similar come to mind. It is important to note, though, that it is not problematic for a participant to provide a response that is both phonologically and semantically related. However, if participants selected responses on the phonological association task strictly from the semantic associates that came to mind, the evidence should support a high rate of semantic-phonological overlap in cue-response pairs.

To address this issue, each unique cue-response pair from the phonological association task was compared to the cue-response semantic association pairs provided by S. De Deyne. Of the 56,754 unique cue-response pairs provided in the phonological association task, only 4,034 pairs were also found in the semantic association data provided by S. De Deyne, or 7.1% of phonological cue-response pairs (e.g., abdomen → abdominal). These findings suggest that although some cue-response pairs were provided on both the phonological and semantic association tasks, participants primarily provided only phonological associates, not semantic associates that also happened to be phonologically related.

Discussion

The purpose of collecting phonological associations was to assess phonological similarity between words using participant-driven data as a new way to construct a phonological network. Studies have shown that a large proportion of phonological association responses tend to differ by only one phoneme (Luce & Large, 2001; Vitevitch, et al., 2014, Vitevitch, et al., 2016). In the present data, only a small proportion of phonological associates differed from the cue word by one phoneme. Instead, a range of phoneme differences from 0 to 11 phonemes was found, with the majority of the cue-response pairs being different by 4 phonemes or less.

There were several factors considered for why there was such a large range in phoneme differences. First, response number may have been a factor, as previous work is conflicted on the validity of using multiple responses in an association task. In the present data, it appears that response number may have had an influence, albeit small. In particular, it was found that second responses had less phonological overlap to the cue word than first responses. Interestingly, though, third responses had the same amount of phonological overlap to the cue word as first responses, but were still more phonologically similar to the second response. These findings are suggestive that earlier responses may have influenced later responses.

Another factor that was considered as an influence on the large range in phoneme differences was the length of cue words. The previous findings showing a high proportion of phonological associates that are only one phoneme different was conducted with cue words that are only three phonemes in length. However, the present study had cue words that ranged from 1 to 14 phonemes. Indeed, when examining the cue words with a length of three phonemes only, the present data is consistent with previous work. In addition, it was also found that as cue word length increased, the number of phonemes different between cue-response pairs also increased. This finding supports the notion that longer cue words require a larger proportion of their phonemes to be maintained, and reduces the number of possible options available. Therefore, having longer cue words in the phonological association task was a contributing factor to the large range in the number of phonemes different between cue and response words.

A final factor was considered in the present study as an influence on the large range in phoneme differences, namely task strategy. In particular, association tasks have typically been used to capture semantic relationships. However, association tasks have also been used to capture phonological similarity between words by modifying the task instructions. In these studies, including the present study, participants are instructed to provide responses that immediately come to mind that also sound similar. This additional instruction may or may not be enough to reduce the production of semantic associates. On the one hand, participants may first have semantic associates come to mind from which they select phonological associates. On the other hand, participants may have phonological associates come to mind, and some may also happen to be semantically related. If the former case is true, a larger proportion of cue-response pairs would be both semantically and phonologically related. However, in the present study, there was only a small proportion (7.1%) of cue-response pairs that were also given in the semantic association task. This suggests that participants were not influenced by semantic associations, and completed the task as instructed.

Although the range in number of phonemes different between cues and responses does raise the question as to how participants completed the task, the findings from this study suggest that participants followed task instructions. In addition, age was found to not be a contributing factor to the number of phonemes different. Time spent on the task did increase with age, but also the proportion of second and third responses. Future research can continue to examine strategies that participants employed in order to complete the phonological association task, as well as other demographic factors, like education. For example, it may have been the case that participants focused on morphology or rhyming, particularly for the longer words. In these instances, only a small portion of the word would be maintained (e.g., a stem or affix), and may have led to some of the higher phoneme differences found.

In sum, association data has been widely used to understand the organization of semantic representations, and will also serve useful in understanding phonological associations. In order to further understand phonological similarity and its representation in the mental lexicon, Network Science tools will be used in the next chapter.

Chapter 3: Comparison of Phonological Association and One-Phoneme Metric Networks

Introduction

Network analyses have been used to better understand the way in which words are represented and structured in the mental lexicon, and the influence of that structure during language processes. In network analyses, words are represented as nodes and edges are placed between words that are related. Associations, provided by participants through an association task, have been used to construct and analyze semantic networks (e.g., De Deyne, et al., 2013; Hills, et al., 2009; Morais, et al., 2013; Nelson, et al., 2000; Steyvers & Tenenbaum, 2005). Phonological associations, however, have not been collected on a large scale, and thus have not been used to create phonological networks.

From the phonological association data analyzed in Chapter 2, a phonological association network was created. This network structure will be compared to the well-studied phonological network of Vitevitch (2008) that defined phonological similarity using a one-phoneme difference. In addition, the phonological association network is constructed using participant-driven data, whereas the phonological network of Vitevitch (2008) is constructed using a corpus. This comparison between different types of phonological networks is important for better understanding how phonological similarity is represented in the mental lexicon. Indeed, the phonological association network might capture novel aspects of phonological structure.

Method

The phonological association data described and analyzed in Chapter 2 was used to construct the phonological association network analyzed in the present chapter. The nodes in this network are the cue and response words, and edges are placed between cue and response pairs. In addition, this data will also be used to construct a one-phoneme metric network that is comparable to the phonological network of Vitevitch (2008). The nodes in this network are the cue and response words, and edges are placed between any two words that differ by one phoneme through addition, substitution, or deletion (Luce & Pisoni, 1998). It's important to note that these networks are defining edges in different ways for different purposes. The phonological association network is only considering participant-driven data, whereas the one-phoneme metric network assumes that any cue word responded to and any response made would exist in the mental lexicon. Therefore, one-phoneme difference edges can be placed between two response words or between two cue words in addition to being placed between cue and response pairs.

To create the networks, two important decisions must be made concerning the way edges are placed between nodes in the network. The first decision is whether to place "arcs" or "edges" between nodes. An arc provides information about directionality. In the case of the association data, an arc could be placed from the node of a cue word *to* the node of a response word since the cue produced the response (and not the other way around). An edge, on the other hand, suggests there is a symmetrical relationship between the two nodes.

The second decision is whether to include weights on edges. Weights would provide information about the strength or frequency of the relationship between two nodes. In the case of the association data, weighting captures the frequency of cue-response pairings. In other words, cue-response pairs that are made by multiple participants would have a weight approaching a value of 1, whereas cue-response pairs made by only one person would have a weight close to 0. On the other hand, unweighted edges are assumed to be of equal strength.

Following Vitevitch (2008), the one-phoneme metric network in this study used undirected, unweighted edges. However, for the phonological association network, directed, weighted edges were used between each unique cue and response pair. Importantly, weighting was calculated by taking the number of a particular response given for a cue word divided by the number of presentations of that cue word. For example, for the cue word ABLE, which was presented 7 times, and the response of LABEL, which was given by 3 participants, the weight between the cue-response pair of ABLE-LABEL would be 3/7, or 0.43. In this dataset, weights ranged from 0.125 to 1.000 (M = 0.217, SD = 0.126), and a histogram of the weights is provided in Figure 6. For reference, a weight less than 0.2 would signify cue-response pairs that were only provided by one participant.

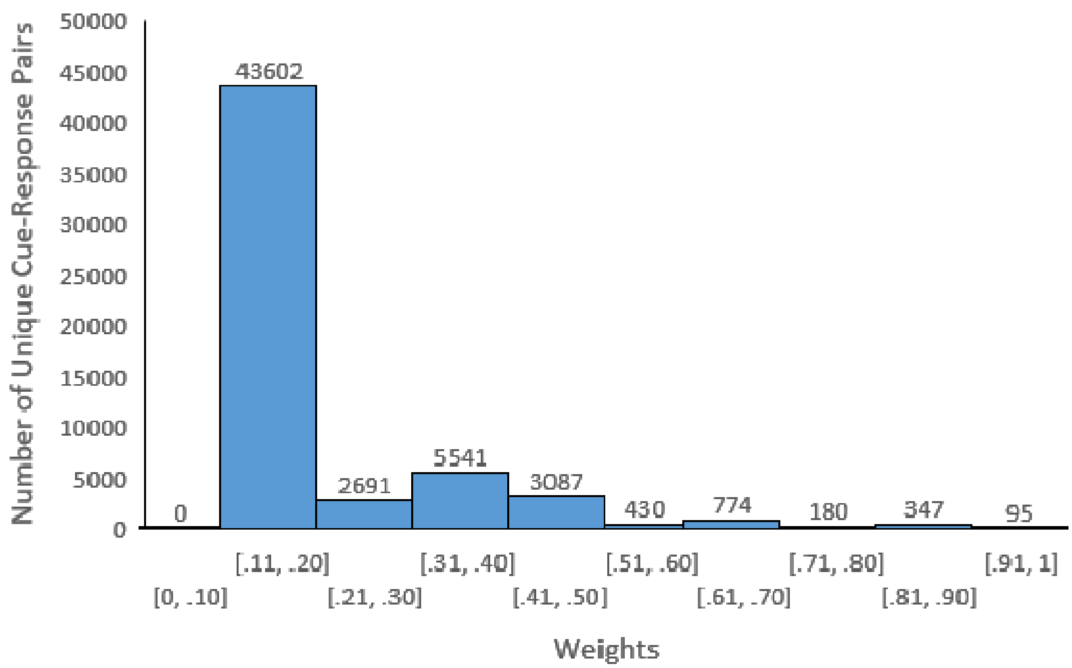

Figure 6. Histogram of weights. Weighting between a cue-response pair is determined by taking the number of times a given response was made for a particular cue word divided by the number of times that cue word was presented.

Finally, common practice in network analysis is to compare the network of interest to a comparably-sized random network, or a network that has the same number of nodes and edges, but where edges are placed randomly. The phonological association and one-phoneme metric networks were compared to comparably-sized random networks (i.e., the same number of nodes and edges). These random networks were only used to determine "small-worldness" of the network of interest. Network generation and analysis were conducted using the igraph package (Csardi & Nepusz, 2006; Ognyanova, K., 2017) in R (R Core Team, 2017).

Analysis

As stated previously, there are several standard measures used to describe and compare networks. This study focused on basic description of the overall network by calculating macro- and meso- level measures, and included weighting of edges in the calculation of these measures. Macro-level descriptions of the network, like small-world and scale-free structure, were determined by analyzing the average shortest path length, average clustering coefficient, and degree distribution of the network. Additional descriptive measures, like location of nodes in the network, mixing by degree, and community structure were used to further describe the macro- and meso- levels of the network. Each of these measures were calculated for the phonological association network, the one-phoneme metric network, and comparably-sized random networks, and compared to the phonological network of Vitevitch (2008). Table 6 presents the results for these network measures for each network type.

Table 6.

Network Structure Measures for the Phonological Association Network, One-Phoneme Metric Network, and Phonological Network of Vitevitch (2008).

Network Measures	Phonological Association Network (PAN)	One-Phoneme Metric Network (1PN)	Vitevitch (2008) One-Phoneme Metric Network[1]
Network Size	Nodes = 20,617 Edges = 56,747	Nodes = 20,575 Edges = 57,042	Nodes = 19,340
Location of Nodes[2]	GC = 20,253 (98.2) Islands = 322 (1.6) Hermits = 42 (0.2)	GC = 10,481 (50.9) Islands = 3,347 (16.3) Hermits = 6,747 (32.8)	GC = 6,508 (33.7) Islands = 2,567 (13.3) Hermits = 10,265 (53.1)
Small World Structure[3]	Avg. Path Len = 9.80 Avg. C = 0.12 S = 724.79	Avg. Path Len -= 6.46 Avg. C = 0.16 S = 1157.50	Avg. Path Len = 6.05 Avg. C = 0.13 S = 1064.38
Scale-Free Structure[4]	P. L. RMSE = 0.64 Exp. RMSE = 0.03	P. L. RMSE = 0.12 Exp. RMSE = 0.02	P. L. RMSE = 0.09 Exp. RMSE = 0.01
Mixing by Degree[5]	$r = 0.44$, $p < 0.0001$	$r = 0.67$, $p < 0.0001$	$r = 0.62$, $p < 0.0001$
Community Structure[6]	70 Communities Mod = 0.86	37 Communities Mod = 0.68	17 Communities Mod = 0.66

[1]Measures reported are obtained from Vitevitch (2008) for all but community size, which was obtained from Siew (2013).
[2]GC = Giant Component, with proportion of nodes in paraentheses
[3]Average Shortest Path Length (Avg. Path Len.) and Average Clustering Coefficient (Avg. C), and Small-world-ness (S) from Humphries & Gurney (2008).
[4]Scale-Free Structure is determined by comparing the Root Mean Square Error (RMSE) of the Power-Law (P.L.) function to the degree distribution, with an alternative fit given with the Exponential (Exp) function
[5]Mixing by Degree is determined by examining the correlation between the degree of a node and each of its neighbors.
[6]Modularity (Mod.) is a measure of the significance of community structure in the network, and values above 0.3 are considered significant (Clauset, Newman, & Moore, 2004).

Phonological Association Network. The analysis of the structure of the phonological association network follows that laid out by Vitevitch (2008), and includes a description of community structure. Recall that the phonological association network is created by representing words as nodes and placing an edge between each cue and response word pair. This definition leads to a network containing 20,615 nodes and 56,754 edges. These nodes resided in one of three places: the giant component, an island, or as a hermit. There were 20,253 nodes (98.2%) in the giant component, 3222 nodes (1.6%) located in islands, and 40 hermit nodes (0.2%). There were 95 islands that ranged in size from 2 to 9 nodes (see Figure 7). Interestingly, islands were organized by phonological overlap in both the initial and rhyme position of words, often overlapping through a suffix (see Table 7).

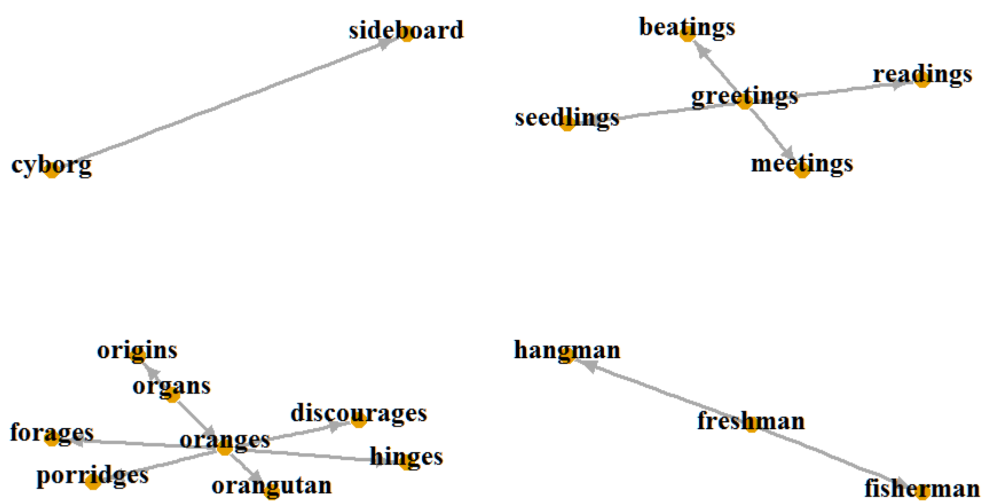

Figure 7. Example islands from the Phonological Association Network.

Table 7.

Proportion of Islands in the Phonological Association Network with Different Types of Overlap.

Type of Overlap	Proportion of Islands
Phonological Overlap	
Alliteration	45.3%
Rhyme	49.5%
Near Rhyme	5.3%
Morphological Overlap	
Prefix	2.1%
Suffix	61.1%
Stem	18.9%

Note Islands can overlap by more than one type.

An analysis was conducted to determine if the phonological association network would be classified as having small-world structure. Recall that having a small-world structure indicates that the network is easy to traverse despite its large size, and is identified by having a similar average shortest path length and larger average clustering coefficient than a comparably-sized random network (Watts & Strogatz, 1998). In order to calculate the average shortest path length and the average clustering coefficient, only those nodes and edges in the giant component were considered, as this is the largest, fully connected component of the network.

The average shortest path length of the phonological association network was 9.80, whereas the average shortest path length of the comparably-sized random network was 9.67. Using network analysis convention, where the difference in values is no greater than 1.5 times in magnitude, the average shortest path length values for the phonological association network and the random network were not significantly different. The average clustering coefficient of the phonological association network was 0.12, whereas the comparably-sized random network had an average clustering coefficient of 0.0002. The average clustering coefficient values for the phonological association network and the random network were significantly different by several magnitudes according to network analysis convention. In addition, a statistical measure of "small-world-ness" was calculated following Humphries & Gurney (2008), where values greater than 1 indicate a small-world network. The phonological association network had a value of 724.79. Therefore, these measures indicate that the phonological association network has a small-world structure.

Next, an analysis was conducted to determine if the phonological association network could be classified as having a scale-free structure. Recall that having a scale-free structure suggests that few nodes have many edges (i.e., hubs) and many nodes have few edges. This is indicated by the degree distribution following a power-law function when plotted on a log-log scale. Figure 8 displays the log-log plot for the degree distribution of the phonological association network. The power-law function was best fit by the equation $y = 4.70x^{-2.62}$, RMSE = 0.64, whereas the exponential curve was best fit by the equation $y = 0.17e^{-0.18x}$, RMSE = .03. Since the exponential curve better fits the data than the power-law function, the phonological association network does not have a scale-free structure.

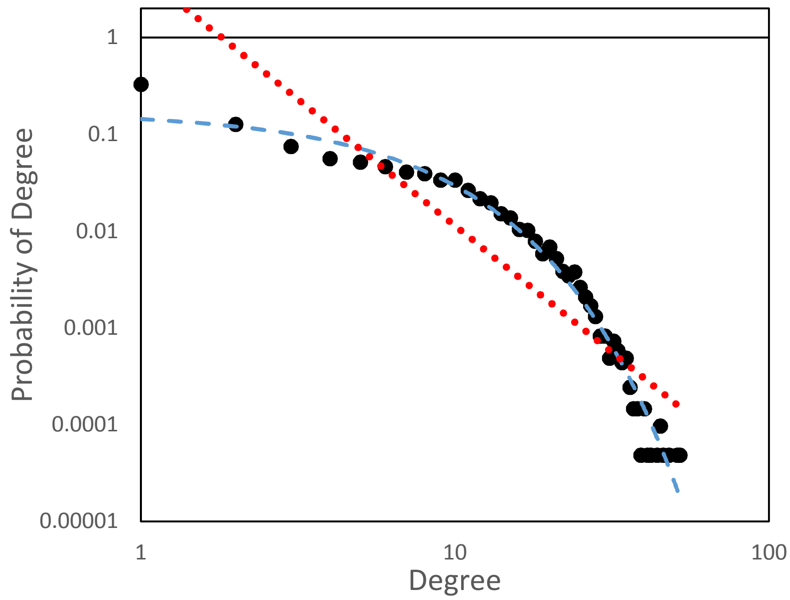

Figure 8. Log-log plot of degree distribution for the Phonological Association Network. The power-law function is represented in red and the exponential curve is represented in blue.

Mixing by degree was also examined in the phonological association network. Recall that assortative mixing by degree occurs when nodes with high degree tend to be connected to other nodes with high degree. On the other hand, disassortative mixing by degree occurs when nodes with high degree tend to be connected to nodes with low degree. To determine the kind of mixing pattern of the phonological association network, a Pearson's correlation between a node's degree and each of its neighbor's degree was examined. A correlation of r (56,515) = 0.44, $p <$ .0001, was found suggesting that an assortative mixing by degree pattern exists in the phonological association network.

Finally, the community structure of the phonological association network was examined. In total, there were 70 communities in the giant component as determined by the Louvain method (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008) with a modularity of 0.86. A modularity value higher than 0.3 is indicative of significant community structure (Clauset, Newman, & Moore, 2004). Figure 9 depicts different communities in the giant component by color. These communities ranged in size from 8 to 1,060 nodes (*M* = 289.33, *SD* = 228.53). Communities overlapped in several ways phonologically and/or morphologically, resulting in smaller groupings of nodes organizing within a community (see Figure 10).

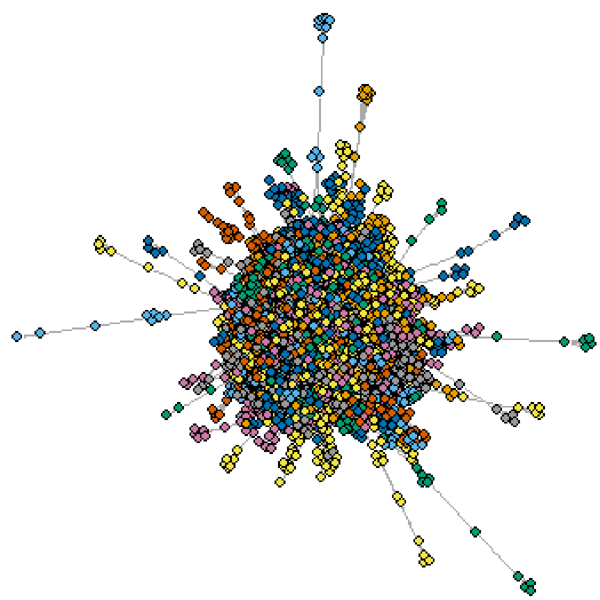

Figure 9. Giant component of the Phonological Association Network. Color represents communities.

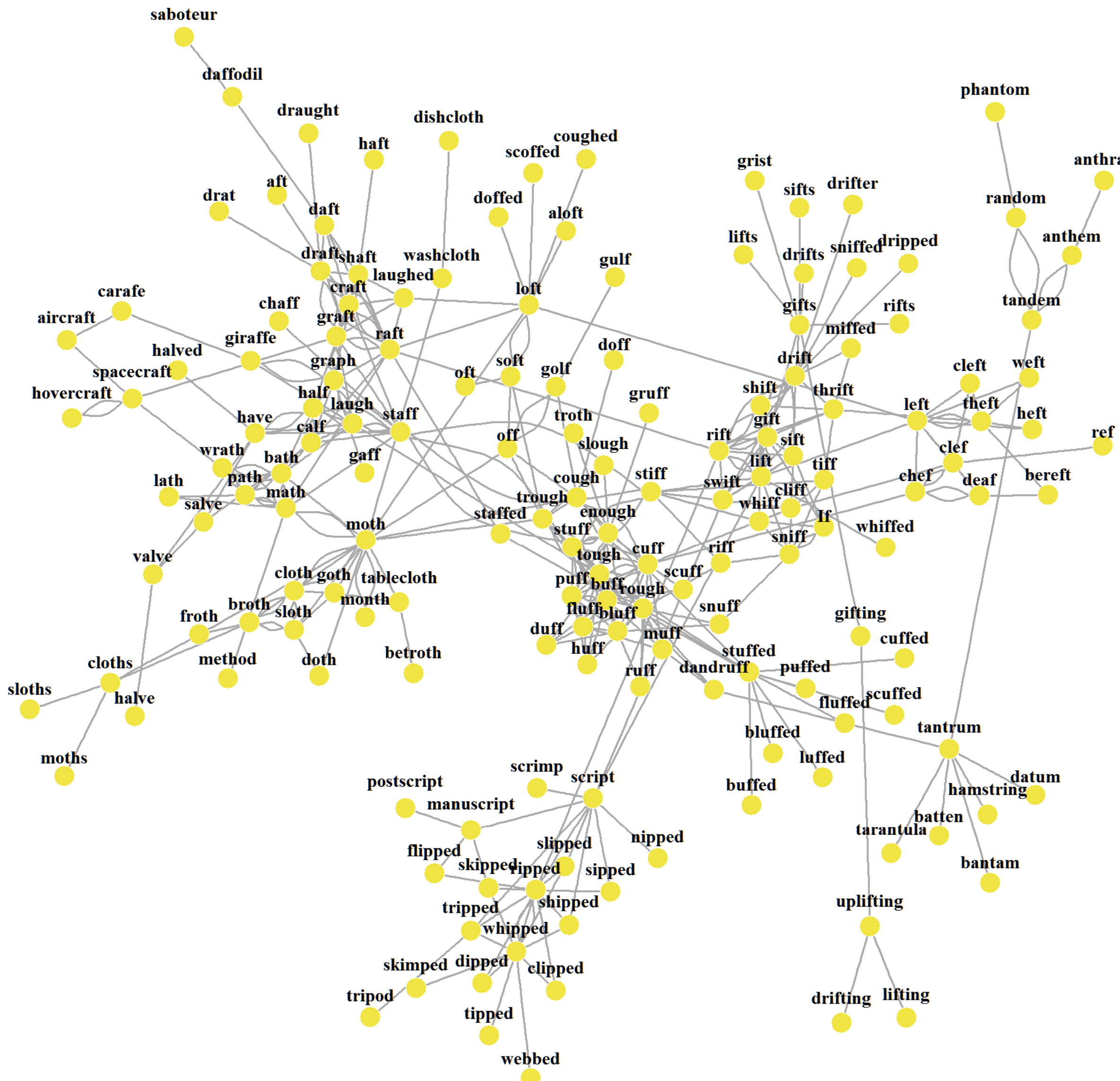

Figure 10. Example community from the Phonological Association Network.

One-Phoneme Metric Network. The analysis of the structure of the one-phoneme metric network follows that laid out by Vitevitch (2008) and as done above for the phonological association network. Recall that the one-phoneme metric network is created by representing words as nodes and placing an edge between each pair of words that differ by one phoneme.

Only cue words that were responded to and all responses given were used as nodes in this network. In other words, the hermit nodes of the phonological association network were not included because there are multiple reasons why participants did not respond to those items during the phonological association task (e.g., did not know the word, did not have any associates come to mind). Therefore, this definition leads to a network containing 20,575 nodes and 57,042 edges. These nodes resided in one of three places: the giant component, an island, or as a hermit. There were 10,481 nodes (50.9%) in the giant component, 3,347 nodes (16.3%) located in islands, and 6,747 hermit nodes (32.8%). There were 1,244 islands that ranged in size from 2 to 77 nodes (see Figure 11). Interestingly, islands were mostly organized by phonological overlap in the initial phoneme position, and often involved a stem that was consistent between words (see Table 8).

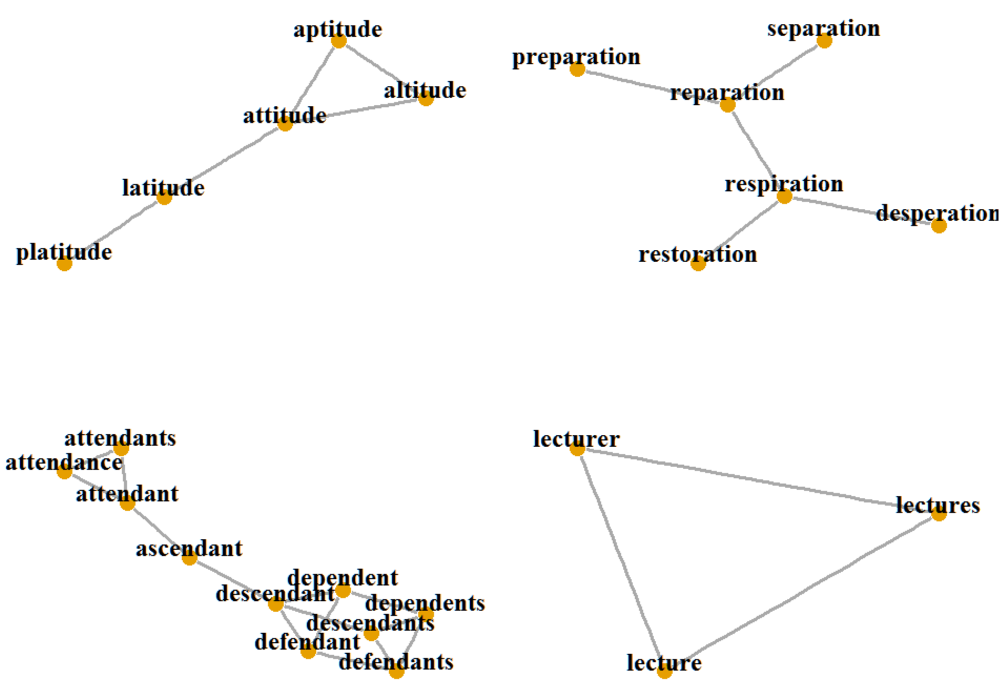

Figure 11. Example islands from the One-Phoneme Metric Network.

Table 8.

Proportion of Islands in the One-Phoneme Metric Network with Different Types of Overlap.

Type of Overlap	Proportion of Islands
Phonological Overlap	
Alliteration	85.7%
Rhyme	26.8%
Morphological Overlap	
Prefix	1.7%
Suffix	20.1%
Stem	66.8%

Note Islands can overlap by more than one type.

An analysis was conducted to determine if the one-phoneme metric network would be classified as having small-world structure. The average shortest path length of the one-phoneme metric network was 6.46, whereas the average shortest path length of the comparably-sized random network was 6.01. Again, using network analysis convention where the difference in values is no greater than 1.5 times in magnitude, the average shortest path length values for the one-phoneme metric network and the random network were not significantly different. The average clustering coefficient of the one-phoneme metric network was 0.16, whereas the comparably-sized random network had an average clustering coefficient of 0.0003. The average clustering coefficient values for the one-phoneme metric network and the random network were significantly different by several magnitudes according to network analysis convention. In addition, "small-world-ness" (Humphries & Gurney, 2008) for the one-phoneme metric network was 1157.50. Therefore, these measures indicate that the one-phoneme metric network has a small-world structure.

Next, an analysis was conducted to determine if the one-phoneme metric network could be classified as having a scale-free structure. Figure 12 displays the log-log plot for the degree distribution of the one-phoneme metric network. The power-law function was best fit by the equation $y = 1.13x^{-1.77}$, RMSE = 0.12, whereas the exponential curve was best fit by the equation $y = 0.10e^{-0.11x}$, RMSE = .02. Since the exponential curve better fits the data than the power-law function, the one-phoneme metric network does not have a scale-free structure.

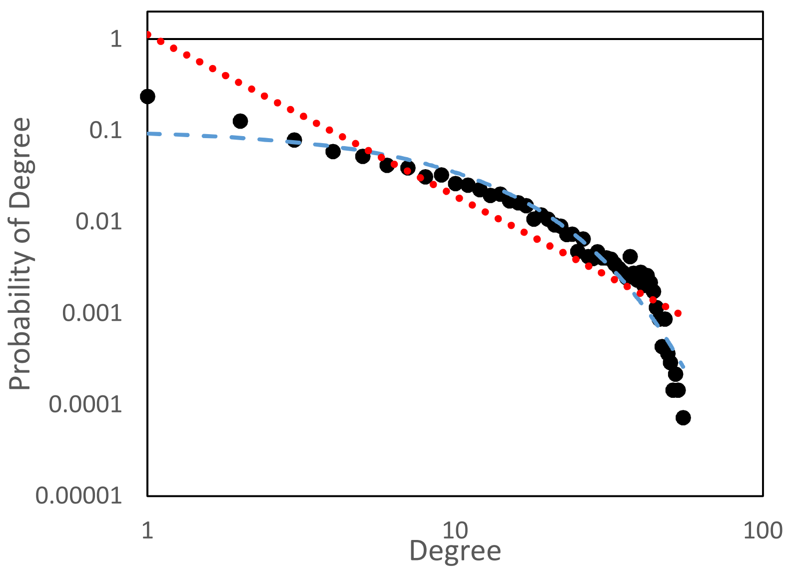

Figure 12. Log-log plot of degree distribution for the One-Phoneme Metric Network. The power-law function is represented in red and the exponential curve is represented in blue.

Mixing by degree was also examined in the one-phoneme metric network. A Pearson's correlation between a node's degree and each of its neighbor's degree was examined. A correlation of r (54,648) = 0.67, $p < .00001$, was found suggesting that an assortative mixing by degree pattern exists in the one-phoneme metric network.

Finally, the community structure of the one-phoneme metric network was examined. In total, there were 37 communities in the giant component as determined by the Louvain method (Blondel, et al., 2008) with a modularity of 0.68. A modularity value higher than 0.3 is indicative of significant community structure (Clauset, et al., 2004). Figure 13 depicts different communities in the giant component by color. These communities ranged in size from 6 to 1,054 nodes (*M* = 283.27, *SD* = 324.99). Similar to the phonological association network, communities overlapped in several ways phonologically and/or morphologically, resulting in smaller groupings of nodes organizing within a community (see Figure 14).

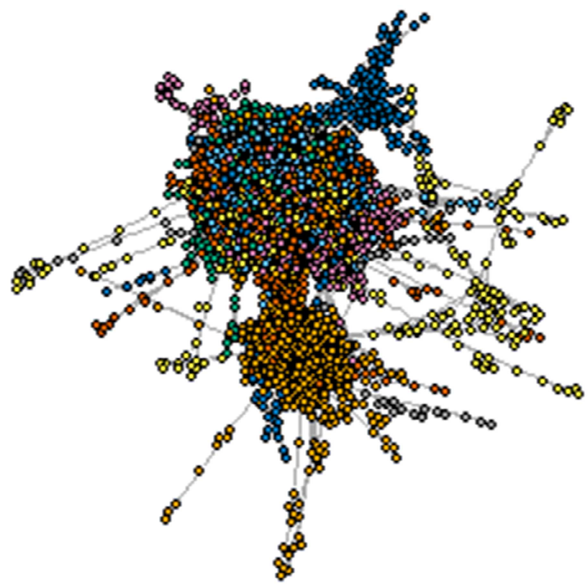

Figure 13. Giant component of the One-Phoneme Metric Network. Color represents communities.

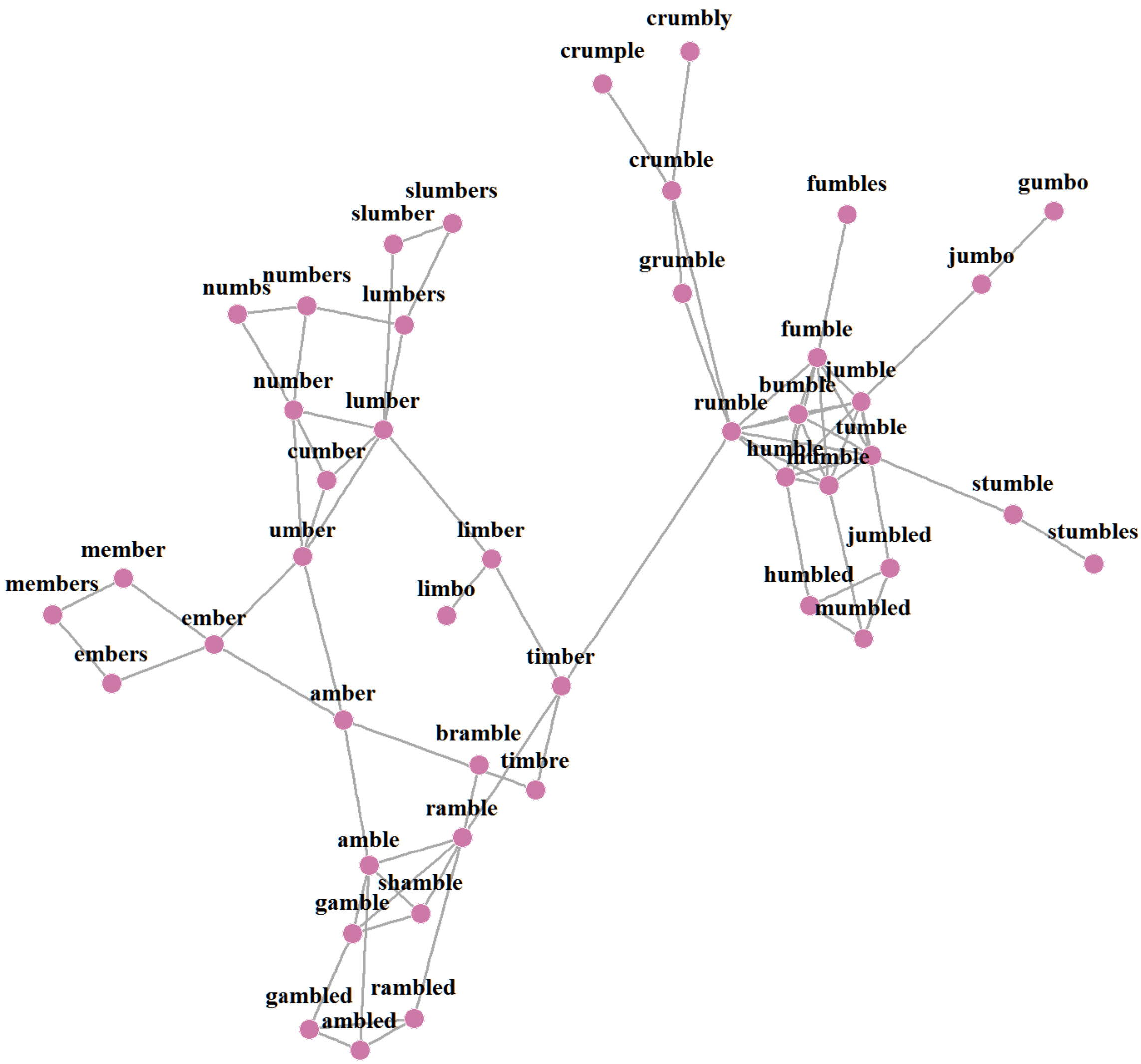

Figure 14. Example community from the One-Phoneme Metric Network.

Correlations between Networks. The previous network examinations focused on macro- and meso- level network structure. However, it is also important to compare an individual word's network structure in the phonological association network to the same word's network structure in the one-phoneme metric network. Differences in the structure of a word between two networks may provide additional insight into how phonological similarity is represented in the mental lexicon. Specifically, the location of a word in each network was examined, as well as a word's degree and clustering coefficient in each network.

First, the location of words in each network was examined, as well as how that location may have differed (or remained the same) between networks. Location of words was categorized into either located in the giant component, an island, or as a hermit for each network. Location of nodes from the phonological association network to the one-phoneme metric network could change in one of five possible ways: from the giant component to an island, from the giant component to a hermit, from an island to the giant component, from an island to a hermit, or remained in the same location in both networks. Note that hermit words in the phonological association network were not included in the one-phoneme metric network, and therefore, there is no possible change of a hermit to an island or of a hermit to the giant component in this analysis. The proportion of nodes for each type of location change is given in Table 9. Interestingly, half of the nodes remained in the same location for the two networks. Not surprisingly, a large portion of nodes “broke away” from the giant component of the phonological association network into islands or hermits in the one-phoneme metric network. The one-phoneme metric network has a “stricter” definition of phonological similarity reducing the likelihood of an edge between two nodes.

Table 9.

Proportion of Nodes for Each Type of Location Change from the Phonological Association Network to the One-Phoneme Metric Network.

Type of Location Change	Count of Nodes	Proportion of Nodes
Giant Component to Island	3280	15.9%
Giant Component to Hermit	6605	32.1%
Island to Giant Component	113	0.5%
Island to Hermit	142	0.7%
Same Location	10435	50.7%

Next, the degree of words in each network was examined. The phonological association network had an average degree of 5.52 (*SD* = 5.84), whereas the one-phoneme metric network had an average degree of 5.54 (*SD* = 8.58). A Pearson's correlation between the degree of a word in the phonological association network and the degree of the same word in the one-phoneme metric network showed that degree between networks was correlated, r (20575) = .51, $p < .0001$ (see Figure 15). Therefore, words have similar degree in each network.

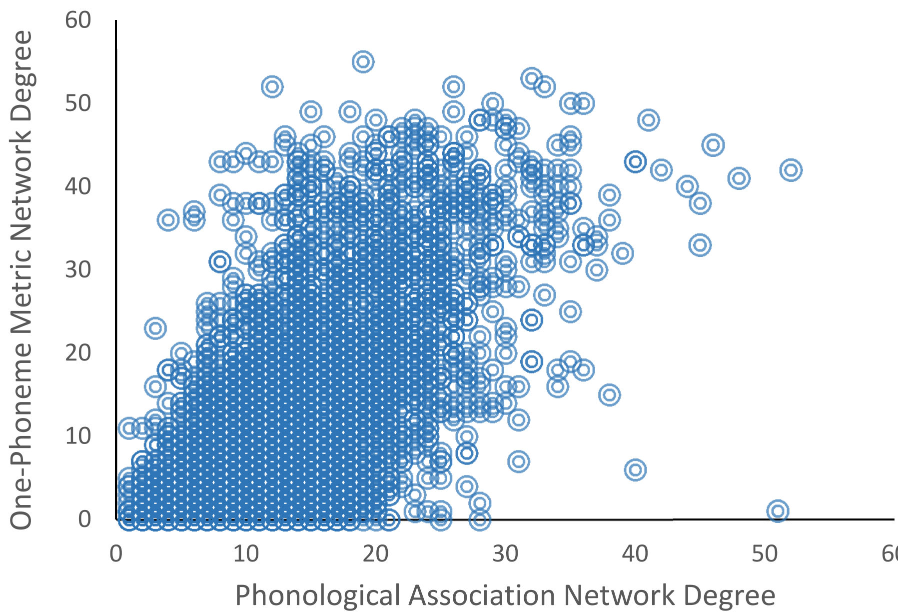

Figure 15. Scatterplot of Phonological Association Network and One-Phoneme Metric Network degrees.

However, thus far degree has been discussed simply as the number of immediate neighbors for a given node. In the phonological association network, though, directed edges were used providing a means to examine two sub-types of degree: in-degree and out-degree. In-degree is the number of edges pointing toward a given node, whereas out-degree is the number of edges pointing from a given node. Since edges were placed from a cue word to a response word, that edge would be considered in the out-degree value for the cue word, but would be considered in the in-degree value for the response word. Therefore, cue words are the only words that would have an out-degree value. However, it is possible that cue words could have in-degree if given in response to another cue word, as well as the novel word responses provided by participants.

This fact is important given that 55% of words used to create the phonological association network would only have one contributing sub-type in the overarching degree value discussed previously, which may bias the results of degree correlation. Therefore, an additional analysis was done examining the correlation of degree between the phonological association network and the one-phoneme metric network for cue words only (whose degree includes the possibility of both in- and out- degree). A Pearson's correlation between the degree of a cue word in the phonological association network and the degree of the same cue word in the one-phoneme metric network showed that degree between networks was correlated, r (9298) = .71, $p < .0001$. Indeed, the r value increased from the previous analysis, supporting the notion that words without the possibility of having both sub-types of degree (i.e., response words) may have influenced the degree correlation findings.

Lastly, the clustering coefficient of words in each network was examined. A Pearson's correlation between the phonological association network and the one-phoneme metric network was small., r (20575) = .14, $p < .0001$. One possibility for this small correlation could be due to the large number of words located as hermits, with no clustering coefficient, in the one-phoneme metric network. An additional Pearson's correlation was conducted excluding hermit words in the networks. Interestingly, the correlation between clustering coefficient of words in the phonological association network and the one-phoneme metric network was smaller in this analysis, r (13828) = .06, $p < .0001$. Therefore, it is interesting to note that the clustering coefficient of words in the phonological association network may only be slightly similar to the clustering coefficient of words in the one-phoneme metric network.

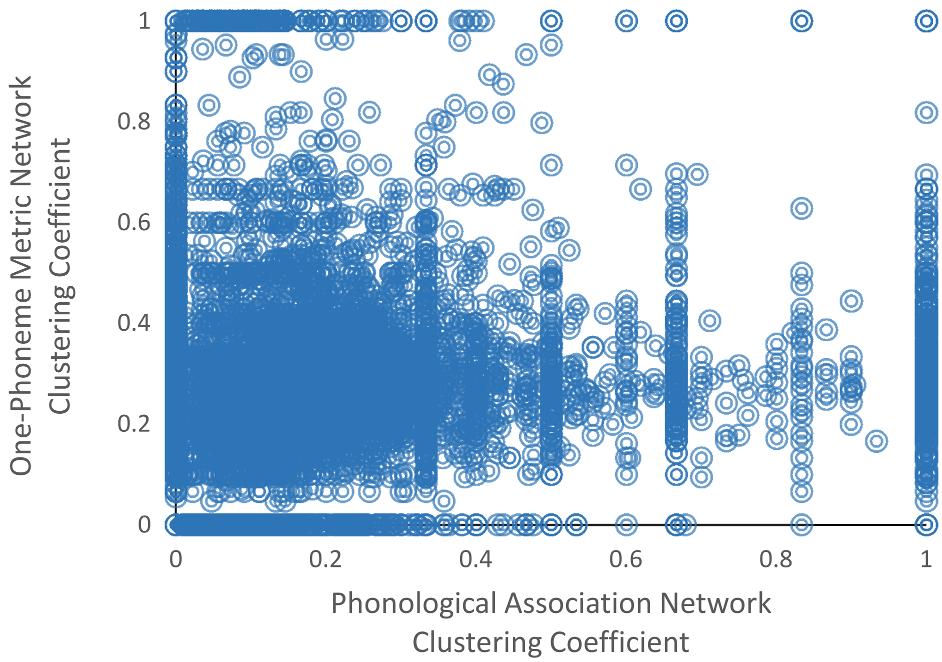

Figure 16. Scatterplot of Phonological Association Network and One-Phoneme Metric Network clustering coefficients.

Discussion

The comparison of different networks provides a means to understand factors that may influence the representation of phonological similarity in the mental lexicon. In this study, three phonological networks were examined that varied in source of data and edge definition: phonological association network, one-phoneme metric network, and phonological network of Vitevitch (2008). The phonological association network and the one-phoneme metric network were derived from collected participant responses on an association task (Chapter 2), whereas the phonological network of Vitevitch (2008) was created using a corpus of words. Both the one-phoneme metric network and the phonological network of Vitevitch (2008) use an edge definition defined by a one phoneme difference (through addition, substitution, or deletion) between a pair of words (Luce & Pisoni, 1998), whereas the phonological association network places edges between cue and response pairs.

It is interesting to note that the phonological association network and the one-phoneme metric resulted in similar macro-level structures. Both networks would be described as having small-world structure, but not scale-free structure. Indeed, the exponential curve better fits both networks with similar RMSE values than a power-law function, which is the signature of a scale-free network. In addition, the phonological association and one-phoneme metric networks have small-world-ness values greater than 1 (724.79 and 1157.50, respectively). These findings for the phonological association and one-phoneme metric networks are consistent with the macro-level structure of the well-studied phonological network of Vitevitch (2008), providing evidence that phonological association data can be used to construct a meaningful representation of phonological similarity.

Despite having some similar macro-level features, the location of nodes differed remarkably between the phonological association network and the one-phoneme metric network, despite using the same data source. The phonological association network had substantially more nodes in the giant component (98.2%) than the one-phoneme metric network (50.9%). This is further supported by the fact that the one-phoneme metric network had substantially more nodes in islands (16.3%) and hermits (32.8%) than the phonological association network (1.6% and 0.2%, respectively). These differences in node location may be due to the way in which edges are defined in each network. In the phonological association network, edges are placed between cue and response pairs resulting in a higher probability of edge placement than in the one-phoneme metric network where edges are more restricted by the one phoneme difference definition. Indeed, examination of how node location changes from the phonological association network to the one-phoneme metric network supports this assertion.

Next, both the phonological association network and the one-phoneme metric network show evidence of assortative mixing by degree, or that high degree nodes tend to be connected to other high degree nodes. However, it appears that the one-phoneme metric network is more influenced by this assortative mixing than the phonological association network as evidenced by its higher Pearson's *r* correlation (0.67 and 0.44, respectively). One explanation for this finding could be that the nodes in the phonological association network had higher degree (i.e., more unique responses) than nodes in the one-phoneme metric network, which could reduce the correlation. However, it was found that node degree was similar between the phonological association network and the one-phoneme metric network, which fails to support this possibility.

Another potential explanation for the difference in Pearson's *r* strength for assortative mixing by degree between the phonological association network and the one-phoneme metric network could be related to the number of nodes lacking the possibility for out-degree in the phonological association network. Recall that only cue words have the potential for both in- and out- degree, whereas response words (that are not cues) will only have in-degree. This means that a degree for response words may be underestimated, and thus influencing differences in degree findings. Indeed, it is often part of the protocol to obtain associations from the responses that were provided in a second run of the task as a way to combat this issue. Although not perfect as new responses can still be generated, this helps to address the concern of edge directionality and degree being biased when examining only one run of the association task. Therefore, continued data collection is needed to determine if the difference in assortative mixing between the phonological association network and the one-phoneme metric network is due to bias in the present data, or if this is a true difference between these two networks.

Lastly, both the phonological association network and the one-phoneme metric network have significant community structure in their giant components with many communities. Interestingly, both networks have more communities than that of the phonological network of Vitevitch (2008). The greater number of communities is likely due to the larger number of nodes located in the giant components of the one-phoneme metric network, and especially the phonological association network. Communities in the one-phoneme metric network and phonological association networks were large, on average, and nodes in each community overlapped in several ways through phonological position (maintenance of initial or rhyme) and morphology (maintenance of stems, prefixes, or suffixes). These findings are important for understanding how phonological similarity is represented in the mental lexicon.

In conclusion, the comparison between the phonological association network, the one-phoneme metric network, and the phonological network of Vitevitch (2008) provide evidence that phonological association data is a viable source for understanding how phonological similarity is represented and organized in the mental lexicon. The differences that do emerge between these networks show that additional factors must be taken into consideration when understanding the representation of and processing in the phonological system of the mental lexicon. For example, the position of phonological overlap and morphology organized islands and communities in the giant components of the networks, which often resulted in greater than a one-phoneme difference.

The results of this study provide new insight into the structure of phonological similarity in the mental lexicon. However, these measures are descriptive in nature. Future research should directly test the influence of these network structures in behavioral experiments. Importantly, the phonological network of Vitevitch (2008) has already been tested with a variety of spoken word recognition and production tasks (Vitevitch & Luce, 2016; Vitevitch, et al., 2014). The same studies could be done with the phonological association network structure measures to determine if the differences in network structure influence language processing. For example, degree of a word was similar between the phonological association network and the one-phoneme metric network, but clustering coefficient was not strongly correlated for a word in the phonological association network and the one-phoneme metric network. Therefore, testing the effect of phonological association clustering coefficient is particularly important for replication of previous findings, and understanding of association data influences the representation of phonological similarity in the mental lexicon.

Chapter 4: Age-Related Phonological Networks

Introduction

The way in which phonological similarity is represented in the mental lexicon may be influenced by several factors. In the previous chapter, an examination of two edge types (associations and one-phoneme differences) provided evidence that phonological similarity can be defined by a range of phoneme differences and may be influenced by location of phonological overlap and morphology. Another factor that may influence the structure of the phonological network is age. Indeed, vocabulary knowledge increases with age (Verhaegen, 2003), which would be predicted to change the underlying structure of the mental lexicon.

Recent work examining semantic networks across adulthood provides evidence that semantic networks change with increasing age (Dubossarsky, De Deyne, & Hills, 2017). Using semantic association data, they found that semantic networks had a U-shaped trajectory across the lifespan with participants aged 10 – 84 years. Specifically, they reported network structure change for in- and out- degree, average shortest path length, and clustering coefficient. In- and out- degree were small in adolescence, increased sharply and remained high across early adulthood, and finally began to decline across mid- to late adulthood. Average shortest path length was high in adolescence, declined sharply and remained low across early adulthood, and finally began to increase across mid- to late adulthood. Finally, clustering coefficient decreased across adolescence, early, and middle adulthood, with a slight increase in late adulthood. Taken together, these findings suggest that the semantic association network is sparse in adolescence, grows increasingly denser across early adulthood, and becomes sparser again into late adulthood (Dubossarsky, et al., 2017).

In addition to examining semantic network change across the lifespan, other types of language networks should also be examined. This study will examine the structure of the phonological association network across adulthood. It is not as clear how phonological network structure might be influenced by age. On the one hand, vocabulary increases with age resulting in more nodes and edges being added to the network. The addition of these nodes and edges could change the overall structure of the entire mental lexicon. However, words added later in life tend to be longer and of lower frequency, and are more likely to reside in islands or as hermits in the phonological network (Siew, 2013). Therefore, the addition of these words would have little influence on macro- and meso- level network measures that mostly only consider the giant component, resulting in the appearance of little network change over time.

Other age-related factors, like hearing loss and cognitive decline, might also affect phonological network structure. For example, older adults are known to perform less well on speech recognition tasks than younger adults, and this difference could be due to changes in auditory perception and cognition with age (Humes & Dubno, 2009; Schneider, Pichora-Fuller, & Daneman, 2009). In addition, evidence already shows that network structure also influences spoken word recognition, such that older adults have more difficulty identifying words with many phonological neighbors than words with few phonological neighbors (Sommers, 1996). These findings suggest that it becomes more difficult to disambiguate similar sounding words with increasing age, which may affect the responses that older adults provide on a phonological association task and the structure of phonological association networks across adulthood.

Understanding how the mental lexicon changes with age is important given the current behavioral findings of language processes that change with age. For example, word retrieval tends to be more disrupted in older adults than younger adults as evidenced by an increase in tip-of-the-tongue (TOT) states (Burke, et al., 1991). TOTs are thought to occur due to a disruption in phonological processing, but not semantic processing. Specifically, the prominent explanation given by the Transmission Deficit Hypothesis is that there is a disruption in processing in the phonological system whereby all of the needed phonological information for word production is not available (Burke, et al., 1991).

Furthermore, studies have shown that priming of phonology before a TOT elicitation task reduces the probability of TOT occurrence, and priming of phonology after indication of being in a TOT state increases word retrieval (James & Burke, 2000). In addition, fewer TOT states have been reported for words with high phonological degree than words with low phonological degree (Vitevitch & Sommers, 2003). Taken together, these results suggest that not just the process, but also the structure of the phonological system is important to the explanation of word retrieval failures. But what is not known is how age may impact the structure of the phonological system, contributing to the increase in TOT states across adulthood. Therefore, this study will compare the phonological network structure of young, early middle, and late middle adults as a starting point for understanding how age impacts language processing at the phonological level.

Method

The previously examined phonological association data was used to construct three age-related phonological association networks. The three age groups were 18-34, 35-54, and 55+ years, representing young, early middle, and late middle adulthood. There were 408 participants in the 18-34 years old age group ($M = 25$, $SD = 5$), 287 participants in the 35-54 years old age group ($M = 44$, $SD = 6$), and 323 participants in the 55+ years old age group ($M = 62$, $SD = 6$).

Only cue words from the previous phonological association task that were seen by at least one participant in each age group were included in the network construction. There were 5,028 of the original cue words that were seen by all three age groups. Of these cue words, participants provided responses to 5,003 words. The number of responses provided by participants to the cue words differed in each age group (see Table 10), with many cue-response pairs provided by only one or two participants.

Table 10.

Proportion of Cues in Each Age Group by Number of Responses Received for a Cue Word.

	Proportion of Cues		
Number of Responses to a Cue Word	Young Adult Data	Early Middle Adult Data	Late Middle Adult Data
1	36.1%	49.9%	42.9%
2	38.1%	35.4%	36.3%
3	20.4%	12.6%	16.3%
4	5.1%	2.1%	4.1%
5	0.2%	0.1%	0.4%
6			<0.1%

The final set of cue and response words for each age group was compared to the set of cue and response words for the aggregated phonological association network (see Table 11). A one-way ANOVA compared word length as measured by the number of phonemes from each age-related network, and the aggregated network was significant, $F(3, 25362) = 300.03$, $p < .0001$. A Tukey's HSD post-hoc analysis indicated that words in the aggregated network were significantly longer than words in all the age-related networks, all *p*s < .0001, with no difference amongst words in the age-related networks. A one-way ANOVA comparing log word frequency from each age-related network and the aggregated network was significant, $F(3, 24975) = 430.11$, $p < .0001$. A Tukey's HSD post-hoc analysis indicated that words in the late middle adult network were significantly higher in frequency than words in the other age-related networks and the aggregated network, all *p*s < .0001. Words in the aggregated network were also significantly lower in frequency than words in the age-related networks, all *p*s < .0001. Words in the young adult and early middle adult networks were not significantly different in frequency.

Table 11.

Word Length and Frequency in Each Age-Related Phonological Association Network and the Aggregated Phonological Association Network.

	Young Adult Network		Early Middle Adult Network		Late Middle Adult Network		Aggregated Network	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Length	5.28	2.21	5.25	2.18	5.29	1.92	5.81	2.02
Word Frequency	1.93	0.93	1.91	0.93	2.15	0.94	1.88	0.92

Construction of the age-related phonological networks was done using igraph (Csardi & Nepusz, 2006; Ognyanova, K., 2017) in R (R Core Team, 2017). The cue words and their responses were used as nodes in the network, and directed, weighted edges were placed between cue-response pairs.

Similar to the previous network comparisons, comparably-sized random networks were also created for each age-related phonological network. The random networks were created with the same number of nodes and edges as its counterpart age-related phonological network, and were used to determine the "small-worldness" of the network of interest.

Analysis

As done in the previous chapter for network comparisons, several standard measures were used to describe and compare the age-related phonological networks, and included weighting of edges in the calculation of these measures. This study focused on macro-level descriptions of the network, like small-world and scale-free structure, determined by measuring the average shortest path length, average clustering coefficient, and degree distribution of the network. In addition, the location of nodes in the network, mixing by degree, and community structure were also determined to further describe the macro- and meso- levels of the network. Each of these measures were calculated for each age-related phonological network and their comparably-sized random networks, and compared to the aggregated phonological association network described in the previous chapter. Table 12 presents the results for these network measures for each network type.

Table 12.

Network Structure Measures for the Young, Early Middle, and Late Middle Adult Phonological Networks.

Network Measures	Young Adult Phonological Network (YPN)	Early Middle Adult Phonological Network (MPN)	Late Middle Adult Phonological Network (OPN)	Phonological Association Network (PAN)
Network Size	Nodes = 10,426 Edges = 15,399	Nodes = 10,404 Edges = 14,318	Nodes = 10,857 Edges = 15,496	Nodes = 20,617 Edges = 56,747
Location of Nodes[1]	GC = 7,966 (76.4) Islands = 2,435 (23.4) Hermits = 25 (0.2)	GC = 7,351 (70.7) Islands = 3,028 (29.1) Hermits = 25 (0.2)	GC = 8,175 (75.3) Islands = 2,657 (24.5) Hermits = 25 (0.2)	GC = 20,253 (98.2) Islands = 322 (1.6) Hermits = 42 (0.2)
Small World Structure[2]	Avg. Path Len = 22.73 Avg. C = 0.09 S = 541.94	Avg. Path Len = 20.01 Avg. C = 0.09 S = 311.97	Avg. Path Len = 21.33 Avg. C = 0.09 S = 529.91	Avg. Path Len = 9.80 Avg. C = 0.12 S = 724.79
Scale-Free Structure[3]	P. L. RMSE = 2.03 Exp. RMSE = 0.14	P. L. RMSE = 1.81 Exp. RMSE = 0.11	P. L. RMSE = 1.69 Exp. RMSE = 0.12	P. L. RMSE = 0.64 Exp. RMSE = 0.03
Mixing by Degree[4]	$r = 0.37, p < 0.0001$	$r = 0.37, p < 0.0001$	$r = 0.35, p < 0.0001$	$r = 0.44, p < 0.0001$
Community Structure[5]	100 Communities Mod = 0.95	104 Communities Mod = 0.96	106 Communities Mod = 0.95	70 Communities Mod = 0.86

[1]GC = Giant Component, with proportion of nodes in parentheses
[3]Average Shortest Path Length (Avg. Path Len.) and Average Clustering Coefficient (Avg. C), and Small-world-ness (S) from Humphries & Gurney (2008).
[4]Scale-Free Structure is determined by comparing the Root Mean Square Error (RMSE) of the Power-Law (P.L.) function to the degree distribution, with an alternative fit given with the Exponential (Exp) function
[5]Mixing by Degree is determined by examining the correlation between the degree of a node and each of its neighbors.
[6]Modularity (Mod.) is a measure of the significance of community structure in the network, and values above 0.3 are considered significant (Clauset, Newman, & Moore, 2004).

Young Adult Phonological Network. The young adult phonological network was created by representing cue words seen by all three age groups and their responses as nodes and placing an edge between each cue and response pair generated by young adults (aged 18-34 years). This definition leads to a network containing 10,426 (5,028 cues + 5,398 unique responses) nodes and 15,399 edges. These nodes resided in one of three places: the giant component, an island, or as a hermit. There were 7,966 nodes (76.4%) in the giant component. 2,435 nodes (23.4%) located in islands, and 25 hermit nodes (0.2%). There were 703 islands that ranged in size from 2 to 27 nodes (see Figure 17). Islands were mostly organized by phonological overlap in the rhyme position, and often involved a suffix that was consistent between words (see Table 13).

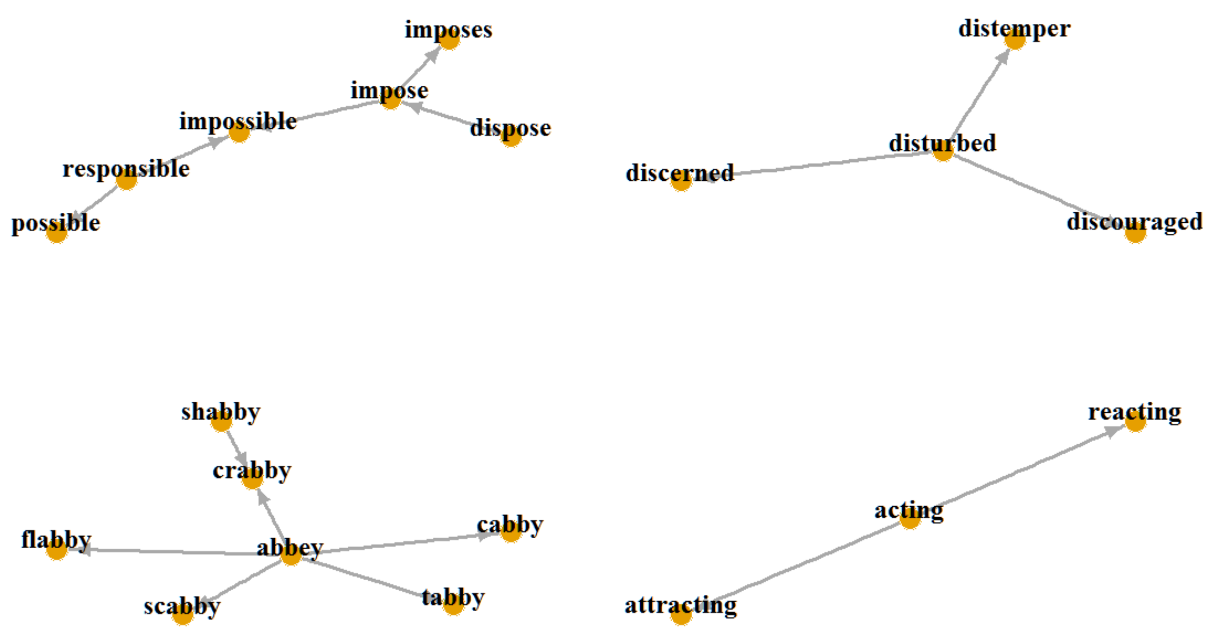

Figure 17. Example islands from the Young Adult Phonological Network.

Table 13.

Proportion of Islands in the Young Adult Phonological Network with Different Types of Overlap.

Type of Overlap	Proportion of Islands
Phonological Overlap	
Alliteration	36.3%
Rhyme	61.8%
Partial Rhyme	21.1%
Morphological Overlap	
Prefix	1.6%
Suffix	52.5%
Stem	12.9%

Note Islands can overlap by more than one type.

The young adult phonological network was examined for small-world structure. Recall that having a small-world structure indicates that the network is easy to traverse despite its large size, and is hallmarked by having a similar average shortest path length and larger average clustering coefficient than a comparably-sized random network (Watts & Strogatz, 1998). The average shortest path length of the young adult phonological network was 22.73, whereas the average shortest path length of the comparably-sized random network was 18.68. Using network analysis convention where the difference in values is no greater than 1.5 times in magnitude, the average shortest path length values for the young adult phonological network and the random network are were significantly different. The average clustering coefficient of the young adult phonological network was 0.09, whereas the comparably-sized random network had an average clustering coefficient of 0.0001. The average clustering coefficient values for the young adult phonological network and the random network were significantly different by several magnitudes according to network analysis convention. In addition, "small-world-ness"

(Humphries & Gurney, 2008) for the young adult phonological network was 541.94. Therefore, these measures indicate that the young adult phonological network has a small-world structure.

Next, an analysis was conducted to determine if the young adult phonological network could be classified as having a scale-free structure. Recall that having a scale-free structure suggests that many nodes have few edges and few nodes have many edges. This is indicated by the degree distribution following a power-law function when plotted on a log-log scale. Figure 18 displays the log-log plot for the degree distribution of the young adult phonological network. The power-law function was best fit by the equation $y = 2.45x^{-2.75}$, RMSE = 2.03, whereas the exponential curve was best fit by the equation $y = 0.43e^{-0.39x}$, RMSE = 0.14. Since the exponential curve better fits the data than the power-law function, the young adult phonological network does not have a scale-free structure.

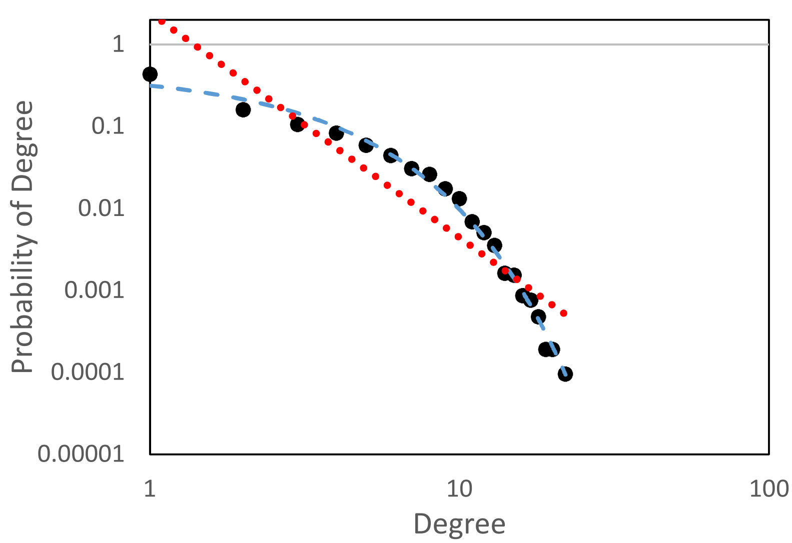

Figure 18. Log-log plot of degree distribution for the Young Adult Phonological Network. The power-law function is represented in red and the exponential curve is represented in blue.

Mixing by degree was also examined in the young adult phonological network. Recall that assortative mixing by degree occurs when nodes with high degree tend to be connected to other nodes with high degree. On the other hand, disassortative mixing by degree occurs when nodes with high degree tend to be connected to nodes with low degree. To determine the kind of mixing pattern of the young adult phonological network, a Pearson's correlation between a node's degree and each of its neighbor's degree was examined. A correlation of r (15399) = 0.37, $p < .0001$, was found suggesting that an assortative mixing by degree pattern exists in the young adult phonological network.

Finally, the community structure of the young adult phonological network was examined. In total, there were 100 communities in the giant component as determined by the Louvain method (Blondel, et al., 2008) with a modularity of 0.95. A modularity value higher than 0.3 is indicative of significant community structure (Clauset, et al., 2004). Figure 19 depicts different communities in the giant component by color. These communities ranged in size from 10 to 212 nodes (M = 79.66, SD = 36.45). Communities overlapped in several ways phonologically and/or morphologically, resulting in smaller groupings of nodes organizing within a community (see Figure 20).

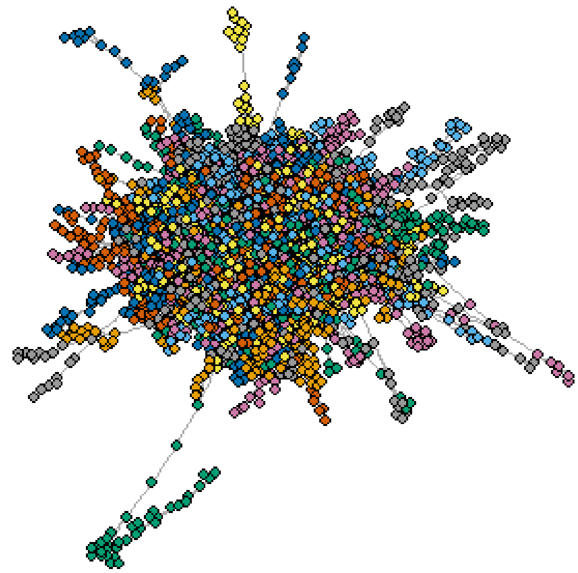

Figure 19. Giant component of the Young Adult Association Network. Color represents communities.

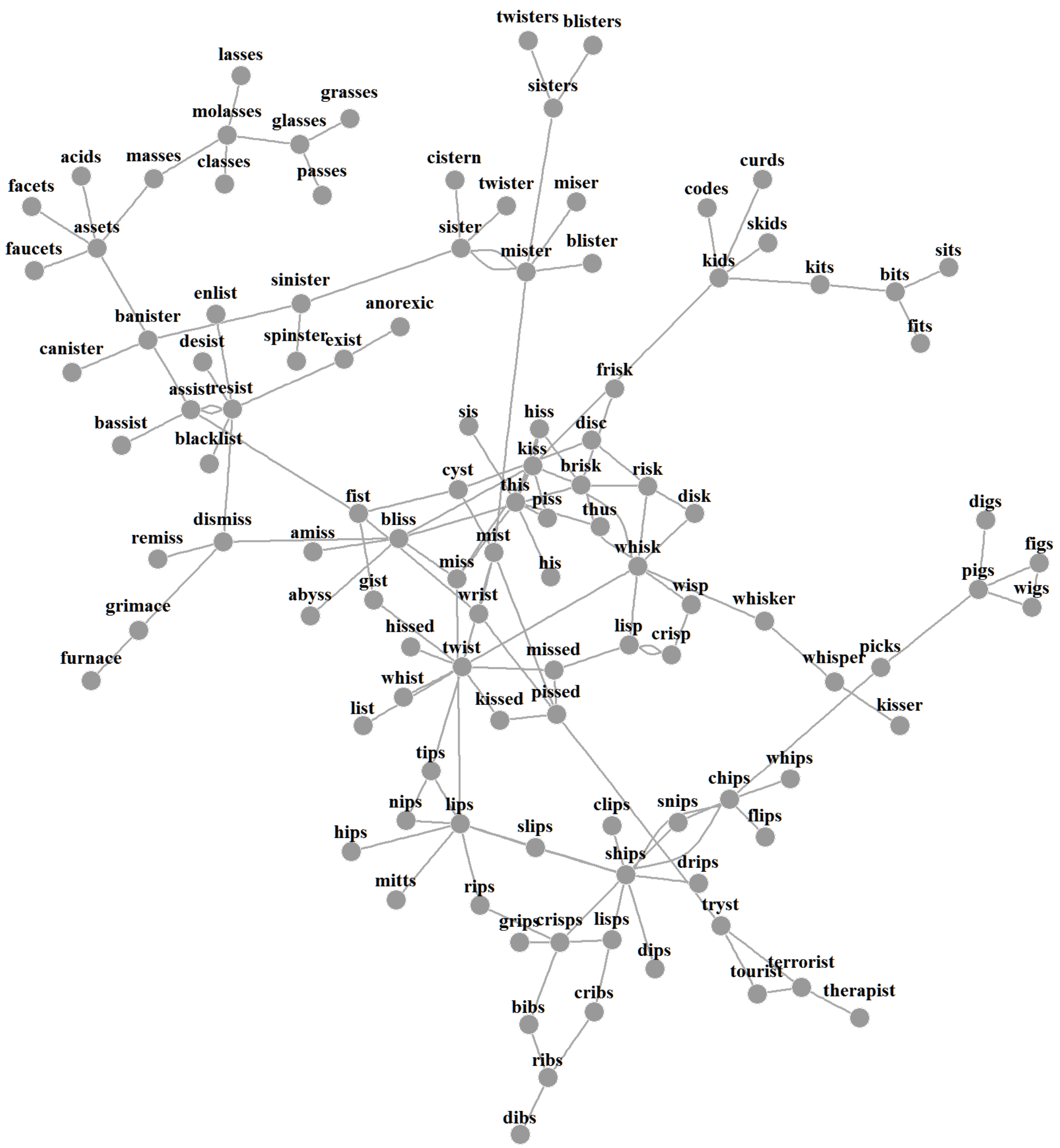

Figure 20. Example community from the Young Adult Phonological Network.

Early Middle Adult Phonological Network. The early middle adult phonological network was created by representing cue words seen by all three age groups and their responses as nodes and placing an edge between each cue and response pair generated by early middle adults (aged 35-54 years). This definition leads to a network containing 10,404 nodes (5,028 cues + 5,376 unique responses) and 14,318 edges. These nodes resided in one of three places: the giant component, an island, or as a hermit. There were 7,351 nodes (70.7%) in the giant component, 3,028 nodes (29.1%) located in islands, and 25 hermit nodes (0.2%). There were 853 islands that ranged in the size from 2 to 25 nodes (see Figure 21). Similar to the young adult phonological network, islands were mostly organized by phonological overlap in the rhyme position, and often involved a suffix that was consistent between words (see Table 14).

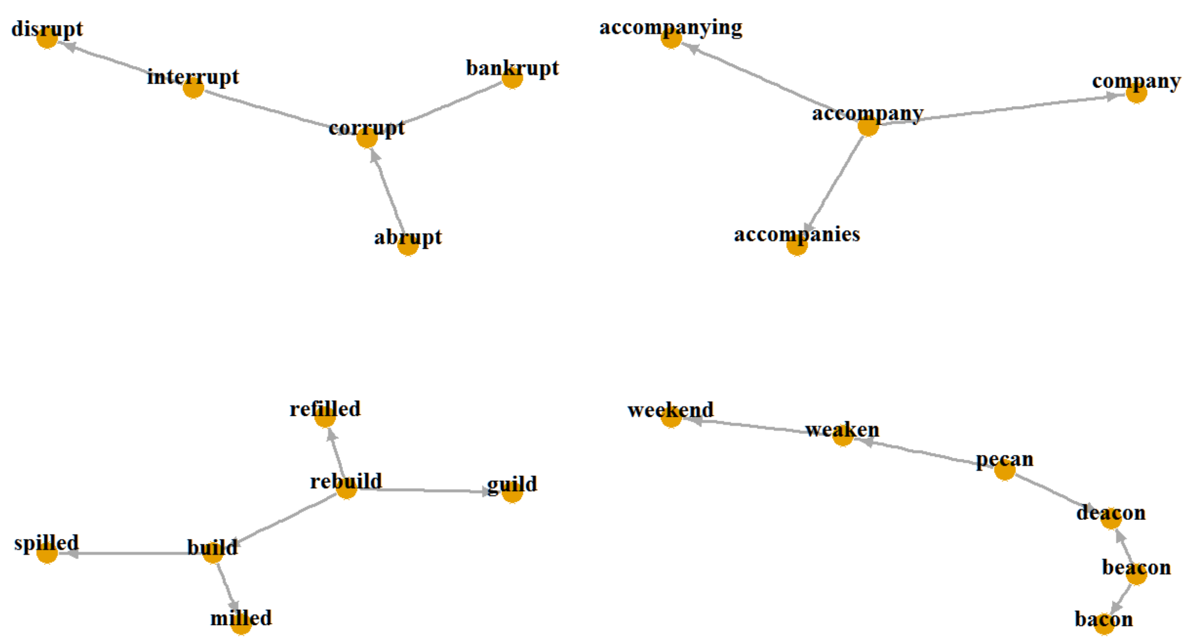

Figure 21. Example islands from the Early Middle Adult Phonological Network.

Table 14.

Proportion of Islands in the Early Middle Adult Phonological Network with Different Types of Overlap.

Type of Overlap	Proportion of Islands
Phonological Overlap	
Alliteration	40.9%
Rhyme	65.6%
Partial Rhyme	16.5%
Morphological Overlap	
Prefix	0.7%
Suffix	55.0%
Stem	12.0%

Note Islands can overlap by more than one type.

The early middle adult phonological network was examined for small-world structure. The average shortest path length of the early middle adult phonological network was 20.01, whereas the average shortest path length of the comparably-sized random network was 21.64. Using network analysis convention where the difference in values is no greater than 1.5 times in magnitude, the average shortest path length values for the early middle adult phonological network and the random network were significantly different. The average clustering coefficient of the early middle adult phonological network was 0.09, whereas the comparably-sized random network had an average clustering coefficient of 0.0003. The average clustering coefficient values for the early middle adult phonological network and the random network were significantly different by several magnitudes according to network analysis convention. In addition, “small-world-ness” (Humphries & Gurney, 2008) for the early middle adult phonological network was 311.97. Therefore, these measures indicate that the early middle adult phonological network has a small-world structure.

Next, an analysis was conducted to determine if the early middle adult phonological network could be classified as having a scale-free structure. Figure 22 displays the log-log plot for the degree distribution of the early middle adult phonological network. The power-law function was best fit by the equation $y = 2.25x^{-1.75}$, RMSE = 1.81, whereas the exponential curve was best fit by the equation $y = 0.57e^{-0.44x}$, RMSE = 0.11. Since the exponential curve better fits the data than the power-law function, the early middle adult phonological network does not have a scale-free structure.

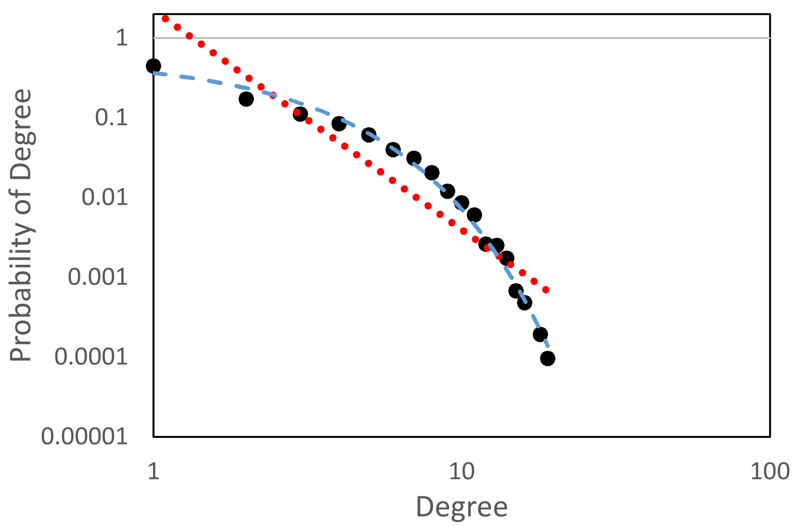

Figure 22. Log-log plot of degree distribution for the Early Middle Adult Phonological Network. The power-law function is represented in red and the exponential curve is represented in blue.

Mixing by degree was also examined in the early middle adult phonological network. A correlation of r (14318) = 0.37, $p < 0.0001$, was found suggesting that an assortative mixing by degree pattern exists in the early middle adult phonological network.

Finally, the community structure of the early middle adult phonological network was examined. In total, there were 104 communities in the giant component as determined by the Louvain method (Blondel, et al., 2008) with a modularity of 0.96. A modularity value higher than 0.3 is indicative of significant community structure (Clauset, et al., 2004). Figure 23 depicts different communities in the giant component by color. These communities ranged in size from 6 to 208 nodes (*M* = 70.67, *SD* = 37.18). Communities overlapped in several ways phonologically and/or morphologically, resulting in smaller groupings of nodes organizing within a community (see Figure 24).

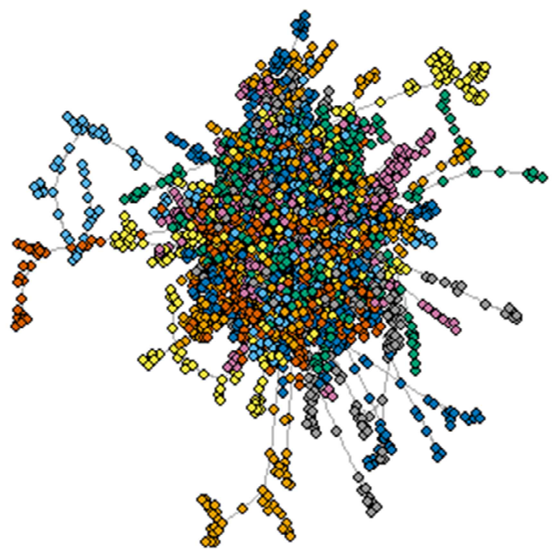

Figure 23. Giant component of the Early Middle Adult Phonological Network. Color represents communities.

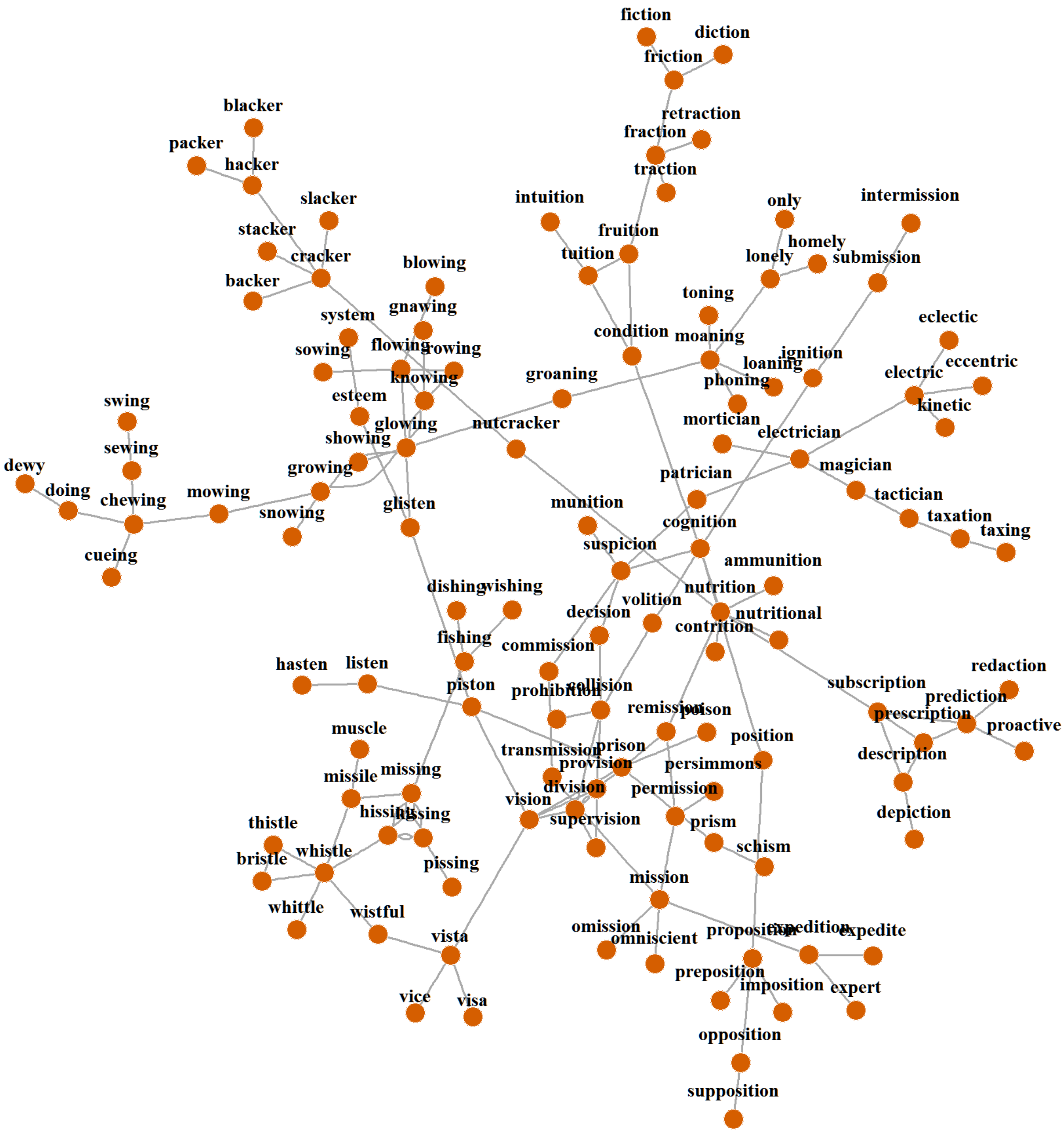

Figure 24. Example community from the Early Middle Adult Phonological Network.

Late Middle Adult Phonological Network. The late middle adult phonological network was created by representing words seen by all three age groups and their responses as nodes and placing an edge between each cue and response pair generated by older adults (aged 55 years and older). This definition leads to a network containing 10,857 nodes (5,028 cues + 5,829 unique responses) and 15,496 edges. These nodes resided in one of three places: the giant component, an island, or as a hermit. There were 8,175 nodes (75.3%) in the giant component, 2,657 nodes (24.5%) located in islands, and 25 hermit nodes (0.2%). There were 712 islands that ranged in size from 2 to 48 nodes (see Figure 25). Like the young and middle adult phonological networks, islands were mostly organized by phonological overlap in the rhyme position, and often involved a suffix that was consistent between words (see Table 15).

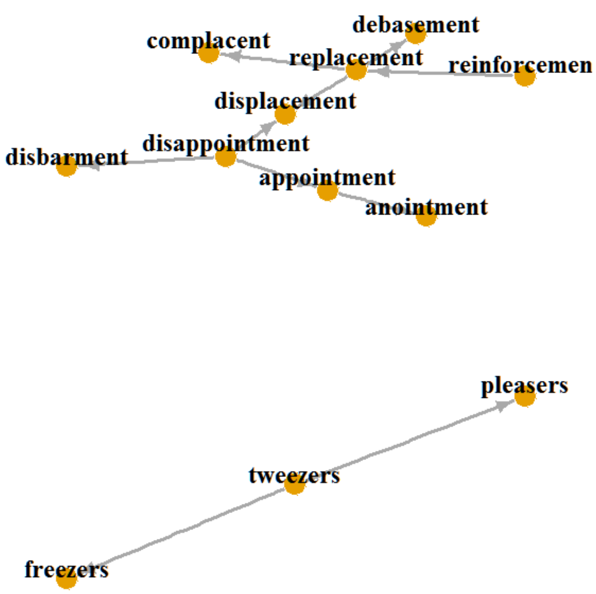

Figure 25. Example islands from the Late Middle Adult Phonological Network.

Table 15.

Proportion of Islands in the Late Middle Adult Phonological Network with Different Types of Overlap.

Type of Overlap	Proportion of Islands
Phonological Overlap	
Alliteration	41.1%
Rhyme	70.2%
Partial Rhyme	16.3%
Morphological Overlap	
Prefix	2.7%
Suffix	57.9%
Stem	13.3%

Note Islands can overlap by more than one type.

The late middle adult phonological network was examined for small-world structure. The average shortest path length of the late middle adult phonological network was 21.33, whereas the average shortest path length of the comparably-sized random network was 21.24. Using network analysis convention where the difference in values is no greater than 1.5 times in magnitude, the average shortest path length values for the late middle adult phonological network and the random network were not significantly different. The average clustering coefficient of the late middle adult phonological network was 0.09, whereas the comparably-sized random network had an average clustering coefficient of 0.0002. The average clustering coefficient values for the late middle adult phonological network and the random network were significantly different by several magnitudes according to network analysis convention. In addition, "small-world-ness" (Humphries & Gurney, 2008) for the late middle adult phonological network was 529.91. Therefore, these measures indicate that the late middle adult phonological network has a small-world structure.

Next, an analysis was conducted to determine if the late middle adult phonological network could be classified as having a scale-free structure. Figure 26 displays the log-log plot for the degree distribution of the late middle adult phonological network. The power-law function was best fit by the equation $y = 2.12x^{-2.69}$, RMSE = 1.69, whereas the exponential curve was best fit by the equation $y = 0.61e^{-0.44x}$, RMSE = 0.12. Since the exponential curve better fits the data than the power-law function, the late middle adult phonological network does not have a scale-free structure.

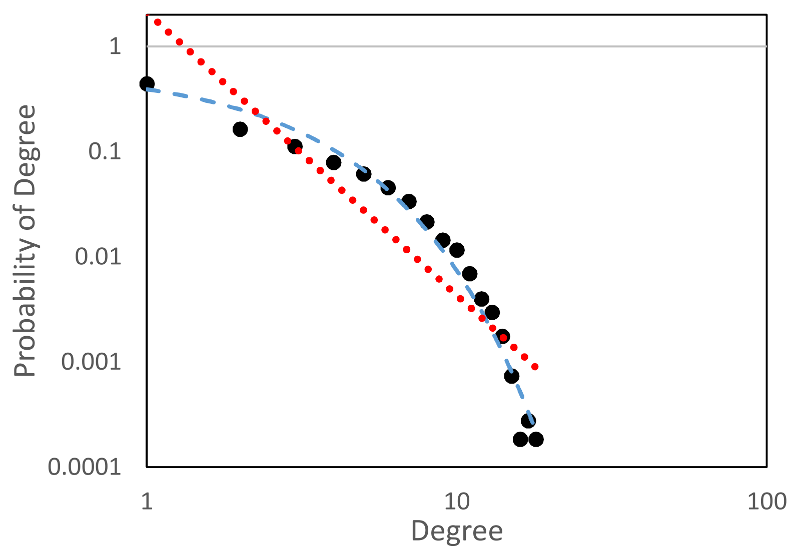

Figure 26. Log-log plot of degree distribution for the Late middle adult Phonological Network. The power-law function is represented in red and the exponential curve is represented in blue.

Mixing by degree was also examined in the late middle adult phonological network. A correlation of r (15496) = 0.35, $p < 0.0001$, was found suggesting that an assortative mixing by degree pattern exists in the late middle adult phonological network.

Finally, the community structure of the late middle adult phonological network was examined. In total, there were 106 communities in the giant component as determined by the Louvain method (Blondel, et al., 2008) with a modularity of 0.95. A modularity value higher than 0.3 is indicative of significant community structure (Clauset, et al., 2004). Figure 27 depicts different communities in the giant component by color. These communities ranged in size from 9 to 197 nodes (M = 77.12, SD = 38.28). Communities overlapped in several ways phonologically and/or morphologically, resulting in smaller groupings of nodes organizing within a community (see Figure 28).

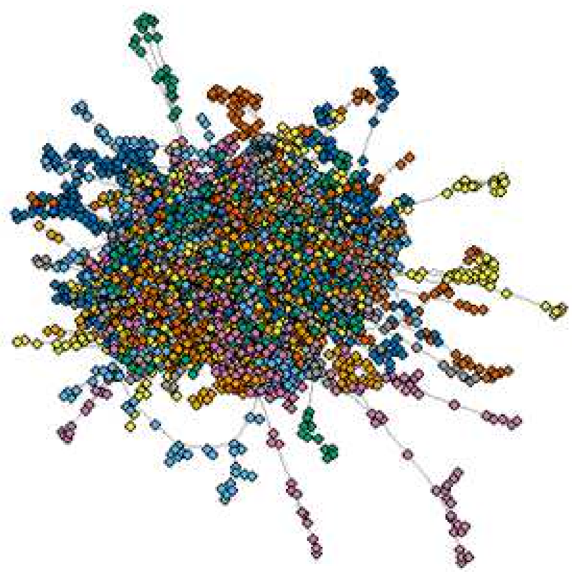

*Figure 27. G*iant component from the Late Middle Adult Phonological Network. Color represents communities.

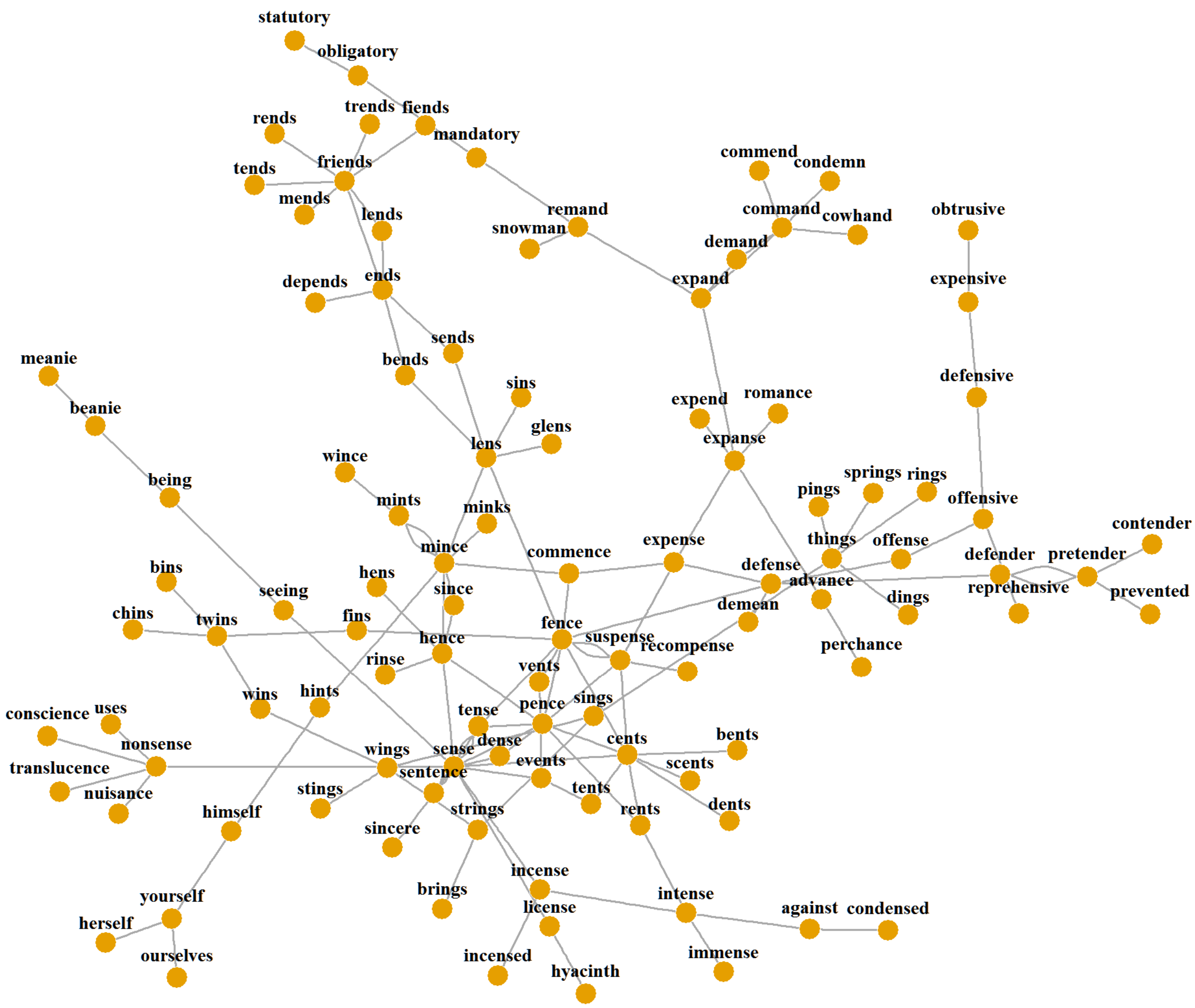

Figure 28. Example community from the Late Middle Adult Phonological Network.

Correlation between Networks. The previous network examinations focused on macro- and meso- level network structure. However, it is also important to identify differences in the structure of a word in each age-related network. These findings may provide additional insight into how phonological similarity is represented in the mental lexicon at different points across adulthood. Specifically, an analysis was done to assess how the age-related networks overlapped with one another, as well as how properties of individual words in each network (location, degree, and clustering coefficient) changed with age.

Although the three age-related networks were similar in size and all three had the same cue words, the responses that were made to those cue words differed. For instance, there were differences in the number of total responses and the number of unique responses that were made to the cue words across age groups. When comparing the young adult network to early middle adult network, there were 3,628 edges in common (23.6% of the young adult edges and 25.3% of early middle adult edges). When comparing the early middle adult network to the late middle adult network, there were 3,724 edges in common (26.0% of early middle adult edges and 24.0% of late middle adult edges).

Furthermore, an examination of the nodes for the overlapping edges between age-related networks was done. The overlapping edges between the young adult and early middle adult networks consisted of 4,428 nodes. These nodes tended to reside in the giant component of the young adult network (82.6%) and in the giant component of the early middle adult network (78.5%). The overlapping edges between the early middle adult and late middle adult networks consisted of 4,513 nodes. These nodes tended to reside in the giant component of the early middle network (77.8%) and in the giant component of the late middle adult network (81.4%). These findings suggest that although about a quarter of the network is consistent across adulthood with that consistency occurring mostly in the giant component. Therefore, it is the periphery of the network (i.e., the islands) that tend to change across adulthood.

The location of nodes in each age-related network was further examined by looking at all cue nodes, rather than just those nodes that were consistent between networks. The location of these cue nodes was categorized into either located in the giant component, an island, or as a hermit for each network, and assessed from the young adult phonological network to the middle adult phonological network and from the middle adult phonological network to the late middle adult phonological network. Given the restriction of data to only those cues seen by all three age groups, the location of nodes could change in only one of three ways: from the giant component to an island, from an island to the giant component, or remained in the same location in both networks. The proportion of nodes for each type of location change and at each assessment are given in Table 16.

Table 16.

Proportion of Nodes for Each Type of Location Change from the Young to Early Middle Adult Network and from the Early Middle to Late Middle Adult Network.

	From the Young Adult Network to Early Middle Adult Network		From the Early Middle Adult Network to the Late Middle Adult Network	
Type of Location Change	Count of Nodes	Proportion of Nodes	Count of Nodes	Proportion of Nodes
Giant Component to Island	1122	14.8%	777	10.0%
Island to Giant Component	734	9.7%	1097	14.2%
Same Location	5725	75.5%	5868	75.8%

Of the 10,401 nodes (cue and response words) in the young adult phonological network, only 7,582 (72.9%) nodes were also in the early middle adult phonological network. The location of most of these nodes remained in the same location from young adulthood to early middle adulthood. Of those nodes that did change in location, more nodes “broke away” from the giant component (i.e., moved from the giant component to an island) than became incorporated into the giant component (i.e., moved from an island to the giant component).

Of the 10,379 nodes (cue and response words) in the early middle adult phonological network, only 7,742 (74.6%) nodes were also in the late middle adult phonological network. Like the previous results, the location of most nodes remained in the same location from early middle adulthood to late middle adulthood. However, of those nodes that did change in location, the opposite result was found. More nodes were incorporated into the giant component than nodes that "broke away" from the giant component. In sum, these location change findings continue to show that most words remain in the same location across adulthood, and that what does change tends to be in the periphery of the network.

Next, the degree of words in each network was examined. The young adult phonological network had an average degree of 2.96 (SD = 2.69). The early middle adult phonological network had an average degree of 2.78 (SD = 2.40). And, the late middle adult phonological network had an average degree of 2.86 (SD = 2.52).

As done in the previous chapter's network comparison, a Pearson's correlation between the degree of a word in one network with the degree of the same word in another network was also calculated. A correlation was determined between the young adult phonological network and the early middle adult phonological, and between the early middle adult phonological network and the late middle adult phonological network. Each correlation included only those nodes (cue and response words) in the former network that were also in the latter network, as done in the previous location change analysis. The degree of words in the young adult phonological network and the degree of words in the early middle adult phonological network were correlated, r (7582) = .62, p < .0001 (see Figure 29). Also, the degree of words in the early middle adult phonological network and the degree of words in the late middle adult phonological network were correlated, r (7742) = .62, p < .0001 (see Figure 30). Therefore, the degree of a node is similar across adulthood.

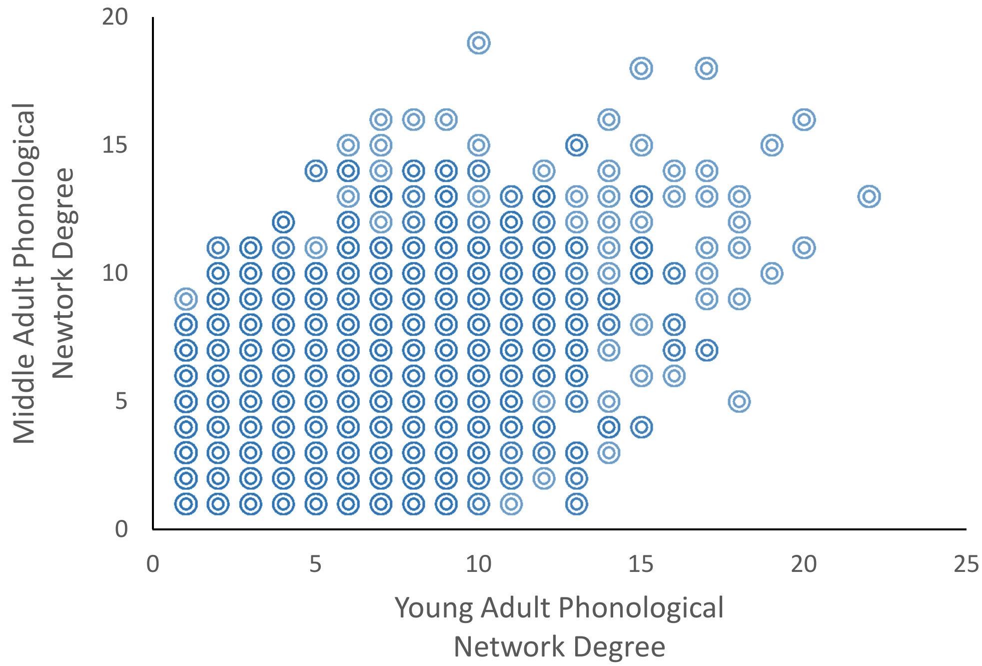

Figure 29. Scatterplot of Young Adult Phonological Network and Early Middle Adult Phonological Network degrees. Darker blues represent a larger proportion of the data.

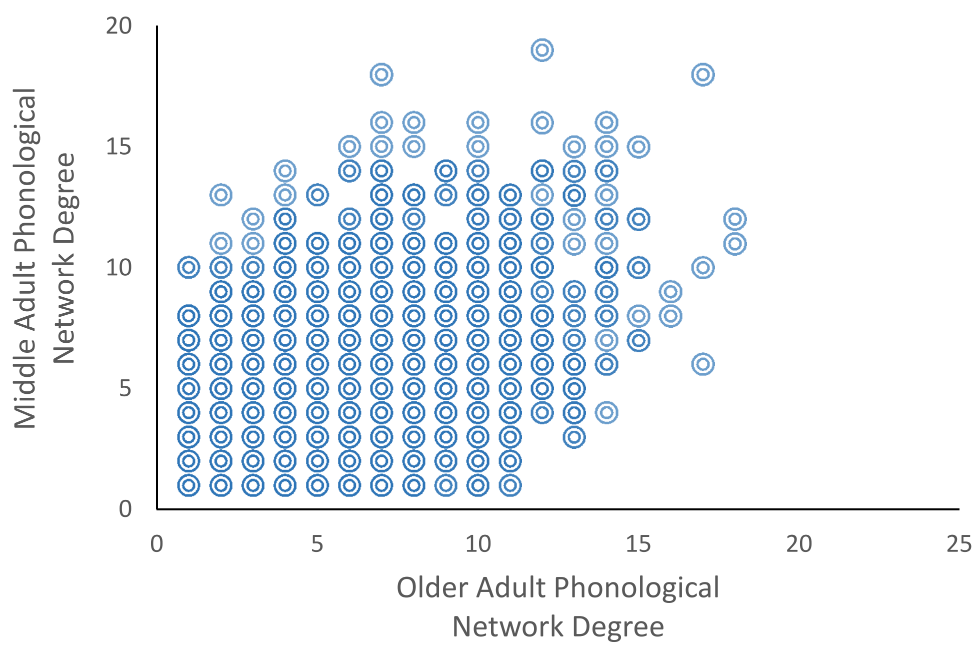

Figure 30. Scatterplot of Early Middle Adult Phonological Network and Late Middle Adult Phonological Network degrees. Darker blues represent a larger proportion of the data.

Lastly, clustering coefficient was also examined using Pearson's correlations between the young adult and early middle adult phonological networks, and between the early middle adult and late middle adult phonological networks. Again, each correlation included only those nodes (cue and response words) in the former network that were also in the latter network. The clustering coefficient of words in the young adult phonological network and the clustering coefficient of words in the early middle adult phonological network were correlated, r (7582) = .22, $p < .0001$ (see Figure 31). The clustering coefficient of words in the early middle adult phonological network and the clustering coefficient of words in the late middle adult phonological network were correlated, r (7742) = .21, $p < .0001$ (see Figure 32). Therefore, the clustering coefficient of a node may be similar across adulthood.

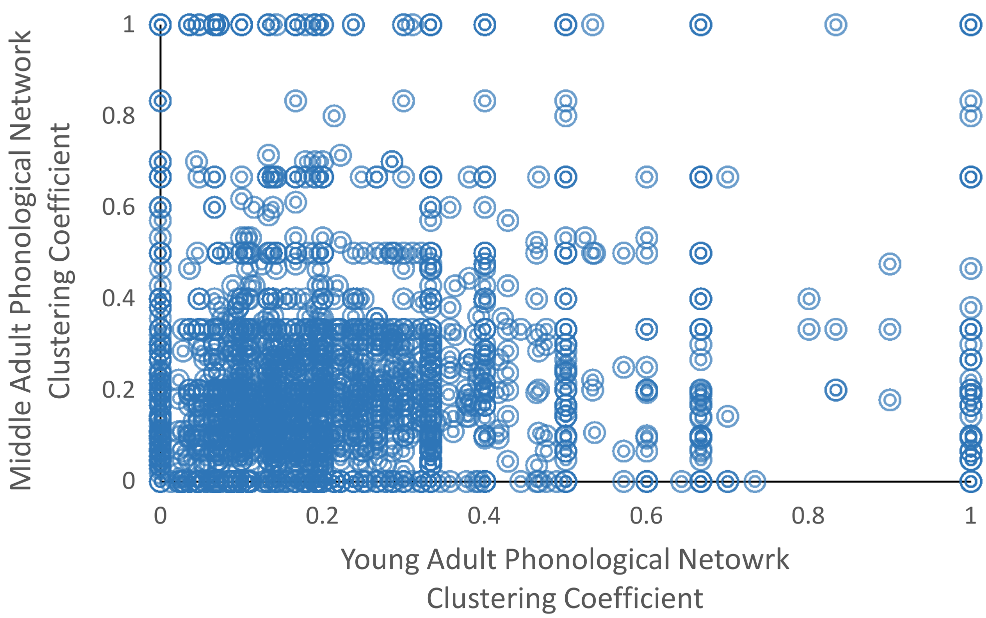

Figure 31. Scatterplot of Young Adult Phonological Network and Early Middle Adult Phonological Network clustering coefficients.

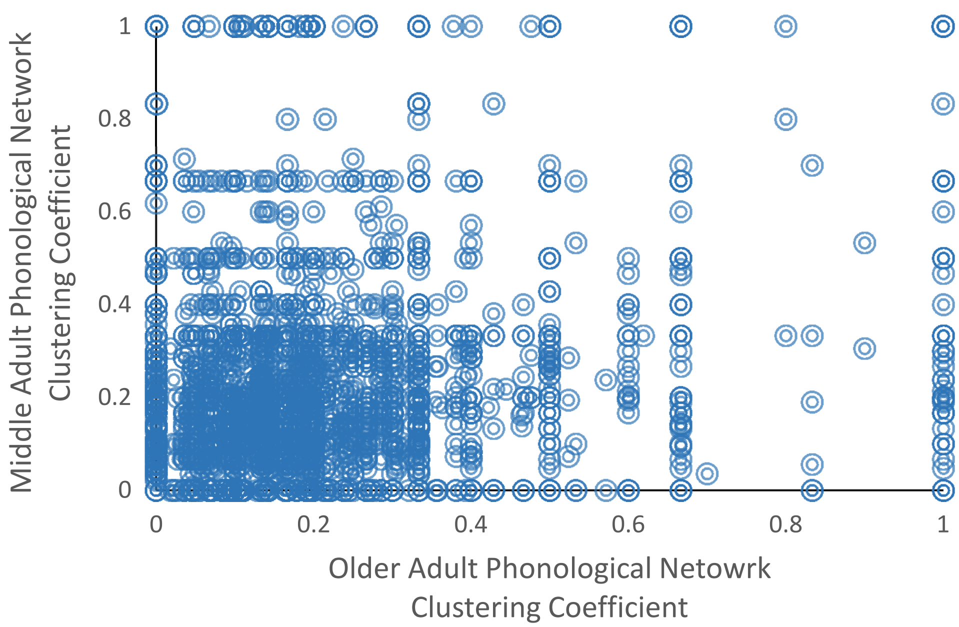

Figure 32. Scatterplot of Early Middle Adult Phonological Network and Late Middle Adult Phonological Network clustering coefficients.

Discussion

The three age-related phonological networks resembled the previous aggregated phonological association network, but to a lesser extent. For example, the phonological association network had more nodes and edges as compared to each age-related network. This larger network size may have also contributed to the aggregated phonological association network having a larger proportion of nodes located in the giant component, with a smaller average path length and higher average clustering coefficient than each age-related phonological network. In addition, all the age-related networks, like the phonological association network, can be described as having small-world, but not scale-free, structure. Finally, assortative mixing by degree and significant community structure were also found in each of the age-related networks, like the aggregated phonological association network.

When comparing the three age-related phonological networks to each other, there appears to be minimal differences in network structure across adulthood. This is in stark contrast to the observation that semantic network structure significantly changes across adulthood (Dubossarsky, et al., 2017). Specifically, Dubossarsky, et al. (2017) found in the semantic association network that in- and out- degree declined across adulthood, average shortest path length increased adulthood, and average clustering coefficient declined across early to middle adulthood with a small increase in late adulthood. In the present phonological association network, degree, average shortest path length, and average clustering coefficient were similar across adulthood. It is intriguing that the semantic network changes, but not the phonological network. One potential explanation for changes in semantic network structure, but not phonological network structure, is that the way in which we associate words through meaning and sound differs. Specifically, meaning-based associations may be more likely to change over time as individuals encounter new words and experiences, leading to changes in semantic associations and their structure across adulthood. However, the way in which words are phonologically constructed must follow certain rules, limiting the likelihood of changing phonological association and their structure across adulthood.

One aspect of phonological network structure that showed evidence of age-related change was the location of nodes across time. In the present findings, most nodes remained in the same location, but those nodes that did change location differed in direction from young to early middle adulthood and from early middle to late middle adulthood. Specifically, the young to early middle adulthood nodes that changed location were likely to move from the giant component to islands, and may be reflective of word learning (e.g., through higher education and/or career training). For example, a new word may make a known word sound more similar to the new word than to other known words, and potentially lead to the formation of an island. On the other hand, the early middle to late middle adulthood nodes that changed location were more likely to move from the islands to the giant component, and may be reflective of age-related sensory or cognitive changes. For example, the notion that older adults have more difficulty identifying words with high phonological degree (Sommers, 1996), may lead to words sounding more similar with increasing age and a shift of nodes from islands to the giant component.

In addition to this age-related network structure change, the set of words used in each network had some differences from each age network and the aggregated network. The age-related networks had shorter words than the aggregated network and had higher word frequency than the aggregated network, especially the late middle adult network. These findings suggest that those items that all individuals, regardless of age, are likely to respond with will be short and of high frequency, with age-related differences in longer, lower frequency words. In addition, the overlap in edges between the age networks also differed with only roughly 25% of edges being consistent from one age network to the next, and these consistent edges are likely to reside in the giant component of the network, where shorter, higher frequency words are often found.

The present findings, though, should be interpreted with caution. Compared to the aggregated phonological association network, the age-related phonological networks had higher average path lengths. smaller clustering coefficients, fewer nodes in the giant component, and more nodes in islands. Indeed, the current age-related phonological network structures may be an underestimate of what is truly represented in the mental lexicon due to the sparse data available. For example, many of the cues were only seen by one participant in each age-related network, whereas cue words were seen by at least six participants in the aggregated phonological association network. Thus, there may not be enough data to capture the true phonological network structure of each age group. Continued collection of phonological association data is needed for a more complete understanding of how phonological network structure may change or remain stable across adulthood.

In addition, the lack of change in phonological network structure might also be due to the lack of data available from participants in older adulthood (75 years and older). The present sample is more representative of early and late middle adulthood, where phonological network structure may not yet be affected by the cognitive changes that typically accompany aging. In addition, the participants that did complete the association task are unlikely to have cognitive deficits due to the computer literacy skills required to complete the task online via Amazon Mechanical Turk. Therefore, it would also be important to collect data not just from healthy older adults, but also to include a more representative sample of adulthood.

In sum, this study provides an initial examination of phonological network structure across adulthood. The current results suggest that phonological network structure, using phonological association data, does not change significantly with age. However, continued collection of data and testing of network structure using behavioral experiments are necessary. For example, even though the structure does not appear to change with age, processing within the network is affected by other age-related changes, like sensory and cognitive declines. In other words, even though the structure does not change, this structure may not be as helpful for language processing in older adulthood. For example, the evidence that phonological degree impacts spoken word recognition in older adults (Sommers, 1996) and that tip-of-the-tongue states increase with age (Burke, et al., 1991), but can be reduced by high phonological degree (Vitevitch & Sommers, 2003), suggests that phonological network structure plays an integral role in the ability to perceive and produce words across adulthood. Behavioral tests can continue to examine how phonological network structure influences language processes in older adulthood to better understand how an un-changing network structure could disrupt or aid in those language processes.

Lastly, although the phonological network does not seem to change significantly across adulthood, the semantic network has been shown to do so. Changes in one type of language network may influence processing not just in that network, but other types of language networks as well. For example, in a word production task, one must access semantic information to select the correct target word and phonological information to produce that selected word. The increasing sparseness of the semantic network with age may result in more difficult or slowed processing. This disruption may then lead to increased difficulty in moving from the semantic system to the phonological system, resulting in slowed production, speech errors, or word retrieval failures. Therefore, understanding how these different types of language networks connect and influence each other requires the need for more complex network analyses that examine multiple layers simultaneously.

Chapter 5: Analysis of a Multiplex

Introduction

Typically, only one network is examined at a time to examine language structure and processes. That is, only the semantic network or only the phonological network would be examined. However, research has shown that both the semantic and phonological systems are important during word retrieval, and that these systems can interact (e.g., Dell & O'Seaghdha, 1992). Therefore, it is important to understand how the semantic and phonological networks connect or overlap.

In Network Science, multiple networks can be examined simultaneously by creating "layers" in the network. Specifically, a multiplex network is one in which nodes are shared between layers, but edges are different in each layer. To date, there is one multiplex network that represents different aspects of the mental lexicon (Stella, et al., 2017). This multiplex includes 529 words with edges placed between words in 4 different layers: 1) semantic free association norms, 2) shared features indicated by synonym relationships, 3) co-occurrence norms, and 4) phonological similarity defined by the one-phoneme metric. Importantly, this multiplex structure has been shown to be a more powerful predictor of word acquisition in children than structural information from a single-layer alone or conventional psycholinguistic measures, like age of acquisition (Stella, et al., 2017). Therefore, examining the whole multiplex provides novel insight into language processes.

This chapter continues to examine multiplex structure of the mental lexicon by using a larger number of words and using a similar edge definition in each layer. Specifically, the present multiplex includes a semantic layer using the semantic association data from S. De Deyne and a phonological layer using the phonological association data from Chapter 2. An analysis of the semantic association network was compared to the phonological association network, and the individual semantic and phonological layers were compared to the aggregated multiplex network. Since multiplex analysis is an emerging, cutting-edge area of Network Science, the tools to fully analyze a large multiplex are not available. Indeed, the multiplex to be examined exceeds the computational limits of the one existing program MuxViz (De Domenico, Porter, & Arenas, 2015) that has been used to analyze small multiplex network, like in Stella, et al. (2017). Therefore, an additional analysis was done looking at degree of words in the individual layers and in the multiplex to further assess the current multiplex structure given that degree has been shown to influence several language processes in single-layer networks.

Method

Cue and response items from S. De Deyne's semantic association data were used. Semantic associations were gathered by presenting a cue word to participants and asking them to provide up to three responses that immediately came to mind. The data provided by S. De Deyne included only the first responses that participants provided to over 10,000 cue words. The number of participants and participant-level data were not currently available at the time of this analysis. However, there were significantly more unique cue-response pairs for first responses only (N = 429,401) than that obtained in the phonological association task in Chapter 2 (N = 32,297), suggesting a significantly larger sample size for the semantic association data than the phonological association data.

For the present semantic association network analysis and multiplex analysis, only data using the same cue and response words from the phonological association task in Chapter 2 were used. This was done for two reasons: 1) to ease interpretation of network comparison to only those words common to both tasks, and 2) to create a multiplex where nodes are identical in each layer. Therefore, there was a total of 9,297 cue words and 5,451 unique response words that matched the phonological association data.

The nodes in the semantic association network were cue and response words, and edges were placed between cue and response pairs. Edges were directed, as done in the previous phonological association network. However, frequency of cue-response pairs was not available, and thus weighting could not be determined.

The semantic association network was compared to a comparably-sized random network. The random network was only used to determine “small-world-ness” of the semantic network. In addition, the semantic association network was also compared to the phonological association network analyzed in Chapter 2. The semantic association and phonological association networks were combined into one network to assess multiplex structure. Network generation and analysis were conducted using the igraph package (Csardi & Nepusz, 2006; Ognyanova, K., 2017) in R (R Core Team, 2017).

Lastly, data from the English Lexicon Project (Balota, et al., 2007) was used to assess the influence of multiplex structure on behavioral data. Specifically, visual lexical decision and naming reaction time were used in this analysis. Previous work has shown that visual lexical decision and picture naming are influenced by single-layer semantic degree (e.g., Duñabeitia, Avilés, & Carreiras, 2008) and phonological degree (e.g., Yates, 2005; Yates, Locker, & Simpson, 2004), providing an opportunity to test the influence of aggregated multiplex degree (semantic + phonological degree) and multidegree (number of multilinks). Only words that have semantic degree, phonological degree, and multidegree were included in this analysis (N = 4,864).

Analysis

As done in the previous chapters for network comparisons, several standard measures were used to describe the semantic association network. This study focused on macro-level descriptions of the network, like small-world and scale-free structure, determined by measuring the average shortest path length, average clustering coefficient, and degree distribution of the network. In addition, the location of nodes in the network, mixing by degree, and community structure were also determined to further describe the macro- and meso- levels of the network. Each of these measures were calculated for the semantic association network and its comparably-sized random network, and compared to the phonological association network described in the Chapter 2. Table 17 presents the results for these network measures for each network type.

Table 17.

Network Structure Measures for the Semantic Association Network and Phonological Association Network.

Network Measures	Semantic Association Network (SAN)	Phonological Association Network (PAN)
Network Size	Nodes = 14,794 Edges = 239,483	Nodes = 20,617 Edges = 56,747
Location of Nodes[1]	GC = 14,794 (100.0) Islands = 0 Hermits = 0	GC = 20,253 (98.2) Islands = 322 (1.6) Hermits = 42 (0.2)
Small-World Structure[2]	Avg. Path Len = 3.77 Avg. C = 0.09 S = 42.12	Avg. Path Len = 9.80 Avg. C = 0.12 S = 724.79
Scale-Free Structure[3]	P. L. RMSE = 1.37 Exp. RMSE = 0.16	P. L. RMSE = 0.64 Exp. RMSE = 0.03
Mixing by Degree[4]	r = 0.03, p < 0.0001	r = 0.44, p < 0.0001
Community Structure[5]	12 Communities Mod = 0.32	70 Communities Mod = 0.86

[1]GC = Giant Component, with proportion of nodes in parentheses
[2]Average Shortest Path Length (Avg. Path Len.) and Average Clustering Coefficient (Avg. C), and Small-world-ness (S) from Humphries & Gurney (2008).
[3]Scale-Free Structure is determined by comparing the Root Mean Square Error (RMSE) of the Power-Law (P.L.) function to an alternative Exponential (Exp) curve.
[4]Mixing by Degree is determined by the correlation between the degree of a node and each of its neighbors.
[5]Modularity (Mod.) is a measure of the significance of community structure in the network, and values above 0.3 are considered significant (Clauset, Newman, & Moore, 2004).

Semantic Association Network. The semantic association network was created by representing words as nodes and placing an edge between each cue and response word pairs. This definition leads to a network containing 14,794 nodes and 239,483 edges. Nodes only resided in one large giant component. Hermits were not expected given the restricted selection of words to match the phonological association network. Islands were likely not present in this analysis given the high number of edges compared to the number of nodes.

The semantic association network was examined for small-world structure. Recall that having a small-world structure indicates that the network is easy to traverse despite its large size, and is hallmarked by having a similar average shortest path length and larger average clustering coefficient than a comparably-sized random network (Watts & Strogatz, 1998). The average shortest path length of the semantic association network was 3.77, whereas the average shortest path length of the comparably-sized random network was 3.74. Using network analysis convention where the difference in values is no greater than 1.5 times in magnitude, the average shortest path length values for the semantic association network and the random network were not significantly different. The average clustering coefficient of the semantic association network was 0.09, whereas the comparably-sized random network had an average clustering coefficient of 0.002. The average clustering coefficient values for the semantic association network and the random network were significantly different by several magnitudes according to network analysis convention. In addition, “small-world-ness” (Humphries & Gurney, 2008) for the semantic association network was 42.12. Therefore, these measures indicate that the semantic association network has a small-world structure.

Next, an analysis was conducted to determine if the semantic association network could be classified as having a scale-free structure. Recall that having a scale-free structure suggests that many nodes have few edges and few nodes have many edges. This is indicated by the degree distribution following a power-law function when plotted on a log-log scale. Figure 33 displays the log-log plot for the degree distribution of the semantic association network. The power-law function was best fit by the equation $y = 1.41x^{-1.65}$, RMSE = 1.37, whereas the exponential curve was best fit by the equation $y = 0.003e^{-0.009x}$, RMSE = 0.16. Since the exponential curve better fits the data than the power-law function, the semantic association network does not have a scale-free structure.

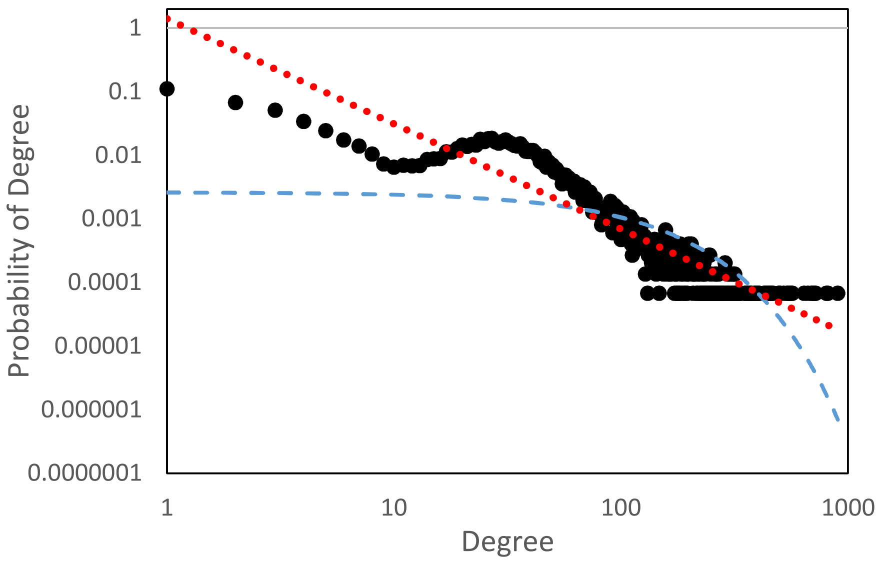

Figure 33. Log-log plot of degree distribution for the Semantic Association Network. The power-law function is represented in red and the exponential curve is represented in blue.

Mixing by degree was also examined in the semantic association network. Recall that assortative mixing by degree occurs when nodes with high degree tend to be connected to other nodes with high degree. On the other hand, disassortative mixing by degree occurs when nodes with high degree tend to be connected to nodes with low degree. A Pearson's correlation of r (239498) = 0.03, $p < .0001$, was found. The r value close to 0 suggests that there was no correlation between a node's degree and each of its neighbor's degree; therefore, the semantic association network does not show evidence of mixing by degree.

Finally, the community structure of the semantic association network was examined. In total, there were 12 communities in the giant component as determined by the Louvain method (Blondel, et al., 2008) with a modularity of 0.32. A modularity value higher than 0.3 is indicative of significant community structure (Clauset, et al., 2004). Figure 28 depicts different communities in the giant component by color. These communities ranged in size from 317 to 2,165 nodes (M = 1232.83, SD = 585.94).

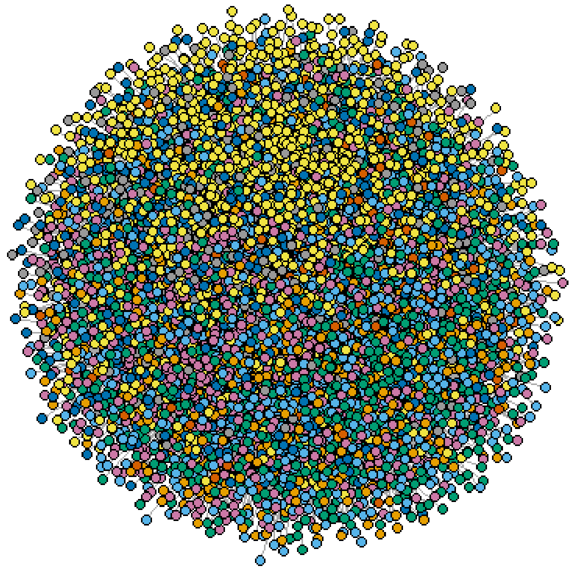

Figure 34. Giant component of the Semantic Association Network. Color represents communities.

Correlation between Networks. The previous network examination focused on macro- and meso- level network structure. However, it is also important to identify differences in the structure of a word in the semantic association network as it compares to the phonological association network. These findings will provide additional insight into how network structure is related in different types of mental lexicon networks. Specifically, the location of a word in each network was examined, as well as a word's degree and clustering coefficient in each network, using the words that are common to both networks.

First, the location of words in each network was examined, as well as how that location may have differed (or remained the same) between networks. Location of words was categorized as being located in the giant component, an island, or as a hermit for each network. Location of nodes from the semantic association network to the phonological association network could change in one of three possible ways: from the giant component of the semantic network to an island of the phonological network, from the giant component of the semantic network to a hermit in the phonological network, or located in the giant component of both networks. (Recall that there were no islands or hermits in the semantic association network). Table 18 provides the proportion of nodes for each type of location change. Most nodes were in the giant component of both networks.

Table 18.

Proportion of Nodes for Each Type of Location Change from the Semantic Association Network to the Phonological Association Network.

Type of Location Change	Count of Nodes	Proportion of Nodes
Giant Component to Island	14558	98.4%
Giant Component to Hermit	196	1.3%
Same Location	40	0.3%

Next, the degree of words in each network was examined. The semantic association network had an average degree of 32.38 (SD = 45.54), whereas the phonological association network had an average degree of 7.02 (SD = 6.22). A Pearson's correlation between the degree of a word in the semantic association network and the degree of the same word in the phonological association network showed that degree between networks was correlated, r (14794) = .46, $p < .0001$ (see Figure 35). Therefore, words have similar degree in each network.

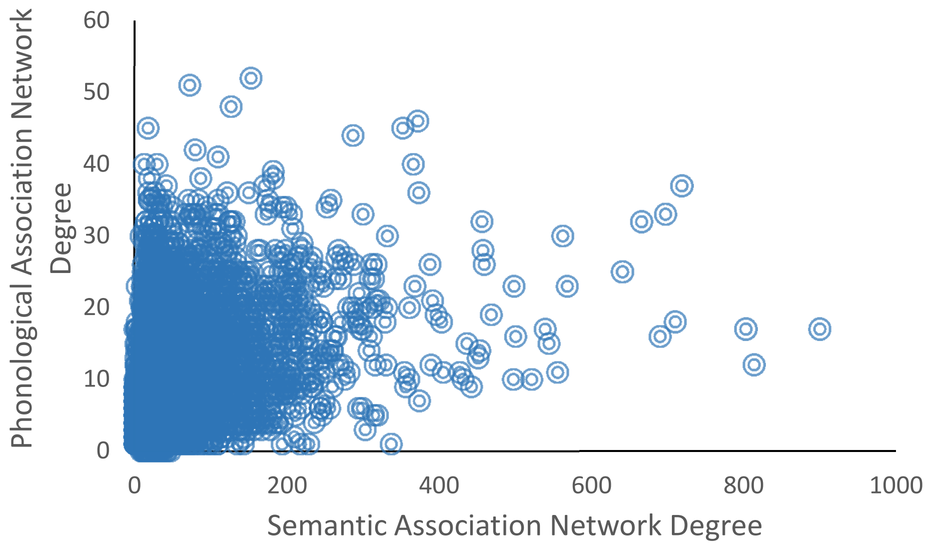

Figure 35. Scatterplot of Semantic Association Network and Phonological Association Network degree.

Lastly, the clustering coefficient of words in each network was examined. A Pearson's r correlation between the clustering coefficient of a word in the semantic association network and the clustering coefficient of the same word in the phonological association network was not correlated, r (14794) = -0.02 p < .0001.

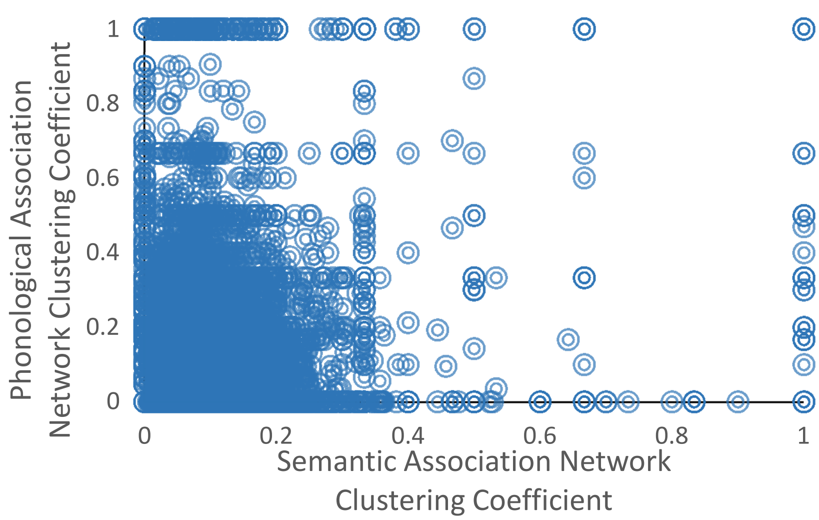

Figure 36. Scatterplot of Semantic Association Network and Phonological Association Network clustering coefficient.

Multiplex Network Analysis. The multiplex structure analysis follows that of Stella, et al. (2017) by comparing average degree, average clustering coefficient, average shortest path length, and the mixing pattern of each individual layer to the aggregated multiplex network (see Table 19). The mean degree of the multiplex was 28.74 (*SD* = 44.63). The average clustering coefficient of the multiplex was 0.09 (*SD* = 0.17). The average shortest path length of the multiplex was 3.83. And, a Pearson's correlation of r (20575) = 0.03, $p < .0001$, was found. The r value close to 0 suggests that there was no correlation between a node's degree and each of its neighbor's degree; therefore, the multiplexn network does not show evidence of mixing by degree.

Table 19.

Network Structure Measures for the Semantic Association Layer, Phonological Association Layer, and Multiplex.

Network Measures	Semantic Layer	Phonological Layer	Multiplex
Average Degree	32.38	5.52	28.74
Average Clustering Coefficient	0.09	0.12	0.09
Average Shortest Path Length	3.77	9.80	3.83
Mixing by Degree	0.03	0.44	0.03

In addition to comparing the aggregated multiplex to each of its individual layers, an analysis assessing edge overlap was also done. Specifically, multilinks and degree of multiplexity were analyzed to assess how much the semantic and phonological layers of the multiplex overlapped. Multilinks are the number of instances where there are multiple edges between a pair of nodes (Bianconi, 2013). In total, there were 4,034 node pairs that had at least one multilink (i.e., both a semantic and phonological edge between them). Examples of node pairs with a multilink include bracelet → anklet, anxiety → anxious, and cake → bake. The degree of multiplexity extends upon multidegree to determine the ratio of node pairs with multilinks to all connected node pairs (Kapferer, 1969). In this analysis, the degree of multiplexity was 0.01, suggesting a very small amount of overlap between the layers.

Individual nodes with multilinks were further assessed by examining how a node's multidegree compared to its semantic degree and its phonological degree. The average multidegree (i.e., number of multilinks) was 3.57 (SD = 4.80). A node's multidegree was compared to the same node's semantic degree and phonological degree using Pearson's correlations. Multidegree for a node was correlated with the semantic degree of the same node, r (5067) = .36, p < .0001, and was also correlated with the phonological degree of the same node, r (5067) = .41, p < .0001 (see Figures 37 and 38).

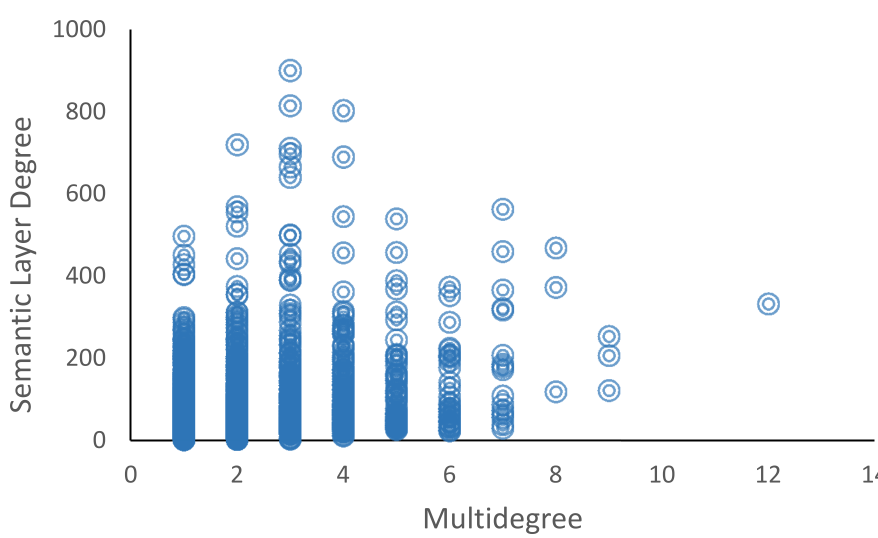

Figure 37. Scatterplot of Multidegree and Semantic Layer degree.

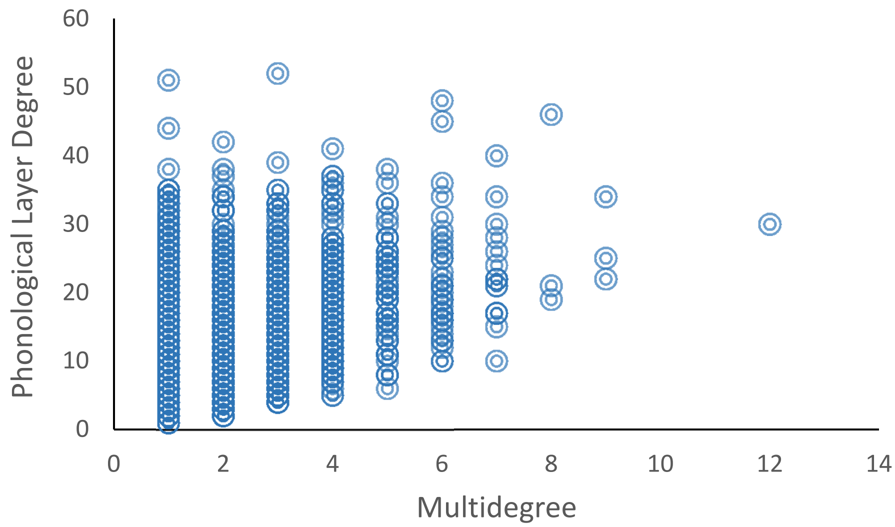

Figure 38. Scatterplot of Multidegree and Phonological Layer Degree.

Lastly, the location of nodes with at least one multilink were assessed. These nodes resided in the giant component of the semantic association network and tended to reside in the giant component of the phonological association network. Specifically, 99.4% of nodes with at least one multilink resided in the giant component of the phonological association network with the remaining 0.6% of nodes residing in an island.

Multiplex Behavioral Analysis. To further assess the multiplex structure, two regression analyses were conducted predicting lexical decision reaction time and naming reaction time from semantic degree, phonological degree, aggregated multiplex degree (semantic + phonological degree), and multidegree (number of multilinks), as well as an interaction between semantic degree and phonological degree. Previous results indicate that words with higher semantic degree were responded to faster than words with lower semantic degree in a visual lexical decision task (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Duñabeitia, Avilés, & Carreiras, 2008; Yates, Locker, & Simpson, 2003) and word naming task (Duñabeitia, et al., 2008). Words with higher phonological degree were also responded to faster than words with lower phonological degree in a visual lexical decision task (Yates, 2005; Yates, Locker, & Simpson, 2004) and word naming task (Yates, 2005). Aggregated multiplex degree, multidegree, and the interaction between semantic and phonological degree will provide new insight into how examination of multiple layers simultaneously impacts language processing. The log of each measure of degree was taken given the skewed distribution of this data.

Stepwise modeling building was conducted using R (R Core Team, 2017) to determine the best model that only includes predictors that contribute significantly to the model. In this procedure, both forward and backward stepwise modeling occurred to determine the predictors of the final model. In both the lexical decision and naming regression analyses, the final models included all degree measures (see Table 20 and 21).

Table 20.

Regression Analysis Predicting Lexical Decision Reaction Time from Different Measures of Degree.

	Estimate	Std. Error	*p*-value
Intercept	858.61	9.76	< .0001
Log Semantic Degree	56.65	8.68	< .0001
Log Phonological Degree	21.64	4.33	< .0001
Log Aggregated Multiplex Degree	-86.88	11.20	< .0001
Log Multidegree	10.74	2.45	< .0001
Log Semantic Degree * Log Phonological Degree	-11.72	1.04	< .0001

Table 21.

Regression Analysis Predicting Naming Reaction Time from Different Measures of Degree.

	Estimate	Std. Error	*p*-value
Intercept	794.81	8.11	< .0001
Log Semantic Degree	62.88	7.22	< .0001
Log Phonological Degree	10.17	3.60	< .01
Log Aggregated Multiplex Degree	-89.26	9.31	< .0001
Log Multidegree	3.11	2.03	0.13
Log Semantic Degree * Log Phonological Degree	-6.11	0.86	< .0001

For the regression analysis predicting lexical decision reaction time, all measures of degree were significant. Interestingly, as the aggregated multiplex degree of a word increased, reaction time also decreased, but as multiplex degree of a word increased, reaction time increased. Having many connections decreased lexical decision time, as long as those connections did not overlap in the multiplex. The interaction between semantic and phonological degree was also significant (see Figure 39). When semantic degree is low, phonological degree did not have a large effect on lexical decision reaction time. However, as semantic degree increased, reaction time also increased, especially when phonological degree was low.

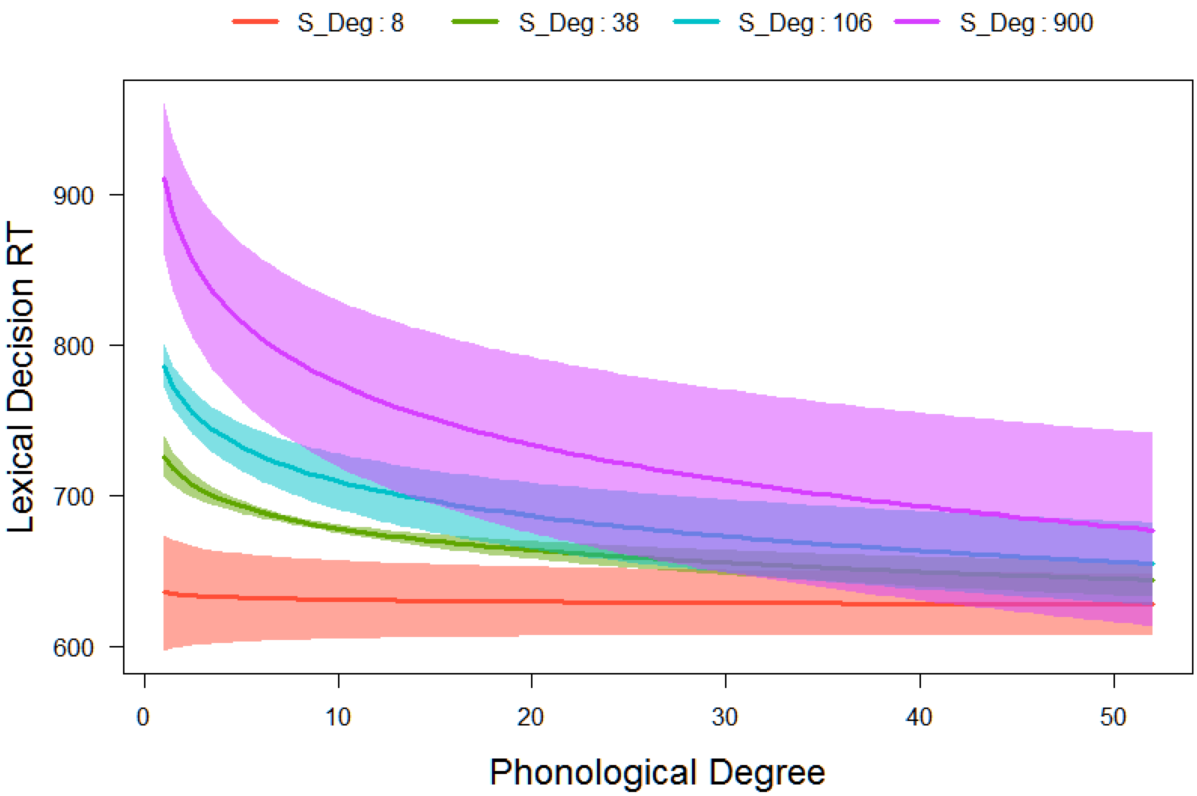

Figure 39. Interaction of semantic and phonological degree on lexical decision reaction time.

For the regression analysis predicting naming reaction time, all measures of degree were significant except for multidegree, although this predictor added significantly to the model. As the aggregated multiplex degree of a word increased, reaction time also decreased. The interaction between semantic and phonological degree was also significant (see Figure 40) following the same interaction pattern as the lexical decision regression analysis. When semantic degree is low, phonological degree did not have a large effect on lexical decision reaction time. However, as semantic degree increased, reaction time also increased, especially when phonological degree was low.

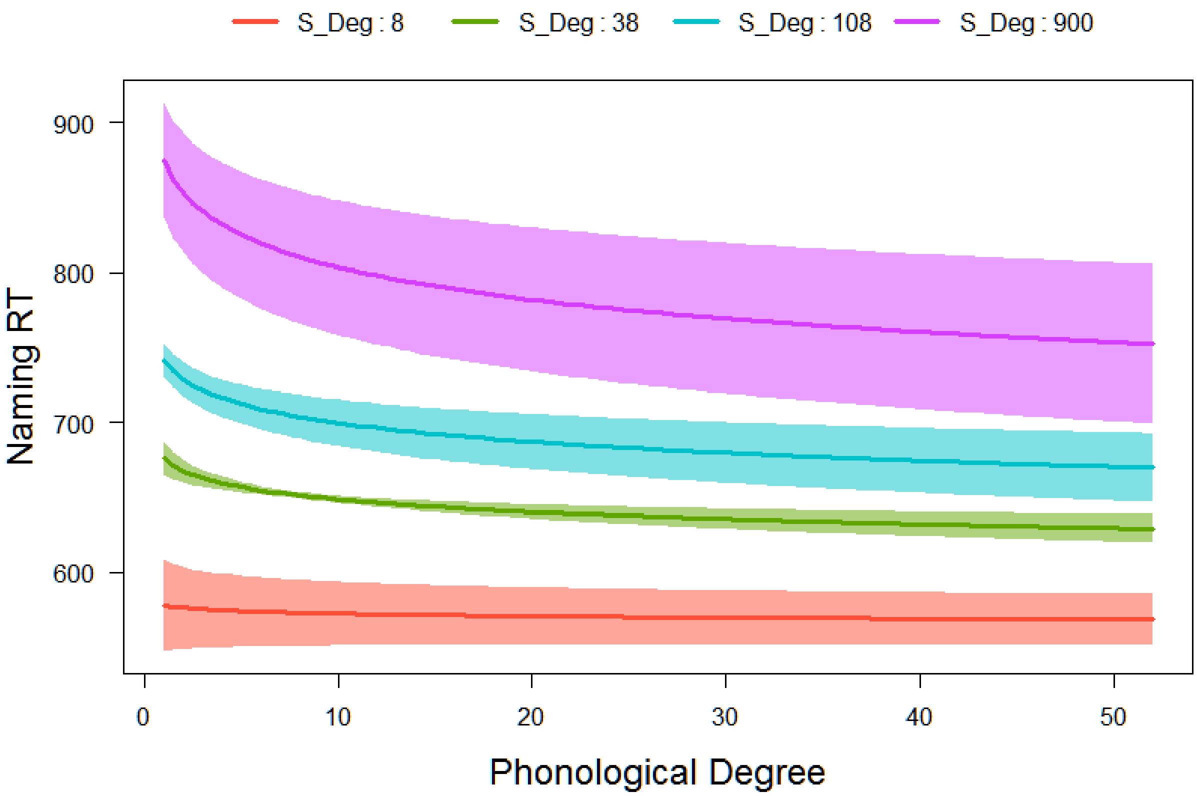

Figure 40. Interaction of semantic and phonological degree on naming reaction time.

Discussion

A semantic association network was compared to a phonological association network using the same nodes, and then combined into a multiplex network to assess the amount of overlap between these two networks. The semantic association network and the phonological association network were different in certain aspects of their structure, although both networks would be described as having small-world, but not scale-free structure.

Despite both networks having all, or almost all, of their nodes in a large giant component, the two networks differed in their average shortest path length and average clustering coefficient. Specifically, the semantic association network had a smaller average path length and a smaller average clustering coefficient than the phonological association network. This is surprising given the significantly larger number of edges placed in the semantic association network than the phonological association network. This finding suggests that having a large number of edges does not necessarily equate to having a more structured network. This is also evident in the smaller number of communities and smaller modularity value for the semantic association network as compared to the phonological association network, as well as the lack of assortative mixing by degree in the semantic association network.

The semantic association network results should be taken with caution. It is possible that the large number of edges in the semantic association network diluted network structure findings. For example, the current dataset did not have available frequency of responses. Including weights to edges might have changed network structure values, or could have been used to eliminate less frequently, and potentially irrelevant, associations. In addition, it is not known whether the sample of participants used to acquire the semantic association data is similar to the sample of participants used to acquire the phonological association data. Differences in participant demographics, like age and education, could substantially impact the kind of responses that were given and change the structure of the network. Another existing semantic association data set, the University of South Florida Free Association Norms (Nelson, McEvoy, & Schreiber, 1998), could be used as a comparison to the semantic association data set provided by S. De Deyne to further assess the reliability of the present network analysis.

Nevertheless, a multiplex was created to assess the amount of overlap between the semantic association and phonological association networks. In the multiplex, a semantic layer and a phonological layer were created where edges connect words according to that layer's association data. The multiplex network structure resembled more closely the semantic association layer than the phonological association layer. In addition, there was a small overlap between the semantic and phonological association layers providing evidence that these layers contribute different information in the representation of the mental lexicon. This corroborates the notion that words meaning and sound are mostly unrelated (with the exception of onomatopoeia) (Hockett & Hockett, 1960). Interestingly, for those nodes with at least one multilink, their multidegree was correlated with their semantic and phonological degree, and these nodes tended to reside in the giant component of each layer.

Taken together, this multiplex that considers associations in both the semantic and phonological layers resembles the multiplex of Stella, et al. (2017) whose multiplex considers different measures of semantic, syntactic, and phonological relationships. Specifically, the present multiplex was like Stella et al.'s (2017) multiplex in average degree, average shortest path length, and mixing pattern. However, the two multiplex networks diverge in average clustering coefficient. The average clustering coefficient of Stella et al.'s multiplex was higher than the average clustering coefficient in the present multiplex. This difference may be due to the additional layers of information that are included in Stella et al.'s (2017) multiplex.

Since analysis of more sophisticated measures of the multiplex structure is not possible due to current computational limitations, an additional analysis was conducted examining how different measures of degree of single-layer networks and the multiplex influence behavioral performance. The regression analyses assessing visual lexical decision and naming of words showed that both single-layer network and multiplex degree measures contribute to the prediction of performance. In previous work, higher degree in the single-layer semantic network and higher degree in the single-layer phonological network led to faster visual lexical decision and naming reaction times. However, in the present analysis, an interaction was found that provides new evidence of how semantic and phonological degree interact. It is interesting to note that increasing semantic degree had the reverse effect in this analysis, such that higher semantic degree resulted in slower lexical decision and naming tasks, especially when phonological degree was small. This contradictory finding may be due to the much larger range in semantic degree in the present analysis (from 1 to 900), whereas previous findings only look at high versus low semantic degree with a much smaller range. For example, in Duñabeitia, Avilés, and Carreiras (2008), high semantic degree ranged from 30 – 39.6 and low semantic degree ranged from 5.6 – 8.1.

In addition, the regression analyses conducted in this study show that consideration of multiple layers of information is necessary to fully understand language processes. The aggregated multiplex degree, multidegree, and the interaction between semantic and phonological degree were all contributed significantly to the regression analyses. Although aggregated multiplex degree and the interaction between semantic and phonological degree could be analyzed without creating a multiplex, multidegree is unique to a multiplex examination. Therefore, the multiplex provides a novel way to assess overlap and interaction between language systems that could not be done with single-layer network analyses alone.

These regression analyses highlight that examination of the entire mental lexicon structure is important to visual word processing. Continued research can explore the effect of multiplex structure on spoken word processing and word retrieval. In addition, other multiplex measures may prove even more predictive of language processes than multidegree, which only assesses the overlap between layers. For example, multiplex closeness centrality was the most predictive variable in Stella, et al. (2017) assessing word acquisition in children.

Given the current data available, the semantic association network and the phonological association network exhibit small-world structures that overlap minimally. However, these networks differ dramatically in their number of edges, which may influence the present network structure findings. In particular, the phonological association network may be too sparse, given the small number of cue word presentations to each participant (ranging from 6-8), leading to potentially missing phonological associations that should be represented and the inability to truly distinguish a viable edge from a spurious edge. On the other hand, the semantic association network may be diluted by edges, given the large number of cue-response pairs provided by S. De Deyne (429,401 pairs) with no indication of frequency of response weight and/or the ability to filter edges. Therefore, the present analysis provides only an initial examination of a multiplex of semantic and phonological associations.

Chapter 6: General Discussion

Language processes are known to involve multiple systems of information, including semantics and phonology. Research from the emerging, interdisciplinary field of Network Science provides evidence that structure is crucial to understanding those processes. The newest frontier in the application of Network Science to psycholinguistics is to move beyond examination of single-layer networks that examine only one system at a time, and instead consider the entire mental lexicon using a multiplex network. A multiplex network provides the ability to understand how different systems overlap and interact during language processes.

The present work is the first to describe the multiplex structure of a network representing semantic and phonological relationships amongst words. This multiplex uses association data to link words in both layers, providing a common measure of similarity, and uses a large dataset. Only one other multiplex language network has been examined that also considers semantic and phonological, as well as syntactic, relationships among words, but does so with a limited set of words and for word acquisition in children (Stella, et al., 2017). Therefore, this work continues to contribute to the investigation of how multiplex structure influences language processes by using a much larger dataset representing the adult mental lexicon.

To construct the present multiplex, phonological association data was collected, while semantic association data was obtained from an existing dataset. Phonological associations have been used by researchers to assess phonological similarity; however, a large dataset has not been made available for research use. The present collection of phonological associations provides an initial dataset that can continue to be expanded and used to better understand how people think about phonological similarity. Indeed, age was examined as one factor that influenced association responses. As age increased, more time was spent on the association task and more secondary and tertiary responses were made. In addition, adults in early and late middle adulthood had to complete the task via Amazon Mechanical Turk. Given the computer literacy skills required to use Amazon Mechanical Turk this suggests that the individuals in this sample were high functioning adults. Therefore, it would be important to continue collecting association data from a larger, more diverse sample of adults, including adults over the age of 75 years.

An alternative method to defining phonological similarity is to assess the amount of overlap in phonemes between words. One method that has been well-studied is a one-phoneme difference through substitution, deletion, or addition of a phoneme (Luce & Pisoni, 1998). Indeed, this definition of phonological similarity has also been used to construct a single-layer phonological network, and the resulting phonological structure has been shown to influence several language processes (Vitevitch, et al., 2014).

The phonological association data collected in the present study was used to construct a phonological network, which was then compared to the network structure of the one-phoneme metric network of Vitevitch (2008). Interestingly, these two networks share small-world properties, assortative mixing by degree, and significant community structure, but represent phonological similarity in different ways. Behavioral tests can be used to further compare the structure of these two types of phonological networks and to better understand how each type of phonological network contributes to language processes. Even if similar effects on language processing are found, the results would suggest that the one-phoneme metric would provide an easier method to achieve the same results, but the phonological association data would provide the opportunity to weight links by frequency, adding an additional piece of information to the network.

The present phonological association data also provided the opportunity to examine how phonological network structure may change across adulthood. Interestingly, this examination found little change in phonological network structure with increasing age, which is in contrast to what has already been shown for the semantic network structure. Phonological processing has been shown to be disrupted with increasing age in word retrieval processes (e.g., increase in tip-of-the-tongue states), so it is necessary to further understand how the lack of change in phonological network structure may be contributing to inefficient phonological processing. It should be noted, though, that the present results may be an underestimate as data is sparse for each age-related network. For example, many cues were only responded to by one person in each age group. In addition, the data does not include many participants over the age of 75, limiting the analyses to adults through late middle adulthood. Significant cognitive changes, like word retrieval difficulties, tend to emerge in older adulthood. Therefore, it may be the case that the lack of change in the structure of the phonological network seen in the present analysis is due to the inability to adequately examine the phonological network structure for individuals over 75 years. Future work can continue to collect phonological association data from adults of all ages, but particularly those over the age of 75, to obtain a better representation of age-related changes in phonological network structure. Behavioral testing and simulations of the phonological network structure across adulthood can also provide insight into how processing is impacted by the lack of change in phonological network structure.

The present work, thus far, has focused on the single-layer phonological network structure. However, as stated previously, examining a multiplex that includes semantic and phonological layers of information is necessary given that these systems are connected and can interact during word processing. The multiplex examined in this study consisted of a semantic association layer and a phonological association layer. The two layers were found to overlap minimally. In addition, a regression analysis using step-wise model building provided evidence that inclusion of predictors that account for single-layer and multiplex structure are needed for prediction of visual lexical decision and naming. The present multiplex analysis focused on degree, but future work can continue to explore the effect of other multiplex measures. For example, in the Stella et al. (2017) multiplex analysis, the measure of closeness centrality was most important in the multiplex, and may serve as a useful predictor in the present multiplex analysis as well. One drawback to the current multiplex analysis is the size of the dataset. Given current computational power the large size of the multiplex network made more complex analyses impossible. However, the field of Network Science (and computational power) continues to develop. Alternative methods of testing the multiplex structure through behavioral experiments will provide a way to continue the effort of understanding how multiplex structure influences language processes.

In sum, Network Science provides a useful method for examining the structure of representations in the mental lexicon. Single--layer network have provided evidence that structure is critical for understanding language processes. Methods and analyses to test the multiplex structure of the mental lexicon are the new frontier in the application of Network Science to psycholinguistics. As the computational methods continue to develop, there will be a better understanding of the overlap and interaction between systems of information. In conjunction with the computational methods, behavioral experiments will provide a way to test theories and explore how network structure can be used to understand changes in language processes across the lifespan. Importantly, understanding processes is important to psycholinguistic research, but in order to understand those processes, one must also fully understand the structure in which those processes take place. And Network Science provides the tools needed to do so.

References

Albert, R., Jeong, H., & Barabási, A. L. (2000). Error and attack tolerance of complex networks. *Nature, 406,* 378-382.

Arbesman, S., Strogatz, S. H., & Vitevitch, M. S. (2010). Comparative analysis of networks of phonologically similar words in English and Spanish. *Entropy*, *12*(3), 327-337.

Baayen, R. H., Piepenbrock, R., & van H, R. (1993). The {CELEX} lexical data base on {CD-ROM}.

Balota, D.A., Yap, M.J., Cortese, M.J., Hutchison, K.A., Kessler, B., Loftis, B., … Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39*, 445-459.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*(2), 283-316.

Beckage, N., Smith, L., & Hills, T. (2011). Small worlds and semantic network growth in typical and late talkers. *PloS One*, *6*(5), e19348. doi: 10.1371/journal.pone.0019348

Bianconi, G. (2013). Statistical mechanics of multiplex networks: Entropy and overlap. *Physical Review E*, *87*(6), 062806.

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

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

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977-990.

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language*, *30*(5), 542-579.

Chan, K. Y., & Vitevitch, M. S. (2010). Network structure influences speech production. *Cognitive Science*, *34*(4), 685-697.

Clauset, A., Newman, M. E., & Moore, C. (2004). Finding community structure in very large networks. *Physical review E*, *70*(6), 066111.

Cozzo, E., Kivelä, M., De Domenico, M., Solé-Ribalta, A., Arenas, A., Gómez, S., ... & Moreno, Y. (2015). Structure of triadic relations in multiplex networks. *New Journal of Physics*, *17*(7), 073029.

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

Damian, M. F., Vigliocco, G., & Levelt, W. J. (2001). Effects of semantic context in the naming of pictures and words. *Cognition*, *81*(3), 77-86.

De Deyne, S., Navarro, D. J., & Storms, G. (2013). Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behavior Research Methods*, *45*(2), 480-498.

De Deyne, S., & Storms, G. (2008). Word associations: Network and semantic properties. *Behavior Research Methods*, *40*(1), 213-231.

De Domenico, M., Porter, M. A., Arenas, A. MuxViz: A tool for multilayer analysis and visualization of networks. *Journal of Complex Networks, 3*(2), 159-176.

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition*, *42*(1), 287-314.

Dubossarsky, H., De Deyne, S., & Hills, T. T. (2017). Quantifying the structure of free association networks across the life span. *Developmental Psychology*. Advanced online publication. doi: 10.1037/dev0000347

Duñabeitia, J. A., Avilés, A., & Carreiras, M. (2008). NoA's ark: Influence of the number of associates in visual word recognition. *Psychonomic Bulletin & Review*, *15*(6), 1072-1077.

Ferrer i Cancho, R., & Solé, R. V. (2001). The small world of human language. *Proceedings of the Royal Society of London B: Biological Sciences*, *268*(1482), 2261-2265.

Gómez, S., Diaz-Guilera, A., Gómez-Gardeñes, J., Pérez-Vicente, C. J., Moreno, Y., & Arenas, A. (2013). Diffusion dynamics on multiplex networks. *Physical review letters*, *110*(2), 028701.

Hantsch, A., Jescheniak, J. D., & Schriefers, H. (2005). Semantic competition between hierarchically related words during speech planning. *Memory & Cognition*, *33*(6), 984-1000.

Harley, T. A., & Bown, H. E. (1998). What causes a tip□of□the□tongue state? Evidence for lexical neighbourhood effects in speech production. *British Journal of Psychology*, *89*(1), 151-174.

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009). Longitudinal analysis of early semantic networks preferential attachment or preferential acquisition? *Psychological Science*, *20*(6), 729-739.

Hockett, C. F., & Hockett, C. D. (1960). The origin of speech. *Scientific American*, *203*(3), 88-97.

Humes, L. E., & Dubno, J. R. (2010). Factors affecting speech understanding in older adults. In S. Gordon-Salant, R. D. Frisina, A. N. Popper, & R. R. Fay (Eds.), *The Aging Auditory System* (pp. 211-257). New York, NY: Springer.

Humphries, M. D., & Gurney, K. (2008). Network 'small-world-ness': a quantitative method for determining canonical network equivalence. *PloS One*, *3*(4), e0002051. doi: 10.1371/journal.pone.0002051

Hutson, J., & Damian, M. F. (2014). Semantic gradients in picture-word interference tasks: is the size of interference effects affected by the degree of semantic overlap? *Frontiers in Psychology*, *5*, 872. doi: 10.3389/fpsyg.2014.00872.

James, L. E., & Burke, D. M. (2000). Phonological priming effects on word retrieval and tip-of-the-tongue experiences in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(6), 1378-1391.

Kapferer B., & Mitchell J. C. (1969). Norms and the manipulation of relationships in a work context. In J. C. Mitchell (Ed.), Social Networks in Urban Situations: Analyses of Personal Relationships in Central African Towns. Manchester: Manchester University Press.

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

Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., & Porter, M. A. (2014). Multilayer networks. *Journal of Complex Networks*, *2*(3), 203-271.

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

Kučera, H., & Francis, W. (1967). Computational analysis of present day American English. Providence, RI: Brown University Press.

Luce, P. A., & Large, N. R. (2001). Phonotactics, density, and entropy in spoken word recognition. *Language and Cognitive Processes*, *16*(5-6), 565-581.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*(1), 1-36.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, *28*(2), 203-208.

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

Morais, A. S., Olsson, H., & Schooler, L. J. (2013). Mapping the structure of semantic memory. *Cognitive Science*, *37*(1), 125-145.

Motter, A. E., De Moura, A. P., Lai, Y. C., & Dasgupta, P. (2002). Topology of the conceptual network of language. *Physical Review E*, *65*(6), 065102.

Nelson, D. L., Bennett, D. J., Gee, N. R., Schreiber, T. A., & McKinney, V. M. (1993). Implicit memory: Effects of network size and interconnectivity on cued recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(4), 747-764.

Nelson, D. L., & McEvoy, C. L. (1979). Encoding context and set size. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(3), 292-314.

Nelson, D. L., McEvoy, C. L., & Dennis, S. (2000). What is free association and what does it measure? *Memory & Cognition*, *28*(6), 887-899.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://www.usf.edu/FreeAssociation/

Newman, M. E. (2002). Assortative mixing in networks. *Physical Review Letters*, *89*(20), 208701.

Newman, M. E. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, *103*(23), 8577-8582.

Newman, M. E., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, *69*(2), 026113.

Newman, M. (2008). The physics of networks. *Physics Today*, *61*(11), 33-38.

Ognyanova, K. (2017) *Network visualization with R.* Retrieved from www.kateto.net/network-visualization

R Core Team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http:/*/www.R-project.org*

Rahman, R. A., & Aristei, S. (2010). Now you see it… and now again: Semantic interference reflects lexical competition in speech production with and without articulation. *Psychonomic Bulletin & Review*, *17*(5), 657-661.

Ravasz, E., & Barabási, A. L. (2003). Hierarchical organization in complex networks. *Physical Review E*, *67*(2), 026112.

Sailor, K., Brooks, P. J., Bruening, P. R., Seiger-Gardner, L., & Guterman, M. (2009). Exploring the time course of semantic interference and associative priming in the picture–word interference task. *The Quarterly Journal of Experimental Psychology*, *62*(4), 789-801.

Schneider, B. A., Pichora-Fuller, K., & Daneman, M. (2010). Effects of senescent changes in audition and cognition on spoken language comprehension. In S. Gordon-Salant, R. D.

Frisina, A. N. Popper, & R. R. Fay (Eds.), *The Aging Auditory System* (pp. 167-210). New York, NY: Springer.

Schreiber, T. A., & Nelson, D. L. (1998). The relation between feelings of knowing and the number of neighboring concepts edgeed to the test cue. *Memory & Cognition*, *26*(5), 869-883.

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

Sommers, M. S. (1996). The structural organization of the mental lexicon and its contribution to age-related declines in spoken-word recognition. *Psychology and Aging*, *11*(2), 333-341.

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

Steyvers, M., & Tenenbaum, J. B. (2005). The large□scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive science*, *29*(1), 41-78.

Strogatz, S. H. (2001). Exploring complex networks. *Nature*, *410*(6825), 268-276.

Van Rensbergen, B., Storms, G., & De Deyne, S. (2015). Examining assortativity in the mental lexicon: Evidence from word associations. *Psychonomic Bulletin & Review*, *22*(6), 1717-1724.

Verhaeghen, P. (2003). Aging and vocabulary score: A meta-analysis. *Psychology and Aging*, *18*(2), 332-339.

Vieth, H. E., McMahon, K. L., & de Zubicaray, G. I. (2014). Feature overlap slows lexical selection: Evidence from the picture–word interference paradigm. *The Quarterly Journal of Experimental Psychology*, *67*(12), 2325-2339.

Vigliocco, G., Vinson, D. P., Damian, M. F., & Levelt, W. (2002). Semantic distance effects on object and action naming. *Cognition*, *85*(3), 61-69.

Vitevitch, M. S. (1997). The neighborhood characteristics of malapropisms. *Language and Speech*, *40*(3), 211-228.

Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(4), 735-747.

Vitevitch, M. S. (2008). What can graph theory tell us about word learning and lexical retrieval? *Journal of Speech, Language, and Hearing Research*, *51*(2), 408-422.

Vitevitch, M. S., & Castro, N. (2015). Using Network Science in the language sciences and clinic. *International Journal of Speech-Language Pathology*, *17*(1), 13-25.

Vitevitch, M. S., Chan, K. Y., & Goldstein, R. (2014). Insights into failed lexical retrieval from Network Science. *Cognitive Psychology*, *68*, 1-32.

Vitevitch, M. S., Goldstein, R., & Johnson, E. (2016). Path-length and the misperception of speech: insights from Network Science and Psycholinguistics. In A. Mehler, A. Lucking, S. Banisch, P. Blanchard, & B. Frank-Job (eds.), *Towards a Theoretical Framework for Analyzing Complex Linguistic Networks* (pp. 29-45). Springer Berlin Heidelberg.

Vitevitch, M. S., Goldstein, R., Siew, C. S., & Castro, N. (2014). Using complex networks to understand the mental lexicon. *Yearbook of the Poznan Linguistic Meeting, 1,* 119-138.

Vitevitch, M. S., & Luce, P. A. (2016). Phonological neighborhood effects in spoken word perception and production. *Annual Review of Linguistics*, *2*, 75-94.

Vitevitch, M. S., & Sommers, M. S. (2003). The facilitative influence of phonological similarity and neighborhood frequency in speech production in younger and older adults. *Memory & Cognition*, *31*(4), 491-504.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *Nature*, *393*(6684), 440-442.

Yates, M. (2005). Phonological neighbors speed visual word processing: evidence from multiple tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1385-1397.

Yates, M., Locker, L., & Simpson, G. B. (2003). Semantic and phonological influences on the processing of words and pseudohomophones. *Memory & Cognition*, *31*(6), 856-866.

Yates, M., Locker, L., & Simpson, G. B. (2004). The influence of phonological neighborhood on visual word perception. *Psychonomic Bulletin & Review*, *11*(3), 452-457.

Appendix A.

Network Measures.

Network Measure	Level of Analysis	Description
Degree	Micro-Level	The number of immediate connections of a given node
Clustering coefficient	Micro-Level	The likelihood that nodes of a particular node are also connected to each other
Shortest Path Length	Micro-Level	The shortest path, or number of edges to be crossed, from a given node to some other node in the network
Communities	Meso-Level	Sub-groupings of nodes that are more connected to each other than to other sub-groupings
Average Clustering Coefficient	Macro-Level	The average of all nodes' clustering coefficient in the network
Average Shortest Path Length	Macro-Level	The average of all shortest path lengths of all possible pairs of nodes in the network
Degree Distribution	Macro-Level	A plot using a logarithmic scale that shows the frequency of each value of degree for all nodes in the network
Location	Macro-Level	Nodes can reside in one of three locations of the network: a giant component, an island (or smaller component), or as an isolated hermit (not connected to any other node), where components are groupings of nodes connected to each other, but not connected to any other component
Small-World Structure	Macro-Level	Similar average shortest path length and higher average clustering coefficient than a comparably-sized random network
Scale-Free Structure	Macro-Level	A degree distribution that follows a power-law, which contrasts with the Poisson distribution of a comparably-sized random network
Aggregated Multiplex Degree	Multiplex	The sum of all layers' degree in the multiplex. In the present analysis, this is the sum of semantic and phonological degree for a given word.
Multidegree	Multiplex	The number of multilinks for a given node, where a multilink is placed between a pair of nodes if they are connected in each layer of the network. In the present analysis, multilinks are placed between pairs of nodes that are connected in both the semantic and phonological layers.

Appendix B.

Homograph Cue Words and Sentences.

absent	She was absent from class today.
abuse	Don't abuse the animals.
addict	The addict needed help.
address	She wrote her address on the paper.
adept	She is an adept leader for our company.
advocate	We advocate for change.
affect	That will affect the results.
aged	He aged very quickly.
alloy	The alloy was stronger than steel.
ally	You are my ally
alternate	Let's alternate between the two teams.
articulate	Please articulate your perspective.
associate	My associate will help you out.
attribute	Sensitivity is his best attribute.
bass	He caught a bass on his fishing trip.
beloved	My beloved toy has broken.
blessed	She felt blessed after the experience.
bow	He shot the bow in the field.
buffet	I was stuffed after eating at the buffet.
certificate	He received a certificate at the meeting.
close	They are about to close.
closer	He moved closer to the screen
combat	The combat waged on for months.
combine	Combine the toys into one basket.
compact	She dropped her compact on the floor.
complex	The math problem was complex.
compliment	He received a nice compliment on his performance.
compress	Compress the material into a ball.
concert	They danced all night at the concert.
conduct	They conduct business together frequently.
conflict	The conflict was resolved.
confound	His explanation will confound you.
conglomerate	The new conglomerate is very powerful.
congress	Congress will discuss the proposed law.
conserve	Conserve your energy.
console	The video game console is broken.

construct	I will construct a new building.
content	He is content with his job.
contents	Include a table of contents in your paper.
contest	She won the contest.
contract	She signed the contract.
contrast	The contrast between the images was stark.
converse	They plan to converse over dinner.
convert	I will convert to the latest upgrade.
convict	The judge will convict the criminal.
coordinate	He will coordinate the event.
crooked	He walked with a crooked cane.
decrease	Submarines decrease rapidly
defect	The toy had a defect.
delegate	They delegate tasks equally in the group.
deliberate	She made a deliberate decision.
desert	The desert was extremely hot.
desolate	The desolate landscape was frightening.
digest	Cows digest food quickly.
document	The document was signed.
documents	The documents were signed.
dove	The dove flew out of the cage.
drawer	The dresser drawer is struck.
drawers	The dresser drawers are stuck.
duplicate	You will receive a duplicate copy.
elaborate	Please elaborate on your reasoning.
entrance	She came through the entrance
escort	The visitor had an escort team.
estimate	They will estimate the cost of service.
evening	This evening is beautiful.
excess	The excess was donated.
excise	The excise tax was very steep.
excuse	Her excuse was accepted.
exploit	They will exploit the services offered.
extract	The machines extract the material.
fragment	The fragment was thrown away.
frequent	Frequent attendance is noticed.
gnome	I gave her a garden gnome.
graduate	They graduate next weekend.
hinder	Do not hinder my momentum.
house	They bought a new house.
implement	That farming implement is broken.

initiate	I will initiate the project.
integral	You are an intergral part of the team.
interest	You will pay interest on the loan.
intimate	The couple had an intimate dinner.
invalid	Your password is invalid.
lead	I will lead the event.
leading	She is leading the event.
legitimate	That is a legitimate excuse.
lineage	Their lineage traces back hundreds of years.
live	The live show was amazing.
lives	She lives peacefully.
lower	The river is lower than usual.
minute	One minute equals 60 seconds.
moderate	That was a moderate amount of money.
mow	He will mow the grass.
multiply	Multiply the numbers to get the answer.
number	Here is my number.
object	The object is round.
offense	The offense scored ten points.
ornament	The ornament glittered on the tree.
pace	My pace improved by two minutes.
pedal	The last flower pedal fell off.
perfect	I will work to perfect my timing.
postulate	I postulate the existence of aliens.
prayer	The congregation said a prayer.
precedent	This serves as a prescedent for future cases.
predicate	The company will predicate a change in policy.
preposition	There is a preposition in this sentence
present	He received a present.
presents	He received many presents.
produce	We produce that computer.
progress	Her progress report had high marks.
project	The class project is due tomorrow.
protest	The protest was peaceful.
raven	The black raven flew overhead.
read	The teacher said to read carefully
rebel	We rebel against that idea.
recall	I recall that memory.
record	The record played all night.
recover	The police will recover the stolen watch.
reflex	Her reflex to the ball was quick.

refuse	We refuse to accept.
reject	I reject your offer.
relay	We won the relay race.
release	He will release the animal
research	He will research that topic.
reside	I reside over there
resume	She will resume after the break.
river	The river was flowing quickly.
row	Row the boat.
rowing	The rowing team won the gold medal.
secretive	He is secretive about his new job.
segment	I have a segment of an orange.
separate	They are a separate group
sewer	The sewer system smelled awful.
shower	He took a long shower.
showers	He is restricted to two showers a day.
singer	That singer has a beautiful voice.
sow	The sow played in the mud pit.
stingy	She is stingy with her money.
subject	That was a hard subject to learn.
subordinate	He is in a subordinate position.
suite	They reserved the honeymoon suite.
supply	Please supply the drinks.
survey	She took the survey in class.
tarry	They tarry for the boat to arrive.
tear	He had a tear in his eye.
tears	He had tears in his eyes.
tower	He is at the top of the tower
unused	That is an unused glass.
use	Let's use this chair.
used	She used up the remaining supplies.
vice	Their vice is gambling.
viola	She played beautiful music with her viola.
wicked	She had a wicked laugh.
wind	A strong wind came with the storm.
wound	The wound needed immediate medical attention.