THE PHONOGRAPHIC NETWORK OF LANGUAGE: USING NETWORK SCIENCE TO INVESTIGATE THE PHONOLOGICAL AND ORTHOGRAPHIC SIMILARITY STRUCTURE OF LANGUAGE

BY

Cynthia S. Q. Siew

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

Chairperson Michael Vitevitch, Ph.D.

Evangelia Chrysikou, Ph.D.

Kelsie Forbush, Ph.D.

Joan Sereno, Ph.D.

Allard Jongman, Ph.D.

Date defended: 8th May 2017

The Dissertation Committee for Cynthia S. Q. Siew certifies that this is the approved version of the following dissertation:

THE PHONOGRAPHIC NETWORK OF LANGUAGE: USING NETWORK SCIENCE TO INVESTIGATE THE PHONOLOGICAL AND ORTHOGRAPHIC SIMILARITY STRUCTURE OF LANGUAGE

____

Chairperson Michael Vitevitch, Ph.D.

Date approved: 7th June 2017

Abstract

Orthographic effects in spoken word recognition and phonological effects in visual word recognition have been observed in a variety of behavioral experimental paradigms, strongly suggesting that a close interrelationship exists between phonology and orthography. However, the metrics used to investigate these effects, such as consistency and neighborhood size, fail to generalize to words of various lengths or syllable structures, and do not take into account the more global similarity structure that exists between phonological and orthographic representations in the language. To address these limitations, the tools of Network Science were used to simultaneously characterize the phonological as well as orthographic similarity structure of words in English within a phonographic multiplex. In this paper, I analyze a section of the phonographic multiplex known as the *phonographic network* of language, where links are placed between words that are *both* phonologically and orthographically similar to each other, i.e., a link would be placed between words such as 'pant' (/p@nt/) and 'punt' (/p^nt/). Conventional psycholinguistic experiments (auditory naming and auditory lexical decision) and an archival analysis of the English Lexicon Project (visual naming and visual lexical decision) were conducted to investigate the influence of two network science metrics derived from the phonographic network—phonographic *degree* and phonographic *clustering coefficient*—on spoken and visual word recognition. Results indicated a facilitatory effect of phonographic degree on visual word recognition, and a facilitatory effect of phonographic clustering coefficient on spoken word recognition. The present findings have implications for theoretical models of spoken and visual word recognition, and for increasing our understanding of language learning and language disorders.

Acknowledgements

The famous Chinese novel, Journey to the West (西游记), tells the story of a Buddhist monk accompanied by three protectors (Monkey King, Pigsy, Sandy), on a perilous journey to the western regions of India to obtain Buddhist sutras. It is an adventure filled with trials and tribulations, highs and lows, as well as fun and light-hearted moments. A fascination with words and language has spurred me to go on my very own multi-year "Journey to the Midwest", a journey that has ultimately ended with me obtaining Permanent Head Damage.

Like the Buddhist monk, I have been very fortunate to have the support of, not just three, but a whole fleet of "protectors" on this great adventure. This dissertation would not have been possible without their encouragement and support. I would like to especially thank and acknowledge the following:

- The Department of Psychology for 5 years of financial support.
- Mike Vitevitch for being awesome. Thank you for being so generous with your time, expertise, and advice, all of which has helped me become a better researcher and educator. Deciding to come to KU and work with you has been one of the best life decisions I have ever made.
- Members of my dissertation committee (Drs. Chrysikou, Sereno, Jongman, Forbush) for taking the time out of your busy schedules to read my dissertation and provide feedback on it. I also want to especially thank Dr. Susan Kemper for serving on my previous committees.
- Winston and Melvin for taking a chance on me way back in 2010 when I was 'shopping' around for a faculty mentor for my undergraduate thesis. I will always look back at this as a defining moment in my academic career.

- Past and present Psychology Graduate Students (especially the Cognitive folks) for your encouragement and support. Special shoutout to Nichol for being a wonderful cohort mate.
- Folks from the Lawrence and Kansas City climbing community for introducing me to the sport of climbing (a much needed distraction from grad school!) and graciously welcoming me into your community.
- Friend-ily in Lawrence/KC (past and present): Sharon, Max, Joanna, Iris, Stephen, Erika, Lynn, Pauying, and many others, for ~~feeding me~~ being my family away from home.
- Dougie “Fre$h” Siew (豆豆) for being an amazing Corg-huahua.
- Finally, Mom and Dad for supporting my personal and professional pursuits no matter what. You will be happy to hear that I am finally done with school!

Table of Contents

List of Figures

List of Tables

List of Appendices

Chapter 1

Phonological and Orthographic Effects in Spoken and Visual Word Recognition

1.1 Introduction

Orthographic effects in spoken word recognition and phonological effects in visual word recognition have been observed in a variety of behavioral experimental paradigms, strongly suggesting that a close interrelationship exists between phonology and orthography. This dissertation uses the tools of network science and conventional psycholinguistic tasks to examine how the phonological and orthographic relationships among words in the language influence word recognition.

Extending our understanding of phonological and orthographic influences on word recognition is of particular importance given the recent movement to ensure the standardization of reading abilities for schoolchildren across the nation. This was exemplified by the enactment of the No Child Left Behind Act in 2002, which requires states to assess the reading skills of students at the end of third grade in order to determine whether students have acquired adequate reading skills. Given that approximately 8.5 million schoolchildren suffer from dyslexia, a specific impairment in reading despite normal intelligence (National Institutes of Neurological Disorders and Stroke, 2015) and that state-of-the-art interventions that are time-consuming and expensive to administer only lead to somewhat modest reading improvements (Torgesen et al., 2001), it is imperative that language researchers expand their horizons and consider how new approaches, such as the network science approach, could lead to new insights into an age-old problem.

1.2 Roadmap of dissertation

Chapter 1 provides an overview of the prior work investigating orthographic effects in spoken word recognition and phonological effects in visual word recognition, in particular focusing on the effects of consistency and neighborhood size, which have emerged as the main metrics used to investigate these effects. Prior work ultimately fails to (1) generalize to words of various lengths or syllable structures, and (2) take into account the more global similarity structure that exists between phonological and orthographic representations in the language.

In Chapter 2, an alternative approach, the network science approach, is proposed to address the limitations of prior work. The tools of network science allow us to characterize, simultaneously, the phonological as well as orthographic similarity structure of words (of all lengths) in the language. A brief introduction to network science, and examples of how this approach has been used to further our understanding of language processes, are provided in this section. The *phonographic network*, where words are connected to both phonologically *and* orthographically similar words, is introduced in Chapter 3. The results of a computational analysis of the structural characteristics at various levels (macro-, meso-, and micro-) of the phonographic network are described. Given that the structure of the phonographic network at varying levels is meaningful (i.e., not merely random), the next section of the dissertation examines whether the structure of the phonographic network has any influence on spoken and visual language processing. In particular, the investigation will focus on the influence of two micro-level network metrics derived from the phonographic network on language processing: (i) phonographic degree and (ii) phonographic clustering coefficient.

Chapters 4 and 5 detail the results of psycholinguistic experiments (speeded naming; auditory lexical decision) that investigated the influence of phonographic degree and phonographic clustering coefficient on *spoken* word recognition. Chapter 6 detail the results of an archival analysis of speeded naming and visual lexical decision data from the English Lexicon Project that was conducted to investigate the influence of phonographic degree and phonographic clustering coefficient on *visual* word recognition. Results indicated a facilitatory effect of phonographic degree on visual word recognition, and a facilitatory effect of phonographic clustering coefficient on spoken word recognition. The present findings have theoretical implications for models of word recognition as well as practical implications for language learning and language disorders. These implications are discussed in the final chapter.

1.3 Phonological effects in visual word recognition

As it is typically presumed that the writing system builds upon spoken language (Liberman, 1992), phonology is expected to play an important role in reading. Therefore the key theoretical debate centers on delineating the extent to which phonology is automatically implicated in visual word processing. In investigating this issue, researchers (either explicitly or implicitly) employ the working hypothesis that increased phonological complexity leads to less efficient lexical processing (Katz & Frost, 1992). Phonological complexity can be generally defined as the extent to which a word conforms to the grapheme-to-phoneme correspondence rules of a language (Frost, 1998; Venezky, 1970) and can be operationalized in various ways. This includes distinguishing whether (non)words are (pseudo)homophones or not, whether words have regular or irregular pronunciations, whether words are consistent or inconsistent, and the size of a word's phonological neighborhood. If reading printed words were indeed mediated by phonology, then one would expect that more phonologically complex words are accessed more slowly and less accurately as compared to less phonologically complex words.

The results from the visual word recognition literature are generally consistent with this hypothesis. For instance, the homophone effect refers to the finding that homophones, words that have different spellings and meanings but sound identical (for instance, “rose”/“rows” and “boar”/”bore”), are more slowly processed in the presence of homophone foils (Ferrand & Grainger, 2003; J. Grainger & Ferrand, 1994; Van Orden, 1987). A somewhat analogous effect also exists for pseudohomophones—letter strings that do not represent any real English words but sound like a real English word. Response times to decide that pseudohomophones such as “brane”, which sounds like “brain”, were nonwords in a lexical decision task were slower as compared to nonword controls like “slint”, which does not map onto any phonological representations of real English words (Besner & Davelaar, 1983; Braun, Hutzler, Ziegler, Dambacher, & Jacobs, 2009; Ferrand & Grainger, 1992).

1.3.1 Feedforward consistency effects

Apart from considering the phonological complexity of whole words, researchers have also characterized the phonological complexity of words at varying grain sizes—at the level of smaller units like letters and phonemes, and larger units like syllables, onsets, and rimes. One of the more commonly used and well-established measures of phonological complexity of words are consistency measures, which focus on the distributions of pronunciations associated with word bodies (Cortese & Simpson, 2000). Bodies are letter patterns that correspond to rimes, and rimes are vowels and consonants that follow the vowel (coda). Feedforward consistent words contain bodies that can only be pronounced in one way (e.g., “-ade” can only be pronounced as /-eId/, as in “wade”|/weId/ and “fade”|/feId/), whereas feedforward inconsistent words contain bodies that can be pronounced in several different ways (e.g., “-ave” can be pronounced as /-eIv/ as in “wave”|/weIv/ or “-æv” as in “have”|/hæv/). The feedforward consistency effect refers to the finding where words containing bodies with inconsistent pronunciations are more slowly processed in a number of visual word recognition tasks (Cortese & Simpson, 2000; Jared, 1997;

Jared, McRae, & Seidenberg, 1990; Stone, Vanhoy, & Van Orden, 1997; Ziegler, Montant, & Jacobs, 1997).

One way to quantify a word's level of (in)consistency is to calculate a *feedforward* consistency ratio that ranges from 0 to 1. Note that the use of the term "feedforward" reflects spelling-to-sound inconsistencies and *not* the sound-to-spelling inconsistencies as described in the spoken word recognition literature. The consistency ratio is the summed frequency of a word's friends (words whose orthographic bodies are pronounced in the same way as that of the target word's body) relative to the summed frequency of a word's friends and enemies (words whose orthographic bodies are pronounced differently from that of the target word's body). Therefore, words with more enemies than friends tend to have low consistency values and words with more friends than enemies tend to have high consistency values. However, the typical approach in these studies is to compare consistent words against highly inconsistent words (Jared et al., 1990; Stone et al., 1997; Ziegler et al., 1997), which disregards the fact that a range of consistency ratios exists among words.

1.3.2 Phonological neighborhood effects

Some recent work has found that the number of phonological neighbors also influences visual word processing tasks. Phonological neighbors are words that differ from a target word by the substitution, addition, or deletion of a single phoneme at any word position (Luce & Pisoni, 1998). However, in an effort to be consistent with the definition of an orthographic neighbor in the visual word recognition literature (where orthographic neighbors are words that differ from a target word by only the substitution of a single letter at any word position; Coltheart, Davelaar, Jonasson, & Besner, 1977), researchers viewed phonological neighbors as words that differ from a target word by the substitution of a single phoneme at any word position (not surprisingly the two measures are highly correlated). These studies reported that the processing of words with many phonological neighbors is facilitated as compared to words with few phonological neighbors (Grainger, Muneaux, Farioli, & Ziegler, 2005; Yates, 2005; Yates, Locker, & Simpson, 2004) Yates (2005) argued that more phonological neighbors lead to an increase in activation within the phonological system, which reduces the time required to produce the phonological code for the target word. The motivation for studying the influence of phonological neighborhood size on visual word recognition rests on the argument that the size of a word's phonological neighborhood better reflects the magnitude of phonological activation during lexical processing, whereas measures like homophony, regularity, and consistency attempt to capture the relationship between orthographic and phonological codes (Yates, 2005; Yates et al., 2004).

1.4 Orthographic effects in spoken word recognition

Seidenberg and Tanenhaus's (1979) paper was one of the first to show that knowledge of orthography does influence the processing of spoken words. Using a rhyme detection task, Seidenberg and Tanenhaus showed that the time taken to decide if two words rhymed was influenced by their orthographic similarity. Participants took a longer time to decide that "tie" and "rye" (orthographically dissimilar pair) rhymed as compared to "tie" and "pie" (orthographically similar pair). This was an important finding at that time because it showed that orthographic information was activated during the processing of auditory stimuli, suggesting that both phonological and orthographic features are encoded in word representations (Seidenberg & Tanenhaus, 1979).

Subsequently, several studies have also found orthographic effects in online tasks such as naming (Ziegler, Ferrand, & Montant, 2004; Ziegler, Muneaux, & Grainger, 2003), auditory lexical decision (Dich, 2011; Roux & Bonin, 2013; Ziegler & Ferrand, 1998; Ziegler et al., 2004, 2003), semantic and gender categorization (Peereman, Dufour, & Burt, 2009), phonological priming (Chéreau, Gaskell, & Dumay, 2007; Slowiaczek, Soltano, Wieting, & Bishop, 2003), and serial recall (Pattamadilok, Lafontaine, Morais, & Kolinsky, 2010). More recently, the use of online measures such as electrophysiological measures in EEG studies (Perre, Midgley, & Ziegler, 2009; Zou, Desroches, Liu, Xia, & Shu, 2012) and eyetracking in the visual world task (Salverda & Tanenhaus, 2010) has provided additional evidence for the co-activation of orthographic information while processing auditory stimuli.

To recapitulate, orthographic effects have been observed in a variety of tasks using a variety of online measures of language processing, such as reaction time, ERPs, and looking time. Together, the presence of orthographic effects obtained across a vast variety of tasks and experimental paradigms, in particular tasks which do not require an explicit analysis of the stimuli (as in Seidenberg and Tanenhaus, 1979), provide converging evidence for the presence of orthographic effects in spoken word recognition. More importantly, these findings raise key questions about the nature of lexical representations that are stored within long-term memory and the cognitive processes that support lexical retrieval.

1.4.1 Feedback consistency effects

One of the most robust orthographic effects found in the studies cited above is the feedback consistency effect. The feedback consistency effect refers to the finding where words containing sound-to-spelling inconsistencies are more slowly and less accurately responded to in lexical decision and word naming tasks (Ziegler & Ferrand, 1998; Ziegler et al., 2004, 2003). Note that the use of the term “feedback” reflects sound-to-spelling inconsistencies (i.e., multiple spellings associated with a particular phonological unit or segment), and *not* the spelling-to-sound (feedforward) inconsistencies as described in the visual word recognition literature.

Consistency can be defined on the basis of a word's onset or rime, although feedback consistency is typically defined as a function of the spelling consistency of phonological *rimes* of words given the dominance of the onset-rime syllable structure in English (Kessler & Treiman, 1997; Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995). Feedback consistent words contain phonological rimes that can only be spelled in one way (e.g., /^k/ can only be spelt as "-uck", as in "duck" and "luck"), whereas feedback inconsistent words contain phonological rimes that can be spelled in several different ways (e.g., /ip/ can be spelt as "-eep" as in "deep" or "-eap" as in "heap").

Similar to feedforward consistency, it is possible to calculate a feedback consistency ratio that reflects a word's degree of feedback inconsistency in a similar manner. The consistency ratio is the summed frequency of a word's friends (friends are words whose phonological rimes are spelled in the same way as that of the target word's rime) relative to the summed frequency of a word's friends and enemies (enemies are words whose phonological rimes are spelled differently from that of the target word's rime).

1.4.2 Orthographic neighborhood effects

Another well-established finding in this area is that the size of the orthographic neighborhood (ON) influences spoken language processing. Orthographic neighborhood size represents the number of words that can be produced by changing a letter in a target word of the same length (i.e., Coltheart's *N*; Coltheart et al., 1977). Based on this definition, "bat", "cot", and "cap" are orthographic neighbors of the word "cat". Studies have shown that when phonological neighborhood (PN) size is controlled for, ON size facilitates spoken word recognition—high ON words are produced and recognized more quickly and accurately than low ON words (Muneaux & Ziegler, 2004; Ziegler et al., 2003).

Ziegler and colleagues (2003) investigated PN and ON effects in auditory lexical decision and shadowing by factorially manipulating PN and ON sizes such that there were four sets of words: Words from dense PNs and dense ONs (PN+ON+), words from dense PNs and sparse ONs (PN+ON-), words from sparse PNs and dense ONs (PN-ON+), and words from sparse PNs and sparse ONs (PN-ON-). They found an inhibitory main effect of PN and a facilitatory main effect of ON. Ziegler and colleagues (2003) proposed that the facilitatory ON effect resulted from the consistency of sublexical mappings between phonology and orthography. Consider PN+ON- words—these words tend to be inconsistent because if a word has many phonological neighbors but only a few orthographic neighbors this implies that that word has a "common" phonology but "rare" spelling. In comparison PN+ON+ words tend to be more consistent because they have a "common" phonology and "common" spelling. This explanation was supported by additional post-hoc analyses, which showed that the ON effect disappeared when feedback consistency was included as a covariate. In addition, it should be noted that Ziegler and colleagues found an inhibitory PN effect in spoken word recognition tasks, whereas Yates and colleagues found a *facilitatory* PN effect in visual word recognition tasks. On the other hand, ON effects appear to be facilitatory in both visual and spoken word recognition tasks. The asymmetric nature of PN and ON effects in visual and auditory tasks is intriguing and may hint at differences in the way that phonological and orthographic information influence lexical processing in different modalities; however, to date, there has not been any attempt to account for and integrate these findings within a single model.

1.5 Limitations of current approaches

Although metrics such as consistency and neighborhood size represent some attempt to capture the relationship between orthography and phonology, they are ostensibly interpreted as lexical characteristics of individual words. The operationalization of these measures appears to capture some aspect of similarity among words in a language—neighborhood size is calculated based on evaluating the orthographic or phonological similarity of a target word to other words in the lexicon (Coltheart et al., 1977; Luce & Pisoni, 1998), and consistency is determined by calculating how often the body of the target word is pronounced or spelt among words that also share the same body or rime (Kessler & Treiman, 1997). Nevertheless these current metrics do not capture the *overall* similarity structure of a language because the way in which these metrics have been operationalized restricts their applicability to a subset of words within the entire mental lexicon and/or represents a particular aspect of language structure (phonological or orthographic similarity) rather than the interrelationship between phonology and orthography.

In order to make continued progress in our understanding of phonological and orthographic influences on language processing the field may need to consider an alternative framework that represents aspects of the relationship between orthography and phonology that are not captured by current approaches and which explicitly considers more global aspects of similarity among words in a language. Such a framework could also promote a more elegant theoretical approach to studying language processing because it considers how lexical processes and mechanisms operate within a complex language structure that emerges from the interrelationships among orthographic and phonological representations. This approach exemplifies the idea that a complete understanding of how a system functions is not possible without taking into account the structural properties of the system (Strogatz, 2001).

In this case, a complete understanding of the cognitive mechanisms that support language processing should seriously consider how these mechanisms occur within the overall linguistic system. As Stone et al. (1997) point out, it is not necessarily a bad thing to try to account for similar results with different models or theories because pluralism of ideas and theories may lead to the discovery of critical common principles that guide lexical processing (Stone & Van Orden, 1994). Moreover, an alternative framework may lead to the development of common metrics that may provide a different perspective on these issues, which may also lead one to ask new questions that one would not ordinarily consider with the contemporary approach.

The asymmetric nature of phonological neighborhood and orthographic neighborhood effects in visual and auditory tasks is one intriguing theoretical question that could be addressed via this alternative framework. Ziegler and colleagues (2003) found an *inhibitory* phonological neighborhood effect in spoken word recognition tasks, whereas Yates and colleagues (2004; 2005; also Grainger et al., 2005) found a *facilitatory* phonological neighborhood effect in visual word recognition tasks. On the other hand, orthographic neighborhood effects appear to be facilitatory in both visual and spoken word recognition tasks (Ziegler et al., 2003).

These findings hint at differences in the way that phonological and orthographic information influence lexical processing in different modalities. However, to date, it is not clear why the effects of similarity on lexical processing differ across different modalities. Furthermore there has been little attempt to account for and integrate these findings within a single model or framework that considers both the phonological and orthographic structure of language. Language researchers have studied the influence of phonological similarity and orthographic similarity on word recognition as if they were two separate distinct influences (i.e., manipulate phonological neighborhood size while controlling for orthographic neighborhood size, and vice versa). However, the evidence for phonological effects in visual word recognition and orthographic effects in spoken word recognition strongly suggests that we should consider the influences of phonological similarity and orthographic similarity in tandem.

In Chapter 2, the emerging field of network science is introduced and I demonstrate how the network science approach can provide language researchers with the tools to characterize the phonological and orthographic structure of language within a single framework. A network of words can be constructed based on a simple operationalization of phonological and orthographic similarity among words, the structure of which can be further investigated and quantified as common metrics of phonological and orthographic similarity.

Chapter 2

Network Science and Language Networks

Network science is an emerging interdisciplinary field, where mathematical techniques are used to characterize and analyze the topology or structure of complex networks in various domains (Barabási, 2009; Watts, 2004). Examples of these complex networks include friendship networks on social media websites (Lewis, Kaufman, Gonzalez, Wimmer, & Christakis, 2008), air transportation networks (Cardillo et al., 2013), the World Wide Web and the Internet (Albert, Jeong, & Barabási, 1999), the human brain (Bullmore & Sporns, 2009), and the mental lexicon (Steyvers & Tenenbaum, 2005; Vitevitch, 2008).

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

Network science offers researchers a theoretical framework as well as a comprehensive suite of methodological tools and techniques to study complex networks. The tools of network science allow researchers to derive a variety of network measures that describe the structure of the network. These include metrics that describe the network's *global* or *macro*-level structure (e.g., average path length, average clustering coefficient, overall degree distribution), *meso*-level structure (i.e., the level that falls between the macro- and micro-levels; e.g., community structure), as well as *local* or *micro*-level structure (e.g., degree, clustering coefficient of individual nodes).

Complex network scientists recognize that studying the relationships between individual entities allows us to understand the global behavior of the larger system in ways that cannot be discerned from studying each entity on its own (Wilson, 1998). Moreover, it is important to study the *structure* of complex networks because network structure affects the *processes* that occur within these networks (Strogatz, 2001). For instance, the structure of the social network affects the way in which information spreads among people, and the structure of air transportation networks affects the way air travel is rerouted when there are major airport closures.

The tools of network science have been used to study the structure of the mental lexicon, which consists of all the words that a person knows that are stored in long-term memory. Language researchers have used these tools to model phonological (Vitevitch, 2008), orthographic (Kello & Beltz, 2009), and semantic (Hills, Maouene, Maouene, Sheya, & Smith, 2009; Solé, Corominas-Murtra, Valverde, & Steels, 2010; Steyvers & Tenenbaum, 2005) language networks of words in the mental lexicon. In these networks, each node represents a word, but the ways in which links are assigned to nodes differ. In the phonological network, links connect pairs of words that are phonologically similar to each other. In the orthographic network, links connect pairs of words that are orthographically similar to each other. Below I briefly discuss prior computational and psycholinguistic work conducted with respect to the phonological network and the orthographic network of language.

2.1 The phonological language network

The phonological network examined in Vitevitch (2008) consisted of the phonological transcriptions of 19,340 words obtained from the 1964 Merriam-Webster Pocket Dictionary. In this network, nodes represented phonological word forms and connections represented phonological similarity between words. Two words were considered phonologically similar if the first word could be transformed to the other by either substituting, adding, or deleting one phoneme in any position (Landauer & Streeter, 1973; Luce & Pisoni, 1998). For instance, the word /kæt/ ("cat") would be connected to /æt/ ("at"), /bæt/ ("bat"), and /skæt/ ("scat"). The phonological network consisted of a giant component (the largest connected component of the network), several lexical islands (smaller connected components of the network), and hermits (nodes that do not connect to any other nodes). Vitevitch (2008) found that the giant component possessed a small-world structure (i.e., short average path length and high average clustering coefficient when compared to a random network of a similar size), a degree distribution that resembled a truncated power law, and assortative mixing by degree (i.e., nodes with several neighbors tend to be connected to nodes with several neighbors; (Newman, 2002).

Prior work by Vitevitch and colleagues have demonstrated that the local or micro-level structure of the phonological network influences spoken word recognition and production, word learning, and short- and long-term memory processes (Chan & Vitevitch, 2009, 2010; Goldstein & Vitevitch, 2014; Vitevitch, Chan, & Roodenrys, 2012). Chan and Vitevitch (2009; 2010) showed that the clustering coefficient, *C*, of a word had measureable effects on a variety of psycholinguistic tasks such as perceptual identification, lexical decision, and picture naming. Clustering coefficient, *C,* is a micro-level network science metric that represents the extent to which neighbors of a word are also neighbors of each other. The phonological neighbors of high *C* words tend to be neighbors of each other, whereas the phonological neighbors of low *C* words do *not* tend to be neighbors of each other. Therefore *C* is one network science metric that can quantify the local or micro-level structure of a node (i.e., by considering the *immediate* neighbors of a node). Chan and Vitevitch (2009; 2010) found a processing advantage for low *C* words as compared to high *C* words.

The structure of the phonological network at levels other than the micro-level has also been found to influence lexical processing. Vitevitch and Goldstein (2014) found a processing advantage for “keywords”—a set of words that, when removed, would cause the network to fracture into several smaller components—as compared to non-keywords with comparable lexical characteristics. Vitevitch, Chan, and Goldstein (2014) analyzed instances of failed lexical retrieval by participants and found that the errors reflected the presence of high assortative mixing by degree (a macro-level metric) in the phonological network. Assortative mixing by degree refers to the tendency for highly connected nodes to be connected to other highly connected nodes in the network (Newman, 2002), and represents one aspect of the macro- level structure of the phonological network. In another study, Siew and Vitevitch (2016) showed that the network component that words resided in (i.e., giant component or lexical islands) affected how quickly words were retrieved from the mental lexicon. Together, these findings suggest that, in addition to the micro-level (as exemplified by the clustering coefficient network metric), the macro-level structure of the phonological network also has important implications for understanding lexical processes.

Ultimately, this body of research has shown that the structure at different levels of the phonological network influences various aspects of lexical processing. The tools of network science have allowed language researchers to explore the structure of the phonological network at different levels of analysis, adding depth to the investigation and providing new insights into the psychological mechanisms that support lexical processing.

2.2 The orthographic language network

In contrast to the phonological network, there has not been as much research that has made use of the tools of network science to study the orthographic language network. One exception is the analysis conducted by Kello and Beltz (2009), who constructed an orthographic word form network whereby links were placed between words that were substrings of other words. For instance, the word "air" would be connected to the words "f<u>air</u>" and "<u>air</u>ed". This network had a tree-like branching structure: Shorter word forms (such as "air") were usually found at the higher sections of the network ("trunks") and longer word forms (such as "faired") were usually found at the lower sections of the network ("branches").

According to Kello and Beltz (see also Ferrer i Cancho & Solé, 2003; Zipf, 1949), language systems need to strike a balance between memory constraints while maximizing the efficiency of disambiguating between different lexical representations. One universal feature of such systems is the presence of scale-free or power laws. With respect to the orthographic language system, in order to maximize distinctiveness between word forms, there should be little substring overlap among word forms. On the other hand, to minimize memory costs of storing word forms in long term memory, substrings should be reused as much as possible. Based on the need to balance between these two competing constraints in the orthographic lexicon, Kello and Beltz predicted, and subsequently found, evidence of a scale-free degree distribution with respect to the number of outgoing links per node (i.e., the number of times a given word form was a substring of another word form) in the orthographic network.

It is important to note that Kello and Beltz's operationalization of orthographic similarity (i.e., placing links between words that were substrings of other words) differs significantly from the way orthographic similarity has been typically operationalized in the psycholinguistic literature, where words are considered to be orthographically similar if they differ by the substitution of a single letter (Coltheart et al., 1977). In the network examined in Vitevitch (2008), the phonological network of language was constructed using a commonly used metric of phonological similarity and no assumptions were made with regards to the ultimate structure of the network. In contrast, the orthographic network described in Kello and Beltz was constructed using a somewhat non-traditional definition of orthographic similarity with an a priori assumption of the existence of a scale-free degree distribution. It is unclear if a similar degree distribution would be observed in the orthographic language network if a more commonly used metric of orthographic similarity were used to specify the links between words instead.

It is also important to note that Kello and Beltz merely conducted a computational analysis of the orthographic network. To date, there has not been any behavioral or experimental work investigating how the network structure of the orthographic lexicon might influence lexical processing. However, the results of a recent paper by Iyengar and colleagues (2012) suggest that the orthographic structure of language could have key implications for navigating the mental lexicon. Participants played a word-morph game where they had to find a sequence of words such that the first word could be transformed to the second word (of the same length) by changing a single letter. For example, the sequence of words to get from "try" to "pot" was "try-toy-ton-tot-pot". Note that this particular definition of orthographic similarity, that is, the substitution of one letter within the word, is more similar to the way in which phonological similarity was defined in the phonological network as described earlier. Iyengar and colleagues found that participants were much faster at the game when they learned to make use of "landmark" words to find the sequence of words. These landmark words were nodes in the word-morph network of three-letter English words that had high closeness centrality—a network science measure indicating the inverse of the sum of distances of a node to all other nodes in the network. High closeness centrality words were "close" to many other words in the network. Iyengar et al.'s findings strongly suggest that the network structure of *orthographic* word forms (albeit one that contained only three-letter words) has behavioral consequences as one navigates the mental lexicon and there could be similar implications for lexical retrieval.

2.3 Introduction to multiplexes

As reviewed above, the application of network science to the field of psycholinguistics has led to several new discoveries and has greatly contributed to our understanding of cognitive and language processes. However, the language networks that have been examined to date do not truly reflect the multiplexity inherent in language. That is, words can be phonologically, orthographically, and semantically related to each other. To date, language networks have been constructed based on a *single* type of relationship among words and analyzed independently of other types of language networks. The network constructed by Vitevitch (2008) was based on phonological similarity among words, whereas the network constructed by Kello and Beltz (2009) considered orthographic relations among words. The recent movement in network science toward characterizing multiple types of relationships among nodes within the same network (i.e., the multiplex) suggests that both phonological as well as orthographic relationships among words should (and could!) be represented within a single language network in order to truly reflect the multiplexity of linguistic structure.

The study of multiplex networks is one of the fastest growing research areas within network science. A *multiplex network* (also known as a multi-layer network or simply a multiplex) consists of multiple layers of networks, whereby the connections within each layer represent a different type of relationship among a common set of nodes (Battiston, Nicosia, & Latora, 2014). Figure 1 shows a simple multiplex. As network scientists are recognizing that real world complex networks are inherently multiplex in nature, that is, nodes can be related to each other in more than one way, much of the recent research activity in this area has been directed toward establishing a consistent mathematical and theoretical framework for modeling multiplex networks (Battiston et al., 2014; Kivelä et al., 2014), as well as developing computational tools for analyzing multiplex networks (De Domenico, Porter, & Arenas, 2015). Network scientists have also adapted and re-conceptualized a number of network metrics such that they would be more appropriate for analyzing the structure of the multiplex (for an overview, see Boccaletti et al., 2014).

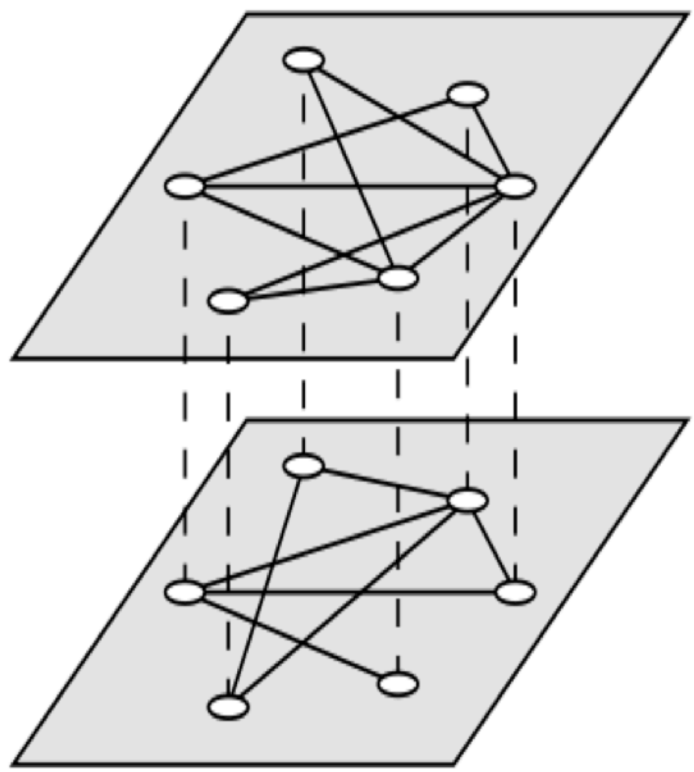

Figure 1. A two-layer multiplex. The same nodes are represented in both layers, and the connections within each layer represent a different type of relationship among nodes (Battiston et al., 2014).

There are various ways in which multiplexity can manifest in real world complex networks. For instance, different kinds of relationships such as platonic, romantic, sexual relationships can exist between people (Lewis et al., 2008). Constructing a social network based on a single type of connection is merely a crude approximation to reality. In an interbank network, banks can be related to each other via different types of financial transactions (Bargigli, di Iasio, Infante, Lillo, & Pierobon, 2015). Considering only a single type of financial transaction can lead to a poor approximation of underlying systemic risk in the banking industry (Bargigli, Iasio, Infante, Lillo, & Pierobon, 2015). In the air transportation network, the flights that connect airports are not identical: some flights are provided by major airlines, others by budget airlines, and yet others by cargo airlines (Cardillo et al., 2013). A more complete picture of the air transportation network can only be obtained by considering the different kinds of flights that connect airports. From these examples, it is clear that multiplexity is an inherent feature of most real world complex networks.

Given that past work has demonstrated how the tools of network science can be used to quantify the structure of language networks and further our understanding of lexical processes, constructing a phonographic multiplex to characterize, simultaneously, the phonological and orthographic structure of language could provide us with new, sophisticated metrics that take into account both the phonological and orthographic structure of language and offer new ways to investigate the interrelationship between phonology and orthography and their influence on lexical processes.

Chapter 3

The Phonographic Network of Language

A phonographic multiplex was constructed with phonological and orthographic layers. The phonological layer was identical to the phonological network constructed by Vitevitch (2008). The orthographic layer consisted of the same words (nodes) in the phonological network, except that the connections in this layer would be based on orthographic similarity. Words would be connected to each other if they differed by the substitution, addition, or deletion of a single *letter* in any word position. Such an operationalization would not only be consistent with the operationalization of phonological similarity in the phonological network, but represents an approach of defining similarity relations among words that has a long history in the field of psycholinguistics (e.g., Coltheart et al., 1977; Greenberg & Jenkins, 1964). Note that this would lead to an orthographic layer in the phonographic multiplex that is quite different from that of the orthographic word form network constructed by Kello and Beltz (2009)—recall that Kello and Beltz placed links between words that were substrings of other words. Placing a link between two words that can be transformed into the other via the substitution, addition, or deletion of a letter would be more consistent with the way in which orthographic similarity has operationalized in the psycholinguistic literature (Coltheart et al., 1977), and is not an ad hoc metric used to impose a scale-free degree distribution on the language network (Kello & Beltz, 2009).

According to the multiplex literature, three different types of *multilinks* are possible in a two-layer multiplex. Nodes can be connected to each other (i) in both layers, (ii) only via layer 1, or (iii) only via layer 2 (Bianconi, 2013; Menichetti, Remondini, Panzarasa, Mondragón, & Bianconi, 2014). In the phonographic multiplex, words can be (i) phonologically and orthographically related to each other, (ii) only phonologically related to each other, or (iii) only orthographically related to each other. Examining the nature of the overlap of links among nodes in the phonological and orthographic layers in the phonographic multiplex could be particularly relevant for studying the interrelationship between phonology and orthography.

For the purposes of the dissertation, I focus on analyzing the part of the phonographic multiplex where the phonological and orthographic links overlap. That is, the section consisting of multilink (i)—links that are found in both the phonological and orthographic layers of the multiplex. The *phonographic network* (named after the *phono*-logical and ortho-*graphic* layers of the multiplex) was constructed by using the links that represented both phonological and orthographic similarity from the phonographic multiplex.

Below I describe the results of a computational analysis of the structure of the phonographic network at the macro- and meso-levels. It is important to first demonstrate that the structure of the phonographic network is indeed meaningful (i.e., not merely random) as compared to other real world networks, before investigating whether the structure of the phonographic network (as exemplified by multiplex metrics) influences spoken and visual word recognition.

3.1 Structural characteristics at the macro-level

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

In the network science literature, most real world networks possess very large giant components, where almost all nodes are connected to form a single connected component (Newman, 2001). In contrast, the proportion of nodes found in the giant components of the phonological networks of various languages such Spanish and Mandarin vary from 34% to 66% (Arbesman, Strogatz, & Vitevitch, 2010). The proportion of nodes residing in the giant component of the phonographic network is small relative to other real-world networks, but comparable to the phonological language networks.

Below, I further examine the structure of the phonographic network. In order to provide a baseline for making comparisons of the structure of the phonographic network, a similarly sized random network was constructed by randomly placing links between nodes (Erdös & Rényi, 1960).

3.1.1 Average path length

The average path length of the (largest component of the) phonographic network was 7.14. On average, approximately seven links had to be traversed to connect any two nodes in the giant component of the phonographic network. The average path length of a random network with the same number of nodes and edges was 4.79. Although the average path length of the phonographic network was somewhat larger than that of a comparably sized random network, the conventions used in network science (with the range of 1.5 of the value obtained from a random graph = 7.18) would consider these values comparable (Watts & Strogatz, 1998).

3.1.2 Clustering coefficient

The average clustering coefficient of the (largest component of the) phonographic network was 0.284. The average clustering coefficient of a random network with the same number of nodes and edges was 0.002. The average clustering coefficient of the phonographic network was much larger by several orders of magnitude than that of a comparably sized random network. This indicates that the neighbors of a given node in the phonographic network are more likely to be neighbors of each other, as compared to the neighbors of a given node in the random network.

According to Watts and Strogatz (1998), a small-world network has (i) an average path length that is comparable to the average path length of a random network, but (ii) a clustering coefficient that is much larger than the average clustering coefficient of a random network with the same number of nodes and edges. Several real world networks, such as the network of scientific collaborations (Newman, 2004a) and the human brain (Bullmore & Sporns, 2009), possess these two characteristics and are said to have a small-world structure. Despite their large sizes, these networks are relatively easy to navigate due to its small average path length and large average clustering coefficient relative to a comparably sized random network with randomly placed links (Kleinberg, 2000). The results of the present analyses suggest that, similar to the phonological (Vitevitch, 2008) and semantic (Steyvers & Tenenbaum, 2005) networks of language, the phonographic network has the features of a small-world network.

3.1.3 Degree distribution

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

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

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

Characterizing the degree distributions of networks can give us clues about the network growth mechanisms that might have led to the present structure observed in the complex networks (however, see D'Souza, Borgs, Chayes, Berger, & Kleinberg (2007) for an alternative to network growth mechanisms that shapes network structure). For instance it has been shown that preferential attachment, the process by which new nodes that are added to the network are more likely to connect to nodes with several connections, leads to the formation of networks with power law degree distributions (Barabási, 2009). Preferential attachment is also known as the "rich gets richer" phenomena—because nodes that already have several connections (these are "rich" nodes) are more likely to gain new connections as the network grows over time (these nodes get "richer") and this leads to a network structure where very few nodes have most of the connections and most nodes only have few connections. Power law degree distributions appear to be a key feature of several real world complex networks (Barabási & Albert, 1999).

Some work has suggested that a log-normal degree distribution might be indicative of a combination of both random attachment and preferential attachment (Jackson & Rogers, 2007). In random attachment nodes are simply added randomly to the network. With respect to the phonographic network, it should be noted that it was constructed by essentially extracting the part of the phonographic multiplex where the phonological and orthographic links overlapped. It is possible that the log-normal degree distribution, which is mathematically closely related to a power law (Mitzenmacher, 2004), was merely a by-product of the way in which the phonographic network was constructed. However, it might also be the case that the log-normal degree distribution reflects the influence of somewhat different network growth mechanisms as compared to the semantic and phonological language networks. For instance, whereas one has to be taught explicitly how to read (i.e., literacy skills are obtained via explicit instruction), the acquisition of speech and meaning are not explicitly taught to native speakers of a language—and this might be reflected in the differences between the degree distributions of various types of language networks.

3.2 Structural characteristics at the meso-level

In addition to delineating the overall topology of a network (i.e., the macro-level), the tools of network science also permit us to investigate the meso-level of a network that is typically exemplified by a network's community structure. Community structure refers to the presence of several smaller groups of nodes within a larger network, where smaller groups form such that there are many connections among nodes within a group, but few connections exist between nodes belonging to different groups (Newman & Girvan, 2004). Communities have been commonly observed in real world networks such as the structure of the human brain (Wu et al., 2011), the World Wide Web (Newman, 2004), as well as the phonological network of language (Siew, 2013). The community structure of networks is of interest to network scientists because it reveals additional information of the network structure that may not be observable at the coarse, top-most level of analysis, nor by examining the individual nodes that comprise the system (Lancichinetti, Kivelä, Saramäki, & Fortunato, 2010; Onnela et al., 2012).

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

The above analyses of the phonographic network at both the macro- and meso-levels reveal that several features of its overall network structure converge with those observed in other real world networks. Similar to the phonological language network, the phonographic network possesses a small-world structure (i.e., short average path length and high average clustering coefficient), a “small” giant component, and robust community structure. However, the degree distribution of the phonographic network appeared to follow a log-normal distribution whereas the degree distribution of the phonological networks of English, Spanish, Basque, and Hawaiian, languages from different language families resembled a truncated power law (Arbesman et al., 2010). Overall, this analysis strongly suggests that the structure of the phonographic network is meaningful (i.e., not merely random) and is worth exploring further.

3.3 Structural characteristics at the micro-level

The above analyses demonstrate how the tools of network science can be used to examine the structure of the phonographic network at varying levels of analysis. The next step is to examine whether the structure of the phonographic network has any influence on spoken and visual language processing. In particular, the present investigation will focus on the influence of two *micro*-level network metrics derived from the phonographic network on language processing: (i) phonographic degree and (ii) phonographic clustering coefficient. This dissertation represents a first step in what will be a continuing line of research that will investigate how the structure of the phonographic network at other levels of analysis (marco-, meso-levels) influences spoken and visual word recognition.

3.3.1 Phonographic degree

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

In the visual word recognition literature, there is a small body of research in the literature investigating the influence of phonographic neighborhood size on language processing (Adelman & Brown, 2007; Muneaux & Ziegler, 2004; Peereman & Content, 1997). The general finding is that the presence of phonographic neighbors facilitates naming of visually presented words (Adelman & Brown, 2007; Peereman & Content, 1997). In other words, there appears to be a processing advantage for words with several phonographic neighbors as compared to words with few phonographic neighbors. To date there has not been any work studying the role of phonographic neighbors in spoken word recognition. Based on the past literature, one would predict a facilitatory effect of phonographic degree on visual word recognition. On the other hand, for spoken word recognition, it is unclear if the presence of more phonographic neighbors would facilitate or inhibit recognition. The presence of more phonographic neighbors could inhibit recognition by contributing greater competition among activated neighbors (Luce & Pisoni, 1998). However, recall that these metrics are obtained from the phonographic network, which represents the part of the phonographic multiplex where the phonological and orthographic layers overlap. Therefore these metrics can be said to quantify the extent to which the structures of the phonological and orthographic neighborhoods of a given word are consistent with each other, and one would predict that the presence of more phonographic neighbors would facilitate processing in the first layer of the multiplex (e.g., phonological) by providing more of a "boost" in activation in the second layer of the multiplex (e.g., orthographic).

3.3.2 Phonographic clustering coefficient

Another micro-level multiplex metric to investigate is the clustering coefficient (*C*) of overlapping (i.e., phonographic) neighbors in the phonographic network. Clustering coefficient is a particularly interesting micro-level metric to explore because it is a measure of the internal *structure* of a word's local neighborhood. If phonographic *C* indeed influences the speed and accuracy of lexical retrieval this would suggest that it is important to further explore how the structure of the phonographic network at the meso- and macro-levels might also influence language processing. Based on the words in the giant component of the phonographic network, the mean phonographic *C* was 0.284 (*SD* = 0.278) with values covering the full range of *C* from 0 to 1. A word with high phonographic *C* would have phonographic neighbors that tend to also be neighbors of each other whereas a word with low phonographic *C* would have phonographic neighbors that do not tend to be neighbors of each other. Consider the following two words: 'mold' and 'pant'. Both 'mold' and 'pant' have 14 phonographic neighbors; however, 'mold' has a higher phonographic *C* (0.440) as compared to 'pant' (phonographic *C* = 0.121). As shown in Figure 2 below, the phonographic neighbors of 'mold' tend to also be phonographic neighbors of each other (greater density of connections within the phonographic neighborhood), whereas the phonographic neighbors of 'pant' do not tend to be phonographic neighbors of each other (lower density of connections within the phonographic neighborhood).

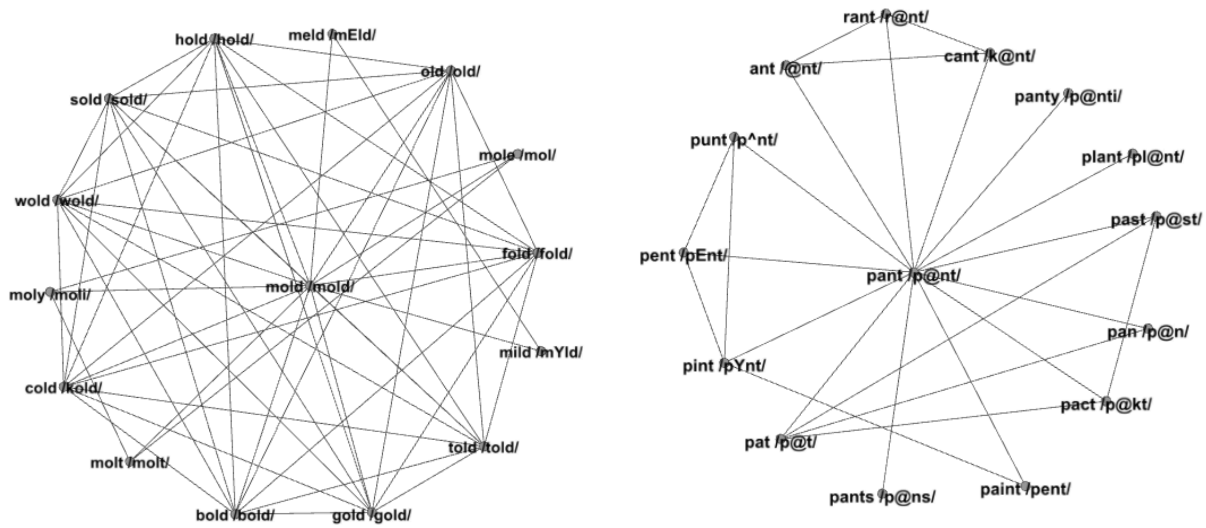

Figure 2. The phonographic neighborhood of ‘mold’ (high phonographic *C*) is shown on the left and the phonographic neighborhood of ‘pant’ (low phonographic *C*) is shown on the right. Both words have the same number of phonographic neighbors, but differ in the level of interconnectivity within their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

Given previous work showing that the local network structure of the phonological network influences lexical retrieval (Chan & Vitevitch, 2010), one might also expect that the internal *structure* of the phonographic neighborhood, in addition to its absolute size, would also affect the speed and accuracy of lexical processes. Specifically, Chan and Vitevitch found an inhibitory effect of phonological *C* on various spoken word recognition tasks. Based on the literature, one might also expect phonographic *C* to exert an inhibitory effect on both visual and spoken word recognition. Nevertheless, it is important to note that phonological *C* metric used by Chan and Vitevitch was based on the (single) phonological layer of the phonographic multiplex—whereas phonographic *C* measures the internal structure of a word’s phonographic neighborhood (based on both layers of the phonographic multiplex). A greater value of phonographic *C* indicates greater similarity in the phonological and orthographic neighborhood structures of a given word. Given past work on "conspiracy models" of word pronunciation which has shown that words with more consistent neighbors tend to be more quickly named (e.g., Taraban & McClelland, 1987), one would expect that the activation dynamics that occur among similar phonological and orthographic network structures would also "conspire" and lead to the facilitation, rather than inhibition, of lexical retrieval. In the following chapters, conventional psycholinguistic experiments and an archival analysis of behavioral data from a language database will be conducted to examine the influence of phonographic degree and phonographic clustering coefficient on spoken and visual word recognition.

Chapter 4

Experiment 1: Auditory Naming Task

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

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

4.1 Method

Participants. Sixty native English speakers were recruited from the Introductory Psychology subject pool at the University of Kansas. All participants had no previous history of speech or hearing disorders and received partial course credit for their participation.

Materials. Two sets of monosyllabic English words were selected as stimuli. The first set consisted of words with either high or low phonographic degree and the second set consisted of words with either high or low phonographic *C*. There were a total of 160 words: 40 high phonographic degree words, 40 low phonographic degree words, 40 high phonographic *C* words, and 40 low phonographic *C* words.

The stimuli were also chosen so as to capture a representative range of lexical variables, while excluding words that had an extreme value of any of the following lexical characteristics: number of phonemes, number of letters, subjective familiarity, word frequency, phonological degree, phonological *C*, phonological neighborhood frequency, two measures of phonotactic probability (positional segment probability and biphone probability), orthographic degree, orthographic *C*, orthographic neighborhood frequency, two measures of bigram frequency (average bigram frequency counts and sum of bigram frequency counts by position). In addition to the key variables of interest (i.e., phonographic degree and phonographic clustering coefficient), these lexical variables will be included as covariates in the multilevel model.

A male native speaker of American English produced the stimuli by speaking at a normal speaking rate into a high-quality microphone in an Industrial Acoustics Company sound-attenuated booth. Individual sound files for each word were edited from the digital recording with SoundEdit16 (Macromedia, Inc). The Normalization function in SoundEdit16 was used to ensure that all sound files were comparable in amplitude. Stimuli durations were equivalent across high and low phonographic degree words, $t(78) < 1$, $p = .36$, and across high and low phonographic *C* words, $t(78) < 1$, $p = .48$.

Table 1 shows the means and standard deviations of various psycholinguistic characteristics of high and low phonographic degree and high and low phonographic *C* words. A list of the stimuli is provided in Appendix A.

Table 1. Lexical characteristics of phonographic degree and phonographic *C* word stimuli.

	Phonographic degree words, N = 80		Phonographic *C* words, N = 80	
Variable	High degree	Low degree	High *C*	Low *C*
Phonographic degree	5.78 (1.00)	3.55 (0.60)	6.85 (3.51)	7.35 (3.88)
Phonographic *C*	0.283 (0.146)	0.321 (0.278)	0.326 (0.123)	0.163 (0.079)
Subjective familiarity	6.83 (0.24)	6.86 (0.25)	6.81 (0.31)	6.83 (0.23)
Word frequency	1.88 (0.55)	2.04 (0.65)	1.95 (0.65)	1.97 (0.73)
Number of phonemes	3.83 (0.59)	3.80 (0.61)	3.60 (0.55)	3.60 (0.55)
Phonological degree	11.83 (4.27)	10.85 (5.06)	14.30 (6.85)	14.98 (7.11)
Phonological *C*	0.294 (0.073)	0.286 (0.085)	0.272 (0.063)	0.249 (0.080)
Phonological neighborhood frequency	1.83 (0.29)	1.88 (0.24)	1.83 (0.29)	1.85 (0.29)
Positional segment probability	0.175 (0.0474)	0.176 (0.0475)	0.171 (0.0451)	0.170 (0.0414)
Biphone probability	0.0120 (0.00687)	0.0116 (0.00693)	0.00940 (0.00606)	0.00942 (0.00576)
Number of letters	4.53 (0.55)	4.50 (0.56)	4.33 (0.53)	4.30 (0.56)
Orthographic degree	6.68 (1.90)	6.65 (1.93)	8.63 (4.00)	9.15 (4.34)
Orthographic *C*	0.277 (0.135)	0.274 (0.141)	0.252 (0.056)	0.247 (0.060)
Orthographic neighborhood frequency	7.13 (1.05)	7.45 (1.13)	7.50 (0.95)	7.40 (1.13)
Average bigram counts	2728.82 (1218.51)	2876.26 (1204.88)	2763.84 (1251.79)	2735.60 (1041.33)
Sum of bigram counts (by position)	1871.23 (662.43)	2061.28 (698.99)	1910.80 (897.78)	1818.35 (722.58)
File duration	570 (58)	584 (67)	584 (74)	572 (73)

Note: *SD*s are provided in parentheses.

Phonographic degree. Phonographic degree refers to the number of words that are *both* phonological and orthographic neighbors of a given word. Therefore, phonographic neighbors differ from the target word by the substitution, deletion, or addition of one phoneme *and* the substitution, deletion, or addition of one letter. High phonographic degree words had a mean phonographic degree of 5.78 (SD = 1.00) and low phonographic degree words had a mean phonographic degree of 3.55 (SD = 0.60), t (78) = 12.09, p < .001. High phonographic C words had a mean phonographic degree of 6.85 (SD = 3.51) and low phonographic C words had a mean phonographic degree of 7.35 (SD = 3.88), t (78) < 1, p = .55.

Phonographic clustering coefficient. The phonographic clustering coefficient, C, refers to the extent to which the phonographic neighbors of a word are also neighbors of each other. To calculate clustering coefficient, the number of connections between neighbors of a target word was counted and divided by the number of possible connections that could exist among the neighbors. Therefore, the clustering coefficient is the ratio of the actual number of connections existing among neighbors to the number of all possible connections among neighbors if every neighbor were connected. The value of the clustering coefficient ranges from 0 to 1; when C = 1 all neighbors of a word are neighbors of each other; when C = 0 no neighbors of the word are neighbors of each other. High phonographic degree words had a mean phonographic clustering coefficient of 0.283 (SD = 0.146) and low phonographic degree words had a mean phonographic clustering coefficient of 0.321 (SD = 0.278), t (78) < 1, p = .44. High phonographic C words had a mean phonographic clustering coefficient of 0.326 (SD = 0.123) and low phonographic C words had a mean phonographic clustering coefficient of 0.163 (SD = 0.079), t (78) = 7.04, p < .001.

Subjective familiarity. Subjective familiarity was measured on a seven-point scale (Nusbaum, Pisoni, & Davis, 1984). The rating scale ranged from 1, *You have never seen the word before*, to 4, *You recognize the word, but don't know the meaning*, to 7, *You recognize the word and are confident that you know the meaning of the word*. High phonographic degree words had a mean familiarity value of 6.83 (*SD* = 0.24) and low phonographic degree words had a mean familiarity value of 6.86 (*SD* = 0.25), *t* (78) < 1, *p* = .56. High phonographic *C* words had a mean familiarity value of 6.81 (*SD* = 0.31) and low phonographic *C* words had a mean familiarity value of 6.83 (*SD* = 0.23), *t* (78) < 1, *p* = .69. Therefore, both sets of words were considered highly familiar.

Word frequency. Word frequency refers to how often a given word occurs in a language. Log-base 10 of the raw frequency counts from Kučera and Francis (1967) were used. High phonographic degree words had a mean word frequency of 1.88 (*SD* = 0.55) and low phonographic degree words had a mean word frequency of 2.04 (*SD* = 0.65), *t* (78) = 1.19, *p* = .24. High phonographic *C* words had a mean word frequency of 1.95 (*SD* = 0.65) and low phonographic *C* words had a mean word frequency of 1.97 (*SD* = 0.73), *t* (78) < 1, *p* = .90. Log-base 10 of the frequency counts from the more current $\text{SUBTLEX}_{\text{US}}$ corpus (Brysbaert & New, 2009) were also obtained from the English Lexicon Project (Balota et al., 2007). Based on these frequency counts, High phonographic degree words had a mean word frequency of 2.40 (*SD* = 0.69) and low phonographic degree words had a mean word frequency of 2.65 (*SD* = 0.77), *t* (78) = -1.50, *p* = .14. High phonographic *C* words had a mean word frequency of 2.68 (*SD* = 0.83) and low phonographic *C* words had a mean word frequency of 2.63 (*SD* = 0.71), *t* (78) < 1, *p* = .80.

Number of phonemes. High phonographic degree words had a mean word length of 3.83 phonemes (*SD* = 0.59) and low phonographic degree words had a mean word length of 3.80 phonemes (*SD* = 0.61), *t* (78) < 1, *p* = .85. High phonographic *C* words had a mean word length of 3.60 phonemes (*SD* = 0.55) and low phonographic *C* words had a mean word length of 3.60 phonemes (*SD* = 0.55), *t* (78) < 1, *p* = 1.00.

Phonological degree. Phonological degree refers to the number of words that are phonologically similar to a given word (also known as phonological neighborhood density in the psycholinguistic literature). Phonological similarity was defined as the substitution, addition, or deletion of one letter in a given word to form a phonological neighbor (Luce & Pisoni, 1998). High phonographic degree words had a mean phonological degree of 11.83 (*SD* = 4.27) and low phonographic degree words had a mean phonological degree of 10.85 (*SD* = 5.06), *t* (78) < 1, *p* = .35. High phonographic *C* words had a mean phonological degree of 14.30 (*SD* = 6.85) and low phonographic *C* words had a mean phonological degree of 14.98 (*SD* = 7.11), *t* (78) < 1, *p* = .67.

Phonological clustering coefficient. The phonological clustering coefficient refers to the extent to which the phonological neighbors of a word are also neighbors of each other. The phonological clustering coefficient was calculated in the same manner as phonographic clustering coefficient (as described above), except that the density of connections was calculated among *phonological*, rather than phonographic, neighbors. High phonographic degree words had a mean phonological *C* of 0.294 (*SD* = 0.073) and low phonographic degree words had a mean phonological *C* of 0.286 (*SD* = 0.085), *t* (78) < 1, *p* = .64. High phonographic *C* words had a mean phonological *C* of 0.272 (*SD* = 0.063) and low phonographic *C* words had a mean phonological *C* of 0.249 (*SD* = 0.080), *t* (78) = 1.45, *p* = .15.

Phonological neighborhood frequency. Phonological neighborhood frequency refers to the average frequency of the phonological neighbors of a target word. High phonographic degree words had a mean phonological neighborhood frequency of 1.83 (*SD* = 0.29) and low phonographic degree words had a mean phonological neighborhood frequency of 1.88 (*SD* = 0.24), t (78) = 1.01, p = .32. High phonographic *C* words had a mean phonological neighborhood frequency of 1.83 (*SD* = 0.29) and low phonographic *C* words had a mean phonological neighborhood frequency of 1.85 (*SD* = 0.29), t (78) < 1, p = .73.

Phonotactic probability. Two measures of phonotactic probability—positional segment probability (the probability that a segment occurs in a certain position of a word) and biphone probability (the probability that two adjacent segments co-occur in a word)—were obtained from the Phonotactic Probability Calculator (Vitevitch & Luce, 2004). High phonographic degree words had a mean positional segment probability of 0.175 (*SD* = 0.0474) and low phonographic degree words had a mean positional segment probability of 0.176 (*SD* = 0.0475), t (78) < 1, p = .93. High phonographic degree words had a mean biphone probability of 0.0120 (*SD* = 0.00687) and low phonographic degree words had a mean biphone probability of 0.0116 (*SD* = 0.00693), t (78) < 1, p = .77. High phonographic *C* words had a mean positional segment probability of 0.171 (*SD* = 0.0451) and low phonographic *C* words had a mean positional segment probability of 0.170 (*SD* = 0.0414), t (78) < 1, p = .87. High phonographic *C* words had a mean biphone probability of 0.00940 (*SD* = 0.00606) and low phonographic *C* words had a mean biphone probability of 0.00942 (*SD* = 0.00576), t (78) < 1, p = .98.

Number of letters. High phonographic degree words had a mean word length of 4.53 letters (*SD* = 0.55) and low phonographic degree words had a mean word length of 4.50 letters (*SD* = 0.56), t (78) < 1, p = .84. High phonographic *C* words had a mean word length of 4.33

letters (*SD* = 0.53) and low phonographic *C* words had a mean word length of 4.30 letters (*SD* = 0.56), t (78) < 1, p = .84.

Orthographic degree. Orthographic degree refers to the number of words that are orthographically similar to a given word (also known as orthographic neighborhood density in the psycholinguistic literature). Orthographic similarity was defined as the substitution, addition, or deletion of one letter in a given word to form an orthographic neighbor. Note that this operationalization is slightly different from that of Coltheart's *N* (Coltheart et al., 1977); the latter refers to the number of orthographic neighbors that could be obtained via the *substitution* of one letter in a given word. To be consistent with the way in which the phonographic language network was constructed, orthographic neighbors were defined based on the substitution, addition, or deletion of one letter in a given word. High phonographic degree words had a mean orthographic degree of 6.68 (*SD* = 1.90) and low phonographic degree words had a mean orthographic degree of 6.65 (*SD* = 1.93), t (78) < 1, p = .95. High phonographic *C* words had a mean orthographic degree of 8.63 (*SD* = 4.00) and low phonographic *C* words had a mean orthographic degree of 9.15 (*SD* = 4.34), t (78) < 1, p = .84.

Orthographic clustering coefficient. The orthographic clustering coefficient refers to the extent to which the orthographic neighbors of a word are also neighbors of each other. The orthographic clustering coefficient was calculated in the same manner as phonographic clustering coefficient (as described above), except that the density of connections was calculated among *orthographic*, rather than phonographic, neighbors. High phonographic degree words had a mean orthographic *C* of 0.277 (*SD* = 0.135) and low phonographic degree words had a mean orthographic *C* of 0.274 (*SD* = 0.141), t (78) < 1, p = .94. High phonographic *C* words had a mean orthographic *C* of 0.252 (*SD* = 0.056) and low phonographic *C* words had a mean orthographic *C* of 0.247 (*SD* = 0.060), *t* (78) < 1, *p* = .72.

Orthographic neighborhood frequency. Orthographic neighborhood frequency refers to the average frequency of the orthographic neighbors of a target word. Values were obtained from the English Lexicon Project, which used log-transformed frequency counts based on the Hyperspace Analogue to Language (HAL) corpus, which consists of approximately 131 million words (Lund & Burgess, 1996), to calculate orthographic neighborhood frequency. High phonographic degree words had a mean orthographic neighborhood frequency of 7.13 (*SD* = 1.05) and low phonographic degree words had a mean orthographic neighborhood frequency of 7.45 (*SD* = 1.13), *t* (78) = 1.29, *p* = .20. High phonographic *C* words had a mean orthographic neighborhood frequency of 7.50 (*SD* = 0.95) and low phonographic *C* words had a mean orthographic neighborhood frequency of 7.40 (*SD* = 1.13), *t* (78) < 1, *p* = .67. To be consistent with phonological neighborhood frequency, orthographic neighborhood frequencies for the word stimuli were also calculated using the Kučera and Francis frequency norms. Based on these norms, high phonographic degree words had a mean orthographic neighborhood frequency of 1.83 (*SD* = 0.40) and low phonographic degree words had a mean orthographic neighborhood frequency of 1.86 (*SD* = 0.31), *t* (78) < 1, *p* = .72. High phonographic *C* words had a mean orthographic neighborhood frequency of 1.84 (*SD* = 0.38) and low phonographic *C* words had a mean orthographic neighborhood frequency of 1.83 (*SD* = 0.32), *t* (78) < 1, *p* = .92.

Bigram frequency. Two measures of bigram frequency—average bigram counts and sum of bigram counts (by position)—were obtained from the ELP (Balota et al., 2007). High phonographic degree words had a mean average bigram count of 2,728.82 (*SD* = 1,218.51) and low phonographic degree words had a mean average bigram count of 2,876.26 (*SD* = 1,204.88), *t*

(78) < 1, *p* = .59. High phonographic degree words had a mean sum of bigram count (by position) of 1,871.23 (*SD* = 662.43) and low phonographic degree words had a mean sum of bigram count (by position) of 2,061.28 (*SD* = 698.99), *t* (78) = 1.25, *p* = .22. High phonographic *C* words had a mean average bigram count of 2,763.84 (*SD* = 1,251.79) and low phonographic *C* words had a mean average bigram count of 2,735.60 (*SD* = 1,041.33), *t* (78) < 1, *p* = .91. High phonographic *C* words had a mean sum of bigram count (by position) of 1,910.80 (*SD* = 897.78) and low phonographic *C* words had a mean sum of bigram count (by position) of 1,818.35 (*SD* = 722.58), *t* (78) < 1, *p* = .61.

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

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

4.2 Results

Accuracy was scored offline by an undergraduate research assistant. Trials containing mispronunciations of the word or responses that triggered the voice-key prematurely (e.g., coughing, “uh”) were coded as errors. The author also independently scored ~10% of the data. There was a high level of agreement between the two independent scorers (~99%).

For the RT data, errors were first excluded, after which responses below 200ms and above 2000ms were eliminated before the overall mean and *SD* of each participant’s RT was calculated. Trials with latencies that were 2 *SD*s above or below each participant’s mean RT were considered outliers and excluded from analysis. This resulted in ~5% of the data being removed.

Trials from two items were excluded from the analysis due to very low overall item accuracies in the naming task (i.e., less than 80%): “lung” (60%) and “mount” (72%). In addition, items with a phonographic degree of 2 were excluded from the analysis, because the phonographic *C* value of these words was either 0 or 1 (i.e., a binary value), and is not an accurate representation of the level of interconnectivity among a word’s phonographic neighbors. On the other hand, for words with more than 2 neighbors, clustering coefficient is a continuous variable that ranges from 0 to 1 that represents the extent to which a word’s neighbors are also neighbors of each other. Indeed, one known limitation of the *C* measure is that its value can be biased by the node’s degree, whereby nodes with few neighbors tend to have a larger clustering coefficient as compared to nodes with several neighbors (see Opsahl & Panzarasa, 2009; Soffer & Vazquez, 2005); although it is important to note that *C* and degree are *not* correlated in the phonological network of language (Vitevitch et al., 2012).

To ensure that the analysis is not biased by words with phonographic *C*s that distort the level of interconnectivity among neighbors, the following items were excluded from the analysis: “balm”, “cue”, “crime” (phonographic $C = 1$), and “bleed”, “slur”, “tomb” (phonographic $C = 0$).

Using the lme4 package in R, linear mixed effects (LME) and generalized linear mixed effects (GLM) models were used to predict RTs and accuracy respectively from the naming data (Bates et al., 2015). The RT model included the following predictors: (a) random effects of participants and items, (b) fixed effects of phonographic degree and phonographic *C.* The Accuracy model included the same predictors: (a) random effects of participants and items, (b) fixed effects of phonographic degree and phonographic *C.* For the RT model, additional lexical variables (i.e., subjective familiarity, word frequency, number of phonemes, phonological degree, phonological *C*, phonological neighborhood frequency, phonotactic probability, number of letters, orthographic degree, orthographic *C*, orthographic neighborhood frequency, bigram frequency) were included as covariates to control for any influences these variables may have on word recognition times. For both models all predictor and covariate variables were standardized—a common practice in regression analysis when variables have very different scales, in order to facilitate comparisons of the relative importance of various predictor variables.

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

Reaction time. Table 2 presents the results of the linear mixed effects model for naming RTs. The following fixed effects were significant: phonographic *C*, familiarity, number of phonemes, number of letters, orthographic degree, orthographic *C*, orthographic neighborhood frequency, average bigram counts, and sum of bigram counts by position. Phonographic degree did not significantly predict naming RTs, standardized $\beta = -14.67$, $t = -1.72$, $p = .09$. The mean RT for high phonographic degree words was 921 ms ($SD = 153$) and the mean RT for low phonographic degree words was 923 ms ($SD = 152$). Phonographic *C* significantly predicted naming RTs, standardized $\beta = -12.39$, $t = -2.40$, $p = .016$, such that words with high phonographic *C* were more quickly named as compared to words with low phonographic *C*. For each standardized unit increase in phonographic *C* (approximately 0.153), the average decrease in naming RTs was 12 ms. The mean RT for high phonographic *C* words was 909 ms ($SD = 149$) and the mean RT for low phonographic *C* words was 913 ms ($SD = 150$).

Table 2. LME model estimates for fixed and random effects for the auditory naming experiment (reaction time; Experiment 1).

Random Effects	Variance	*SD*
Items		
Intercept	2096.00	45.78
Participants		
Intercept	22115.00	148.71

Fixed Effects	β	*SE*	*t*	*p*-value
Intercept	917.57	19.57	46.88	<.001***
Phonographic degree	-14.67	8.55	-1.72	.086+
Phonographic *C*	-12.34	5.13	-2.40	.016*
Number of phonemes	24.21	11.01	2.20	.028*
Phonological degree	2.66	7.69	0.35	.73
Phonological *C*	-1.37	4.49	-0.31	.76
Phonological neighborhood frequency	-6.18	5.11	-1.21	.23
Positional segment probability	-0.37	6.94	-0.05	.96
Biphone probability	-8.62	7.18	-1.20	.23
Number of letters	15.22	7.05	2.16	.031*
Orthographic degree	23.81	8.48	2.81	.005**
Orthographic *C*	17.75	5.25	3.38	<.001***
Orthographic neighborhood frequency	11.36	4.30	2.64	.008**
Average bigram counts	14.53	5.11	2.85	.004**
Sum of bigram counts by position	-16.22	6.65	-2.44	.014*
Subjective familiarity	8.60	4.28	2.01	.044*
Word frequency	-1.76	4.50	-0.39	.70

Note: + $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Accuracy. Table 3 presents the results of the generalized linear mixed effects model for naming accuracies. No fixed effects were significant. Phonographic degree did not significantly predict naming accuracies, standardized $\beta = -0.012$, $z < 1$, $p = .93$. The mean accuracy for high phonographic degree words was 98.4% ($SD = 2.02$) and the mean accuracy for low phonographic degree words was 97.9% ($SD = 2.50$). Phonographic *C* did not significantly predict naming accuracies, standardized $\beta = -0.205$, $z = -1.53$, $p = .13$. The mean accuracy for high phonographic *C* words was 97.5% ($SD = 2.92$) and the mean accuracy for low phonographic *C* words was 98.6% ($SD = 2.38$).

Table 3. GLM model estimates for fixed and random effects for the auditory naming experiment (accuracy; Experiment 1).

Random Effects	Variance	*SD*
Items		
Intercept	1.74	1.32
Participants		
Intercept	0.32	0.56

Fixed Effects	β	*SE*	*z*	*p*-value
Intercept	4.74	0.19	24.78	<.001***
Phonographic degree	-0.01	0.14	-0.08	.93
Phonographic *C*	-0.21	0.13	-1.53	.13

Note: $^{+}$ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

4.3 Discussion

The results of Experiment 1 showed that phonographic *C*, but not phonographic degree, predicted naming RTs. High phonographic *C* words were named more quickly than low phonographic *C* words, after taking into account the variance contributed by several lexical variables known to influence language processing.

Recall that phonographic *C* refers to the extent to which the phonographic neighbors of a word are also phonographic neighbors of each other, and that phonographic neighbors are words that both phonologically and orthographically similar to a target word. In the present study, a facilitatory effect was observed for words with a *high* level of interconnectivity among its phonographic neighbors. At first glance, this result appears to contradict previous work investigating the influence of the phonological clustering coefficient on spoken word recognition, which found that words with high phonological clustering coefficients were more slowly and less accurately processed (Chan & Vitevitch, 2009; 2010; see also Siew, 2016 for a similar finding with the network density measure). A simple diffusion framework was used to account for this finding. In this framework, activation spreads back and forth between the target word, its neighbors, and other words in the network (see also the computer simulation reported in Vitevitch, Ercal, & Adagarla, 2011). For words with highly interconnected neighborhoods, over time a greater amount of activation will remain within the neighborhood, instead of diffusing to the rest of the network. On the other hand, for words with less interconnected neighborhoods, over time most of the activation will be spread to the rest of the network. Based on this account, it is more difficult for words with highly interconnected neighborhoods (i.e., words with high phonological clustering coefficients) to “stand out” from its competitors as compared to words with less interconnected neighborhoods (i.e., words with low phonological clustering coefficients).

However, it is important to note that although both phonological *C* and phonographic *C* represent the amount of interconnectivity among a word's neighbors, these two measures are different in that phonological *C* represents the structure of a word's phonological neighborhood (i.e., the phonological layer in the phonographic multiplex), whereas phonographic *C* represents the structure of a word's phonological and orthographic neighborhoods (i.e., the phonological and orthographic layers of the phonographic multiplex). More specifically, phonographic *C* can be viewed as a metric that represents the internal structure of the area where the phonological and orthographic neighborhoods of words overlap. Therefore, a facilitatory effect might be expected in this case because higher phonographic *C* values indicate greater overlap in the similarity structures in the phonological and orthographic neighborhoods of words. Based on the activation diffusion framework described earlier, one might expect that for high phonographic *C* words, similar, overlapping patterns of activation occur in the phonological and orthographic layers of the phonographic multiplex, which reinforce each other during processing and facilitate the recognition of the target word.

The present results are significant because it is the first to demonstrate that a network science metric—the phonographic clustering coefficient—which simultaneously represents the phonological *and* orthographic structure of language, influences spoken word recognition. It is important to ensure that these findings are not task-specific and that they can indeed be replicated using a different psycholinguistic task. The next experiment sought to replicate the present findings using another traditional task from psycholinguistics—auditory lexical decision task.

Chapter 5

Experiment 2: Auditory Lexical Decision

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

5.1 Method

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

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

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

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

5.2 Results

The trimming procedure is identical to that used in Experiment 1. For the RT data, errors were first excluded, after which responses below 200ms and above 2000ms were eliminated before the overall mean and *SD* of each participant’s RT was calculated. Trials with latencies that were 2 *SD*s above or below each participant’s mean RT were removed. This resulted in ~7% of the data being removed.

Trials from four items were excluded from the analysis due to very low overall item accuracies in the lexical decision task (i.e., less than 50%): “clod” (23%), “balk” (32%), “plume” (38%), and “posh” (46%). In addition, items with a phonographic degree of 2 (the same ones listed in Experiment 1) were excluded from the analysis to ensure that the analysis is not biased by words with phonographic *C*s that distort the level of interconnectivity among neighbors.

As in Experiment 1, linear mixed effects (LME) and generalized linear mixed effects (GLM) models were used to predict RTs and accuracy respectively from the lexical decision data. The RT model included the following predictors: (a) random effects of participants and items, (b) fixed effects of phonographic degree and phonographic *C*. The Accuracy model included the same predictors: (a) random effects of participants and items, (b) fixed effects of phonographic degree and phonographic *C*. For the RT model, additional lexical variables (i.e., subjective familiarity, word frequency, number of phonemes, phonological degree, phonological *C*, phonological neighborhood frequency, phonotactic probability, number of letters, orthographic degree, orthographic *C*, orthographic neighborhood frequency, bigram frequency) were included as covariates to control for any influences these variables may have on word recognition times. For both models all predictor and covariate variables were standardized—a common practice in regression analysis when variables have very different scales, in order to facilitate comparisons of the relative importance of various predictor variables.

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

Reaction time. Table 4 presents the results of the linear mixed effects model for lexical decision RTs. The following fixed effects were significant: phonographic *C*, frequency, familiarity, orthographic *C*, and average bigram counts. Phonographic degree did not significantly predict lexical decision RTs, standardized $\beta = -17.73$, $t = -1.59$, $p = .11$. The mean RT for high phonographic degree words was 890 ms ($SD = 84$) and the mean RT for low phonographic degree words was 901 ms (*SD* = 86). Phonographic *C* significantly predicted lexical decision RTs, standardized $\beta = -18.41$, $t = -2.60$, $p = .009$, such that words with high phonographic *C* were more quickly responded to as compared to words with low phonographic *C*. For each standardized unit increase in phonographic *C* (approximately 0.153), the average decrease in naming RTs was 18 ms. The mean RT for high phonographic *C* words was 906 ms (*SD* = 87) and the mean RT for low phonographic *C* words was 901 ms (*SD* = 101).

Table 4. LME model estimates for fixed and random effects for the auditory lexical decision experiment (reaction time; Experiment 2).

Random Effects	Variance	*SD*
Items		
Intercept	3127.00	55.92
Participants		
Intercept	7138.00	84.49

Fixed Effects	β	*SE*	*t*	*p*-value
Intercept	902.55	11.54	78.19	<.001***
Phonographic degree	-17.73	11.18	-1.59	.11
Phonographic *C*	-18.41	7.07	-2.60	.009**
Number of phonemes	12.06	14.00	0.86	.39
Phonological degree	-0.46	9.76	-0.05	.96
Phonological *C*	-2.64	5.68	-0.47	.64
Phonological neighborhood frequency	5.14	6.45	0.80	.43
Positional segment probability	7.37	8.87	0.83	.41
Biphone probability	-10.00	9.14	-1.09	.27
Number of letters	6.16	8.96	0.69	.49
Orthographic degree	21.80	11.17	1.95	.051⁺
Orthographic *C*	22.33	7.22	3.09	.002**
Orthographic neighborhood frequency	5.36	5.39	0.99	.32
Average bigram counts	15.61	6.40	2.44	.015*
Sum of bigram counts by position	-13.42	8.53	-1.57	.12
Subjective familiarity	-13.08	5.01	-2.61	.009**
Word frequency	-11.72	5.60	-2.09	.036*

Note: ⁺ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Accuracy. Table 5 presents the results of the generalized linear mixed effects model for lexical decision accuracies. No fixed effects were significant. Phonographic degree did not significantly predict lexical decision accuracies, standardized $\beta = 0.117$, $z = 1.08$, $p = .28$. The mean accuracy for high phonographic degree words was 90.7% ($SD = 6.72$) and the mean accuracy for low phonographic degree words was 92.1% ($SD = 6.52$). Phonographic *C* did not significantly predict lexical decision accuracies, standardized $\beta = 0.068$, $z < 1$, $p = .53$. The mean accuracy for high phonographic *C* words was 88.8% ($SD = 6.40$) and the mean accuracy for low phonographic *C* words was 91.2% ($SD = 6.20$).

Table 5. GLM model estimates for fixed and random effects for the auditory lexical decision experiment (accuracy; Experiment 2).

Random Effects	Variance	*SD*
Items		
Intercept	1.52	1.23
Participants		
Intercept	0.37	0.61

Fixed Effects	β	*SE*	*z*	*p*-value
Intercept	2.75	0.13	20.63	<.001***
Phonographic degree	0.12	0.11	1.08	.28
Phonographic *C*	0.07	0.11	0.64	.53

Note: $^{+}$ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

5.3 Discussion

The results of Experiment 2 showed that phonographic *C*, but not phonographic degree, predicted lexical decision RTs, replicating the results of Experiment 1. High phonographic *C* words were recognized more quickly than low phonographic *C* words, after taking into account the variance contributed by several lexical variables known to influence language processing.

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

In addition, it is worth noting that phonographic *C*, but not phonographic degree, was a significant predictor in both experiments. Both measures capture somewhat different aspects of the phonological and orthographic similarity structure of language. Phonographic degree can be viewed as a more "coarse-grained" measure of phonological and orthographic similarity, as it simply represents the number of words that are both phonologically and orthographically similar to a given word, whereas phonographic *C* captures more subtle aspects of the similarity structure—namely, the internal connectivity among these phonographic neighbors. Overall, the results suggest that phonographic *C* is a better predictor of spoken word recognition performance than phonographic degree.   Together the results of Experiments 1 and 2 demonstrate that the phonographic clustering coefficient, a network science metric that simultaneously represents the phonological *and* orthographic structure of language, influences spoken word recognition. In the next chapter, I analyze speeded naming and lexical decision behavioral data from the English Lexicon Project to determine if the results can be replicated in visual word recognition.

Chapter 6

English Lexicon Project Analyses

The availability of databases containing item-level behavioral data and lexical variables for a large set of words has afforded large-scale, megastudies of visual word recognition (New, Ferrand, Pallier, & Brysbaert, 2006; Yap & Balota, 2009). The megastudy approach can complement the conventional psycholinguistic approach of factorial experiments whereby specific variables are targeted in a small-scale study by overcoming some of the limitations associated with the traditional approach (Balota, Yap, Hutchison, & Cortese, 2012). In a factorial experiment, psycholinguists typically carefully select word stimuli such that groups of words are matched on a variety of lexical characteristics while manipulating the lexical variable of interest whereas in the megastudy approach, extraneous lexical variables can be statistically controlled for. In addition, whereas lab-based experiments with carefully controlled stimuli can answer the question of whether phonographic degree and phonographic *C* influences word recognition, the large-database approach can answer the slightly different question of *how much influence* phonographic degree and phonographic *C* have on word recognition performance, after taking into account the influence of other lexical variables on word recognition. The database approach also allows for replication using a larger set of stimuli. In the present chapter, I conducted a regression analysis of words in the ELP to determine if phonographic degree and phonographic clustering coefficient are significant predictors of performance of speeded naming and visual lexical decision for a large set of words, after taking into account the contributions of other lexical variables.

6.1 Method

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

Dataset/Materials. ELP behavioral data exist for 2,914 of the 3,292 (~90%) words in the giant component of the phonographic network. It is important to note that some of the words in the phonographic network do not have a "meaningful" phonographic clustering coefficient value. For instance, it is not possible to calculate the clustering coefficient for words with either 0 or 1 phonographic neighbor(s) (i.e., phonographic *C* for these words is undefined). As discussed earlier, for words with more than 2 neighbors, clustering coefficient is a continuous variable that ranges from 0 to 1 that represents the extent to which a word's neighbors are also neighbors of each other. However, the phonographic clustering coefficient for words with 2 phonographic neighbors is binary (i.e., either 0 or 1), and does not accurately represent the level of interconnectivity among a word's phonographic neighbors. Therefore, to ensure that the analysis was not biased by the presence of several words with an undefined phonographic *C* (i.e., words with a phonographic degree of 1), or by words with phonographic *C*s that distort the level of interconnectivity among neighbors (i.e., words with a phonographic degree of 2), words with 2 or fewer phonographic neighbors were excluded, resulting in a total of 2,120 words for the subsequent regression analyses.

6.2 Results

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

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

6.2.1 Speeded naming

Reaction times. Table 6 shows the results of the regression analysis on *z*-scored naming RTs. In Step 1, the following variables significantly predicted naming RTs: number of phonemes, phonological degree, positional segment probability, biphone probability, number of letters, orthographic *C*, average bigram counts, sum of bigram counts by position, familiarity, and frequency. Together, the variables entered at Step 1 explained 28.0% of the variance in naming RTs, accounting for a significant proportion of the variance in naming RTs, $R^2 = .280$, $F(14, 2103) = 58.27$, $p < .001$.

In Step 2, the following variables significantly predicted RTs: positional segment probability, biphone probability, number of letters, orthographic degree, average bigram counts, sum of bigram counts by position, familiarity, frequency, and phonographic degree. Phonographic degree significantly predicted naming RTs, standardized $\beta = -0.0763$, $t(2101) = -7.72$, $p < .001$, such that words with high phonographic degree were more quickly named as compared to words with low phonographic degree. For each standardized unit increase in phonographic degree (approximately 4.31), the average decrease in *z*-scored naming RTs was 0.076 *SD*s. Phonographic *C* did not significantly predict naming RTs, standardized $\beta = -0.0102$, $t(2101) = -1.37$, $p = .17$. The influence of phonographic variables accounted for an additional 1.9% of the variance, $\Delta R^2 = .019$, $F(2, 2101) = 29.85$, $p < .001$. Together, the variables entered at both steps explained 29.9% of the variance in naming RTs, accounting for a significant proportion of the variance in naming RTs, $R^2 = .299$, $F(16, 2101) = 56.12$, $p < .001$.

Table 6. Regression results for ELP naming reaction times.

Variable	β	*SE*	*t*	*p*	R^2	ΔR^2
Step 1						
Number of phonemes	-0.0254	0.0103	-2.46	.014*		
Phonological degree	-0.0191	0.00724	-2.64	.008**		
Phonological *C*	0.00742	0.00465	1.60	.11		
Phonological neighborhood frequency	0.00356	0.00529	0.673	.50		
Positional segment probability	0.0435	0.00734	5.92	<.001***		
Biphone probability	-0.0252	0.00710	-3.55	<.001***		
Number of letters	0.0723	0.00832	8.69	<.001***		
Orthographic degree	-0.00843	0.00669	-1.26	.21		
Orthographic *C*	-0.0132	0.00453	-2.92	.004**		
Orthographic neighborhood frequency	0.00734	0.00525	1.40	.16		
Average bigram counts	0.0235	0.00501	4.70	<.001***		
Sum of bigram counts by position	-0.0323	0.00648	-4.99	<.001***		
Subjective familiarity	-0.0516	0.00451	-11.44	<.001***		
Word frequency	-0.0450	0.00486	-9.26	<.001***		
					.280***	
Step 2						
Number of phonemes	-0.0138	0.0103	-1.34	.18		
Phonological degree	-0.00337	0.00773	0.436	.66		
Phonological *C*	0.00786	0.00468	1.68	.09		
Phonological neighborhood frequency	0.000791	0.00524	0.151	.88		
Positional segment probability	0.0447	0.00726	6.16	<.001***		
Biphone probability	-0.0227	0.00701	-3.24	.001**		
Number of letters	0.0758	0.00822	9.22	<.001***		
Orthographic degree	0.0523	0.0103	5.10	<.001***		
Orthographic *C*	-0.000485	0.00739	-0.066	.95		
Orthographic neighborhood frequency	0.00799	0.00518	1.54	.12		
Average bigram counts	0.0260	0.00495	5.25	<.001***		

Sum of bigram counts by position	-0.0420	0.00652	-6.44	<.001***		
Subjective familiarity	-0.0501	0.00445	-11.26	<.001***		
Word frequency	-0.0489	0.00482	-10.15	<.001***		
Phonographic degree	-0.0763	0.00989	-7.72	<.001***		
Phonographic *C*	-0.0102	0.00744	-1.37	.17		
					.299***	.019***

Note: N = 2,120; $^{+}$ p < .10, * p < .05, ** p < .01, *** p < .001

Accuracy. Table 7 shows the results of the regression analysis on naming accuracies. In Step 1, the following variables significantly predicted accuracies: number of phonemes, phonological neighborhood frequency, biphone probability, orthographic *C*, orthographic neighborhood frequency, sum of bigram counts by position, familiarity, and frequency. Together, the variables entered at Step 1 explained 27.5% of the variance in naming accuracies, accounting for a significant proportion of the variance in naming accuracies, $R^2 = .275$, $F(14, 2103) = 56.86$, $p < .001$.

Table 7. Regression results for ELP naming accuracies.

Variable	β	*SE*	*t*	*p*	R^2	ΔR^2
Step 1						
Number of phonemes	0.00605	0.00275	2.20	.028*		
Phonological degree	0.00293	0.00193	1.52	.13		
Phonological *C*	0.000633	0.00124	0.512	.61		
Phonological neighborhood frequency	-0.00327	0.00141	-2.32	.020*		
Positional segment probability	-0.00349	0.00195	-1.79	.074^{+}		
Biphone probability	0.00534	0.00189	2.83	.005**		
Number of letters	0.00111	0.00221	0.502	.62		
Orthographic degree	0.00159	0.00178	0.892	.37		
Orthographic *C*	0.00243	0.00120	2.02	.043*		
Orthographic neighborhood frequency	0.00290	0.00140	2.08	.038*		
Average bigram counts	-0.00130	0.00133	-0.973	.33		
Sum of bigram counts by position	-0.00516	0.00172	-3.00	.003**		
Subjective familiarity	0.0274	0.00120	22.86	<.001***		
Word frequency	0.00353	0.00129	2.73	.006**		
					.275***	
Step 2						
Number of phonemes	0.00432	0.00277	1.56	.12		
Phonological degree	-0.000402	0.00208	-0.194	.85		
Phonological *C*	0.000977	0.00126	0.777	.44		
Phonological neighborhood frequency	-0.00274	0.00141	-1.95	.051^{+}		
Positional segment probability	-0.00349	0.00195	-1.79	.073^{+}		
Biphone probability	0.00490	0.00188	2.60	.009**		
Number of letters	0.000615	0.00221	0.279	.78		
Orthographic degree	-0.00679	0.00276	-2.46	.014*		
Orthographic *C*	0.00289	0.00198	1.46	.15		
Orthographic neighborhood frequency	0.00282	0.00139	2.03	.043*		
Average bigram counts	-0.00169	0.00133	-1.27	.20		

Sum of bigram counts by position	-0.00376	0.00175	-2.15	.032*		
Subjective familiarity	0.0272	0.00120	22.78	<.001***		
Word frequency	0.00411	0.00129	3.17	.002**		
Phonographic degree	0.0108	0.00265	4.07	<.001***		
Phonographic *C*	-0.00167	0.00200	-0.83	.40		
					.281***	.006***

Note: N = 2,120; $^{+}$ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

In Step 2, the following variables significantly predicted naming accuracies: biphone probability, orthographic degree, orthographic neighborhood frequency, sum of bigram counts by position, familiarity, frequency, and phonographic degree. Phonographic degree significantly predicted naming accuracies, standardized $\beta = 0.0108$, $t(2101) = 4.07$, $p < .001$, such that words with high phonographic degree were more accurately named as compared to words with low phonographic degree. For each standardized unit increase in phonographic degree (approximately 4.31), the average increase in naming accuracies was 1.09%. Phonographic *C* did not significantly predict naming accuracies, standardized $\beta = -0.00167$, $t(2101) < 1$, $p = .40$. The influence of phonographic variables accounted for an additional 0.6% of the variance, $\Delta R^2 = .006$, $F(2, 2101) = 9.22$, $p < .001$. Together, the variables entered at both steps explained 28.1% of the variance in naming accuracies, accounting for a significant proportion of the variance in naming accuracies, $R^2 = .281$, $F(16, 2101) = 51.29$, $p < .001$.

6.2.2 Visual lexical decision

Reaction times. Table 8 shows the results of the regression analysis on *z*-scored lexical decision RTs. In Step 1, the following variables significantly predicted lexical decision RTs: number of letters, orthographic degree, orthographic *C*, familiarity, and frequency. Together, the variables entered at Step 1 explained 50.5% of the variance in lexical decision RTs, accounting for a significant proportion of the variance in lexical decision RTs, $R^2 = .505$, $F(14, 2103) = 153.5$, $p < .001$.

Table 8. Regression results for ELP visual lexical decision reaction times.

Variable	β	*SE*	*t*	*p*	R^2	ΔR^2
Step 1						
Number of phonemes	0.00948	0.00119	0.795	.43		
Phonological degree	0.000773	0.00836	0.092	.93		
Phonological *C*	0.00787	0.00537	1.47	.14		
Phonological neighborhood frequency	0.00906	0.00611	1.48	.14		
Positional segment probability	-0.0100	0.00848	-1.185	.24		
Biphone probability	0.00173	0.00820	0.210	.83		
Number of letters	-0.0206	0.00961	-2.15	.032*		
Orthographic degree	-0.0242	0.00772	-3.14	.002**		
Orthographic *C*	-0.0119	0.00523	-2.27	.023*		
Orthographic neighborhood frequency	-0.00159	0.00606	-0.261	.79		
Average bigram counts	0.0113	0.00579	1.96	.050^{+}		
Sum of bigram counts by position	0.00240	0.00748	0.321	.75		
Subjective familiarity	-0.133	0.00521	-25.45	<.001***		
Word frequency	-0.121	0.00561	-21.49	<.001***		
					.505***	
Step 2						
Number of phonemes	0.0149	0.0120	1.24	.22		
Phonological degree	0.0113	0.00903	1.25	.21		
Phonological *C*	0.00733	0.00548	1.34	.18		
Phonological neighborhood frequency	0.00756	0.00613	1.23	.22		
Positional segment probability	-0.00979	0.00848	-1.15	.25		
Biphone probability	0.00302	0.00820	0.369	.71		
Number of letters	-0.0190	0.00961	-1.98	.048*		
Orthographic degree	0.00307	0.0120	0.255	.80		
Orthographic *C*	-0.00999	0.00863	-1.16	.25		
Orthographic neighborhood frequency	-0.00130	0.00605	-0.215	.83		
Average bigram counts	0.0125	0.00579	2.17	.030*		
Sum of bigram counts by position	-0.00206	0.00761	-0.271	.79		

Subjective familiarity	-0.132	0.00520	-25.35	<.001***		
Word frequency	-0.122	0.00563	-21.72	<.001***		
Phonographic degree	-0.0348	0.0116	-3.01	.003**		
Phonographic *C*	0.000994	0.00870	0.114	.91		
					.508***	.003**

Note: $N = 2{,}120$; $^{+}p < .10$, $*p < .05$, $**p < .01$, $***p < .001$

In Step 2, the following variables significantly predicted lexical decision RTs: number of letters, average bigram counts, familiarity, frequency, and phonographic degree. Phonographic degree significantly predicted lexical decision RTs, standardized β = -0.0348, t (2101) = -3.01, p = .003, such that words with high phonographic degree were more quickly responded to as compared to words with low phonographic degree. For each standardized unit increase in phonographic degree (approximately 4.31), the average decrease in z-scored lexical decision RTs was 0.035 *SD*s. Phonographic *C* did not significantly predict lexical decision RTs, standardized β = 0.000994, t (2101) < 1, p = .91. The influence of phonographic variables accounted for an additional 0.3% of the variance, $\Delta R^2 = .003$, F (2, 2101) = 4.66, p = .01. Together, the variables entered at both steps explained 50.8% of the variance in lexical decision RTs, accounting for a significant proportion of the variance in lexical decision RTs, $R^2 = .508$, F (16, 2101) = 135.4, p < .001.

Accuracy. Table 9 shows the results of the regression analysis on lexical decision accuracies. In Step 1, the following variables significantly predicted lexical decision accuracies: phonological neighborhood frequency, number of letters, orthographic degree, sum of bigram counts by position, familiarity, and frequency. Together, the variables entered at Step 1 explained 65.0% of the variance in lexical decision accuracies, accounting for a significant proportion of the variance in lexical decision accuracies, $R^2 = .650$, F (14, 2103) = 279.2, p < .001.

Table 9. Regression results for ELP visual lexical decision accuracies.

Variable	β	*SE*	*t*	*p*	R^2	ΔR^2
Step 1						
Number of phonemes	0.000508	0.00597	0.085	.93		
Phonological degree	-0.00514	0.00418	-1.23	.22		
Phonological *C*	-0.0000376	0.00268	-0.014	.99		
Phonological neighborhood frequency	-0.00806	0.00306	-2.64	.008**		
Positional segment probability	-0.000768	0.00424	-0.181	.86		
Biphone probability	0.00329	0.00410	0.802	.42		
Number of letters	0.0243	0.00481	5.06	<.001***		
Orthographic degree	0.0156	0.00386	4.05	<.001***		
Orthographic *C*	0.00295	0.00261	1.13	.26		
Orthographic neighborhood frequency	0.00145	0.00303	0.478	.63		
Average bigram counts	0.000382	0.00289	0.132	.90		
Sum of bigram counts by position	-0.00751	0.00374	-2.01	.045*		
Subjective familiarity	0.130	0.00260	49.93	<.001***		
Word frequency	0.0294	0.00281	10.46	<.001***		
					.650***	
Step 2						
Number of phonemes	-0.00218	0.00602	-0.363	.72		
Phonological degree	-0.0104	0.00452	-2.30	.022*		
Phonological *C*	-0.0000581	0.00274	-0.021	.98		
Phonological neighborhood frequency	-0.00739	0.00306	-2.41	.016*		
Positional segment probability	-0.00102	0.00424	-0.241	.81		
Biphone probability	0.00270	0.00410	0.658	.51		
Number of letters	0.0235	0.00481	4.89	<.001***		
Orthographic degree	0.00165	0.00600	0.274	.78		
Orthographic *C*	0.000269	0.00432	0.062	.95		
Orthographic neighborhood frequency	0.00130	0.00303	0.430	.67		
Average bigram counts	-0.000200	0.00290	-0.069	.95		
Sum of bigram counts by position	-0.00528	0.00381	-1.39	.17		

Subjective familiarity	0.130	0.00260	49.84	<.001***		
Word frequency	0.0303	0.00282	10.74	<.001***		
Phonographic degree	0.0176	0.00578	3.05	.002**		
Phonographic *C*	0.00172	0.00435	0.396	.69		
					.652***	.002*

Note: N = 2,120; $^{+}$ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

In Step 2, the following variables significantly predicted lexical decision accuracies: phonological degree, phonological neighborhood frequency, number of letters, familiarity, frequency, and phonographic degree. Phonographic degree significantly predicted lexical decision accuracies, standardized $\beta = 0.0176$, $t\ (2101) = 3.05$, $p = .002$, such that words with high phonographic degree were more accurately responded to as compared to words with low phonographic degree. For each standardized unit increase in phonographic degree (approximately 4.31), the average increase in lexical decision accuracies was 1.76%. Phonographic *C* did not significantly predict lexical decision accuracies, standardized $\beta = 0.00172$, $t\ (2101) < 1$, $p = .69$. The influence of phonographic variables accounted for an additional 0.2% of the variance, $\Delta R^2 = .002$, $F\ (2, 2101) = 4.66$, $p = .01$. Together, the variables entered at both steps explained 65.2% of the variance in lexical decision accuracies, accounting for a significant proportion of the variance in lexical decision accuracies, $R^2 = .652$, $F\ (16, 2101) = 245.7$, $p < .001$.

6.3 Discussion

The results of the ELP analyses showed that phonographic degree, but not phonographic *C*, predicted naming and lexical decision RTs in visual word recognition. High phonographic degree words were named and recognized more quickly than low phonographic degree words, after taking into account the variance contributed by several lexical variables known to influence language processing. These analyses replicated previous work showing that the presence of phonographic neighbors facilitates naming of visually presented words (Adelman & Brown, 2007; Peereman & Content, 1997) and extended to include lexical decision.

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

Chapter 7

General Discussion

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

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

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

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

7.1 Similarity Effects in Spoken and Visual Word Recognition

An intriguing aspect of the present findings is that phonographic degree facilitated visual word recognition but not spoken word recognition, and phonographic *C* facilitated spoken word recognition but not visual word recognition. This divergence may reflect differences in the way that written and spoken words are processed. A long-standing question within psycholinguistics is whether similarity among phonological and orthographic word forms facilitates or hinders word recognition. Below I briefly review the literature related to this issue and then discuss how considering the phonological and orthographic similarity among words as a single multiplex network structure could contribute to this debate.

7.1.1 Degree (neighborhood density)

In visual word recognition, the most commonly used measure of orthographic similarity is Coltheart's *N*, or orthographic neighborhood density (Coltheart et al., 1977), which is the number of words that can be obtained by substituting one letter from a target word of the same length. The presence of more orthographic neighbors facilitates recognition of the target word (for a review, see Andrews, 1997)—a finding that was contrary to predictions made by interactive-activation models of word recognition (McClelland & Rumelhart, 1981), which predicted that more neighbors would lead to greater interference when trying to recognize a target word. The nature of the influence of phonological neighborhood density or degree on spoken word recognition appears to be more consistent. Words with several phonological neighbors are more slowly and less accurately recognized as compared to words with few phonological neighbors (Goh, Suárez, Yap, & Tan, 2009; Luce & Pisoni, 1998). The presence of several phonological neighbors inhibits recognition of the target due to greater competition among activated neighbors.

As discussed in Chapters 1 and 2, past work has found that phonological neighborhood density influences visual word recognition and orthographic neighborhood density influence spoken word recognition. Specifically, in visual word recognition, the processing of words with many phonological neighbors is facilitated as compared to words with few phonological neighbors (Grainger et al., 2005; Yates et al., 2004). In spoken word recognition, the processing of words with many orthographic neighbors is also facilitated as compared to words with few orthographic neighbors (Muneaux & Ziegler, 2004; Ziegler et al., 2003).

7.1.2 Levenshtein distance

One limitation of Coltheart's *N* is that the measure does not consider words of different lengths into its calculation of similarity. According to this measure, 'cat' and 'hat' are orthographic neighbors, but not 'cat' and 'chat'. To overcome length restrictions of Coltheart's *N*, a different measure of orthographic similarity, orthographic Levenshtein distance, was developed by Yarkoni, Balota, and Yap (2008). Levenshtein distance refers to the number of substitutions, additions, or deletions (of letters or phonemes) required to transform one string to another (e.g., the Levenshtein distance between 'co**nd**i**t**ion' and 'co**ll**i**s**ion' is 3). To calculate a Levenshtein distance-based measure of orthographic similarity, Yarkoni et al. calculated orthographic Levenshtein distance-20 (OLD20), which is the average Levenshtein distance of the 20 "closest" orthographic neighbors of a target word. The phonological counterpart is the phonological Levenshtein distance-20 (PLD20), which is the average Levenshtein distance of the

20 "closest" phonological neighbors of a target word (Suárez, Tan, Yap, & Goh, 2011). A smaller Levenshtein distance indicates that several words are similar, or 'close' to, a given word (i.e., fewer substitutions, additions, or deletions (of letters or phonemes) required to transform one word to another). In visual word recognition, close OLD20 words were recognized more quickly than distant OLD20 words (Yarkoni et al., 2008), whereas in spoken word recognition, close PLD20 words were recognized more slowly than distant PLD20 words, indicating lexical competition (Suárez et al., 2011). Greater similarity (quantified as increased "closeness" to other words) facilitated visual word recognition but not spoken word recognition.

7.1.3 Clustering coefficient

The application of network science to model the structure of the mental lexicon (e.g., Vitevitch, 2008) has provided researchers with metrics to quantify the internal similarity structure of a word's neighborhood, namely clustering coefficient, *C*, which represents the level of interconnectivity among a word's neighbors. In spoken word recognition, high phonological *C* words are less quickly recognized as compared to low phonological *C* words (e.g., Chan & Vitevitch, 2010). A diffusion of activation framework was used to account for this finding—activation tends to be "trapped" within more densely clustered neighborhoods of words, making it harder for high *C*s word to stand out as compared to low *C*s words. On the other hand, in visual word recognition, high orthographic *C* words are more quickly recognized as compared to low orthographic *C* words (Siew, submitted). A similar framework was adopted to account for this finding—again, activation tends to be "trapped" within more densely clustered neighborhoods of words, however, the high overall sum of activation more readily pushes the system pass the recognition threshold, allowing high orthographic *C* words to be more readily retrieved (Note: this is analogous to the way the Multiple Read-Out Model accounts for visual lexical decision

RTs; Jonathan Grainger & Jacobs, 1996). To my knowledge no study has examined the influence of orthographic clustering coefficient on spoken word recognition, and only one study has examined influence of phonological clustering coefficient on visual word recognition—Yates (2013) found that high phonological *C* words were *less* quickly recognized as compared to low phonological *C* words in visual lexical decision.

Overall, across three different measures of similarity (degree/neighborhood density, Levenshtein distance, clustering coefficient), the results indicate that greater similarity among orthographic representations facilitates visual word recognition and greater similarity among phonological representations inhibits spoken word recognition. On the other hand, the pattern of results is less clear in studies that investigated the influence of phonological similarity on visual word recognition and the influence of orthographic similarity on spoken word recognition.

It is important to note that the dissertation differs from all of these previous studies in a fundamental way. The network metrics investigated in this dissertation simultaneously represented the phonological and orthographic similarity structure of language. However, in previous work, phonology and orthography were treated as separate influences to be manipulated or controlled for. This makes it difficult to assess the seemingly contradictory effects of phonological similarity on word recognition.

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

In contrast, the network metrics investigated in this paper, degree and clustering coefficient represent micro-level network measures of both phonological and orthographic similarity, with each metric representing different structural aspects of the phonographic neighborhood of a particular target word. Recall that degree simply refers to the number of phonographic neighbors, whereas clustering coefficient refers to the extent to which phonographic neighbors are also phonographic neighbors of each other. Hence, for a given word, degree can be viewed as a coarse, “brute-force” measure of the magnitude of similarity whereas clustering coefficient captures more subtle aspects of the internal similarity structure of the word’s phonographic neighborhood.

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

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

The general argument from the above study is that when processing is difficult or effortful in some way (such as low frequency words), it permits more time for additional influences (such as phonology) to come into play. A variant of this argument could be applied to explain the present set of findings: Processing spoken words, which are more ambiguous than written words due to the nature of auditory input, may permit more time for more subtle similarity effects such as *C* to develop and subsequently affect recognition. In other words, ambiguity in the bottom up signal may lead to greater sensitivity to nuances in the similarity space.

While these explanations are somewhat speculative in nature, they could be tested and investigated in future experimental work. For instance, while it would not be possible to present or process auditory stimuli in a "parallel" fashion, one could deliberately slow down visual word recognition (via some form of masking or degradation of the visual stimuli, or through the use of low frequency stimuli) to induce more effortful processing of the visual stimuli, and see if the manipulation reduces the influence of degree but increases the influence of *C* in visual word recognition. Such an approach has precedence in the field where researchers orthogonally manipulated variables of interest (e.g., stimulus quality, word frequency, neighborhood density) to examine the nature of interaction among lexical variables in order to distinguish between competing models of word recognition (Borowsky & Besner, 1993; Siew, Yap, & Goh, submitted; Yap & Balota, 2007).

Another way to further test these explanations is to conduct computer simulations. To investigate the clustering coefficient effect in spoken word recognition (Chan & Vitevitch, 2009), Vitevitch, Ercal, and Adagarla (2011) simulated the diffusion of activation among connected nodes in a network structure. In the initial time step, the target node arbitrarily received 100 units of activation. A portion of this initial activation was retained by the target word, and the remaining amount of activation was evenly spread among neighbors of the target word. In the next time step, each neighbor retained a portion of its activation and the remaining activation was spread to nodes to which it was connected to (including the target, other neighbors, and words not connected to the target word, etc.). This process was repeated for a number of time steps. The final activation value of the target was inversely mapped to response times and directly mapped to accuracy, such that higher activation values indicated that lexical retrieval occurred rapidly and more accurately. Using this simple model, Vitevitch et al. (2011) were able to account for the finding that low clustering coefficient words were more quickly and accurately recognized as compared to high clustering coefficient words (Chan & Vitevitch, 2009).

This simple diffusion model could be readily implemented in the phonographic multiplex to examine if the present findings can be replicated. Instead of simulating the diffusion of activation in a single layered network (as implemented in Vitevitch et al., 2011), activation will spread to connected nodes in *both* layers of the multiplex, with phonographic neighbors receiving a greater share of the activation as compared to only phonological or orthographic neighbors. In addition, the number of time steps which the simulation is allowed to run for can be manipulated to mimic differences in processing of visual and auditory information: The simulation is allowed to continue for a greater number of time steps for spoken word recognition (i.e., to permit more "time" for auditory information to unfold over time) than for visual word recognition (i.e., recognition occurs more immediately). High phonographic degree words and high phonographic *C* words are expected to have higher final activation values as compared to low phonographic degree words and low phonographic *C* words, with degree having a greater effect for simulations over few time steps (i.e., visual word recognition) and *C* having a greater effect for simulations over several time steps (i.e., spoken word recognition).

Overall, the take home message from this section is that the network science approach provides researchers with common metrics of similarity in phonology and orthography that can be used to gain a more holistic understanding of spoken and visual word recognition. Future work will continue to examine how metrics at various levels of the phonographic multiplex (e.g., macro-, meso-levels) influence spoken and visual word recognition.

7.2 Theoretical Implications for Models of Word Recognition

Several well-established theories have been put forth to account for multiple aspects of visual and spoken word recognition. The leading models of visual word recognition can be broadly classified into two groups: dual route models, which posit the presence of two distinct, independent pathways in visual word recognition (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler (2001)'s Dual Route Cascaded (DRC) model) and parallel, distributed models, which consist of orthographic units, phonological units, and a set of hidden units that interface between the orthographic and phonological units (e.g., Seidenberg & McClelland (1989)'s Parallel Distributed Processing (PDP) model). The cohort model (Marslen-Wilson, 1987), TRACE (McClelland & Elman, 1986), Shortlist B (Norris & McQueen, 2008), and Neighborhood Activation Model (Luce & Pisoni, 1998) represent the prominent models of spoken word recognition. As it is beyond the scope of the dissertation to provide a detailed discussion of each of the above models, I will focus on the Seidenberg and McClelland (1989) PDP model of visual word recognition and pronunciation as an example, and consider how it may or may not be able to account for the phonographic effects observed in the present studies.

The PDP model consists of orthographic units, phonological units, and a set of hidden units that interface between the orthographic and phonological units. One key feature of distributed models is the ability of the model to "learn"—and thereby approximate the language acquisition process in children—by modifying connection weights between units via a back propagation learning algorithm during training (Seidenberg & McClelland, 1989). In a connectionist model, the relative influence of orthography and phonology on lexical retrieval depends on the extent to which orthographic and phonological codes overlap (Harm & Seidenberg, 2004). A greater amount of overlap in orthography and phonology would be expected to speed up processing and lead to faster access to a word's meaning (Harm &

Seidenberg, 2004); this is consistent with the present finding that phonographic similarity generally facilitates processing in both visual and spoken word recognition. However, it is not entirely clear how more subtle aspects of the phonographic similarity structure (i.e., the level of interconnectivity among similar words) would be implemented in the model. Perhaps a new theoretical framework, one that explicitly considers the linguistic structure of language, is necessary to account for the present findings and advance the field.

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

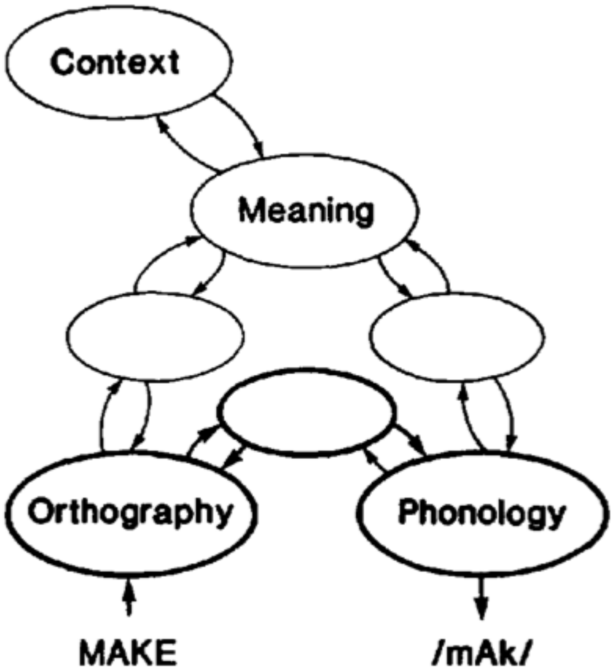

Figure 3. The general framework of the Parallel Distributed Processing model (Seidenberg & McClelland, 1989). Each ellipse represents a set of primitive units, with empty ellipses representing hidden units.

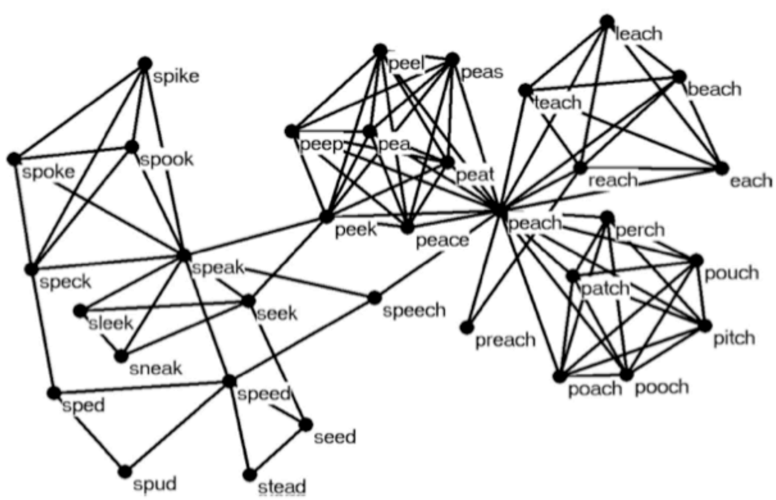

Figure 1. A portion of the phonological network examined in Vitevitch (2008). Depicted are the word *speech*, phonological neighbors of the word *speech*, and the phonological neighbors of those neighbors.

Figure 4. A section of the phonological network of language showing the word 'speech', its phonological neighbors, and the phonological neighbors of its phonological neighbors. Links are placed between words that differ by the substitution, addition, or deletion of 1 phoneme. From Vitevitch (2008).

The connectionist approach as exemplified by the PDP appears to be computationally expensive given the number of primitive units represented in the model—indeed, Coltheart, Curtis, Atkins and Haller (1993) point out that Rumelhart and McClelland (1986) had to drastically reduce the number of output units from over 1,000 units to about 400 units to make the simulations tenable (no clear rationale was provided otherwise). The PDP model is also quite complicated and values had to be arbitrarily set for a number of parameters, such as the level of excitation for feature-to-letter connections (.005), level of inhibition for letter-to-word connections (.04), among others (McClelland & Rumelhart, 1981). The network science approach, despite being based on simple assumptions, reveals a complex language network structure whereby a simple diffusion of activation mechanism can be implemented in order to account for behavioral findings such as the clustering coefficient effect (Vitevitch et al., 2011). Simulations conducted by Chan and Vitevitch (2009) using jTRACE, the computational implementation of the TRACE model of speech perception (Strauss, Harris, & Magnuson, 2007), were unable to account for the clustering coefficient effect. Furthermore, even though sublexical units (such as letters or phonemes) are not explicitly represented as individual nodes or entities within the language network one could potentially account for both lexical and sublexical effects by examining the language network at differing levels of the network (micro-, meso-, macro-). For instance, Siew (2013) speculated that (sublexical) phonotactic probability effects could emerge at the meso-level (community) structure of the network and (lexical) neighborhood density effects could arise at the micro-level of the network.

Without explicitly considering how the *overall* phonological and orthographic similarity structure of language affects lexical retrieval, it is unclear how current models of word recognition would be able to account for the present findings. In addition, models of spoken word recognition do not consider the role of orthographic information on speech processing and would not predict orthographic effects in spoken word recognition in the first place. Certainly future work will need to address whether these models can indeed replicate the present findings of an effect of phonographic degree on visual word recognition and an effect of phonographic clustering coefficient on spoken word recognition via computer simulations. Nevertheless the present findings suggest that models of word recognition will need to explicitly consider how the cognitive processes that underlying lexical retrieval operate within a complex language structure as represented by the phonographic multiplex.

7.3 Limitations and Future Directions

In this section, I briefly point out some limitations of the present work and suggest some future directions. First, the phonographic network consists of approximately 6,000 words, and this dissertation specifically focused on analyzing and investigating words from the giant component of the phonographic network (~3000 words). This represents a smaller set of words as compared to the 19,340 words used to construct the original phonological network in Vitevitch (2008). In addition, the 160 word stimuli selected for Experiments 1 and 2 were mostly short, monosyllabic words that also tended to be feedforward and feedback consistent. Therefore, this might lead one to question if the network science approach and the multiplex metrics described in previous chapters are indeed useful and applicable to a wider range and variety of words.

However, I would argue that the metrics are indeed applicable to a broader variety of words than implied in the stimuli lists for Experiments 1 and 2. The giant component of the phonographic network consisted of a greater variety of words other than the consistent, monosyllabic CVC words that are typically studied in the literature—this included words such as 'stub' and 'grub' (CCVC structure), 'dreary' and 'heavy' (2 syllables), and 'scrimp' and 'whisky' (6 phonemes). Various network metrics such as degree and clustering coefficient can be calculated for these words. Indeed, the regression analyses of the ELP described in Chapter 6 consisted of over 2,000 words, several of which were not consistent, monosyllabic CVC words.

It is also important to re-emphasize that this dissertation represents a first step in a future line of research on how the structure of the phonographic multiplex influences spoken and visual word recognition. Future work could investigate how the structure of the phonographic multiplex at various levels (macro-, meso-, micro-) influence processing and begin to look at other areas of the multiplex beyond the phonographic network. For instance, future investigations could include longer, less frequent, multisyllabic words that are not found in the giant component of the phonographic network. These words tend to be "hermits" as they do not have any phonological or orthographic neighbors. However, words that are hermits in one layer of the multiplex might be found in connected network components (i.e., lexical islands) in the other layer (i.e., phonological hermits with orthographic neighbors, and orthographic hermits with phonological neighbors) and one could investigate whether the presence of connections/neighbors in one modality might facilitate processing in the other modality. By adopting a network science approach to study the interrelationship between phonology and orthographic, language researchers can begin to address important questions regarding the role of phonographic similarity on spoken and visual word recognition with respect to words that are not typically studied in the previous literature.

7.4 Practical Implications and Applications

Characterizing the phonological and orthographic similarity structure of language as a multiplex could have important applications for understanding (i) language disorders related to reading and writing and (ii) first and second language acquisition. In this final section, I speculate the various ways in which the network science approach can contribute towards the current research on dyslexia and shed new light on language learning.

7.4.1 Language Disorders

Dyslexia is a language-based learning disability characterized by marked difficulties in spelling words, writing words, reading aloud, or understanding what was read (Goswami, 2000). Given that the main deficit of individuals with dyslexia is a slow and inefficient grapheme-phoneme recoding process, researchers have focused on studying the factors that contribute toward the inefficiency of this process (Ziegler & Goswami, 2005). Recall, however, that a major tenet of complex systems and network science is that in order to achieve a fuller understanding of any process, it is important to consider the structure of the system that the process occurs in. As shown in this dissertation, the structure of language networks has important, measurable influences on lexical processes and mechanisms. Researchers should explicitly consider how both the structures of the phonological and orthographic language networks (as represented in the phonographic multiplex) might influence grapheme-to-phoneme recoding processes.

Researchers have begun using network science measures to analyze language networks of individuals with language disorders or delays and comparing these against the language networks of typically developing individuals (Beckage, Smith, & Hills, 2011). Below I describe one example of this approach and discuss how this method could be applied to compare the phonographic multiplexes of individuals with and without dyslexia.

Beckage, Smith, and Hills (2011) analyzed the semantic language networks of typically developing children and “late talkers”, that is, children who start talking later and who also have a slower rate of vocabulary growth (Moyle, Weismer, Evans, & Lindstrom, 2007). After controlling for vocabulary size, Beckage and colleagues found that the network measures of semantic networks of typical children indicate a higher level of lexical connectivity, as compared to the semantic networks of late talkers. Based on these results, Beckage et al. suggested that late talkers might have a maladaptive bias to acquire particularly novel new words that are not semantically related to words that they already know, leading to a semantic network that is lacking in the “small-world” network structure that is important for facilitating lexical retrieval.

This study is a noteworthy example of how studying the underlying structure of the mental lexicon can lead to important insights into language disorders and/or delays—insights that would not have been possible without using the network science approach to represent and characterize the structure of the mental lexicon. Furthermore, network science can offer new ways of identifying children with language delays through the application of more sensitive, nuanced measures that take into account the structural cohesiveness of the child’s mental lexicon (e.g., level of connectivity among known word forms), which could represent an improvement over the “blunt”, conventional approaches used to identify children with language delays (e.g., how many words does the child know). With respect to dyslexia, it is possible that the pronounced difficulties that individuals with dyslexia face with spelling and reading are due to a reduced overall connectivity of the orthographic layer of the phonographic multiplex, or due to greater mismatch between the structural patterns of the phonological and orthographic layers of the multiplex (as compared to the phonographic multiplexes of typically developing individuals). Various macro-level network metrics (e.g., average path length, average clustering coefficient) could be used to quantify the structural differences between the phonographic multiplexes of individuals with and without dyslexia.

Should structural differences exist between the phonographic multiplexes of individuals with and without dyslexia, this suggests that reading difficulties experienced by individuals with dyslexia might be due to the fact that phonological recoding processes are occurring within a phonographic multiplex that is inefficiently organized in the first place. Furthermore, the network science approach could inspire new ways of helping these individuals. One possibility is to conduct a community detection analysis to uncover the community structure of words in the multiplex. This could lead to the discovery of words that reside in different (i.e., non-overlapping) communities in the phonological and orthographic layers of the phonographic multiplex. The orthographic representations of these words might be more challenging to acquire because they have dissimilar orthographic and phonological structures. Dyslexic learners may require additional support when learning the orthographic forms of these words because the phonological structure of these words may fail to provide the necessary scaffold for the development of the corresponding orthographic structure.

7.4.2 First and Second Language Acquisition

Learning a second language (L2) is very different from acquiring one's native language (L1) in various ways, particularly with respect to the time course in the development of the phonological and orthographic structure of the language. When acquiring one's native language, phonological representations are acquired first, and then orthographic knowledge is acquired when one learns to read. On the other hand, second language learners generally acquire the spoken and written representations of new words at the same time, especially with the prevalence of web-based language learning applications (Golonka, Bowles, Frank, Richardson, & Freynik, 2014).

Several have argued that learning the orthographic representations of words reorganizes and restructures the L1 phonological system (Burnham, 2003; Perre, Pattamadilok, Montant, & Ziegler, 2009), leading to the co-activation of orthography and phonology in language processing (e.g., Ziegler & Ferrand, 1998) and changes in the functional organization of the brain (Castro-Caldas, Petersson, Reis, Stone-Elander, & Ingvar, 1998). In contrast, such "re-structuring" does not occur for second language acquisition as the phonological and orthographic systems develop at the same time.

However, in addition to orthographic knowledge "restructuring" the phonological system over the course of language development (Burnham, 2003), it is possible that the pre-existing structure of the phonological system scaffolds the development of orthographic knowledge. As for second language acquisition, there are certainly bidirectional influences of phonology and orthography on the development of each system. Furthermore, perhaps some of the difficulty in learning a second language may be due to learners having to assemble the L2 phonological-orthographic multiplex from scratch, or due to interference from their L1 network (L1 orthographic transfer; see Sun-Alperin & Wang, 2011; Vokic, 2011). These theoretically important questions regarding the nature of language acquisition and learning could be addressed via the network science approach. By conceptualizing the phonological and orthographic systems of language as the phonological and orthographic layers in a phonographic multiplex, network science can provide computational tools and methods to quantify the phonological and orthographic structures, track how these structures interact and develop over time, and how the development of the phonographic multiplex differs when acquiring a L1 or L2.

The network science approach can not only provide us with insights into the ways first and second language acquisition differ, but also point to new ways of maximizing the success of second language instruction. For instance, the approach can be used to identify areas of the phonographic multiplex where L2 learners might particularly struggle with (e.g., words with little overlap in phonological and orthographic structures of the L2 multiplex), or identify an optimal sequence of words to be taught that is likely to lead to the development of an efficiently, well-connected multiplex structure.

7.5 Conclusion

This dissertation uses the tools of Network Science to simultaneously characterize the phonological as well as orthographic similarity structure of words in English within a phonographic multiplex. Specifically, it focuses on the section of the phonographic multiplex known as the phonographic network of language, where links are placed between words that are both phonologically and orthographically similar to each other. Conventional psycholinguistic experiments and an archival analysis of the English Lexicon Project were conducted to investigate the influence of phonographic degree and phonographic clustering coefficient—network science metrics derived from the phonographic network—on spoken and visual word recognition. Overall, results indicated a facilitatory effect of phonographic degree on visual word recognition, and a facilitatory effect of phonographic clustering coefficient on spoken word recognition. The present findings showed that greater phonographic similarity facilitates word recognition in both auditory and visual modalities, and have important implications for theoretical models of spoken and visual word recognition, as well as for increasing our understanding of language disorders and language acquisition. Ultimately, this dissertation demonstrates how simultaneously representing the phonological and orthographic similarity of words within a phonographic multiplex can lead to a deeper understanding of how similarity influences spoken and visual word recognition.

References

Adelman, J. S., & Brown, G. D. (2007). Phonographic neighbors, not orthographic neighbors, determine word naming latencies. *Psychonomic Bulletin & Review*, *14*(3), 455–459.

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, *74*(1), 47–97. https://doi.org/10.1103/RevModPhys.74.47

Albert, R., Jeong, H., & Barabási, A.-L. (1999). Internet: Diameter of the World-Wide Web. *Nature*, *401*(6749), 130–131. https://doi.org/10.1038/43601

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory & Cognition*, *10*(6), 565–575.

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

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

Balota, D. A., Yap, M. J., Hutchison, K. A., & Cortese, M. J. (2012). 5 Megastudies. *Visual Word Recognition Volume 1: Models and Methods, Orthography and Phonology*, 90.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., … Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, *39*(3), 445–459.

Barabási, A.-L. (2009). Scale-free networks: A decade and beyond. *Science*, *325*(5939), 412–413. https://doi.org/10.1126/science.1173299

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*(5439), 509–512.

Bargigli, L., Iasio, G. di, Infante, L., Lillo, F., & Pierobon, F. (2015). The multiplex structure of interbank networks. *Quantitative Finance*, *15*(4), 673–691. https://doi.org/10.1080/14697688.2014.968356

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., & others. (2015). lme4: Linear mixed-effects models using Eigen and S4, 2014. *R Package Version*, 1–1.

Battiston, F., Nicosia, V., & Latora, V. (2014). Structural measures for multiplex networks. *Physical Review E*, *89*(3), 032804. https://doi.org/10.1103/PhysRevE.89.032804

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

Besner, D., & Davelaar, E. (1983). Suedohomofoan effects in visual word recognition: Evidence for phonological processing. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *37*(2), 300–305. https://doi.org/10.1037/h0080719

Bianconi, G. (2013). Statistical mechanics of multiplex networks: Entropy and overlap. *Physical Review E*, *87*(6), 062806. https://doi.org/10.1103/PhysRevE.87.062806

Boccaletti, S., Bianconi, G., Criado, R., Del Genio, C. I., Gómez-Gardeñes, J., Romance, M., … Zanin, M. (2014). The structure and dynamics of multilayer networks. *Physics Reports*, *544*(1), 1–122. https://doi.org/10.1016/j.physrep.2014.07.001

Borowsky, R., & Besner, D. (1993). Visual word recognition: a multistage activation model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(4), 813.

Braun, M., Hutzler, F., Ziegler, J. C., Dambacher, M., & Jacobs, A. M. (2009). Pseudohomophone effects provide evidence of early lexico-phonological processing in visual word recognition. *Human Brain Mapping*, *30*(7), 1977–1989. https://doi.org/10.1002/hbm.20643

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.

Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, *10*(3), 186–198. https://doi.org/10.1038/nrn2575

Burnham, D. (2003). Language specific speech perception and the onset of reading. *Reading and Writing*, *16*(6), 573–609.

Cardillo, A., Zanin, M., Gómez-Gardeñes, J., Romance, M., Amo, A. J. G. del, & Boccaletti, S. (2013). Modeling the multi-layer nature of the European Air Transport Network: Resilience and passengers re-scheduling under random failures. *The European Physical Journal Special Topics*, *215*(1), 23–33. https://doi.org/10.1140/epjst/e2013-01712-8

Castro-Caldas, A., Petersson, K. M., Reis, A., Stone-Elander, S., & Ingvar, M. (1998). The illiterate brain: Learning to read and write during childhood influences the functional organization of the adult brain. *Brain*, *121*, 1053–1063.

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

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

Chéreau, C., Gaskell, M. G., & Dumay, N. (2007). Reading spoken words: Orthographic effects in auditory priming. *Cognition*, *102*(3), 341–360. https://doi.org/10.1016/j.cognition.2006.01.001

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100*(4), 589.

Coltheart, M., Davelaar, E., Jonasson, T., & Besner, D. (1977). Access to the internal lexicon. Lawrence Erlbaum Associates.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204.

Cortese, M. J., & Simpson, G. B. (2000). Regularity effects in word naming: What are they? *Memory & Cognition*, *28*(8), 1269–1276. https://doi.org/10.3758/BF03211827

D'Souza, R. M., Borgs, C., Chayes, J. T., Berger, N., & Kleinberg, R. D. (2007). Emergence of tempered preferential attachment from optimization. *Proceedings of the National Academy of Sciences, 104*(15), 6112–6117.

De Domenico, M., Porter, M. A., & Arenas, A. (2015). MuxViz: a tool for multilayer analysis and visualization of networks. *Journal of Complex Networks*, *3*(2), 159–176. https://doi.org/10.1093/comnet/cnu038

Dich, N. (2011). Individual differences in the size of orthographic effects in spoken word recognition: The role of listeners' orthographic skills. *Applied Psycholinguistics*, *32*(1), 169–186. https://doi.org/10.1017/S0142716410000330

Eager, C., & Roy, J. (2017). Mixed Effects Models are Sometimes Terrible. *arXiv Preprint arXiv:1701.04858*. Retrieved from https://arxiv.org/abs/1701.04858

Erdös, P., & Rényi, A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, *5*(17–61), 43.

Ferrand, L., & Grainger, J. (1992). Phonology and orthography in visual word recognition: Evidence from masked non-word priming. *The Quarterly Journal of Experimental Psychology Section A*, *45*(3), 353–372. https://doi.org/10.1080/02724989208250619

Ferrand, L., & Grainger, J. (2003). Homophone interference effects in visual word recognition. *The Quarterly Journal of Experimental Psychology: Section A*, *56*(3), 403–419. https://doi.org/10.1080/02724980244000422

Ferrer i Cancho, R., & Solé, R. V. (2003). Least effort and the origins of scaling in human language. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(3), 788–91.

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

Frauenfelder, U. H., Baayen, R. H., & Hellwig, F. M. (1993). Neighborhood density and frequency across languages and modalities. *Journal of Memory and Language*, *32*(6), 781.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin*, *123*(1), 71–99. https://doi.org/10.1037/0033-2909.123.1.71

Goh, W. D., Suárez, L., Yap, M. J., & Tan, S. H. (2009). Distributional analyses in auditory lexical decision: Neighborhood density and word-frequency effects. *Psychonomic Bulletin & Review*, *16*(5), 882–887.

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

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer Assisted Language Learning*, *27*(1), 70–105.

Goswami, U. (2000). Phonological representations, reading development and dyslexia: Towards a cross-linguistic theoretical framework. *Dyslexia*, *6*(2), 133–151.

Grainger, J., & Ferrand, L. (1994). Phonology and orthography in visual word recognition: Effects of masked homophone primes. *Journal of Memory and Language*, *33*(2), 218–233. https://doi.org/10.1006/jmla.1994.1011

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: a multiple read-out model. *Psychological Review*, *103*(3), 518.

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

Greenberg, J. H., & Jenkins, J. J. (1964). Studies in the psychological correlates of the sound system of American English. *Word*, *20*(2), 157–177. https://doi.org/10.1080/00437956.1964.11659816

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, *111*(3), 662.

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009). Longitudinal analysis of early semantic networks. *Psychological Science*, *20*(6), 729–739. https://doi.org/10.1111/j.1467-9280.2009.02365.x

Iyengar, S. R. S., Madhavan, C. E. V., Zweig, K. A., & Natarajan, A. (2012). Understanding human navigation using network analysis. *Topics in Cognitive Science*, *4*(1), 121–134. https://doi.org/10.1111/j.1756-8765.2011.01178.x

Jackson, M. O., & Rogers, B. W. (2007). Meeting strangers and friends of friends: How random are social networks? *The American Economic Review*, *97*(3), 890–915.

Jared, D. (1997). Spelling–sound consistency affects the naming of high-frequency words. *Journal of Memory and Language*, *36*(4), 505–529. https://doi.org/10.1006/jmla.1997.2496

Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language*, *29*(6), 687–715. https://doi.org/10.1016/0749-596X(90)90044-Z

Katz, L., & Frost, R. (1992). The Reading Process is Different for Different Orthographies: The Orthographic Depth Hypothesis. *Advances in Psychology*, *94*, 67–84. https://doi.org/10.1016/S0166-4115(08)62789-2

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

Kessler, B., & Treiman, R. (1997). Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language*, *37*(3), 295–311. https://doi.org/10.1006/jmla.1997.2522

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. https://doi.org/10.1093/comnet/cnu016

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

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Dartmouth Publishing Group.

Lancichinetti, A., Kivelä, M., Saramäki, J., & Fortunato, S. (2010). Characterizing the community structure of complex networks. *PloS One*, *5*(8), e11976.

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

Lewis, K., Kaufman, J., Gonzalez, M., Wimmer, A., & Christakis, N. (2008). Tastes, ties, and time: A new social network dataset using Facebook.com. *Social Networks*, *30*(4), 330–342. https://doi.org/10.1016/j.socnet.2008.07.002

Liberman, A. M. (1992). The relation of speech to reading and writing. *Advances in Psychology*, *94*, 167–178.

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.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1), 71–102.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, *88*(5), 375.

Menichetti, G., Remondini, D., Panzarasa, P., Mondragón, R., & Bianconi, G. (2014). Weighted multiplex networks. *PLoS One*, *9*(6). https://doi.org/10.1371/journal.pone.0097857

Mitzenmacher, M. (2004). A brief history of generative models for power law and lognormal distributions. *Internet Mathematics*, *1*(2), 226–251.

Moyle, M. J., Weismer, S. E., Evans, J. L., & Lindstrom, M. J. (2007). Longitudinal relationships between lexical and grammatical development in typical and late-talking children. *Journal of Speech, Language, and Hearing Research*, *50*(2), 508–528.

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

National Institute of Neurological Disorders and Stroke. (2015). Retrieved April 25, 2016, from http://www.ninds.nih.gov/disorders/dyslexia/dyslexia.htm

New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review*, *13*(1), 45–52. https://doi.org/10.3758/BF03193811

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

Newman, M. E. J. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(2), 404–409.

Newman, M. E. J. (2002). Assortative mixing in networks. *Physical Review Letters*, *89*(20). https://doi.org/10.1103/PhysRevLett.89.208701

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

Newman, M. E. J. (2004b). Fast algorithm for detecting community structure in networks. *Physical Review E*, *69*(6), 066133.

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

Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, *115*(2), 357.

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

Onnela, J.-P., Fenn, D. J., Reid, S., Porter, M. A., Mucha, P. J., Fricker, M. D., & Jones, N. S. (2012). Taxonomies of networks from community structure. *Physical Review E*, *86*(3), 036104.

Opsahl, T., & Panzarasa, P. (2009). Clustering in weighted networks. *Social Networks*, *31*(2), 155–163.

Pattamadilok, C., Lafontaine, H., Morais, J., & Kolinsky, R. (2010). Auditory word serial recall benefits from orthographic dissimilarity. *Language and Speech*, *53*(3), 321–341. https://doi.org/10.1177/0023830910371450

Peereman, R., & Content, A. (1997). Orthographic and Phonological Neighborhoods in Naming: Not All Neighbors Are Equally Influential in Orthographic Space. *Journal of Memory and Language*, *37*(3), 382–410. https://doi.org/10.1006/jmla.1997.2516

Peereman, R., Dufour, S., & Burt, J. S. (2009). Orthographic influences in spoken word recognition: The consistency effect in semantic and gender categorization tasks. *Psychonomic Bulletin & Review*, *16*(2), 363–368. https://doi.org/10.3758/PBR.16.2.363

Perre, L., Midgley, K., & Ziegler, J. C. (2009). When beef primes reef more than leaf: Orthographic information affects phonological priming in spoken word recognition. *Psychophysiology*, *46*(4), 739–746. https://doi.org/10.1111/j.1469-8986.2009.00813.x

Perre, L., Pattamadilok, C., Montant, M., & Ziegler, J. C. (2009). Orthographic effects in spoken language: on-line activation or phonological restructuring? *Brain Research*, *1275*, 73–80.

Roux, S., & Bonin, P. (2013). "With a little help from my friends": Orthographic influences in spoken word recognition. *L'Année Psychologique*, *113*(01), 35–48. https://doi.org/10.4074/S0003503313001024

Salverda, A. P., & Tanenhaus, M. K. (2010). Tracking the time course of orthographic information in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(5), 1108–1117. https://doi.org/10.1037/a0019901

Seidenberg, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition*, *19*(1), 1–30.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*(4), 523.

Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(6), 546–554. https://doi.org/10.1037/0278-7393.5.6.546

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior*, *23*(3), 383–404.

Siew, C. S. Q. (2016). The influence of 2-hop network density on spoken word recognition. *Psychonomic Bulletin & Review*, 1–7.

Siew, C. S. Q. (2013). Community structure in the phonological network. *Frontiers in Psychology*, *4*, 553.

Siew, C. S. Q. (submitted). The influence of orthographic clustering coefficient on visual word recognition.

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

Siew, C. S. Q., Yap, M. J., & Goh, W. D. (submitted). Joint effects of word-frequency and stimulus quality in auditory lexical decision.

Slowiaczek, L. M., Soltano, E. G., Wieting, S. J., & Bishop, K. L. (2003). An investigation of phonology and orthography in spoken-word recognition. *The Quarterly Journal of Experimental Psychology Section A*, *56*(2), 233–262. https://doi.org/10.1080/02724980244000323

Soffer, S. N., & Vazquez, A. (2005). Network clustering coefficient without degree-correlation biases. *Physical Review E*, *71*(5), 057101.

Solé, R. V., Corominas-Murtra, B., Valverde, S., & Steels, L. (2010). Language networks: Their structure, function, and evolution. *Complexity*, *15*(6), 20–26. https://doi.org/10.1002/cplx.20305

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

Stone, G. O., & Van Orden, G. C. (1994). Building a resonance framework for word recognition using design and system principles. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(6), 1248–1268. https://doi.org/10.1037/0096-1523.20.6.1248

Stone, G. O., Vanhoy, M., & Van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language*, *36*(3), 337–359. https://doi.org/10.1006/jmla.1996.2487

Strauss, T. J., Harris, H. D., & Magnuson, J. S. (2007). jTRACE: A reimplementation and extension of the TRACE model of speech perception and spoken word recognition. *Behavior Research Methods*, *39*(1), 19–30. https://doi.org/10.3758/BF03192840

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

Suárez, L., Tan, S. H., Yap, M. J., & Goh, W. D. (2011). Observing neighborhood effects without neighbors. *Psychonomic Bulletin & Review*, *18*(3), 605–611.

Sun-Alperin, M. K., & Wang, M. (2011). Cross-language transfer of phonological and orthographic processing skills from Spanish L1 to English L2. *Reading and Writing*, *24*(5), 591–614.

Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory and Language, 26*(6), 608-631.

Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K. S., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities*, *34*(1), 33–58. https://doi.org/10.1177/002221940103400104

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General*, *124*(2), 107–136. https://doi.org/10.1037/0096-3445.124.2.107

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition*, *15*(3), 181–198. https://doi.org/10.3758/BF03197716

Venezky, R. L. (1970). *The Structure of English Orthography*. Walter de Gruyter.

Vitevitch, M. S. (2008). What can graph theory tell us about word learning and lexical retrieval? *Journal of Speech, Language, and Hearing Research : JSLHR*, *51*(2), 408–422. https://doi.org/10.1044/1092-4388(2008/030)

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

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

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

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

Vitevitch, M. S., & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods*, *36*(3), 481–487.

Vokic, G. (2011). When alphabets collide: Alphabetic first-language speakers' approach to speech production in an alphabetic second language. *Second Language Research*, *27*(3), 391–417.

Watts, D. J. (2004). The "new" science of networks. *Annual Review of Sociology*, *30*(1), 243–270. https://doi.org/10.1146/annurev.soc.30.020404.104342

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

Wilson, E. O. (1998). *Consilience: The unity of knowledge* (1st ed..). New York: Knopf : Distributed by Random House.

Wu, K., Taki, Y., Sato, K., Sassa, Y., Inoue, K., Goto, R., … others. (2011). The overlapping community structure of structural brain network in young healthy individuals. *PLoS One*, *6*(5), e19608.

Yap, M. J., & Balota, D. A. (2007). Additive and interactive effects on response time distributions in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(2), 274.

Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language*, *60*(4), 502–529.

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, *15*(5), 971–979.

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. https://doi.org/10.1037/0278-7393.31.6.1385

Yates, M. (2013). How the clustering of phonological neighbors affects visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(5), 1649.

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

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review*, *5*(4), 683–689. https://doi.org/10.3758/BF03208845

Ziegler, J. C., Ferrand, L., & Montant, M. (2004). Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. *Memory & Cognition*, *32*(5), 732–741. https://doi.org/10.3758/BF03195863

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3.

Ziegler, J. C., Montant, M., & Jacobs, A. M. (1997). The feedback consistency effect in lexical decision and naming. *Journal of Memory and Language*, *37*(4), 533–554. https://doi.org/10.1006/jmla.1997.2525

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

Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology.* Cambridge, Mass.: Addison-Wesley Press.

Zou, L., Desroches, A. S., Liu, Y., Xia, Z., & Shu, H. (2012). Orthographic facilitation in Chinese spoken word recognition: An ERP study. *Brain and Language*, *123*(3), 164–173. https://doi.org/10.1016/j.bandl.2012.09.006

Appendix A
List of word stimuli used in Experiments 1 and 2.

Phonographic degree words		Phonographic *C* words	
High degree	Low degree	High *C*	Low *C*
bloat	balm	balk	bleed
brace	blame	bland	blob
brew	blow	bleak	bread
chart	brood	blown	brink
clip	brook	boss	broom
deck	charm	bride	cave
draft	chase	clot	clod
drag	clean	count	drive
drew	cleat	crime	drove
drip	clove	dream	duct
flake	clump	drown	flat
flew	cue	duke	hack
flick	doom	dwell	hawk
float	dorm	flip	hive
flush	dread	flop	hoop
gripe	drum	haze	lab
gum	fled	hook	lobe
gust	food	husk	mile
hulk	grab	loaf	moat
hurl	halt	lunch	mount
loud	hurt	lung	nose
moist	limb	mean	pluck
mood	plea	paid	plum
pleat	porch	posh	plume
scoop	range	pulp	raft
shame	roof	pump	ramp

shock	scab	rack	ripe
slid	scan	reef	rose
slum	scorn	rope	run
slump	sheep	save	side
spool	short	skin	slap
spunk	shout	slam	slur
stall	smack	snip	snap
steep	spear	spice	space
stub	stark	stock	swell
swim	start	tab	tomb
swing	swap	tile	tool
swoop	sweep	tote	trail
trust	wand	trace	tread
weep	yarn	truce	wheat

Appendix B
List of nonword stimuli used in Experiment 2.

Nonword foils for phonographic degree words		Nonword foils for phonographic *C* words	
bles	bɪm	blɪk	blaɪd
blet	blɛ	blɝn	bɹab
bɹɛ	blʊk	blund	bɹɔɪm
tʃeɹt	bɹem	blaɪd	bɹʌd
dɔk	bɹɪd	bɔɪk	bɹʌŋk
dɹɛft	tʃis	bos	dɑkt
dɹɪg	tʃoɹm	dek	dɹɑv
dɹop	dɔɪm	dɹom	dɹɛv
dɹaʊ	dɹem	dɹaɪn	flet
floʃ	dɹʌd	dwɑl	hik
flʊk	dʊɹm	flup	hok
flʌt	flud	fɹɑp	hov
flɪk	fʊd	hɑsk	hʊp
fɹu	glæb	hɑz	hwɑt
gæm	hɛt	hek	kɛv
gɑst	hilt	klʊt	kɹɑd
gɹʌp	klɔɪmp	kɔɪnt	lɛb
hɔɪl	klut	kɹaʊm	lʊb
holk	klʊv	lɑŋ	mont
klup	kɹin	lef	mɔɪt
lʌd	kjɑ	lʊntʃ	maʊl
mɝd	lom	mɔn	nɪz
mʊst	piɹtʃ	pɑmp	plɛm
plut	plʊ	peʃ	pɹum
ʃem	ɹæf	pilp	pɹʌk
skɔɪp	ɹinʤ	pɔɪd	ɹemp

slimp	ʃɑɹt	ɹof	ɹɪn
slom	skɪb	ɹuk	ɹɪz
slaʊd	skʊn	ɹaʊp	ɹoft
ʃok	sloɹn	skɔɪn	ɹʊp
spæl	smɛk	slɑk	sles
spɪŋk	ʃɔɪt	slem	slɔɪ
spʌb	spaʊɹ	snop	snɝp
stop	steɹt	sov	stɑp
stʊl	stiɹk	spos	stɛl
swɔŋ	ʃup	tɛt	saʊd
swɝp	swɝp	tɪb	tɛb
swum	swʌp	tɹɑs	tɹɛl
tɹɔɪst	wɔɪnd	tɹɪs	tɹʌd
wæp	joɹn	tɝl	tʌl