The multiplex structure of the mental lexicon influences picture naming in people with aphasia

Nichol Castro[†,1]

School of Psychology, Georgia Institute of Technology, 654 Cherry Street, Atlanta, Georgia, 30332 USA

AND

Massimo Stella[†,1]

Institute for Complex Systems Simulation, University of Southampton, University Road 4, SO17 1BJ, Southampton, UK and Complex Science Consulting, Via Amilcare Foscarini 2, 73100, Lecce, Italy
[†]Corresponding author. Email: ncastro1@uw.edu; massimo.stella@inbox.com

Edited by: Danielle Bassett

An emerging area of research in cognitive science is the utilization of networks to model the structure and processes of the mental lexicon in healthy and clinical populations, like aphasia. Previous research has focused on only one type of word similarity at a time (e.g., semantic relationships), even though words are multi-faceted. Here, we investigate lexical retrieval in a picture naming task from people with Broca's and Wernicke's aphasia and healthy controls by utilizing a multiplex network structure that accounts for the interplay between multiple semantic and phonological relationships among words in the mental lexicon. Extending upon previous work, we focused on the global network measure of closeness centrality which is known to capture spreading activation, an important process supporting lexical retrieval. We conducted a series of logistic regression models predicting the probability of correct picture naming. We tested whether multiplex closeness centrality was a better predictor of picture naming performance than single-layer closeness centralities, other network measures assessing local and meso-scale structure, psycholinguistic variables and group differences. We also examined production gaps, or the difference between the likelihood of producing a word with the lowest and highest closeness centralities. Our results indicated that multiplex closeness centrality was a significant predictor of picture naming performance, where words with high closeness centrality were more likely to be produced than words with low closeness centrality. Additionally, multiplex closeness centrality outperformed single-layer closeness centralities and other multiplex network measures, and remained a significant predictor after controlling for psycholinguistic variables and group differences. Furthermore, we found that the facilitative effect of closeness centrality was similar for both types of aphasia. Our results underline the importance of integrating multiple measures of word similarities in cognitive language networks for better understanding lexical retrieval in aphasia, with an eye towards future clinical applications.

Keywords: aphasia; language sciences; network science; complex systems; multiplex networks; mental lexicon; picture naming; lexical retrieval.

---

[1]These authors contributed equally to this work.

1. Introduction

Network science provides mathematical and computational tools to model the structure and processes of human language. Two such types of network approaches are left hemispheric language networks and cognitive language networks. Left hemispheric language networks utilize neuroimaging data and computational approaches from neuroscience to represent the structure of neural networks in the brain [1]. On the other hand, cognitive language networks utilize corpora or behavioural language performance data to represent the structure of the mental lexicon (e.g., [2, 3]), or the place in long-term memory where words are stored [4]. Specifically, we focus on cognitive language networks, where nodes represent words and edges encode different relationships between words, including semantics, syntax, phonology and orthography. Previous work on cognitive language networks has typically examined several types of 'single-layer networks' were only one type of edge is represented in the network at a time in order to more carefully consider different aspects of language. For example, semantic networks of word associations [5, 6] have been used to identify writing styles [7], to quantify creativity [8] and for modelling word acquisition in young children [9]. On the other hand, phonological networks capturing phonological overlap between words [10] have been used to better understand word acquisition [11, 12] and spoken word recognition and production [13].

Much of the work examining cognitive language networks has focused on the influence of the local structure (or immediate connections) of words in the mental lexicon of adults (e.g., [6, 13, 14]). However, the local structure of words in both phonological and semantic networks has also been shown to influence what words children know and what new words children are likely to learn. Two measures that have been closely examined are degree (the number of immediately connected nodes) and clustering coefficient (the amount of interconnectivity among immediately connected nodes). In phonological networks, words with high phonological degree or low clustering coefficient were more likely to be known and to be learned by infants and children than words with low phonological degree or high clustering coefficient [11, 12]. Furthermore, the influence of these local phonological network measures disappeared after approximately 30 months of age [9, 12]. In semantic networks, words with high semantic degree and high semantic clustering coefficient were also more likely to be known than words with low semantic degree and low semantic clustering coefficient [11]. Furthermore, Hills *et al*. [9] showed that words with high semantic degree were also more likely to be learned than words with low semantic degree. Taken together, these results highlight that word acquisition is strongly influenced by both the semantic and phonological network structure of words.

In addition to word acquisition, the local network structure of the mental lexicon has also been examined in populations with language disorders, like aphasia. For example, phonological degree has been shown to facilitate lexical retrieval and production by people with aphasia [15, 16]. That is, during a production task, a single lexical item is activated from which its sub-lexical units must be activated (e.g., syllables, phonemes). A word with many phonological neighbours, which share the phonological units needed for production of the selected lexical item, will be produced more quickly than a word with few phonological neighbours [15, 16]. Additional support for the facilitative influence of phonological degree on speech production also comes from work examining speech errors and tip-of-the-tongue states in unimpaired adults [17], where words with high phonological degree elicited fewer errors than words with low phonological degree. Despite a language impairment due to stroke, the language network structure still had a similar effect on lexical retrieval processes by people with aphasia as seen in unimpaired adults.

More recent investigations of structural influence on language processes has turned to examination of global network measures, or measures that assess how a node is connected to the rest of the network. For example, word acquisition is strongly influenced by the global topology of the semantic network [18],

where new words are likely to fill existing gaps (see also [19]). However, in the case of children with a language delay, new words learned tend to be 'odd-balls', or atypical words given what the child already knows [20]. For example, children with a language delay may be more likely to learn the word *telephone* rather than *dog* after learning the word *cat*, even though *dog* and *cat* are more semantically similar than *telephone* and *cat*. Critically, the creation of these somewhat arbitrary edges in the semantic network of children with a language delay led to less small-world structure than what was found for semantic networks of typically developing children. Furthermore, a lack of significant small-world structure, which is important for the ease of language processing and comprehension [21], likely reflects the language difficulties of these children. Thus, knowing what words children are learning and how learned words influence the global structure of the network will be particularly useful in developing novel diagnostic and intervention strategies, which would not have been possible by just considering the immediate, local structure of words.

Two other global network measures currently being applied to cognitive language networks are network distance and closeness centrality. Network distance between a pair of nodes is simply the shortest path connecting them [22], while closeness centrality captures all of the shortest paths connecting a single node to all other nodes in the network [23]. These global network measures have implications for how processes move through a network, specifically spreading activation and navigation in the case of cognitive language networks. For example, cued recall was more successful when studied prime words were relatively close, in terms of semantic network distance, to the target word [22]. This finding suggests that spreading activation may only travel a limited distance in the network. Additionally, closeness centrality of words in a phonological network serves to aid navigation during picture naming [16] and spoken word recognition [24]. Specifically, words with high closeness centrality 'pop out' like a landmark in the network, making these words easier to access. Taken together, these results suggest that global network measures may serve as powerful proxies for operationalizing spreading activation and furthering our understanding of how network structure influences processes like lexical retrieval.

Although single-layer language networks have been successful in understanding the influence of structure on language processes, their main limitation is that these networks only consider one type of relationship between words at a time. In turn, this implies neglecting any interplay among different aspects of language, such as semantics and phonology, which are generally considered strongly entwined in the human mind [4, 25, 26]. An example of the interactive nature of semantics and phonology are mixed errors in speech production [27–29] (e.g., saying *rat* for *cat* as they both sound similar and are semantically related). To account for these errors, one prominent psycholinguistic theory is Dell's [30] connectionist model through which lexical retrieval occurs over two stages: lemma selection and phonological access. During the selection phase, conceptual representations (e.g., semantic features) activate an appropriate lemma, or a non-phonological form of a word, that best matches input received (e.g., a picture). After lemma selection, sub-lexical units (e.g., phonemes) related to the lemma selected activate to allow for word production. Importantly though, Dells' model is *interactive*, in that connections between semantic and lemma nodes and between lemma and phoneme nodes are bi-directional. Thus, it is possible for phonological information to have an upward influence on lemma selection just as semantic information can have a downward influence on phonological access, allowing for the production of mixed errors in the model.

Given that the interaction of semantics and phonology can occur during lexical retrieval, it is imperative that the most appropriate network representation is one that considers semantics and phonology simultaneously. Fortunately, a new movement in network science focuses on how multiple relationships between nodes interact and influence one another with the use of a *multiplex network representation* [31–34]. In this type of representation, the network is composed of several layers of information, with each layer representing a different type of relationship (i.e., edge) between nodes. The same nodes exist in all

layers, but nodes are connected differently within each layer. As applied to language, a multiplex network representation could consist of layers of information capturing semantic and phonological relationships between words. Importantly, the multiplex representation of the mental lexicon differs from Dell's interactive model in construction. Nodes in Dell's model represent concepts, lemmas and phonemes, with links connecting either concepts and lemmas or lemmas and phonemes. No links between lemmas were considered in Dell's model. On the other hand, nodes in the multiplex network are akin to lemmas, with multiple types of semantic and phonological links directly connecting lemma nodes. We propose that these differences would make the multiplex network more parsimonious since it does not depend on the choice between semantic or phoneme nodes for every single word, while still capturing the interactive nature of semantics and phonology.

Indeed, work by Stella *et al*. has examined the multiplex network representation of the mental lexicon (henceforth called the *multiplex lexical network* and its influence on word acquisition). For example, [35] examined a multiplex lexical network composed of four layers capturing semantic, syntactic and phonological relationships among words. This multiplex structure was more predictive than any single-layer network feature in determining which words were newly acquired on which month during normative learning (see also [36]). Additionally, not all layers were found to be significant, with the phonological layer having the smallest influence on predicting word acquisition and a layer of free associations having the most influence [35]. Subsequent analyses in [37] used a larger multiplex lexical network of free associations, synsets, taxonomic relations and phonological similarities between words (Fig. 1a and b). With this multiplex lexical network, the authors identified a core, or a group of densely connected nodes across all layers, that was robust to random failure/removal of words. The core contained more frequent, more concrete, earlier acquired, shorter and more polysemous words compared to words outside the core. Additionally, the core appeared around 7–8 years of age, a well-documented phase of increased operational [38], reading [39] and communicative [4] skills. Lastly, [40] also showed that the entropy of the multiplex network distance and the multiplex closeness centralities of words were both powerful features for predicting early word acquisition. Taken together, these findings indicate that the multiplex lexical network can account for novel aspects of word acquisition that could not be captured by examining only single-layer networks.

In sum, the recent application of multiplex networks to the mental lexicon, along with the ability of global network measures to quantify how processes operate within a network, provides an opportunity to better understand lexical retrieval. We extend upon [16] who examined the influence of phonological closeness centrality on picture naming accuracy by people with aphasia and healthy controls. However, we utilized the multiplex framework which increases the number of available paths for navigation in the mental lexicon [16, 24] and has the potential to capture the spread of activation between semantics and phonology [27]. Therefore, we predict that multiplex closeness centrality will better capture picture naming performance than single-layer measures of closeness centrality, continuing to highlight that multiplex lexical relationships are cognitively relevant. Furthermore, we predict that closeness centrality will outperform other network measures capturing local and meso-scale structure. And, in conjunction with [16], we also predict that the effect of multiplex closeness centrality will be similar across types of aphasia and healthy controls.

2. Methods

2.1 *Linguistic dataset for picture naming*

Picture naming data from people with aphasia and healthy controls were obtained from the Moss Aphasia Psycholinguistics Project Database (mappd.org), a freely available online database [41]. Specifically, we
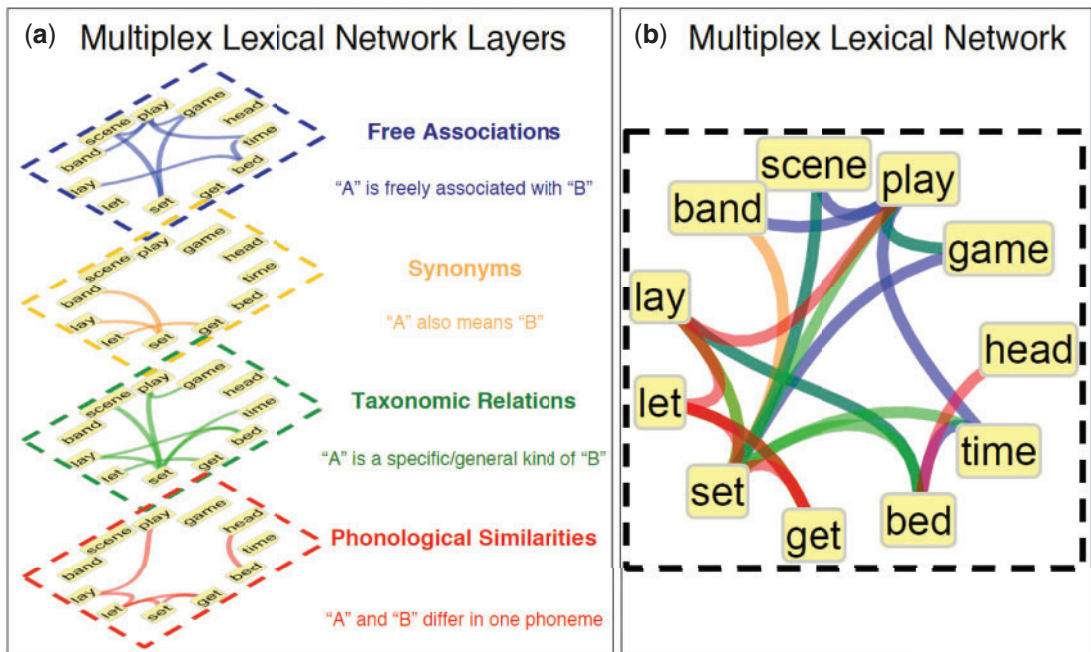
FIG. 1. (a) Visual representation of a subset of the multiplex lexical network representing semantic and phonological relationships in the mental lexicon. The multiplex lexical network is made of four layers: free associations, synonyms, taxonomic relations and phonological similarities. (b) Compact visualization of the subset of the multiplex lexical network.

report analyses using the Philadelphia Naming Test, where participants attempted to name 175 pictures whose names were nouns varying in word frequency [42]. Although the data contains 175 picture names, only 142 picture names were found in the multiplex lexical network described below, and were analysed (see Appendix A). Participants' data came from 58 people with Broca's aphasia, 36 people with Wernicke's aphasia and 20 healthy controls. Naming responses were coded for accuracy and, in the case of inaccurate responses, coded for type of speech error. Only picture naming accuracy was considered and coded as a binomial response (0 = incorrect and 1 = correct). The average accuracy on picture naming by group was the following: 52.98% for people with Broca's aphasia, 39.43% for people with Wernicke's aphasia and 97.42% for the control group.

2.2 Linguistic datasets for the multiplex lexical network

The multiplex lexical network from Stella *et al.* [37] was used as a representation of the mental lexicon of unimpaired adult English speakers. This representation combines the most prominent semantic and phonological network structures of the mental lexicon, namely free associations, synonyms, taxonomic relations and phonological similarities. Figure 1a and b provides an example of the combination of layers in this network. Note that the words chosen in the multiplex lexical network were sampled from the union of the lexical inventories used to construct each layer (described below) and then selected based on their connectivity: Words included in the multiplex lexical network were required to have at least one connection on at least one layer. This choice led to discarding words that were disconnected across all layers of the multiplex lexical network, leading to 8531 words/nodes.

To define one of the semantic layers, we used free associations which do not rely on rigid criteria for defining word relations. Typically free associations are collected empirically from subjects who report the first word that comes to mind when presented with a given cue (e.g., for the cue 'dog', responses might include 'cat', 'pet' and 'ball'), and have been used for decades in memory and language studies (e.g., [43]). Furthermore, free associations have been used for network analyses of semantic memory [6] and to capture cognitive processes (e.g., [7–9]). We obtained free associations from the Edinburgh Associative Thesaurus [44], where only free associations provided by at least two different subjects were considered in order to filter out idiosyncratic associations [43, 45].

In addition to free associations, which do not always capture ordered or well-defined relationships among words, we also included synonyms (e.g., 'bad' and 'vicious') and taxonomic dependencies (e.g., 'dove' is a type of 'bird') as additional semantic layers. Both of these linguistic constructs have been shown to play a highly important role in influencing lexical processing [2, 14, 46]. We obtained synsets from *WordNet* 3.0 [47] and taxonomic relationships from *WordData* by Wolfram Research [48].

Lastly, we included a phonological layer based on American English, relying on the measure of phonological similarity originally introduced by Luce and Pisoni [49] and tested within a network science perspective by Vitevitch [10]. Pairs of words were connected on the phonological layer if they differed by only one phoneme through addition, substitution, or deletion (e.g., 'cat' would be connected to the words 'mat' and 'at'). Several studies provide evidence that phonological network structure influences word acquisition [11, 12] and spoken word recognition and production [13]. We obtained phonological transcriptions from *WordData* by Wolfram Research [48].

Finally, it is important to point out that edge directionality and weighting of edges in the multiplex lexical network were not adopted. This was particularly relevant to the free association layer where one could model directed, weighted edges given the nature of the free association data. That is, a cue word leads to a particular response word (i.e., directed) and some cue-response pairs may be more common across the sample than other pairs (i.e., weighted). Previous approaches found that using directed, weighted edges for a free association network did not lead to a substantially different network than a non-directed, unweighted network [8, 22]. Therefore, to keep all layers equivalent, to maintain the simplest model possible and to remain analogous with previous research, we used non-directed, unweighted edges for all layers in the multiplex lexical network. We additionally compared the closeness centrality of words in an unweighted network versus a weighted network, and found that they were strongly correlated (Kendall Tau $\tau = 0.90$, p < 0.0001), further supporting use of the simpler network model.

2.3 *Modelling approach*

Our main aim was to test how well the multiplex closeness centrality of words predicts picture naming performance by people with aphasia and healthy controls. Therefore, we tested and compared several logistic regression analyses to predict the probability of correct picture naming. All model building and analyses were done using *Mathematica 10*. Craig-Uhler pseudo-R^2 was used to asses the amount of variance explained by the predictors in a model, and to compare models to one another.

Recall that closeness centrality captures how well connected a node is to the rest of network. Notice that in a given network with N connected nodes, the closeness centrality c_i of a node i is defined as [23]:

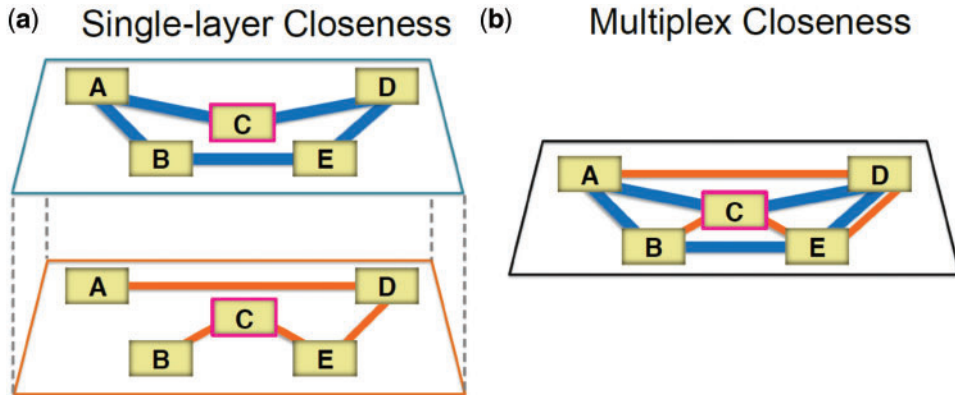$$c_i = \frac{N}{\sum_j d_{ij}} \tag{1}$$

FIG. 2. A depiction of single-layer versus multiplex closeness centrality using a network of five nodes connected on two layers. (a) Node C has a single-layer closeness centrality value less than one on each layer. (b) Node C has a multiplex closeness centrality value of one, as it is now fully connected to all nodes when both layers are considered simultaneously.

where d_{ij} is the shortest path connecting nodes i and j. High closeness centrality indicates that a given node is relatively 'close' (in terms of edges) to all other connected nodes in the network, while low closeness centrality indicates that a given node is relatively 'far' from all other connected nodes in the network. The whole multiplex lexical network was fully connected [37], so that there was always a shortest path connecting any two nodes. However, the individual layers were not fully connected. Couples of disconnected words i and j provided a value $d_{ij} = \infty$ and a contribution $1/d_{ij} = 0$ to closeness centrality, as usually assumed when computing closeness centrality [23]. Importantly, when considering the multiplex structure, the availability of more links can dramatically alter the closeness centrality of nodes as compared to when considering individual layers separately. For instance, as reported in Fig. 2, node C has a closeness of $c_C^{blue} = 5/6 < 1$ on the blue layer and a closeness of $c_C^{red} = 5/7 < 1$ on the red layer. In contrast, when the multiplex structure is considered, node 'C' is one link away from every other node in the multiplex and its closeness becomes $c_C^{mult} = 1$. Thus, for each word in the multiplex lexical network, we obtained a closeness centrality value based on each single-layer (free associations, synsets, taxonomic relations and phonological similarities) and based on the multiplex lexical network.

Importantly, we expect that the multiplex closeness centrality of words is not identical to a trivial combination of single-layer closeness centralities. Combining multiple single-layer closeness centralities would be considered a linear, multi-layer measure since it would consider multiple aspects of the mental lexicon, but would not consider inter-layer transitions. In contrast, multiplex closeness centrality would be considered a non-linear, multi-layer measure since it also considers multiple aspects of the mental lexicon while allowing for jumps between semantics and phonology. Therefore, we tested whether multiplex closeness centrality better predicts picture naming accuracy than the combination of single-layer closeness centralities in Models 1a–e.

We also aimed to determine if global network measures better account for picture naming performance than local or meso-scale measures. Therefore, we compared the measure of multiplex closeness centrality to multidegree (Model 2a) and multiplex versatility PageRank (Model 2b). Recall that degree captures the local structure of a word through the number of immediate neighbours, with multidegree as the sum of degree across all layers [34]. PageRank, on the other hand, captures aspects of meso-scale structure by determining the probability of randomly visiting a given node [23], with multiplex versatility PageRank also considering all layers of the network in its calculation [50].

Even if multiplex closeness centrality was a significant predictor of picture naming performance, we wanted to test whether its influence was greater than other predictors known to influence language processes. Thus, we also assessed the influence of multiplex closeness centrality above and beyond traditional psycholinguistic variables and group differences in Models 3a–c. Specifically, we included predictors of word frequency measured from English subtitles [51], word length measured by the number of phonemes and age of acquisition measured from Amazon Turk [52] and compared healthy controls to people with Broca's aphasia and people with Wernicke's aphasia. In addition, we also tested the interaction of group and multiplex closeness centrality in Model 3d, following the previous work of [16].

Finally, after analysing the goodness-of-fit of individual logistic regression models, we calculated the estimated parameters of single-layer and multiplex closeness centralities for predicting the likelihood of correct picture naming. This analysis allowed us to further examine the relative influence of each type of closeness centrality for each group (i.e., healthy controls, people with Broca's aphasia and people with Wernicke's aphasia). As a rather simple quantification we considered the concept of production gaps, or the absolute difference between the production likelihood of the word with the lowest and the highest closeness centralities estimated from a given regression model:

$$\Delta p = |p(W) - p(w)| \tag{2}$$

where Δp captures the difference in the probability of correct production between the word W (word with highest closeness centrality) and w (word with lowest closeness centrality).

3. Results

Recall that the first set of model comparisons tested whether multiplex closeness centrality was a better predictor of picture naming performance than the combination of the four single-layer closeness centralities. Table 1 provides the models' output and the pseudo-R^2 values. Of note, there were two possible ways in which to assess the single-layer closeness centralities: as four separate predictors in the model and as a single summed predictor. We tested both instances with Models 1a and 1b, and compared these models to another model containing only multiplex closeness centrality (Model 1c). Finally, we also tested models containing both single-layer and multiplex closeness centralities simultaneously (Models 1d and 1e).

The comparison between Models 1a and 1b indicated that the combination of single-layer closeness centralities was not the same as the sum of single-layer closeness centralities. Specifically, Model 1a containing the combination of single-layer closeness centralities explained more variance in predicting the likelihood of correct picture naming than Model 1b containing only the sum of single-layer closeness centralities ($R^2 = 0.013$ and 0.009, respectively). In contrast to Models 1a and 1b, we found that Model 1c containing only multiplex closeness centrality explained more variance in predicting the likelihood of correct picture naming ($R^2 = 0.024$). Our results indicated that multiplex closeness centrality, which considers paths transitioning across layers, plays a greater role in explaining picture naming performance than single-layer closeness centralities.

In addition, we also tested whether inclusion of single-layer and multiplex closeness centralities in the same model better predicts picture naming performance than multiplex closeness centrality alone. Thus, we tested Models 1d (the combination of single-layer closeness centralities plus multiplex closeness centrality) and 1e (the sum of single-layer closeness centralities plus multiplex closeness centrality). Although Model 1d had a nearly equivalent psuedo-R^2 as Model 1c ($R^2 = 0.026$), Model 1d contained five predictor variables with relatively small coefficients for the single-layer closeness centralities as compared to multiplex closeness centrality. Additionally, Model 1e explained less variance ($R^2 = 0.018$)

TABLE 1 *Logistic regression model outputs predicting the probability of correct picture naming from single-layer and multiplex closeness centralities*

	Coefficient	Standard error	z-statistic	p-value	Pseudo-R^2
Model 1a					0.0130
Intercept	0.2037	0.0131	15.5488	<0.0001	
Phonological CC	0.1212	0.0142	8.5218	<0.0001	
Taxonomic CC	0.0831	0.0136	6.1309	<0.0001	
Synsets CC	−0.0155	0.0137	−1.1281	0.2592	
Associations CC	0.1046	0.0132	7.8830	<0.0001	
Model 1b					0.0097
Intercept	0.2051	0.0130	15.6767	<0.0001	
Sum of CCs	0.0750	0.0057	13.1539	<0.0001	
Model 1c					0.0246
Intercept	0.2033	0.0131	15.4592	<0.0001	
Multiplex CC	0.2789	0.0135	20.6673	<0.0001	
Model 1d					0.0266
Intercept	0.2035	0.0131	15.4577	<0.0001	
Phonological CC	0.0539	0.0142	3.7909	<0.0001	
Taxonomic CC	−0.0585	0.0164	−3.5682	<0.0001	
Synsets CC	0.0114	0.0139	0.8206	0.4118	
Associations CC	−0.0524	0.0139	−3.1296	0.0017	
Multiplex CC	0.3229	0.0208	15.5187	<0.0001	
Model 1e					0.0185
Intercept	0.2033	0.0131	15.4537	<0.0001	
Sum of CCs	−0.0031	0.0007	−0.4303	<0.6669	
Multiplex CC	0.2837	0.0174	16.2456	<0.0001	

All variables were standardized. CC, closeness centrality.

than Model 1c. Given these results, we utilized only the single predictor of multiplex closeness centrality in the remainder of this article.

In the next set of model comparisons, we examined whether other commonly studied network measures assessing local and meso-scale structure captured greater variance in predicting the likelihood of correct picture naming than the global network measure of closeness centrality. Recall, we specifically chose to assess multidegree in Model 2a and multiplex versatility PageRank in Model 2b. Table 2 provides the models' output and the pseudo-R^2 values. We found that both Models 2a ($R^2 = 0.019$) and 2b ($R^2 = 0.020$) explained less variance than Model 1c, indicating that a global network measure is better suited for understanding lexical retrieval than local or meso-scale network measures.

We also chose next to test whether the effect of multiplex closeness centrality remained significant in predicting the likelihood of correct picture naming after controlling for traditionally examined psycholinguistic variables and group differences. Table 3 provides the models' output and the pseudo-R^2 values. We first tested a model containing only psycholinguistic variables (Model 3a), followed by a model containing the psycholinguistic variables plus group (Model 3b). Model 3c then included multiplex closeness centrality and Model 3d tested the interaction of group and multiplex closeness centrality.

TABLE 2 *Logistic regression model outputs predicting the probability of correct picture naming from multidegree and multiplex versatility PageRank*

	Coefficient	Standard error	z-statistic	p-value	Pseudo-R^2
Model 2a					0.0199
Intercept	0.0637	0.0154	4.1313	<0.0001	
Multidegree	0.1759	0.0093	18.8626	<0.0001	
Model 2b					0.0204
Intercept	0.0728	0.0151	4.224	<0.0001	
Multiplex versatility PageRank	0.1815	0.0094	19.1478	<0.0001	

All variables were standardized.

Recall that Model 1c, containing only multiplex closeness centrality, had a pseudo-R^2 = 0.024. Interestingly, Model 3a which only included psycholinguistic variables explained more variance than Model 1c (R^2 = 0.032). Furthermore, adding group in Model 3b substantially increased the amount of variance explained (R^2 = 0.209). As expected, people with Broca's and Wernicke's aphasia had a lower probability of correct picture naming than healthy controls, with the odds of correct picture naming at 0.02 for people with Broca's aphasia and at 0.01 for people with Wernicke's aphasia as compared to healthy controls. Therefore, it remains imperative to include psycholinguistic variables and group differences when examining lexical retrieval.

Critically, we also found that including multiplex closeness centrality (Model 3c) remained a significant predictor of picture naming accuracy, even when controlling for psycholinguistic variables and group differences (R^2= 0.214). Holding all other variables constant, the odds of correct picture naming increased by 1.18 for each one standard deviation increase in multiplex closeness centrality. Stated alternatively, words with higher multiplex closeness centrality were more likely to be named correctly than words with lower multiplex closeness centrality. Additionally, the coefficient for multiplex closeness centrality was greater than the coefficients for word frequency and length, and slightly less than the coefficient for age of acquisition, further indicating that consideration of this network measure is important for lexical retrieval. Lastly, Model 3d containing the interaction between multiplex closeness centrality and group did not explain more variance in picture naming performance (R^2 = 0.212) than Model 3c, which suggests that multiplex closeness centrality had a facilitative effect for both types of aphasia.

Finally, we report the production gaps for single-layer and multiplex closeness centralities for each group as an additional way to examine the measures' influence on picture naming performance. Recall that production gaps were calculated as the absolute difference between the production likelihood of the word with the lowest and the highest closeness centralities. Larger gaps in the production likelihood indicate that the probability of producing a given word is sensitive to the closeness centrality of that word on a given lexical structure. Smaller gaps indicate that words, even with very different closeness centrality values, possess similar probabilities of being produced (see also Figure 3). Table 4 provides the estimated production gaps for each type of closeness centrality (all single-layer variants and the multiplex variant) and for each group (healthy controls, people with Broca's aphasia and people with Wernicke's aphasia).

As seen in Table 4, we found that multiplex closeness centrality had a substantially larger estimated production gap than any of the single-layer closeness centralities for both types of aphasia. Figure 2 provides a visualization of the estimated production gaps for multiplex closeness centrality for each group, which supports what was found in Model 3d, where the interaction of multiplex closeness centrality and

TABLE 3 *Logistic regression model outputs predicting the probability of correct picture naming psycholinguistic variables, group and multiplex closeness centrality*

	Coefficient	Standard error	z-statistic	p-value	Pseudo-R^2
Model 3a					0.0320
Intercept	0.2066	0.0132	15.6548	<0.0001	
Word frequency	0.0953	0.0151	6.2927	<0.0001	
Word length	−0.1373	0.0144	−9.5057	<0.0001	
Age of acquisition	−0.1882	0.0143	−13.1182	<0.0001	
Model 3b					0.2095
Intercept	4.0136	0.1416	28.3348	<0.0001	
Word frequency	0.1069	0.0160	6.6579	<0.0001	
Word length	−0.1584	0.0155	−10.1590	<0.0001	
Age of acquisition	−0.2153	0.0154	−13.9699	<0.0001	
Broca's aphasia	−3.7502	0.1430	−26.2240	<0.0001	
Wernicke's aphasia	−4.3267	0.1431	−30.2330	<0.0001	
Model 3c					0.2140
Intercept	4.0492	0.1416	28.5916	<0.0001	
Word frequency	0.0273	0.0198	1.1377	0.1683	
Word length	−0.1001	0.0175	−5.7054	<0.0001	
Age of acquisition	−0.1944	0.0156	−12.4552	<0.0001	
Broca's aphasia	−3.7556	0.1430	−26.2553	<0.0001	
Wernicke's aphasia	−4.3334	0.1431	−30.2715	<0.0001	
Multiplex CC	0.1649	0.023	7.0768	<0.0001	
Model 3d					0.2120
Intercept	4.0387	0.1426	28.3027	<0.001	
Word frequency	0.0225	0.0200	1.1249	0.2606	
Word length	−0.0992	0.0175	−5.6573	<0.0001	
Age of acquisition	−0.1978	0.01565	−12.7303	<0.0001	
Broca's aphasia	−3.787	0.1429	−26.4889	<0.0001	
Wernicke's aphasia	−4.3598	0.1430	−30.4791	<0.0001	
Multiplex CC	0.1589	0.0231	6.8513	<0.0001	
Interaction group × multiplex CC	0.0410	0.0277	1.4813	0.1385	

Healthy controls were used as the reference group to each type of aphasia. All other variables were standardized.
CC, closeness centrality.

group was not significant. That is, multiplex closeness centrality had a facilitative influence on picture naming performance that did not differ between types of aphasia. Additionally, we found that there were relatively small estimated production gaps for the healthy controls compared to people with aphasia, regardless of single-layer or multiplex closeness centralities. Visually, healthy controls also seemed to exhibit a facilitative trend for multiplex closeness centrality on picture naming accuracy. However, the high performance of healthy controls (i.e., relatively few incorrect picture names produced, with an overall accuracy of 97.42%) likely led to a ceiling effect. A more sensitive measure, like reaction time, may be necessary to test the effect of multiplex closeness centrality on picture naming performance by healthy adults.

TABLE 4 *Production gaps for single-layer and multiplex closeness centrality measures for each group*

Estimated production gaps	Healthy controls	Wernicke's aphasia	Broca's aphasia
Free associations	0.04 (0.01)	0.17 (0.01)	0.18 (0.01)
Synonyms	0.002 (0.01)	0.0002 (0.02)	0.005 (0.01)
Taxonomic relations	0.02 (0.02)	0.18 (0.01)	0.14 (0.01)
Phonological similarities	0.005 (0.03)	0.06 (0.02)	0.04 (0.01)
Multiplex lexical representation	0.06 (0.02)	0.49 (0.03)	0.46 (0.02)

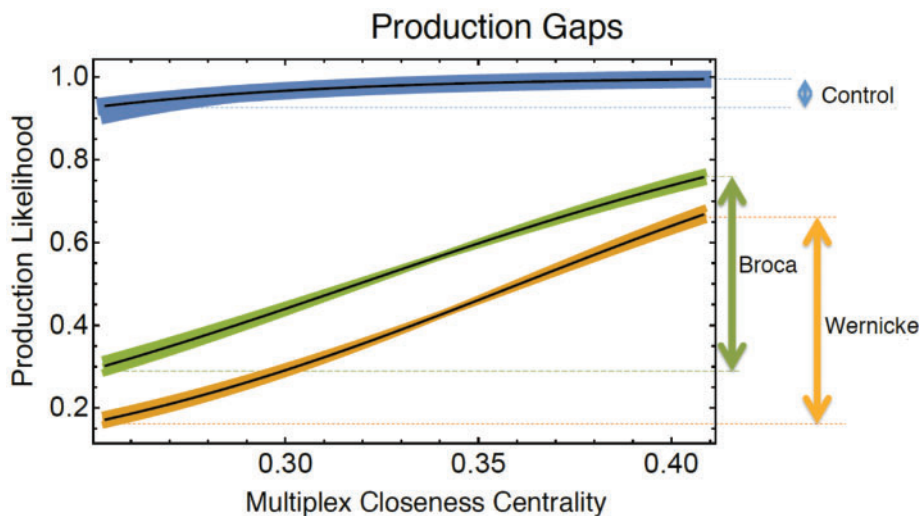Error margins are indicated in parenthesis.

FIG. 3. The estimated production gaps between words with the lowest and highest multiplex closeness centralities for each group. People with aphasia had substantially larger production gaps than healthy controls. The black lines represent the best fit to the actual data, with the coloured curves indicating error margins.

4. Discussion

We applied the multiplex lexical network of [37] to better understand lexical retrieval processes in healthy adults and people with aphasia. We found that multiplex closeness centrality (a measure capturing how well connected a node is to the network using all layers) was a better predictor of picture naming performance than the combination of single-layer closeness centralities. Specifically, picture names with high multiplex closeness centrality were more likely to be accurately produced than picture names with low multiplex closeness centrality, even when controlling for standard psycholinguistic measures and group differences. These logistic regression results were further substantiated by the production gaps analysis, where single-layer closeness centralities had substantially smaller estimated production gaps than multiplex closeness centrality. It is important to note that [16] found *phonological* closeness centrality negatively impacted picture naming accuracy in the same populations (i.e., high phonological closeness centrality words were less likely to be produced correctly than low phonological closeness centrality words). However, that study differs in several ways from the present investigation, including the way

in which closeness centrality was calculated and the lack of controlling for other types of closeness centrality. In contrast, our results strongly indicate that considering the multi-relational nature of the mental lexicon, in terms of both phonology and semantics combined [4, 26], is of utmost importance for modelling lexical retrieval processes.

Furthermore, even when considering the multiplex structure, we found that the global network measure of closeness centrality outperformed the local measure of degree [34]) and the meso-scale measure of PageRank [50, 53]) in predicting picture naming performance. We conjecture that the higher predictive power of closeness centrality compared to the other network measures is due to its sensitivity to short-cuts (i.e., paths connecting two lexical items with relatively few edges) in the mental lexicon [2, 14, 20, 22, 35, 37, 40, 54], although simulations of spreading activation within these language networks are still needed. Importantly, words with high closeness centrality connected by these short-cuts will receive more spreading activation than words with low closeness centrality, thereby increasing the ease of access and probability of successful lexical retrieval. Additionally, the multiplex nature of closeness centrality tested here further exemplifies the importance of being well connected on short-cut paths that traverse the entire (semantic and phonological) lexicon, not just a single layer. That is, words with high multiplex closeness centrality likely receive activation from multiple sources (or layers) than words with low multiplex closeness centrality, and may serve as critical points of interactivity between semantics and phonology. Importantly, the ability to capture the spread of activation in the entire network is critical for understanding how network structure influences lexical retrieval, which cannot be captured with local or meso-scale network measures.

Additionally, our results indicated that the effect of multiplex closeness centrality did not differ between types of aphasia, and perhaps even from healthy controls. This finding was somewhat unexpected given the differences in language difficulties between people with Broca's aphasia and Wernicke's aphasia. Broca's aphasia is characterized by non-fluent speech, spared comprehension and difficulties with repetition, whereas Wernicke's aphasia is characterized by fluent speech with comprehension and repetition difficulties. However, when considering the processes of picture naming, multiple regions of the brain are needed [55], including Broca's and Wernicke's areas. Any impairment in either or both of those sites will lead to picture naming difficulties, although the type of difficulties (e.g., as evidenced by the type of errors produced) will differ (e.g., [56]). Thus, it is quite interesting that a measure of multiplex closeness centrality which considers the location of words in the mental lexicon, not brain regions, was a significant predictor of picture naming performance for both types of aphasia. One interpretation of this result is that the structure of the mental lexicon is not affected by aphasia, but rather the processes acting on the structure (e.g., spread of activation to support lexical retrieval). Furthermore, it may be the case that the type of aphasia influences where in the mental lexicon disruptions in processing occur, opening new important challenges for future studies characterizing different types of aphasia through ad-hoc networks and clinical data.

Although, our results are in agreement with Dell's [27] connectionist model, we want to emphasize that the multiplex lexical network is distinct and propose that the multiplex lexical network could be particularly beneficial to furthering our understanding of aphasia. First, note the differences in how each model is constructed. Dell's model would best be described as a tripartite network containing three types of nodes and two types of edges: semantic/conceptual nodes connected to lemma nodes and lemma nodes connected to phoneme nodes. In contrast, the multiplex lexical network only contains lemma nodes connected via four types of edges: free associations, synonyms, taxonomic dependencies and phonological overlap. Rather than explicitly linking lemmas to their semantic and phonological features, we encapsulate these features through network links and layers (e.g., the fact that two words share meaning is represented as a link between the nodes, rather than as connected through a conceptual node), making the multiplex lexical network more parsimonious in node specification.

Second, the multiplex lexical network is more descriptive in defining word-word similarities than Dell's model, namely through the inclusion of multiple types of semantic relatedness and a stricter definition of phonological overlap. Indeed, the extensive research on semantic networks shows the diversity in types of single-layer semantic networks, including synonym networks [14], free association networks [6] and co-occurrence networks [21], all contributing to the understanding of how network structure influences language processes. Furthermore, Stella *et al*. show that each layer in the multiplex lexical network captures unique features of language (see also Appendix B, where single-layer closeness centralities are not correlated with one another), namely for predicting language learning [35], providing robustness to lexical degradation [37] and highlighting priming effects [54]. Taken together, these results highlight the need for consideration of multiple types of semantic relatedness beyond a single-layer abstract semantic representation, like in Dell's model [27]. Additionally, the multiplex lexical network also accounts for the commonly found phonological neighbourhood density effects [10, 49, 57], by placing edges between words that differ by only one phoneme. This is in contrast to Dell's model where words would be connected through as many phonological nodes as the number of overlapping phonemes. Although Dell's model can use proportionally incremental increases in phonological relatedness to account for priming effects [27], the multiplex lexical network can also account for analogous phonological priming effects [54] even with a stricter definition of phonological overlap. Given the greater descriptiveness of the multiplex lexical network as compared to Dell's model, we believe that this approach better specifies how several semantic and phonological word-word similarities in the mental lexicon affect lexical retrieval.

Lastly, stemming from Dell's interactive model are two possible theories of how aphasia impacts the mental lexicon: (i) the weight-decay model [27] and (ii) the semantic-phonological model [58, 59]. In the weight-decay model, aphasia impacts how much activation spreads in the network (global connection weight) and how quickly that activation dissipates (decay rate). Alternatively, in the semantic-phonological model, aphasia impacts how much activation spreads from conceptual nodes to lemma nodes (semantic weight) and how much activation spreads from lemma nodes to phonological nodes (phonological weight). We propose that with the use of the multiplex lexical network, these two theories could be tested in a unique way. For example, the present use of the multiplex lexical network considers all layers simultaneously and *equivalently*. However, it would be possible to weight individual layers to test whether aphasia impacts 'global connection weights' or separate 'semantic and phonological weights'. Indeed, recent work by Stella *et al*. [35] used machine learning techniques to quantify the role of semantic and phonological layers on word acquisition. A similar approach could be used to identify specific types of aphasia based on the weightings of layers in the multiplex lexical network. Additionally, through simulation studies and dynamical network models, researchers could also quantify the spread of activation and its decay rate within the multiplex lexical network. Indeed, we conjecture that aphasia may alter both structure and process of multiplex lexical networks.

In sum, we found that a multiplex lexical network accounted for picture naming accuracy of healthy adults and people with aphasia, particularly when utilizing a network measure of closeness centrality. This work represents the first step in utilizing multiplex lexical networks to better understand lexical retrieval by people with aphasia and paves the way for future research in this area. We conjecture that this application of network science will be particularly useful for developing not just psycholinguistic theory, but also clinical applications, like diagnostic measures and treatment interventions.

Acknowledgements

REFERENCES

1. FEDORENKO, E. & THOMPSON-SCHILL, S. L. (2014) Reworking the language network. *Trends Cogn. Sci.,* **18**, 120–126.

2. STEYVERS, M. & TENENBAUM, J. B. (2005) The large-scale structure of semantic networks: statistical analyses and a model of semantic growth. *Cogn. Sci.*, **29**, 41–78.

3. BARONCHELLI, A., FERRER-I CANCHO, R., PASTOR-SATORRAS, R., CHATER, N. & CHRISTIANSEN, M. H. (2013) Networks in cognitive science. *Trends Cogn. Sci.*, **17**, 348–360.

4. AITCHISON, J. (2012) *Words in the Mind: An Introduction to the Mental Lexicon*. New York, United States: John Wiley & Sons.

5. DEYNE, S. DE & STORMS, G. (2008) Word associations: network and semantic properties. *Behav. Res. Meth.*, **40**, 213–231.

6. DE DEYNE, S., NAVARRO, D. J. & STORMS, G. (2013) Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behav. Res. Meth.*, **45**, 480–498.

7. AMANCIO, D. R., OLIVEIRA, O. N. JR. & DA F COSTA, L. (2012) Using complex networks to quantify consistency in the use of words. *J. Stat. Mech. Theor. Exp.*, 2012:P01004.

8. KENETT, Y. N., ANAKI, D. & FAUST, M. (2014) Investigating the structure of semantic networks in low and high creative persons. *Front. Human Neurosci.*, **8**, 407.

9. HILLS, T. T., MAOUENE, M., MAOUENE, J., SHEYA, A. & SMITH, L. (2009) Longitudinal analysis of early semantic networks preferential attachment or preferential acquisition? *Psychol. Sci.*, **20**, 729–739.

10. VITEVITCH, M. S. (2008) What can graph theory tell us about word learning and lexical retrieval? *J. Speech Lang. Hear. Res.*, **51**, 408–422.

11. STORKEL, H. L. (2009) Developmental differences in the effects of phonological, lexical and semantic variables on word learning in infants. *J. Child Lang.*, **36**, 291–321.

12. CARLSON, M. T., SONDEREGGER, M. & BANE, M. (2014) How children explore the phonological network in child-directed speech: a survival analysis of children's first word productions. *J. Mem. Lang.*, 75:159–180.

13. VITEVITCH, M. S., GOLDSTEIN, R., SIEW, C. S. Q. & CASTRO, N. (2014) Using complex networks to understand the mental lexicon. *Yearb. Pozn. Linguist. Meet.*, **1**, 119–138.

14. SIGMAN, M. & CECCHI, G. A. (2002) Global organization of the Wordnet lexicon. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 1742–1747.

15. GORDON, J. K. (2002) Phonological neighborhood effects in aphasic speech errors: spontaneous and structured contexts. *Brain Lang.*, **82** 113–145.

16. VITEVITCH, M. S. & CASTRO, N. (2015) Using network science in the language sciences and clinic. *Int. J. Speech Lang. Pathol.*, **17**, 13–25.

17. VITEVITCH, M. S. & SOMMERS, M. S. (2003) The facilitative influence of phonological similarity and neighborhood frequency in speech production in younger and older adults. *Mem. Cogn.*, **31**, 491–504.

18. SIZEMORE, A. E., KARUZA, E. A., GIUSTI, C. & BASSETT, D. S. (2018) Knowledge gaps in the early growth of semantic feature networks. *Nat. Human Behav.*, **2**, 682.

19. THOMAS, H. T. & SIEW, C. S. Q. (2018) Filling gaps in early word learning. *Nat. Human Behav.*, **2**, 622–623.

20. BECKAGE, N., SMITH, L. & HILLS, T. (2011) Small worlds and semantic network growth in typical and late talkers. *PLoS One*, **6**, e19348.

21. FERRER I CANCHO, R. & SOLÉ, R. V. (2001) The small world of human language. *Proc. Biol. Sci.*, **268**, 2261–2265.

22. KENETT, Y. N., LEVI, E., ANAKI, D. & FAUST, M. (2017) The semantic distance task: quantifying semantic distance with semantic network path length. *J. Exp. Psychol. Learn. Mem. Cognit.*, **43**, 1470.

23. NEWMAN, M. (2010)*Networks: An Introduction*. Oxford, UK: Oxford University Press.

24. Goldstein, R. & Vitevitch, M. S. (2017) The influence of closeness centrality on lexical processing. *Front. Psychol.*, **8**, 1683.

25. Collins, A. M. and Loftus, E. F. (1975) A spreading-activation theory of semantic processing. *Psychol. Rev.*, **82**, 407.

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

27. Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M. & Gagnon, D. A. (1997) Lexical access in aphasic and nonaphasic speakers. *Psychol. Rev.*, **104**, 801–838.

28. Martin, N., Gagnon, D. A., Schwartz, M. F., Dell, G. S. & Saffran, E. M. (1996) Phonological facilitation of semantic errors in normal and aphasic speakers. *Lang. Cogn. Process.*, **11**, 257–282.

29. Rapp, B. & Goldrick, M. (2000) Discreteness and interactivity in spoken word production. *Psychol. Rev.*, **107**, 460–499.

30. Dell, G. S. (1986) A spreading-activation theory of retrieval in sentence production. *Psychol. Rev.*, **93**, 283.

31. De Domenico, M., Solé-Ribalta, A., Cozzo, E., Kivelä, M., Moreno, Y., Porter, M. A., Gómez, S. & Arenas, A. (2013) Mathematical formulation of multilayer networks. *Phys. Rev. X*, **3**, 041022, 2013.

32. Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y. & Porter, M. A. (2014) Multilayer networks. *J. Complex Netw.*, **2**, 203–271.

33. Boccaletti, S., Bianconi, G., Criado, R., Del Genio, C. I. Gómez-Gardeñes, J., Romance, M., Sendina-Nadal, I., Wang, Z. & Zanin, M. (2014) The structure and dynamics of multilayer networks. *Phys. Rep.*, **544**, 1–122.

34. Battiston, F., Nicosia, V. & Latora, V. (2017) The new challenges of multiplex networks: measures and models. *Eur. Phys. J. Spec. Top.*, **226**, 401–416.

35. Stella, M., Beckage, N. M. & Brede, M. (2017) Multiplex lexical networks reveal patterns in early word acquisition in children. *Sci. Rep.*, **7**.

36. Stella, M. & Brede, M. (2016) Mental lexicon growth modelling reveals the multiplexity of the English language. *Complex Networks VII* (Aiello, L. M., Cherifi, C., Cherifi, H., Lambiotte, R., Lio, P. & Rocha, L. M. eds). New York, United States: Springer, pp. 267–279.

37. Stella, M., Beckage, N. M., Brede, M. & De Domenico, M. (2018) Multiplex model of mental lexicon reveals explosive learning in humans. *Sci. Rep.*, **8**, 2259.

38. Ginsburg, H. P. and Opper, S. (1988) *Piaget's Theory of Intellectual Development*. Engelwood Cliffs, United States: Prentice-Hall, Inc., https://psycnet.apa.org/record/1987-98474-000.

39. Frith, U. (1985) Beneath the surface of developmental dyslexia. *Surf. Dys.*, **32**, 301–330.

40. Stella, M. & De Domenico, M. (2018) Distance entropy cartography characterises centrality in complex networks. *Entropy*, **20**, 268.

41. Mirman, D., Strauss, T. J., Brecher, A., Walker, G. M., Sobel, P. Dell, G. S. & Schwartz, M. F. (2010) A large, searchable, web-based database of aphasic performance on picture naming and other tests of cognitive function. *Cogn. Neuropsychol.*, **27**, 495–504.

42. Roach, A., Schwartz, M. F., Martin, N., Grewal, R. S. & Brecher, A. (1996) The Philadelphia naming test: scoring and rationale. *Clin. Aphasiol.*, **24**, 121–133.

43. Nelson, D. L., McEvoy, C. L. & Dennis, S. (2000) What is free association and what does it measure? *Mem. Cogn.*, **28**, 887–899.

44. Kiss, G. R., Armstrong, C. A. & Milroy, R. (1972) (1972) *An Associative Thesaurus of English*. Medical Research Council, Speech and Communication Unit, Edinburgh, UK: University of Edinburgh.

45. Nelson, D. L., McEvoy, C. L. & Schreiber, T. A. (2004) The University of South Florida free association, rhyme, and word fragment norms. *Behav. Res. Methods Instrum. Comput.*, **36**, 402–407.

46. Collins, A. M. & Quillian, M. R. (1969) Retrieval time from semantic memory. *J. Verbal Learning Verbal Behav.*, **8**, 240–247.

47. Miller, G. A. (1995) WordNet: a lexical database for English. *Commun. ACM*, **38**, 39–41.

48. WolframResearch. WordData Source Information. https://reference.wolfram.com/language/ref/WordData.html (last accessed 14 May 2017).

49. LUCE, P. A. AND PISONI, D. B. (1998) Recognizing spoken words: the neighborhood activation model. *Ear Hear.*, **19**, 1.

50. DE DOMENICO, M., SOLÉ-RIBALTA, A., GÓMEZ, S. & ARENAS, A. (2014) Navigability of interconnected networks under random failures. *Proc. Natl. Acad. Sci. USA.*, **111**, 8351–8356.

51. BARBARESI, A. (2014) Language-classified open subtitles (LACLOS): download, extraction, and quality assessment. *Ph.D. Thesis*, BBAW.

52. KUPERMAN, V., STADTHAGEN-GONZALEZ, H. & BRYSBAERT, M. (2012) Age-of-acquisition ratings for 30,000 english words. *Behav. Res. Methods*, **44**, 978–990.

53. GRIFFITHS, T. L., STEYVERS, M. & FIRL, A. (2007) Google and the mind predicting fluency with PageRank. *Psychol. Sci.*, **18**, 1069–1076.

54. STELLA, M. (2018) Cohort and rhyme priming emerge from the multiplex network structure of the mental lexicon. *Complexity*, 2018, Article ID 6438702, 14 pages, https://doi.org/10.1155/2018/6438702.

55. DELEON, J., GOTTESMAN, R. F., KLEINMAN, J. T., NEWHART, M. DAVIS, C., HEIDLER-GARY, J., LEE, A. & HILLIS, A. E. (2007) Neural regions essential for distinct cognitive processes underlying picture naming. Oxford, UK, *Brain*, **130**, 1408–1422.

56. KOHN, S. E. & GOODGLASS, H. (1985) Picture-naming in aphasia. *Brain Lang.*, **24**, 266–283.

57. VITEVITCH, M. S. & LUCE, P. A. (2016) Phonological neighborhood effects in spoken word perception and production. *Annu. Rev. Linguist.*, **2**, 75–94.

58. FOYGEL, D. & DELL, G. S. (2000) Models of impaired lexical access in speech production. *J. Mem. Lang.*, **43**, 182–216.

59. SCHWARTZ, M. F., DELL, G. S., MARTIN, N., GAHL, S. & SOBEL, P. (2006) A case-series test of the interactive two-step model of lexical access: evidence from picture naming. *J. Mem. Lang.*, **54**, 228–264.

Appendix A. List of words tested in the multiplex network

Only 142 out of 175 words from the Philadelphia Naming Task (PNT) were found in the multiplex lexical network and analysed (see Table A.1 for a list of the missing words). We noticed two important characteristics in the 33 missing words as compared to the 142 tested words. First, many of the missing words include many specific nouns, such as types of animals (e.g., zebra, camel, elephant, squirrel) or types of tools (e.g., crutches, binoculars, thermometer, towel). Additionally, we found that the missing words had significantly lower word frequency (based on the OpenSubtitles dataset; Barbaresi, 2014 [51]) than the tested words using a Kruskal–Wallis test (6.35, p = 0.01). In order to test the resilience of our results despite missing 33 of the PNT test items, we performed the logistic regression of our best performing model (Model 2c) over subsamples of our data where 33 words were removed at random. Averaging over 20 randomizations (where 33 words were removed at random and the logistic regression was performed over the remaining 109 words) led to models with an average $R^2 = 0.2144 \pm 0.006$, which is compatible with the R^2 of the whole dataset including 142 words. Hence, we expect that the exclusion of the 33 missing words did not influence our results.

TABLE A.1 *Words from the Philadelphia Naming Task that were not present in the multiplex lexical network or included in analyses*

ambulance	binoculars	bride	calendar	camel	cheerleaders	crutches	desk
dinosaur	elephant	Eskimo	flashlight	kitchen	kite	lamp	microscope
mustache	necklace	pencil	pie	pumpkin	pyramid	sandwich	skis
skull	squirrel	stethoscope	thermometer	towel	tractor	typewriter	vase
zebra							

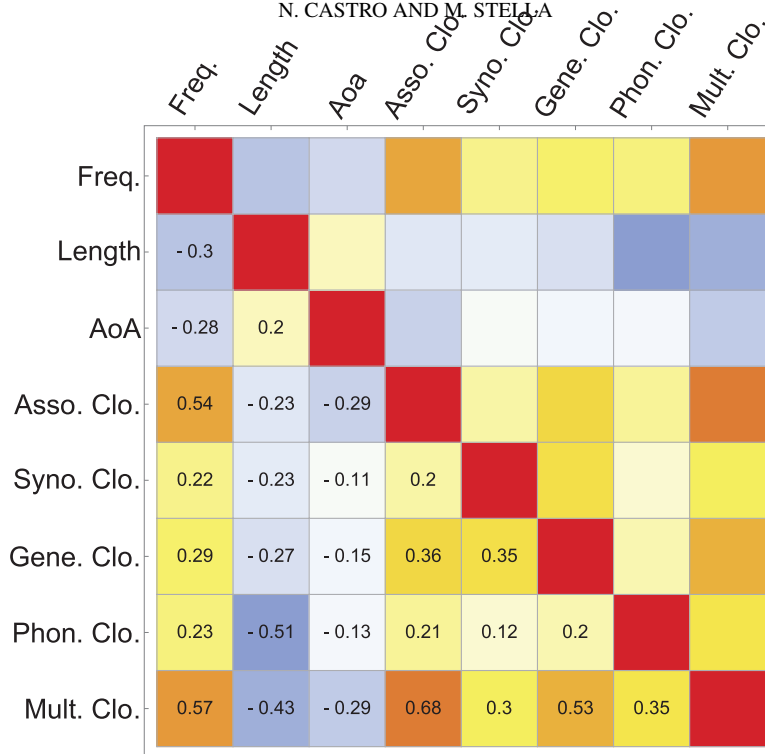|	Freq.	Length	Aoa	Asso. Clo.	Syno. Clo.	Gene. Clo.	Phon. Clo.	Mult. Clo.
Freq.								
Length	- 0.3							
AoA	- 0.28	0.2						
Asso. Clo.	0.54	- 0.23	- 0.29					
Syno. Clo.	0.22	- 0.23	- 0.11	0.2				
Gene. Clo.	0.29	- 0.27	- 0.15	0.36	0.35			
Phon. Clo.	0.23	- 0.51	- 0.13	0.21	0.12	0.2		
Mult. Clo.	0.57	- 0.43	- 0.29	0.68	0.3	0.53	0.35	

Fɪɢ. B.1. Kendall Tau correlations between psycholinguistic variables and closeness centrality measures of words tested in the main text. Negative correlations are highlighted in blue and positive correlations are highlighted in gold, with darker colours representing larger absolute values. All correlations were statistically significant at a 0.1 significance level.

Appendix B. Ranges and correlations of predictor variables

We report in Fig. B.1 the Kendall Tau correlations between traditional psycholinguistic variables (word frequency, word length and age of acquisition) and network measures of closeness centrality (single-layer and multiplex) used as predictors in the main text. All correlations were statistically significant at a 0.1 significance level. Table B.2 reports ranges, means and standard deviations of the non-standardized predictor variables used in the article. As evident from the standard deviations, closeness centrality in the synonym and generalization layers is highly skewed. On the multiplex network, the distribution of closeness centrality is compatible with a Gaussian distribution at a 0.1 significance level (Kolmogorov–Smirnov test, $s = 0.0735$, $p = 0.0581$).

TABLE B.2 *Descriptors of predictor variables: lower and upper bound, mean and standard deviation*

	Minimum	Maximum	Mean	Standard
Frequency	2.5	8.85	4.768	1.254
Length	3	10	4.863	1.47
AoA	5.193	13.586	8.3	1.445
Asso. Clo.	0	0.372	0.285	0.06
Syno. Clo.	0	1	0.208	0.252
Gene. Clo.	0	0.322	0.251	0.064
Phon. Clo.	0	1	0.146	0.115
Mult. Clo.	0.253	0.408	0.339	0.023

AoA, age of acquisition; Asso. Clo., closeness in the association layer; Syno. Clo., closeness in the synonyms layer; Gene. Clo., closeness in the generalisation layer; Phon. Clo., closeness in the phonological layer; Mult. Clo., closeness in the whole multiplex network.