languages

MDPI

Review

Organizational and Formational Structures of Networks in the Mental Lexicon: A State-Of-The-Art through Systematic Review

Luke McCarthy [1,*] and Imma Miralpeix [2]

1 Department of Languages and Literature of Europe and Americas, University of Hawaii at Manoa, Honolulu, HI 96822, USA
2 Department of English, Universitat de Barcelona, 08007 Barcelona, Spain; imiralpeix@ub.edu
* Correspondence: luke603@hawaii.edu

Received: 18 July 2019; Accepted: 13 December 2019; Published: 24 December 2019

Abstract: This state-of-the-art presents a systematic exploration on the use of network patterns in global research efforts to understand, organize and represent the mental lexicon. Results have shown an increase over recent years in the usage of complex, small-world and scale-free network patterns within the literature. With the increasing complexity of network patterns, we see more potential in the inter-disciplinary exploration of the mental lexicon through universal and mathematically-describable, behavioral patterns in small-world and scale-free networks. A systematic review of 36 items of methodologically-selected literature serve as a means to explore how the greater literary body understands network structures within the mental lexicon. Network-based approaches are discriminated between three contrasting varieties. These include: 'simple networks', characterized by arbitrarily organized graph patterns of metaphorical importance; 'connectionist networks', a broad category of networks which explore the structural features of a system through the analysis of emergent properties; and lastly 'complex networks', distinguished as small-world, scale-free networks which follow a strict and mathematically-describable structure in agreement with the Barabási–Albert model. Each network approach is explored in terms of their discernible differences which relate to their parameters and affect their implications. A final evaluation of observed patterns within the selected literature is offered, as well as an elaboration on the sense of trajectory beheld in the research in order to offer insight and orientation for future research.

Keywords: mental lexicon; simple networks; connectionist networks; complex networks; vocabulary

1. Aims and Structures

Throughout the past half century, the complexities of the organizational and formational properties of the mental lexicon have been rigorously explored and challenged by a variety of linguists from diverse backgrounds, utilizing a wide array of approaches (Miller 1986; Elman 2004). With such progress, the field may benefit from a push to evaluate and mature our understanding of this intricate question further. Over the past several decades, numerous researchers have expatiated on the budding use of graphs and consequential network models, which seek to elaborate on the characteristics of structures governing the mental lexicon. The aim of this research is to execute a systematic exploration and structured review in order to present an accurate representation and evaluation of where global research efforts have led us in terms of how linguists understand, organize and represent the mental lexicon, as it relates to networks, and to offer some orientation for further perspectives. To do this, the following aims guide our research:

(1) To identify the types of network-based approaches utilized in research on the organization and formation of the mental lexicon.
(2) To identify the defining characteristics of each type of organizational and formational network.
(3) To explore patterns and trajectory as it relates to the organizational traits of the networks which researchers have characterized to the mental lexicon. We may infer direction and offer perspective on the future of the field.

Oldfield (1966) was the first to mention the mental lexicon by name, defining it as a sort of 'mental dictionary', in which information pertaining to a word's meaning is stored and retrieved. Today, we recognize the elaborate nature of the mental lexicon; it serves as the cornerstone of language ability, a theoretical characterization of linguistic knowledge which operates with rapid and effortless functionality, with an ability to self-organize after continuous experience and with the capacity to store information over a lifetime (Libben and Jarema 2002). An established definition for the mental lexicon, as stated by Ying (2017, p. 24) is paraphrased from Carroll (1999), who defines the term as "a mental representation of words stored in memory, including information about a word's meaning, pronunciation, syntactic characteristics, and so on"; just precisely "what" and "and so on" refers to, and exactly how it is stored in memory, has yet to be resolved by the scientific community.

In recent global efforts to expand and elaborate on the organization of the mental lexicon, a surge of network patterns being applied to the mental lexicon has been acknowledged by many linguists and psychologists (Steyvers and Tenenbaum 2005; Zhao and Li 2010; Vitevitch and Goldstein 2014. In this systematic review, we explore the various network-based approaches taken to investigate the organizational patterns in the mental lexicon; we elaborate on the prevalent notion of networks in the mental lexicon, their uses and impact, and how this term has meant and continues to mean different things for different researchers. Three distinct types of networks are identified, delineated and assessed: simple, connectionist and complex networks. Finally, we end with an evaluation on the sense of trajectory within the selected literature, and draw upon observed patterns to offer direction and insight on the future of the field.

2. Methodology

The selected literature contains 36 academic articles systematically selected in order to create an externally valid sample which represents the greater body of research written on the organizational and formational properties of the mental lexicon as it pertains to networks. These 36 articles were obtained from a larger selected literature of 100 articles used to explore other aspects and points of contention within the global canon of research concerning the mental lexicon. The original 100 articles consisted of a wide array of approaches towards understanding the mental lexicon, 36 of which, all those used in this research, employ the explicit utilization of networks as a methodology to explore the mental lexicon.

To create this selected literature, six databases were used, including Journal Storage (JSTOR), Linguistics and Language Behavioral Abstracts (LLBA), Education Resources Information Center (ERIC), Scopus, Vocabulary Acquisition Research Group Archive (VARGA) and Google Scholar. Each database was used three times, under the searches Mental Lexicon and words or mind or network, Lexical Representations and words or mind or network and Lexical Organization and words or mind or network. The top 50 results were examined for each search of each database, resulting in the inspection of nearly 900 pieces of literature. These 900 were reduced to 100, utilizing the following three criteria: modernity, as we decided to only consider material published in and after 1984, as to ensure the research was up to date with current academic understandings; medium, as we decided to only include academic articles and book chapters; and relevancy, as we only included literature relating to linguistics and language that has at least two words from the appropriate search in the abstract or title. On this last point, in order to avoid bias from the wording of the before mentioned searches, we allowed substitute phrases, such as models, lexical structure, graphs theory, word frequency and lexical storage. The goal of this procedure was to allow a large breadth of perspectives on modern research to be examined.

The 100 pieces of selected literature span the 34 years dating from 1984 to 2018 with a median year of 2005/06, and include a roughly half-and-half ratio (48:52) of articles written from multilingual and monolingual perspectives. From this group of literature, all items containing approaches utilizing networks, 36 in total, were taken for the purpose of this systematic review. In methodological similarity to Libben and Jarema (2002), this selected literature is used as a snapshot representation of how researchers negotiate and approach certain pivotal questions regarding networks in the mental lexicon. As each piece of literature is read under the lens of our research aims, we offer a systematic exploration of the employed network methodologies, their consequential understandings of the mental lexicon, as well as their implications as it relates to its organizational features. Furthermore, in an attempt to elaborate on the themes presented in the selected literature, we cite additional resources outside of the selected literature. These additional resources are pertinent to the explored concepts, as they often provide historical, mathematical or contextual information. All items in the selected literature are given in the list of references.

3. Networks as Organizational Approaches of the Mental Lexicon

The approach taken on how the mental lexicon is organized not only frames individual investigations within those approaches' certain implications and limitations, but also directs the discussion between researchers at large. Of the original selected literature containing 100 items, network-based approaches comprise 36 items; this loosely suggests that approximately one-third of research pertaining to the mental lexicon employs networks methodologically. As the network metaphor can sometimes seem ubiquitous, it is imperative to distinguish and discriminate between the numerous usages of this commonly used word.

Three broad types of networks have been identified in the data, as illustrated in Figure 1, namely: simple, connectionist and small-world scale-free complex networks; these types of networks are represented as 22.2% (n = 8), 50% (n = 18) and 27.8% (n = 10) of the selected literature, respectively. These network types have discernible differences which relate to their parameters and affect their implications. It also must be noted that the differences between these network types are not solely mathematical in nature; the differences are represented also in variation of scope and how their findings can be implicated into our understanding of the mental lexicon.

While simple and connectionist networks constitute a majority of the selected literature, the historical lens of this systematic review notes that they have in some regard resulted in today's research on modern interdisciplinary complex networks. Due to the more encompassing and holistic nature of complex networks, as well as their greater potential implications, these networks will be presented in different sections. In an attempt to expiate on our third research aim, Section 4 of this review will give particular emphasis on future research and resulting implications as it relates to complex networks.

Another manner in which complex networks stand apart from their predecessors is their fluent mix of microscopic, mesoscopic and macroscopic approaches. These different scales in network scope play a major role in the consequences and implications of what can be learned about linguistic behavior in the mental lexicon. The macroscopic approach sets the experimental boundaries on a large scale in attempt to observe over-arching, linguistic behavior at the expense of exploring focused, specific linguistic phenomena. This approach may be seen in many connectionist networks, including Boolean networks (Meara 2006), which explore questions over the selective or non-selective nature of the multilingual mental lexicon. Conversely, the microscopic approach explores the specific differences in behavior in the mental lexicon which can be analyzed on a smaller basis, such as on the level of individual words, morphemes, or sounds; this approach may be employed in studies such as those exploring the priming effects of specific characteristics of vocabulary information, such as rhyme or specific morphological features. Mesoscopic approaches find a space between the large and small scales of the beforementioned approaches. Complex networks are unique amongst the three types of explored networks in this review, as they fluidly offer a portrait of the organizational features of the mental lexicon incorporating each of the three scales in scope.

An example of this would be in Vitevitch et al. (2014), where links within the mental lexicon are explored as it relates to phonetic differences between words, while the patterns found in the network at large are compared to those found in other networks which incorporate different linguistic perspectives, such as semantics or morphosyntactic information.

Relating to the mechanical aspects of networks in the mental lexicon, it is important to acknowledge and remember that networks reflect the structure of linked items within the lexicon, and therefore most networked systems require an attached linguistic approach or perspective to link the nodes. In other words, certain networks postulate links between nodes by way of phonetic similarities (Vitevitch 2008), word association data (Wilks and Meara 2002), or other discernible properties of language which can be used to organize the lexicon, such as word frequency or semantic categories. In most of the networks in the selected literature, nodes represent words, while the links between nodes represent the linguistic perspective(s) which bind the words within the mental lexicon.

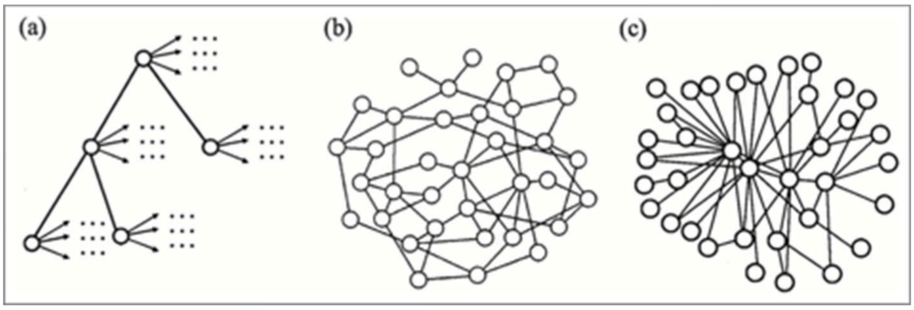

Figure 1. Three illustrations of different types of networks from the selected literature. We can see (**a**) simple networks, (**b**) one possible example of a connectionist network, and (**c**) a complex network. Source: (Steyvers and Tenenbaum 2005).

3.1. Simple Networks and Connectionist Networks

3.1.1. Simple Networks

The first type of network refers to the arbitrarily organized, small-scale graph patterns which appear in the early literature (Sandra and Rice 1995) and argue for the general structure of theoretical network models over the list-structures often found in semantic theory (Brugman and Lakoff 1988); they comprise slightly over 22% of the selected literature pertaining to networks. These simple networks started to appear in greater numbers after the seminal work of Collins and Quillian (1969), which expatiated on the interaction between one type of simply structured semantic network and its corresponding processes. This work offers a theoretical model on how semantic information is stored in the mind; their Hierarchical Network Model of Semantic Theory allows for the prediction of faster or slower reaction times in semantic recognition tasks. Collins and Quillian, and later continued with Collins and Loftus (1975), laid the groundwork for most future work on linguistic networks, as their ideas included notions such as spreading activation, a crucial feature for many connectionist network models.

Simple networks can be viewed as a new way of exploring human cognition, as seen in approaches including artificial neural networks (ANNs) (Rosenblatt 1958) and semantic theory (Quillian 1967). Additionally, these networks have been used to explore basic concepts involving the intersection of graph theory and linguistics (Meara 1992), and have been useful as a metaphor and visual aid. Additional features set them apart from the other two general types of networks delineated in the article, chiefly their limited scale and subsequent limitations of application. Simple networks do not have the weight of a large scale which permits exploration on such macroscopic linguistic phenomena which the other network types may allow, namely small-world structures.

3.1.2. Connectionist Networks

The second broad category of networks refers to the numerous computational networks, which explore specific characteristics of the structure of the network and analyze emergent properties. They include 50% of the selected literature pertaining to networks, and include simple recurrent networks (Elman 1991) and networks utilizing WordNet (Ma 2013) and Dev-Lex I (Li et al. 2004) and II (Li et al. 2007). These networks have sought to elaborate on certain intricacies such as nodal spreading activation (Ying 2017), network density (Wilks et al. 2005) and unsupervised self-organizing patterns, which seek to replicate learning algorithms, such as Hebbian learning principles, which explain how repeated behavior leads to strengthened connections in memory (Zhao and Li 2010). For more information on the variety of connectionist networks employed by the mainstream literature, see DeAnda et al. (2016). One principle element of a connectionist network is the mathematical or computational consequences of the network's internal structure in the exploration of governing patterns within the mental lexicon (Wilks and Meara 2002). A predominant example within this broad category seen in the selected literature is Meara's (2006) Boolean networks.

Random autonomous Boolean network models are an attempt to solve elaborate issues pertaining to the mental lexicon by describing a simplified model of linguistic behavior. This simplified model removes the complexity of the multitude of dynamic variables which affect linguistic behavior on a microscopic level in favor of exploring macroscopic themes pertaining to, among other things, the multilingual mental lexicon. Pioneered by Glass and Kauffman (1973), Boolean networks trade away intricate information regarding the linguistic aspects of each item and in return simplify the nature of each lexical unit in terms of a simple binary distinction. Meara (2006) utilizes Boolean networks with the purpose of exploring to what extent the multilingual mental lexicon is shared or separate. Boolean Networks work as such: the network consists of many words, each in one of two states, activated and unactivated. The current state of each word depends on the way it interacts with the other words in the network, similar to productive and receptive states. Each word is connected in a one-way direction to two other random words. Words receive input from their two connections. Some words, described as AND words, become activated when both its connecting words become activated. Other words, OR words, need only one active neighboring word to become activated. This creates two levels of activation threshold. The system begins completely unactivated. External factors activate one item, and then the network begins to light up. Every 'moment', certain words are activated, give input and subsequently become unactivated. Eventually, the system settles into a stable state, where it oscillates minimally at a stable level of activation. Despite moments where many new stimuli might be presented, called forcing, the network will always fall into the same attractor state. For example, when multiple languages are added to the network using a careful and dynamic methodology, properties emerge which partially support Grosjean's (1989) argument for a mixed multilingual lexicon. When focusing intently on a simplified, non-linguistic view of lexical structure, certain properties emerge which can give insight as to what we can expect from real lexical networks. Meara presents these explorations as thought experiments, or mirrors which reflect a simplified version of what occurs in reality (Meara 2006).

Within the realm of connectionist networks, there is an important distinction between localist and distributed networks. This distinction elaborates on what nodes represent in the network; in a localist network a node represents a singular concept, and therefore has precise meaning and interpretation, while nodes in a distributed networks share concepts, therefore implying that a concept is represented by a pattern of nodal activity (Plate 2002). This discrimination is important as it changes the nature of the network, as well as its operational processes. This can include a list of assumptions that are different from a traditional network approach. The methodology employed in this systematic review does not limit nor possess bias in the inclusion of localist and distributed networks. Both localist networks (Kenett et al. 2011; Li et al. 2007) and distributed networks (Elman 1991; Miikkulainen and Dyer 1991; Plaut 1997) can be found in the selected literature.

To place all these vastly different types of connectionist networks into one category does not do them nor their implications justice; connectionist networks are wide-ranging and vary greatly in their methodology and aims. They are unified in a certain methodological quality: they use or create a network system with its own rules and restrictions, with which they execute mathematical procedures, in an attempt to better understand certain behavioral characteristics of the mental lexicon, based on the observed emergent characteristics of their experiments. These connectionist network systems explore specific circumstances and conditions, as is the case with the Boolean network models, where the network serves as an imitation or mirror of one behavioral aspect of a real system. These connectionist networks attempt to triangulate the mechanics of specified operating patterns, and have no common mathematical language with which to communicate with other networks which utilize different circumstances and conditions. Importantly, they are not small-world, scale-free complex networks, which are naturally occurring phenomena observed in naturally formed network patterns. These characteristics are further explained in Section 3.2.

3.2. Small-World Scale-Free Complex Networks

The third and final type identified in the selected literature, comprising nearly 28% of this same selected literature, is small-world, scale-free complex networks which follow a specific and strict mathematically describable structure, as first explored by physicist Barabási and Albert (1999) with his Barabási–Albert model of networks. Complex systems harbor a vast sea of individual parts interacting in relatively simple ways, which perhaps appear chaotic or even unpredictable when viewed from afar. Due to the certain global principles which govern all small-world and scale-free systems, network theory allows researchers to extend the discoveries of emergent properties from one complex system to similar networks within other domains of research. For example, exploring complex social networks can inform our predictions of an analogous network of a cognitive system. Therefore, we understand complex networks as interdisciplinary network systems which share a common mathematical language across individual networks, described as small-world and scale-free.

Several key characteristics mathematically describe all complex networks. These key characteristics include average path length, clustering coefficient, degree of distribution and assortative mixing. Average path length describes the average shortest distance between every node (or word) in the network and every other node in the network, measured in links, as approximated here, where n = total number of nodes, k = mean degree of nodes:

$$l_{ran} \approx \frac{\ln(n)}{ln\langle k\rangle} \tag{1}$$

A small-world network is described as having a small average path length, approximately 6, often due to the existence of key players, or nodes which are few yet highly connected to many other nodes. This coincides with the popular culture's notion of six degrees of separation, where supposedly any person on the planet can be traced to any other person via six or less acquaintances. Complex networks in linguistics follow this pattern, and are considered 'small-world' (Vitevitch et al. 2012; Solé et al. 2010).

A complex network's clustering coefficient is a value attributed to each node between zero and one which describes to what extent the neighboring nodes are also connected to each other, as approximated below:

$$C_{ran} \approx \frac{\langle k\rangle}{n} \tag{2}$$

A value of one describes a node where neighboring nodes are all connected to each other, while a value of zero describes a node where neighboring nodes are completely unconnected. Small-world networks, as described by Watts and Strogatz (1998) and Newman and Watts (1999), not only possess an average path length of 6, but also have a much higher clustering coefficient than that of a completely random network, while having the same number of nodes and average links (i.e., degree) per node.

This results in small-world networks being not only intricate but also easy to traverse, giving the impression of a small environment and highlighting their functional efficiency. Small-world, complex networks in linguistics possess clustering coefficients comparable to those in other small-world networks (Albert and Barabási 2002; Batagelj and Zaveršnik 2002). It should be noted that some research suggests that small-world and scale-free networks are not as common as the selected literature may suggest (Arbesman et al. 2010; Broido and Clauset 2019).

As the degree of any given node refers to quantifying the links connecting a node to neighboring nodes, the degree distribution of any given complex network refers to a mathematical object from probability theory which marks the areas over which the values from a random variable will fall, and is calculated by marking the proportion of nodes [P(k)] on a graph in order to observe the distribution of degrees across all nodes. In accordance with the Barabási–Albert model (Albert and Barabási 2002), a scale-free network pattern is described by a power-law function in the degree distribution. In a log–log plot of the degree distribution, a power-law relationship appears as a straight line.

The slope of this line, called the degree exponent, delineated as g, in most scale-free networks is $2 < g < 3$, although not always (Montoya and Solé 2002). All scale-free networks inherently share two essential and internal mechanisms: growth and preferential attachment (Vitevitch 2008). These mechanisms enable a fluid process for the automatic self-organization of the network; growth refers to the addition of new nodes over time, and preferential attachment refers to the tendency and constraint of new nodes to connect to highly connected nodes. These two internal mechanisms, akin to Hebbian learning, make scale-free complex networks very relevant for those studying language acquisition within the lexicon.

Lastly, assortative mixing by degrees, as explained by Vitevitch and Goldstein (2014), explores how nodes attach themselves to others; a system where nodes attach themselves to similar nodes (e.g., when a teenager's friends are all teenagers) can be described through assortative mixing, while the opposite, where the pattern is inversed, can be described through disassortative mixing. This pertains to the preferential attachment patterns of a growing complex system, and can be statistically described as a correlation between the degree of each node with the degree of each of its neighbors; a positive correlation describes an association, while a negative correlation describes disassociation.

Just as these small-world, small-scale networks have been found in a wide array of disciplines regarding the economy, technology, and biology (Barabási 2009), brain connectivity (Sporns 2010), collective behavior (Mason et al. 2008) and more, they are also found within different approaches taken to understand the mental lexicon. Complex networks have been explored in relation to phonetics (Vitevitch 2008), orthography (Iyengar et al. 2012), syntax (Cancho et al. 2004) and co-occurrence (Masucci and Rodg 2006), amongst others. Additionally, complex networks have been attribute to broad array of spoken languages, including Hebrew (Kenett et al. 2011), Spanish, Mandarin, Hawaiian and Basque (Arbesman et al. 2010), cementing their conclusive findings in the literature.

While not a discernible mathematical characteristic of solely complex networks, another concept which universally binds all naturally appearing language networks approached through word frequency or co-frequency is the phenomenon of Zipf's law. As it relates to the selected literature, we see that Zipf's law, as well as the Pareto principle it is based on, is an important concept and instrument relating to the exploration of complex frequency-driven network systems across disciplines. Zipf's law is defined as:

$$f(r) \propto r^{-\gamma} \quad (3)$$

where f refers to the frequency of any given ranked word, and r to frequency rank, while $\gamma \approx 1$. Zipf's law explains hermetically that every word in any given lexicon will be less frequent than the most frequent word by a ratio equal to one over its frequency rank (Allahverdyan et al. 2013). Zipf's law is an elaboration of the Pareto distribution, which inherently explains that 20% of the causes (i.e., words) are responsible for 80% of the outcome (i.e., frequency), a phenomenon called the Pareto principle. These universal happenings are found all throughout nature and society (Newman 2005) and also relate to latent semantic analysis, which explains that words which have already been used are more likely to be used again, while words that have not been used recently are less likely to appear, a phenomenon referred to as prior density (Allahverdyan et al. 2013). The principles surrounding Zipf's law give evidence to the stability and efficiency observed within systems serving the mental lexicon, akin to other stable and efficient system found in the natural world, including small-world and scale-free complex networks.

4. Looking Ahead: Conclusions and Further Research

Within this state-of-the-art we have presented a systematic exploration of how global research efforts involving network patterns have led us to understand, organize and represent the mental lexicon. Results have shown that while a majority of employed networks are connectionist, more recently there has been an increase in the usage of complex small-world and scale-free network patterns in the literature. In addition, we have delineated certain distinctions between all three types of network patterns which relate to their parameters and potential implications.

After exploring the selected literature, certain patterns and trajectory emerge to offer insight into the orientation of future research within the field. Since Miller's 1986 work delineating the computational abilities of the vocabulary-storing mind, we have seen a deep connection rooted between the functional mechanics of computers and our understanding of organizational and formational systems within the mental lexicon; historically, these fields are entwined, as seen through the parallel development of technology and the observable computational patterns which govern the lexicon. What began as encyclopedic, dictionary-styled lists of information framed as computational and hierarchical networks, have shifted into more dynamic and holistic networks of disseminated information, as seen in Cancho and Solé (2001) and Solé et al. (2010); we more frequently see notions of multi-disciplinarily understood principles, such as Hebbian learning and neural networks, all of which are understood through modern computational technology. In this vein, we have seen a recently budding use of graphs and network models which seek to expand on the governing structures of the mental lexicon. These structures are increasingly being understood as more universal and interdisciplinary. While some literature contains conflicting views on preferred styles of networks which investigate different scopes and aspects of the formation of the mental lexicon, such as Meara (2006), which participates in the debate between specific yet simple versus intricate yet under-described connectionist network models, other literature has started examining the vast and largely unexplored realm of complex networks and their implications. As Meara explained in 2006:

> "Few linguists working in this area have explored just how far this metaphor can be pushed, and few have looked in detail at the wider implications of adopting the network metaphor. Few of us have a really thorough grasp of what the intrinsic properties of a network are, and … what its behaviors might be." (Meara 2006, p. 624)

In this regard, the current review may offer a starting point of reference as it relates to understanding "what the intrinsic properties of a network are". Moving forward, it is imperative for linguists who study the mental lexicon to become conversant on the limitations and strengths of each type of network. As Meara says, "the basic problem is that we use metaphors such as 'lexical network' without really following through with what the implications of these metaphors are (Meara 2006, p. 641). While the simple networks of the past have had a "broad and important influence" on cognitive science (Vitevitch et al. 2012, p. 41), they have over time become an obsolete instrument, and even have led to confusion over what precisely complex networks can show us. Additionally, the difference between emergentist or connectionist networks which are not small-world and scale-free, and those which are must be remembered and utilized. This discrimination is vital, as the former is used to explore specific circumstances and conditions (e.g., Boolean networks in biology), while the latter can be used inter-disciplinarily to examine the structural patterns of complex cognitive systems, to elaborate on the organization of language in the mind, and to test new hypotheses about cognition and language. Already, there is an increasing body of literature providing evidence that suggests that the structure of these complex systems attribute to the rapidity and exactitude of aspects of lexical processing (Chan and Vitevitch 2009, 2010; Sudarshan Iyengar et al. 2012; Vitevitch et al. 2014; Vitevitch et al. 2011)

In addition to these differences which fundamentally separate complex networks from connectionist and simple networks, we see another defining characteristic which is its holistic emphasis on wide-ranging scale. Particularly, complex networks simultaneously investigate the microscopic, mesoscopic and macroscopic scales of the mental lexicon. Simple networks are limited in scale, while connectionist networks are forced to investigate the mental lexicon on a scale related to the investigation's parameters which are not necessarily universal. Additionally, as connectionist networks are limited to their parameters, their results and implications are applicable to our understanding of the mental lexicon only within the confines of the scale of the experiment. In other words, if a connectionist network explores behavior on a macroscopic approach, such as Boolean networks, it cannot necessarily give us insight into the microscopic behavior of the mental lexicon.

In many recent years certain research has set a foundation for the future role that small-world and scale-free complex networks will most likely play. Researchers such as de Deyne et al. (2013) have created tested corpora ready to be used for investigations involving complex networks. Dubossarsky et al. (2017) have carried out studies with complex networks concerning semantic word associations in which age is observed as a factor; the study suggests that the complex network holds its integrity throughout life—this only emphasizes the universal nature and illuminative prospectives of small-world and scale-free complex networks. Moving into interdisciplinary territory, Bruza et al. (2009) explore the common mathematical ground between complex networks based on word association data and quantum mechanics from physics; the authors open the door to theoretical similarities between these seemingly different phenomena as a way to understand unusual behavior within the mental lexicon. Zipf's law and the Pareto principle also make gains in a similar regard (Allahverdyan et al. 2013).

As research utilizing complex networks matures, there have been new studies involving the use of multiplex networks, which are layered complex networks where nodes belong to different networks, and are therefore interconnected through multiple perspectives, such as semantically, phonologically, or syntactically. These works started to appear across several disciplines around 2013 (Gomez et al. 2013; Kivelä et al. 2014; Nicosia et al. 2013). This thrust of work utilizing multiplex networks blends together previous attempts at complex networks carried out using singular perspectives and pushes for a more holistic and complete attempt at understanding the mental lexicon through network science. More recently, we have seen multiplex networks employed to explore patterns in early word acquisition in children (Stella et al. 2017). While multiplex networks layer different linguistic perspectives on what connects nodes in a language-based network, the real progress lies in the combination of these perspectives as it leads to new qualitatively distinct patterns. An example of this is case of the vocabulary spurt in Stella et al. (2017). Certainly multiplex networks, which advance the use of complex networks, allow us to move forward with network science in our investigations of the mental lexicon.

Looking to the future, as we see a shift towards a more multi-disciplinary effort into explorations on the organizational and formational systems governing the lexicon, doors are open for more investigations spanning several branches of expertise and knowledge. Just as researchers such as Vitevitch and Solé suggest in their work, there is a gap in the research in terms of explorations of mathematically describable universal patterns, such as those seen in complex, small-world and scale-free patterns. Such explorations have been carried out in other non-linguistic fields. This includes the potential significance of fractal dimensions within the organization of the mental lexicon as it pertains to word frequency and association, as an extension of what has been observed in the phenomenon of preferential attachment in the growth of scale-free complex networks. An example of such an endeavor, albeit not relating to linguistics, is Guida and Maria (2007), which identify the topology of an Italian airport network, designed to be as efficient as possible, as a small-world and scale-free network with a fractal structure. Owing to the fact that the conclusion drawn by analysis of complex networks are applicable across disciplines, such interdisciplinary explorations into the mental lexicon could result in a deepening understanding of certain universal patterns governing human language; this could lead to many implications within the field of second language acquisition, psycholinguists and cognitive linguistics. As Vitevitch explains:

> "Cross-disciplinary analyses of the mental lexicon might lead to the discovery of various parameters that influence the development of network structures in many real-world systems." (Vitevitch 2008, p. 419)

While cross-disciplinary analysis may lead to such discoveries, it is also worth noting how the emerging properties of such networks already offer us some insight into much debated contentions in the field of linguistics. One such example is the dialog, or lack of, between emergentist and Chomskyan nativists.

Emergentism posits language as more than the sum of its parts, or a collage of emergentist properties which as explained by Meara (2006) are features, characteristics or behavioral patterns which appear from simple interactions, seemingly from nowhere yet interconnected to all factors, rather than being native to the fundamental assumptions of the structure in question. Nativism, the belief that all humans are born with an innate instinctual knowledge of grammar that serves as the basis for all language acquisition (Chomsky 1965), is underrepresented in the selected literature. This emergentist style of understanding the mental lexicon is widespread in the selected literature: while some authors avoid summoning the debate between nativist theory and connectionist, none of the authors within the selected literature exploring networks advocate the Chomskyan perspective. This unbalanced distribution of perspectives in the selected literature seems to offer support for the emergentist tradition over nativism. Chiefly, this support comes from what these small-world, scale-free complex networks have begun to show us: that the incredibly efficient, seemingly innate linguistic knowledge and subsequent ability to acquire new language-based information may not be native to humans—in fact, these power capabilities are possessed by small-world, scale-free complex networks found all throughout nature, both within the human brain and without. The complex system's ability to grow and self-organize through preferential attachment coupled with its ability to interact at acutely rapid speeds as explored through the Barabási–Albert model not only pertains to economies, technology and biology (Barabási 2009; Solé et al. 2010) but also, and most certainly, to human linguistic knowledge, just as does Zipf's law and the Pareto principle.

Moving forward, we see an under-developed subfield within complex networks pertaining to linguistic knowledge concerning studies on multilinguals. Within the selected literature, no studies involving complex networks mention bilinguals or multilinguals. Within the canon of work utilizing connectionist networks, we see a long tradition of investigating the multilingual lexicon (Haastrup and Henriksen 2000; Wilks and Meara 2002; Elman 2004; Meara 2006; Zhao and Li 2013), however this is nonexistent within studies involving small-world and scale-free networks. As Grosjean (1989) explained nearly thirty years ago, a thrust of studies exploring bilinguals and especially multilinguals is imperative as the mental lexicon is more aptly understood first within the dynamic and complex mind-space of a multilingual; tested hypothesis and theories can then be applied to the more straightforward monolingual mental lexicon, opposed to the reverse, where monolingual structures are forced upon nuanced multilinguals, as has been common. Thus, the field requires a push in studies involving complex networks and multilinguals.

There are many unexplored implications regarding the mental lexicon and complex networks on subfields within second language acquisition, cognitive linguistics and psycholinguistics. One of such implications pertains to how linguists perceive language processing; while it is outside the scope of this research to elaborate on the vast array of proposed language processors within the greater literature, we can note that language processing models and representations of the mental lexicon are implicitly linked and reflect similar cognitive concepts understood within both fields. To date, we have not seen an inquiry into understanding language processing through specifically complex networks and through their emergent properties, namely their Pareto-structured small-world form.

In general, the field has struggled with questions concerning organization versus access; by applying complex networks and their universally emergent properties into our understanding of comprehensible language processors we can hope to bridge this gap.

More untrodden territory concerning the implications of complex networks on fields pertaining to applied linguistics includes vocabulary and language acquisition. As we have learned through preferential attachment in growth patterns in complex network, similar to Hebbian principles, words do not simply materialize in the mental lexicon; rather, they self-organize and connect to the lexicon following a specific pattern. By applying our knowledge of how these growth patterns unfold, as observed universally in scale-free networks, we can garner insight into how best to teach and learn vocabulary. Today we see a cause and effect relationship between certain effective strategies or practices and their results in language learners; properties of complex networks may hold they keys to the underlying structures which give efficacy to such advantageous vocabulary learning practices.

While this review has explicated the characteristics of complex networks as reflected in the selected literature, it must be noted that recent studies have challenged these notions. For one, the small-world scale-free structure of complex networks has often been disputed (Broido and Clauset 2019). Some researchers are particularly skeptical of the notion that complex networks are scale-free, often challenging it on the premise of statistics and theory (Lima-Mendez and van Helden 2009). As the nature of complex networks is so vastly interdisciplinary, the doors are wide open for further investigation into the structural characteristics of complex networks both non-linguistically and linguistically. As is explained by Brodio and Clauset in relation to their challenges in studying language-based complex networks: "This conflict in perspective has persisted because past work has typically relied upon small, often domain-specific data sets, less rigorous statistical methods, differing definitions of 'scale-free' structure, and unclear standards of what counts as evidence for or against the scale-free hypothesis (Broido and Clauset 2019, p. 1)."

Research utilizing complex networks has been restrained by its share of limitations: for example, underdeveloped or over-specific network systems can possess a large number of hermits and islands which do not connect to the main network structure and result in indecisive results as did happen with Vitevitch (2008) in his investigation of phonetic networks, although this can perhaps be attributed to implementational issues from the limited data set and structural decisions. Furthermore, as we see with Vitevitch, these issues are not weaknesses of the approach, but rather calls for more systematic methodological research in the field, as it highlights the need to explore how and from what data we construct the networks used to explore the organizational and formational patterns in the mental lexicon.

Author Contributions: Conceptualization, L.M. and I.M.; methodology, L.M. and I.M.; investigation, L.M.; writing—original draft preparation, L.M.; writing—review and editing, L.M., I.M.; supervision, I.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors would like to thank the two anonymous reviewers for their valuable suggestions and comments. Any shortcomings remain the responsibility of the authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

Albert, Réka, and Albert-László Barabási. 2002. Statistical mechanics of complex networks. *Reviews of Modern Physics* 74: 47–97. [CrossRef]

Allahverdyan, Armen E., Weibing Deng, and Qiuping A. Wang. 2013. Explaining Zipf's law via a mental lexicon. *Physical Review E* 88: 062804. [CrossRef] [PubMed]

Arbesman, Samuel, Steven H. Strogatz, and Michael S. Vitevitch. 2010. The Structure of Phonological Networks Across Multiple Languages. *International Journal of Bifurcation and Chaos* 20: 679–85. [CrossRef]

Barabási, Albert-László. 2009. Scale-free networks: a decade and beyond. *Science* 325: 412–13. [CrossRef]

Barabási, Albert-László, and Réka Albert. 1999. Emergence of scaling in random networks. *Science* 286: 509–12. [CrossRef]

Batagelj, Vladimir, and Matjaz Zaveršnik. 2002. Generalized Cores. *arXiv* arXiv:cs/0202039.

Broido, Anna D., and Aaron Clauset. 2019. Scale-free networks are rare. *Nature Communications* 10: 1017. [CrossRef]

Brugman, Claudia, and George Lakoff. 1988. Cognitive Topology and Lexical Networks. In *Lexical Ambiguity Resolution*. Amsterdam: Elsevier, pp. 477–508.

Bruza, Peter, Kirsty Kitto, Douglas Nelson, and Cathy McEvoy. 2009. Is there something quantum-like about the human mental lexicon? *Journal of Mathematical Psychology* 53: 363–77. [CrossRef]

Cancho, Ramon Ferrer I, and Richard V. Solé. 2001. The small world of human language. *Proceedings of Biological Sciences* 268: 2261–65. [CrossRef]

Cancho, Ramon Ferrer I., Ricard V. Solé, and Reinhard Köhler. 2004. Patterns in syntactic dependency networks. *Physical Review E* 69: 051915. [CrossRef]

Carroll, David. 1999. *Psychology of Language*. New York: Brooks/Cole Publishing Company.

Chan, Kim-Yin, and Michael Vitevitch. 2009. The influence of the phonological neighborhood clustering coefficient on spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance* 35: 1934. [PubMed]

Chan, Kim-Yin, and Michael Vitevitch. 2010. Network structure influences speech production. *Cognitive Science* 34: 685–97. [CrossRef] [PubMed]

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge: MIT Press.

Collins, Allan M., and Elizabeth F. Loftus. 1975. A spreading-activation theory of semantic processing. *Psychological Review* 82: 407–28. [CrossRef]

Collins, Allan M., and M. Ross Quillian. 1969. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior* 8: 240–47. [CrossRef]

de Deyne, Simon, Daniel J. Navarro, and Gert Storms. 2013. Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behavior Research Methods* 45: 480–98. [CrossRef]

DeAnda, Stephanie, Diane Poulin-Dubois, Pascal Zesiger, and Margaret Friend. 2016. Lexical processing and organization in bilingual first language acquisition: Guiding future research. *Psychological Bulletin* 142: 655–67. [CrossRef]

Dubossarsky, Haim, Simon De Deyne, and Thomas Hills. 2017. Quantifying the structure of free association networks across the life span. *Developmental Psychology* 53: 1560–70. [CrossRef]

Elman, Jeffrey L. 1991. Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning* 7: 195–225. [CrossRef]

Elman, Jeffrey L. 2004. An alternative view of the mental lexicon. *Trends in Cognitive Sciences* 8: 301–6. [CrossRef]

Glass, Leon, and Stuart A. Kauffman. 1973. The logical analysis of continuous, non-linear biochemical control networks. *Journal of Theoretical Biology* 39: 103–29. [CrossRef]

Gomez, Sergio, Albert Díaz-Guilera, Jesus Gómez-Gardeñes, CJ Pérez-Vicente, Yamir Moreno, and Alex Arenas. 2013. Diffusion dynamics on multiplex networks. *Physical Review Letters* 110: 028701. [CrossRef]

Grosjean, François. 1989. Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language* 36: 3–15. [CrossRef]

Guida, Michele, and Funaro Maria. 2007. Topology of the Italian airport network: A scale-free small-world network with a fractal structure? *Chaos, Solitons & Fractals* 31: 527–36.

Haastrup, Kirsten, and Birgit Henriksen. 2000. Vocabulary acquisition: acquiring depth of knowledge through network building. *International Journal of Applied Linguistics* 10: 221–40. [CrossRef]

Iyengar, S. R. Sudarshan, C. E. Veni Madhavan, Katharina Anna Zweig, and Abhiram Natarajan. 2012. Understanding Human Navigation Using Network Analysis. *Topics in Cognitive Science* 4: 121–34. [CrossRef] [PubMed]

Kenett, Yoed N., Dror Y. Kenett, Eshel Ben-Jacob, and Miriam Faust. 2011. Global and Local Features of Semantic Networks: Evidence from the Hebrew Mental Lexicon. *PLoS ONE* 6: e23912. [CrossRef]

Kivelä, Mikko, Alex Arenas, Marc Barthelemy, James P. Gleeson, Yamir Moreno, and Mason A. Porter. 2014. Multilayer networks. *Journal of Complex Networks* 2: 203–71. [CrossRef]

Li, Ping, Igor Farkas, and Brian MacWhinney. 2004. Early lexical development in a self-organizing neural network. *Neural Networks* 17: 1345–62. [CrossRef]

Li, Ping, Xiaowei Zhao, and Brian Mac Whinney. 2007. Dynamic Self-Organization and Early Lexical Development in Children. *Cognitive Science* 31: 581–612. [CrossRef]

Libben, Gary, and Gonia Jarema. 2002. Mental lexicon research in the new millennium. *Brain and Language* 81: 2–11. [CrossRef]

Lima-Mendez, Gipsi, and Jacques van Helden. 2009. The powerful law of the power law and other myths in network biology. *Molecular BioSystems* 5: 1482–93. [CrossRef] [PubMed]

Ma, Xiaojuan. 2013. Evocation: analyzing and propagating a semantic link based on free word association. *Language Resources and Evaluation* 47: 819–37. [CrossRef]

Mason, Winter A., Andy Jones, and Robert L. Goldstone. 2008. Propagation of innovations in networked groups. *Journal of Experimental Psychology* 137: 422. [CrossRef] [PubMed]

Masucci, Adolfo Paolo, and Geoff J. Rodg. 2006. Network properties of written human language. *Physical Review E* 74: 026102. [CrossRef]

Meara, Paul. 1992. Network Structures and Vocabulary Acquisition in a Foreign Language. In *Vocabulary and Applied Linguistics*. London: Palgrave Macmillan UK, pp. 62–70.

Meara, Paul. 2006. Emergent Properties of Multilingual Lexicons. *Applied Linguistics* 27: 620–44. [CrossRef]

Miikkulainen, Risto, and Michael G. Dyer. 1991. Natural Language Processing with Modular PDP Networks and Distributed Lexicon. *Cognitive Science* 15: 343–99. [CrossRef]

Miller, George A. 1986. Dictionaries in the mind. *Language and Cognitive Processes* 1: 171–85. [CrossRef]

Montoya, Jose M., and Ricard V. Solé. 2002. Small World Patterns in Food Webs. *Journal of Theoretical Biology* 214: 405–12. [CrossRef]

Newman, Mark EJ. 2005. Power laws, Pareto distributions and Zipf's law. *Contemporary Physics* 46: 323–51. [CrossRef]

Newman, Mark, and Duncan Watts. 1999. Renormalization group analysis of the small-world network model. *Physics Letters A* 263: 341–46. [CrossRef]

Nicosia, Vincenzo, Ginestra Bianconi, Vito Latora, and Marc Barthelemy. 2013. Growing multiplex networks. *Physical Review Letters* 111: 058701. [CrossRef] [PubMed]

Oldfield, Richard. 1966. Things, Words and the Brain. *Quarterly Journal of Experimental Psychology* 18: 340–53. [CrossRef] [PubMed]

Plate, Tony. 2002. Distributed representations. In *Encyclopedia of Cognitive Science*. Edited by Nadel Lynn. London: Palgrave Macmillan.

Plaut, David C. 1997. Structure and Function in the Lexical System: Insights from Distributed Models of Word Reading and Lexical Decision. *Language and Cognitive Processes* 12: 765–806. [CrossRef]

Quillian, M. Ross. 1967. Word concepts: A theory and simulation of some basic semantic capabilities. *Behavioral Science* 12: 410–30. [CrossRef]

Rosenblatt, Frank F. 1958. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review* 65: 386–408. [CrossRef]

Sandra, Dominiek, and Sally Rice. 1995. Network analyses of prepositional meaning: Mirroring whose mind—The linguist's or the language user's? *Cognitive Linguistics* 6: 89–130. [CrossRef]

Solé, Ricard V., Bernat Corominas-Murtra, Sergi Valverde, and Luc Steels. 2010. Language networks: Their structure, function, and evolution. *Complexity* 15: 20–26. [CrossRef]

Sporns, Olaf. 2010. *Networks of the Brain*. Cambridge: MIT Press.

Stella, Massimo, Nicole M. Beckage, and Markus Brede. 2017. Multiplex lexical networks reveal patterns in early word acquisition in children. *Scientific Reports* 7: 46730. [CrossRef]

Steyvers, Mark, and Joshua B. Tenenbaum. 2005. The Large-Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth. *Cognitive Science* 29: 41–78. [CrossRef]

Vitevitch, Michael S. 2008. What Can Graph Theory Tell Us About Word Learning and Lexical Retrieval. *Journal of Speech, Language, and Hearing Research* 51: 408. [CrossRef]

Vitevitch, Michael S., and Rutherford Goldstein. 2014. Keywords in the mental lexicon. *Journal of Memory and Language* 73: 131–47. [CrossRef] [PubMed]

Vitevitch, Michael S., Gunes Ercal, and Bhargav Adagarla. 2011. Simulating Retrieval from a Highly Clustered Network: Implications for Spoken Word Recognition. *Frontiers in Psychology* 2: 369. [CrossRef] [PubMed]

Vitevitch, Michael S., Kit Ying Chan, and Steven Roodenrys. 2012. Complex network structure influences processing in long-term and short-term memory. *Journal of Memory and Language* 67: 30–44. [CrossRef] [PubMed]

Vitevitch, Michael S., Chan Ying, and Rutherford Goldstein. 2014. Insights into failed lexical retrieval from network science. *Cognitive Psychology* 68: 1–32. [CrossRef] [PubMed]

Watts, Duncan J., and Steven H. Strogatz. 1998. Collective dynamics of 'small-world' networks. *Nature* 393: 440–42. [CrossRef] [PubMed]

Wilks, Clarissa, and Paul Meara. 2002. Untangling word webs: Graph theory and the notion of density in second language word association networks. *Second Language Research* 18: 303–24. [CrossRef]

Wilks, Clarissa, Paul Meara, and Brent Wolter. 2005. A further note on simulating word association behaviour in a second language. *Second Language Research* 21: 359–72. [CrossRef]

Ying, Zhang. 2017. The Representation of Bilingual Mental Lexicon and English Vocabulary Acquisition. *English Language Teaching* 10: 24–27. [CrossRef]

Zhao, Xiaowei, and Ping Li. 2010. Bilingual lexical interactions in an unsupervised neural network model. *International Journal of Bilingual Education and Bilingualism* 13: 505–24. [CrossRef]

Zhao, Xiaowei, and Ping Li. 2013. Simulating cross-language priming with a dynamic computational model of the lexicon. *Bilingualism Language and Cognition* 16: 288–303. [CrossRef]