CLINICAL LINGUISTICS & PHONETICS
2026, VOL. 40, NO. 2, 187–205
https://doi.org/10.1080/02699206.2025.2507049

Lexical network properties in aphasic discourse production

Juqiang Chen[a,b] and Hui Chang[a,b]

[a]School of Foreign Languages, Shanghai Jiao Tong University, Shanghai, P. R. China; [b]National Research Centre for Language and Well-being, Shanghai Jiao Tong University, Shanghai, P. R. China

ABSTRACT
Evaluating discourse production provides an effective means of assessing the language competence of persons with aphasia (PWA) in naturalistic and ecologically valid communication contexts. Existing research on PWA discourse has focused on discrete linguistic measures such as sample length or type-token ratios. To provide a holistic view of the interconnections of linguistic entities in PWA discourse, this study compared the lexical networks of discourse produced by PWA with those by healthy controls. At the macro level, PWA networks showed a smaller size and diameter but higher density compared to those of healthy controls. In addition, networks of both groups showed small-world characteristics. However, the degree of small-worldness was lower for the networks of PWA compared to those of healthy controls. At the meso level, network modularity measures of healthy controls were larger than those of PWA in the story-narrative task. At the micro level, the degrees of nodes in the networks of both groups followed a power-law distribution, suggesting that their networks were scale-free. Our findings reveal that lexical networks of PWA discourse are constrained and less efficient but still retain some basic characteristics of typical language networks.

ARTICLE HISTORY
Received 22 July 2024
Revised 21 April 2025
Accepted 11 May 2025

KEYWORDS
Aphasic discourse; lexical networks; small-world networks; modularity

Introduction

Persons with aphasia (PWA) experience varying degrees of difficulty in lexical retrieval in discourse production, which significantly impacts their ability to effectively communicate in everyday life. Given the importance of discourse production in aphasia research (Dipper et al., 2021; Dipper et al., 2017; Pritchard et al., 2017, 2018), a variety of linguistic measures have been employed to investigate discourse production in PWA, such as language productivity (e.g. sample length, lexical diversity, Fergadiotis & Wright, 2011; Fergadiotis et al., 2013), information content (e.g. correct information unit analysis or core lexicon analysis, Chen & Chang, 2024; Dalton et al., 2022; Kim & Wright, 2020) and grammatical complexity (e.g. morphological and syntactic complexity measures) (see Bryant et al., 2016, for a review of linguistic analysis of discourse in aphasia).

Although existing aphasic discourse analyses have revealed important characteristics of PWA language deficits using these discrete measures, there is a lack of a holistic view of the interconnection of entities in the PWA language system. Language is a system of interdependent terms in which the value of each term results solely from the simultaneous presence of the others (Saussure, 1959). Thus, it is argued that the human language system

CONTACT Hui Chang ch9647@sjtu.edu.cn School of Foreign Languages, Shanghai Jiao Tong University, No. 800 Dongchuan Road, Minhang District, Shanghai 200240, P. R. China

should be modelled as an interconnected complex system in which the whole is often greater than the sum of its parts (Vitevitch & Castro, 2015).

Modern network science offers techniques to model human language as a complex system with nodes being words (Amancio, 2015) or paragraphs (de Arruda et al., 2019) and links being their phonological (Vitevitch & Castro, 2015), semantic (de Arruda et al., 2019) and syntactic (Abramov & Mehler, 2011; Fan & Jiang, 2021; Liu, 2008) relations. Measures of these networks can reveal important interactions among entities in the language system and are useful for studying language classification (Abramov & Mehler, 2011), detecting stylistic variations (Amancio, 2015; de Arruda et al., 2019; Fan & Jiang, 2021) and extracting keywords (Blanco & Lioma, 2012; Yang et al., 2018). Recently, the use of network analysis has been extended to understanding language development (Beckage et al., 2011), language impairment (Castro et al., 2020) and age-related cognitive decline (Liu et al., 2021).

Building on these advancements, lexical networks constructed from discourse data have emerged as a novel approach to studying discourse production by people with cognitive decline (e.g. Alzheimer's disease or AD, Liu et al., 2021; Malcorra et al., 2021), attention-deficit hyperactivity symptoms (Coelho et al., 2021) and children with developmental disorders (Beckage et al., 2011). In such lexical networks, words (as nodes) are connected by edges that denote semantic relationships derived from co-occurrence within a specific span of discourse (Coelho et al., 2021; Mota et al., 2023; Wießner et al., 2023). In one study with children (Beckage et al., 2011), lexical networks were built with nodes representing words used by children as reported by their parents and links between them derived from the co-occurrence statistics of these words in a child-directed speech corpus (i.e. CHILDES, MacWhinney, 2016). The lexical network of the typically-developing children had a higher clustering coefficient, a shorter average path length and displayed greater small-world characteristics relative to the late-talking children. In another study, lexical networks created from narratives produced by AD patients were less connected than cognitively healthy older adults (Malcorra et al., 2021). Nevertheless, very few aphasia studies have used network modelling.

Applying this network-based approach to aphasic discourse can reveal patterns of word retrieval deficits or preserved connections. Specifically, network measures, such as network size, density, diameter, small-world properties, modularity and scale-free properties, are expected to complement previous linguistic metrics by capturing how lexical items interact and form meaningful patterns in discourse production (for an overview of network measures applied to cognitive and/or language networks, see Siew et al., 2019; Steyvers & Tenenbaum, 2005). For instance, network density, which measures the proportion of actual connections between words relative to all possible connections, offers insights into the strength of word connections, potentially highlighting weakened or disrupted semantic linkages in PWA discourse. In addition, small-world networks are characterised by strong local connectivity and efficient long-distance connections, which may be reduced in PWA discourse, indicating impairments in the efficient organisation of lexical resources. Modularity, the degree to which the network can be divided into distinct clusters or communities of words, reflects thematic or semantic organisation of discourse. Low modularity may suggest difficulties in organising the content of a picture or story into meaningful lexical clusters. Scale-free networks, where a small number of highly connected nodes (hub words) coexist with many less connected nodes, indicate optimised language processing. Deviations from this in the PWA network may reflect deficits in the accessibility of critical lexical hubs.

In one aphasic network study (Castro et al., 2020), a multiplex lexical network was built on free association, synonym, general-specific and phonological similarity relations. Network measures, such as degree (the number of connections per node) and closeness centrality (the inverse of the average of the length of the shortest path between the node and all other nodes in the network), were then used as predictors for picture naming performance of aphasic patients (Castro et al., 2020). They found that both naming accuracy and naming error types were affected by the degree and closeness centrality of the target word. In addition, error types also varied as a function of the network distance between the target word and the error word. It should be noted that their networks were not based on discourse production data. Thus, the properties of PWA lexical networks based on discourse data remain an unresolved issue.

In addition, discourse types, such as picture descriptions, story narratives and procedural descriptions, are reported to affect the linguistic characteristics of discourse production. Single-picture description tasks, for example, elicit discourse production with very limited lexical diversity relative to multi-picture description tasks for both aphasia and control groups (Stark, 2019). Storytelling narrative tasks based on familiar stories (such as Cinderella) elicit more complex language content compared to picture description tasks as well as procedure description tasks (Stark, 2019). According to the Triadic Componential Framework (Robinson & Gilabert, 2007), these tasks differ in their task complexity, which can be used to account for task-induced variations in discourse production. Task complexity has two dimensions: resource-directing and resource-dispersing. The resource-directing dimension includes task characteristics related to concepts, such as, relative time, spatial location, causal relationships and intentionality. Different tasks, for example, vary in the degree of reasoning required for causal events and internal states. Storytelling involves extensive reasoning about motivations and consequences, while picture description requires less, and procedural discourse is primarily objective. These differences influence lexical organisation and how speakers structure their discourse production.

In addition, resource-dispersing variables, such as planning time or background knowledge, can also affect task performance. The picture description task directly provides background knowledge for the task, while storytelling and procedural tasks require speakers to rely more on their existing knowledge. The storytelling task is expected to be especially demanding, as it offers less background information and imposes higher cognitive and memory loads. PWA is predicted to struggle more with the storytelling task relative to the control group, resulting in less effective use of their lexical resources.

However, there has been no comprehensive research on lexical networks of aphasic discourse production and how different elicitation stimuli/tasks affect lexical network properties. The present study used data from the AphasiaBank (MacWhinney & Fromm, 2016; MacWhinney et al., 2011) to construct lexical networks for three different discourse types and compared network properties between control and aphasia groups.

Method

Data

Discourse production data utilised in this study were retrieved from the AphasiaBank (MacWhinney & Fromm, 2016; MacWhinney et al., 2011). We selected data of picture description, story narrative and procedural discourse out of the five task types in the original protocol because these tasks are more informative and comparable among participants than the greeting part (which provides very limited data) and free speech (which shows huge variations among participants). The Broken Window multi-picture description task, which is more comparable to the story narrative task, was selected over other single-picture tasks such as Cat Rescue and Flood. The story narrative task required participants to recount the Cinderella story, while the procedural discourse task required participants to describe how to make a peanut butter sandwich. Detailed guidelines pertaining to task administration and transcription of discourse samples are available on the AphasiaBank website.

Transcripts of 78 English-speaking PWAs were retrieved from the AphasiaBank database (see Table 1). Participants included in the study were diagnosed using WAB-R (Kertesz, 2006) and were identified as anomic ($n = 34$), Broca's ($n = 21$), conduction ($n = 16$), transcortical motor ($n = 2$), transcortical sensory ($n = 1$) and Wernicke's ($n = 4$) aphasia. Transcripts from a cohort of 78 healthy English speakers were chosen from the control group within the same database to serve as a baseline for comparison with PWAs. Statistical analysis revealed no significant differences between the two groups in age and years of education.

Data pre-processing

In the present study, we used the Computerized Language Analysis (CLAN) program to extract transcriptions of clear speech (no repetitions, repairs or other types of disfluencies as annotated) (MacWhinney et al., 2011) and these transcriptions were tokenised into words via the *tidytext* package (Silge & Robinson, 2016) in *R* (R Core Team, 2018). These words were then used as nodes in the network construction.

To establish the semantic connections (or vertices) between words, we first utilised a word embedding model from *fasttext* (Grave et al., 2018) pre-trained on large-scale data to represent each word in a 300-dimension space. Words that co-occur in comparable contexts tend to have closer proximity in the vector space. This property allows the word embedding model to effectively capture semantic relationships between words. Then, we calculated the cosine similarity between two words as vertices, which measures the cosine of the angle between two vectors in the 300-dimension space. To make our models

Table 1. Demographic information for all individuals. AQ = aphasia quotient.

Group	n	Age (yrs)		WAB AQ		Years of education	
		M	*SD*	*M*	*SD*	*M*	*SD*
Control	78	51.9	11.1	–	–	15.8	2.5
PWA	78	54.3	7.1	74.1	14.7	15.1	2.4

computationally tractable and efficient, words with low (less than 0.7) and negative cosine values (as in Xu et al., 2021) were not linked.

Network properties

In brain science, three dimensions of brain networks, i.e. the order, hierarchy and degree diversity dimensions, are considered important for differentiating between normal and impaired brain networks (Stam, 2014). For the order dimension, normal brain networks have strong local connectivity and efficient long-distance connections (i.e. small-world networks). For the hierarchy dimension, normal brain networks show hierarchical modularity (Meunier et al., 2010), in which different modules align with major functional systems. For the degree diversity dimension, normal brain networks are scale-free, with a number of highly connected hubs. Assuming the brain network is the biological basis for language, we examined lexical networks based on discourse production for these three dimensions by measuring network small-worldness, modularity and scale-free properties.

Small-world index

Many language networks, such as semantic networks (Steyvers & Tenenbaum, 2005), are reported to show small-world characteristics. In order to measure network small-world characteristics, we first constructed random networks with the same number of nodes and edges for each discourse production task. Two measures are essential in quantifying small-worldness: Average Shortest Path Length (ASPL) and global clustering coefficient (C). ASPL refers to the mean of the shortest possible path between all possible pairs of nodes in the network. The global clustering coefficient is calculated as the average of the local clustering coefficients of all nodes (i.e. the proportion of connections between its neighbours divided by the total possible connections among them). A small-world network has a shorter ASPL and a larger clustering coefficient than its reference random network with the same number of nodes and edges.

The index we used to quantify small-worldness is the small-world quotient (Q) (Davis et al., 2003; Uzzi & Spiro, 2005) and is widely used in different fields (Bassett et al., 2008; Bullmore & Sporns, 2009). If Q is greater than 1, the network is considered small-world (Humphries et al., 2008). In other words, the network's ASPL is comparable to that of a random reference graph and/or that the network's clustering coefficient is larger than a random reference graph.

$$Q = \frac{C/C_r}{ASPL/ASPL_r}$$

C_r and $ASPL_r$ refer to the global clustering coefficient and mean average shortest path length of a random reference graph.

Modularity

In many real-world networks (e.g. lexical semantic network, Wulff et al., 2022), community structures are observed, in which nodes are grouped into distinct communities based on the relatively high density of connections within the same community compared to connections between nodes from different communities (Newman, 2006). Modularity serves as a quantitative measure of the density of connections within communities relative to that of connections across different communities (Fortunato, 2010). Networks with higher modularity have a strong community structure.

Various community detection techniques have been devised to identify communities in the network. In the present study, four different methods were used for community detection via functions in the *igraph* package in R (Csardi & Nepusz, 2006). The utilisation of different methods enables us to triangulate data, enhancing the validity of the findings through corroborative evidence.

The edge-betweenness method operates on the principle that edges connecting distinct communities typically have high edge-betweenness values, as they are on numerous shortest paths between nodes of different communities. The algorithm begins by calculating the edge-betweenness for all edges in the graph. It then removes the edge with the highest edge-betweenness score and recalculates the edge-betweenness for the remaining edges. This iterative process of removing the edge with the highest edge-betweenness and recalculating continues until no further improvement in modularity is possible.

The Louvain method is based on the principle that communities can be viewed as aggregates of smaller communities (Blondel et al., 2008). Initially, each node is assigned to its own unique community, resulting in as many communities as there are nodes. The algorithm then iteratively evaluates each node, relocating it to the neighbouring community that provides the greatest increase in modularity. Following this reallocation, a new network is constructed where the nodes represent the communities identified in the previous phase. This process repeats until no further improvement in modularity is achievable.

The random walker method is based on the idea that a random walker spends more time within a community rather than between distinct communities. The algorithm thus groups nodes together based on the similarities of the paths traversed by a random walker starting from each node.

The Infomap method uses the probability flow of random walks on a network as a proxy for information flow and identifies modules by compressing the probability flow (Rosvall & Bergstrom, 2008).

Scale-free characteristics

In scale-free networks, the degrees of nodes follow a power law distribution. To test whether the degrees of nodes in our language networks follow a power law distribution, we used *fit_power_law* function in the *igraph* package. The Kolmogorov–Smirnov test was used to compare the fitted distribution with the input vector. Smaller scores indicate a better fit. The *p*-values of the Kolmogorov–Smirnov test were also reported. Small *p*-values (less than 0.05) indicate the rejection of the null hypothesis (i.e. the data followed a power-law distribution). The emergence of the scale-free property appears to be influenced by the type of linguistic network, with evidence supporting its presence in semantic association networks (Steyvers & Tenenbaum, 2005) but not in phonological networks (Vitevitch, 2008).

Results

Lexical networks of discourse production by PWA and healthy controls

First, we aggregated discourse production data and constructed lexical semantic networks for each task and each group (Figure 1). Visual inspection reveals that story narrative networks exhibited overall higher density and complexity than those of picture descriptions, which in turn displayed higher density and complexity than those of procedural discourse. There were also noticeable structural differences between PWA and control groups.

In order to formally compare different networks, we first calculated network measures, i.e. network size, density, diameter, for each participant and each task. Then these network properties were selected as dependent variables and each fitted with a separate linear mixed-effects model using the *lme4* package (Bates et al., 2015) in (R Core Team, 2018). Participant groups (PWA versus healthy controls) and task types (picture description, story narrative, procedural discourse) were selected as fixed factors. Participants were specified as a random factor (intercept). Three models were built to test all the main effects and interactions. For each model, we used the Kenward-Roger degrees of freedom approximation (Halekoh & Hojsgaard, 2014) to estimate p-values for the fixed effects of interest. Here, we report main effects for multilevel factors or interactions using F-tests via the Anova function from the *car* package (Fox & Weisberg, 2019) in R (see Table 2), rather than comparisons with the baseline level using t-tests. Thus, the reported main and interaction effects were averaged across all levels of the other effects.

There were significant main effects of the group, $F(1, 154) = 155$, $p < 0.001$ and task types $F(2, 308) = 519$, $p < 0.001$ for the network size. In addition, there was a group × task interaction, $F(2, 308) = 129$, $p < 0.001$. To further examine the main effects and the interaction, we ran pairwise multiple t-tests with Tukey adjustments. For the main effect of the group, lexical networks of healthy controls ($M = 92.4$, 95% CI [86.6, 98.1]) were larger than those of PWA ($M = 40.9$, 95% CI [35.2, 46.7]), $\beta = 51.40$, $SE = 4.13$, $t(154) = 12.46$, $p < 0.001$. For the main effect of task type, lexical networks of story narratives ($M = 130.0$, 95% CI [124.4, 135.7]) were larger than those of picture descriptions ($M = 39.1$, 95% CI [33.4, 44.7]), $\beta = 90.99$, $SE = 3.42$, $t(308) = 26.6$, $p < 0.001$, which in turn were larger than those of procedural discourse ($M = 30.8$, 95% CI [25.2, 36.4]), $\beta = -8.23$, $SE = 3.42$, $t(308) = -2.41$, $p = 0.04$. For the task type by group interaction, we report here the comparisons between two groups within the same task. It was found that lexical networks of healthy controls were larger than those of PWA for story narratives, $\beta = 114.65$, $SE = 5.71$, $t(98) = -20.07$, $p < 0.01$, and picture descriptions, $\beta = 24.65$, $SE = 5.71$, $t(398) = 4.32$, $p < 0.01$, but not for procedural discourse, $\beta = 15.04$, $SE = 5.71$, $t(398) = 2.63$, $p = 0.09$.

For the density of networks, we found significant main effects of group, $F(1, 154) = 94$, $p < 0.001$ and task type, $F(2, 308) = 75$, $p < 0.001$, but no group × task interaction. For the main effect of the group, lexical networks of healthy controls ($M = 0.15$, 95% CI [0.14, 0.16]) were less dense than those of PWA ($M = 0.21$, 95% CI [0.20, 0.22]), $\beta = -0.06$, $SE = 0.006$, $t(154) = -9.69$, $p < 0.001$. For the main effect of task type, the network density of story narratives ($M = 0.13$, 95% CI [0.12, 0.14]) was smaller than that of picture descriptions ($M = 0.20$, 95% CI [0.19, 0.21]), $\beta = -0.06$, $SE = 0.007$, $t(308) = -8.72$, $p < 0.001$, which in turn was smaller than that of procedural discourse ($M = 0.22$, 95% CI [0.21, 0.23]), $\beta = 0.02$, $SE = 0.007$, $t(308) = 3.13$, $p = 0.005$. These findings seemingly contradict the visual inspection of the group-level networks and need to be interpreted properly. First, the

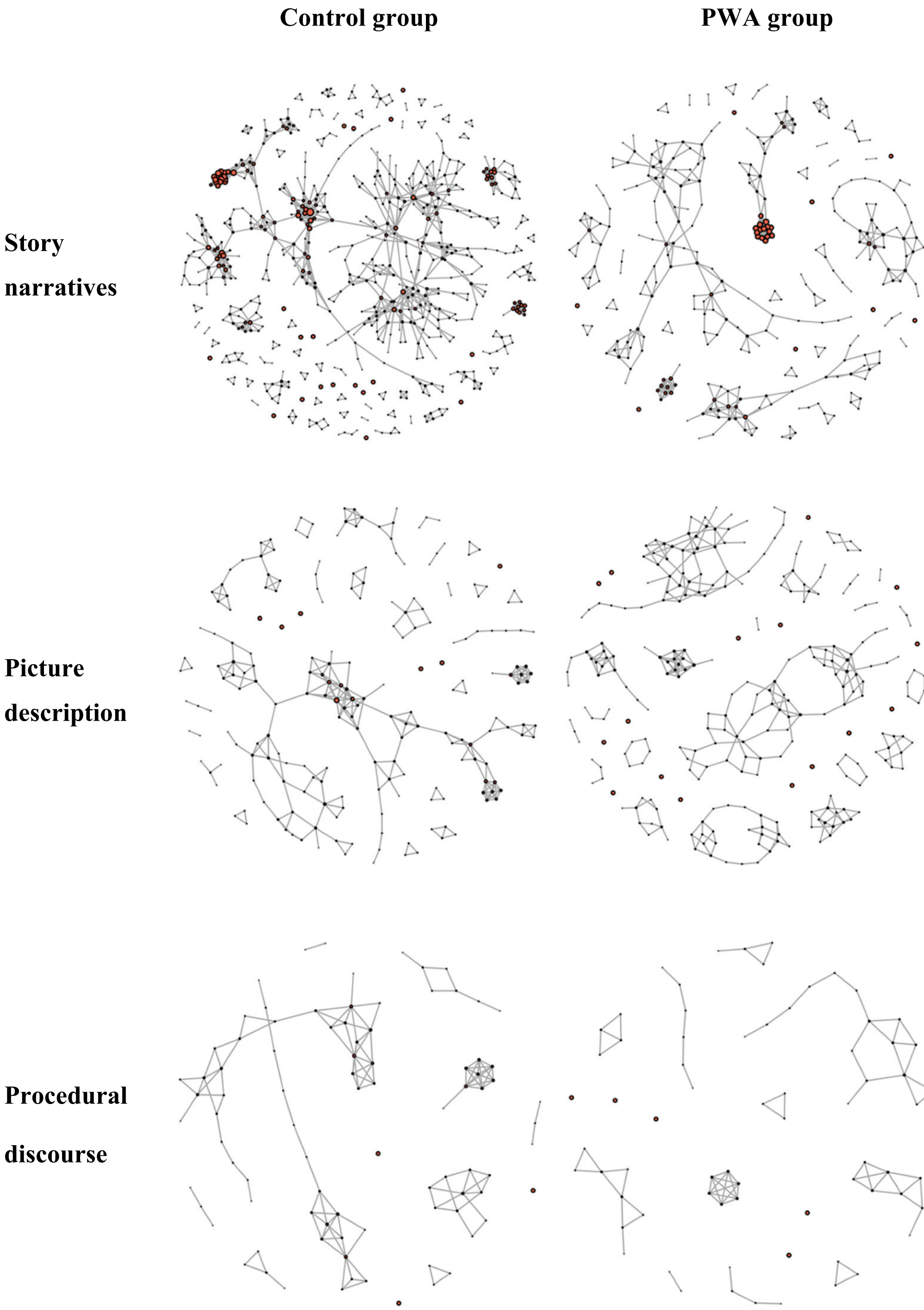

Figure 1. Lexical networks of discourse production by PWA and healthy controls. Node size represents node degree.

Table 2. Sizes, density and diameters of lexical networks.

		Size		Density		Diameter	
Tasks	Groups	Mean	SD	Mean	SD	Mean	SD
Story narratives	Control	187.37	66.77	0.10	0.02	7.95	1.34
	PWA	72.72	42.29	0.16	0.06	6.45	1.70
Picture description	Control	51.39	25.07	0.16	0.03	7.46	1.60
	PWA	26.73	17.44	0.23	0.12	5.03	1.87
Procedural discourse	Control	38.35	16.29	0.19	0.04	5.85	2.10
	PWA	23.31	13.83	0.24	0.08	4.18	1.52

density of a graph is the ratio of the number of edges to the number of possible edges. Given that networks of participants in the control group were larger in size, their networks had a larger number of possible edges, which could make the overall density smaller. In other words, PWA produced small networks with highly related words at the individual level. Similarly, networks of story narratives were larger in size than those of picture descriptions and procedural discourse and thus also had a larger number of possible edges, rendering the overall network density smaller.

For the network diameters, there were significant main effects of the group, $F(1, 154) = 101$, $p < 0.001$, and the task type, $F(2, 308) = 80$, $p < 0.001$. In addition, there was a group × task interaction, $F(2, 308) = 4$, $p = 0.02$. For the main effect of the group, the network diameters of healthy controls ($M = 7.09$, 95% CI [6.83, 7.34]) were larger than those of PWA ($M = 5.22$, 95% CI [4.96, 5.48]), $\beta = 1.87$, $SE = 0.19$, $t(154) = 10.06$, $p < 0.001$. For the main effect of task type, the network diameters of story narratives ($M = 7.20$, 95% CI [6.93, 7.47]) were larger than those of picture descriptions ($M = 6.24$, 95% CI [5.98, 6.51]), $\beta = 0.96$, $SE = 0.17$, $t(308) = 5.51$, $p < 0.001$, which in turn were larger than those of procedural discourse ($M = 5.01$, 95% CI [4.74, 5.28]), $\beta = -1.23$, $SE = 0.17$, $t(308) = -7.10$, $p < 0.001$. For the task type by group interaction, comparisons between two groups for the same task revealed that network diameters of healthy controls were larger than those of PWA for all three task types: the story narrative task ($\beta = 1.50$, $SE = 0.27$, $t(430) = 5.49$, $p < 0.001$), the picture description task ($\beta = 2.44$, $SE = 0.27$, $t(430) = 8.92$, $p < 0.001$) and the procedural description task ($\beta = 1.67$, $SE = 0.27$, $t(430) = 6.10$, $p < 0.001$).

Small-world properties of lexical networks

In order to compute small-world indices for different networks, we first built a reference random network with the same number of nodes and edges for each participant and each task. Then, we calculated the small-world indices as defined in the method section. The average small-world indices were all larger than 1 for both groups across different tasks, suggesting that their lexical networks were small-world networks (see Table 3).

The small-world indices were then fitted as the dependent variable for a linear mixed-effects model. Participant groups (PWA versus healthy controls) and task types (picture description, story narrative, procedural discourse) were selected as fixed factors. Participants were specified as a random factor (intercept).

There were no significant main effects of group nor task type in terms of small-world characteristics. However, there was a group × task two-way interaction, $F(2, 303) = 5.9$,

Table 3. Small-world indices for lexical networks in discourse production by persons with aphasia and healthy controls.

Tasks	Groups	Clustering Coefficient M	Clustering Coefficient SD	Average path length M	Average path length SD	SWI M	SWI SD
Story narratives	Control	0.50	0.03	2.80	0.18	3.71	0.53
	PWA	0.51	0.06	2.51	0.36	2.81	0.81
Picture description	Control	0.56	0.06	2.72	0.37	2.91	0.83
	PWA	0.55	0.14	2.12	0.50	3.54	4.24
Procedural discourse	Control	0.65	0.09	2.30	0.66	3.29	1.20
	PWA	0.57	0.16	1.90	0.46	3.26	1.62

p = 0.003. To further examine the interaction, we ran pairwise multiple t-tests with Tukey adjustments. Comparisons between the two groups for the same task indicated that the small-world scores of healthy control networks were marginally significantly larger than those of PWA networks only in the story narrative task, $\beta = 0.90$, $SE = 0.31$, $t(450) = 2.88$, $p = 0.047$.

Modularity

In order to compute modularity measures for different networks, we first need to identify communities in the networks. We used four different methods for community detection and obtained modularity measures for each participant and each task as defined in the method section (Table 4). Then these modularity measures were each selected as dependent variables and fitted with a separate linear mixed-effects model. Participant groups and task types were selected as fixed factors. Participants were specified as a random factor (intercept).

There were significant main effects of group, $F(1, 154) = 10$, $p = 0.001$, and task type, $F(2, 308) = 12$, $p < 0.001$ for the Infomap modularity of networks. In addition, there was a group × task interaction, $F(2, 308) = 14$, $p < 0.001$. For the main effect of the group, network Infomap modularity of healthy controls ($M = 0.29$, 95% CI [0.27, 0.31]) was larger than that of PWA ($M = 0.25$, 95% CI [0.23, 0.27]), $\beta = 0.05$, $SE = 0.01$, $t(154) = 3.16$, p = 0.002. For the main effect of task types, the network Infomap modularity of story narratives ($M = 0.30$, 95% CI [0.28, 0.33]) was larger than that of picture descriptions ($M = 0.27$, 95% CI [0.25, 0.29]), $\beta = 0.04$, $SE = 0.01$, $t(308) = 2.91$, $p = 0.01$, and that of procedural discourse ($M = 0.24$, 95% CI [0.22, 0.26]), $\beta = 0.06$, $SE = 0.01$, $t(308) = 4.94$, $p < 0.001$. For the task type by group interaction, comparisons between the two groups

Table 4. Modularity measures for lexical networks in discourse production by persons with aphasia and healthy controls.

		Infomap M	Infomap SD	Louvain M	Louvain SD	Walktrap M	Walktrap SD	Edge betweenness M	Edge betweenness SD
Story narratives	Control	0.37	0.06	0.41	0.04	0.36	0.05	0.32	0.07
	PWA	0.24	0.14	0.36	0.07	0.31	0.08	0.25	0.10
Picture description	Control	0.28	0.12	0.36	0.06	0.31	0.08	0.28	0.10
	PWA	0.25	0.18	0.35	0.13	0.31	0.14	0.28	0.17
Procedural discourse	Control	0.23	0.11	0.31	0.07	0.26	0.08	0.24	0.09
	PWA	0.25	0.15	0.33	0.10	0.29	0.12	0.27	0.14

for the same task revealed that network modularity of healthy controls was larger than that of PWA only in the story narrative task, $\beta = 0.12$, $SE = 0.02$, $t(422) = 5.88$, $p < 0.001$.

There were significant main effects of group, $F(1, 154) = 5$, $p = 0.02$ and task type, $F(2, 308) = 24$, $p < 0.001$ for the Louvain modularity. In addition, there was a group × task two-way interaction, $F(2, 308) = 7$, $p = 0.001$. For the main effect of the group, Louvain network modularity of healthy controls ($M = 0.36$, 95% CI [0.35, 0.37]) was larger than that of PWA ($M = 0.34$, 95% CI [0.33, 0.36]), $\beta = 0.02$, $SE = 0.01$, $t(154) = 2.28$, $p = 0.02$. For the main effect of task types, the Louvain network modularity of story narratives ($M = 0.38$, 95% CI [0.37, 0.40]) was larger than that of picture descriptions ($M = 0.35$, 95% CI [0.34, 0.37]), $\beta = 0.03$, $SE = 0.01$, $t(308) = 3.35$, $p = 0.003$, which was in turn larger than that of procedural discourse, ($M = 0.32$, 95% CI [0.31, 0.33]), $\beta = -0.03$, $SE = 0.01$, $t(308) = -3.58$, $p = 0.001$. For the task type by group interaction, comparisons between two groups within the same task indicated that Louvain network modularity of healthy controls was larger than that of PWA only in the story narrative task, $\beta = 0.05$, $SE = 0.01$, $t(460) = 3.98$, $p = 0.001$.

For the Walktrap modularity of networks, there were no significant main effects of group but significant main effect of task type, $F(2, 308) = 14$, $p < 0.001$, and a group × task type interaction, $F(2, 308) = 7$, $p < 0.001$. For the main effect of task type, the Walktrap network modularity of story narratives ($M = 0.33$, 95% CI [0.32, 0.35]) was marginally larger than that of picture descriptions ($M = 0.31$, 95% CI [0.29, 0.33]), $\beta = 0.02$, $SE = 0.01$, $t(308) = 2.19$, $p = 0.07$, which was in turn larger than that of procedural discourse ($M = 0.28$, 95% CI [0.26, 0.29]), $\beta = -0.03$, $SE = 0.01$, $t(308) = -3.08$, $p = 0.006$. For the task type by group interaction, comparisons between the two groups for the same task indicated that Walktrap network modularity of healthy controls was larger than that of PWA only in the story narrative task, $\beta = 0.05$, $SE = 0.02$, $t(453) = 3.12$, $p = 0.02$.

For the edge-betweenness modularity of networks, there was no main effect of group but a main effect of task type, $F(2, 308) = 4$, $p = 0.02$, and a group × task interaction, $F(2, 308) = 10$, $p < 0.001$. For the main effect of task types, edge-betweenness network modularity of story narratives ($M = 0.28$, 95% CI [0.27, 0.30]) was comparable to that of picture descriptions ($M = 0.28$, 95% CI [0.26, 0.30]), but was larger than that of procedural discourse ($M = 0.26$, 95% CI [0.24, 0.27]), $\beta = 0.03$, $SE = 0.01$, $t(308) = 2.56$, $p = 0.03$. For the task type by group interaction, comparisons between the two groups for the same task indicated that edge-betweenness network modularity of healthy controls was larger than that of PWA only in the story narrative task, $\beta = 0.07$, $SE = 0.02$, $t(417) = 3.58$, $p = 0.005$.

In general, in networks with high modularity, such as those observed in healthy controls and in the story narrative task, words within the same semantic category or conceptual (topic) domain tend to cluster together to form modules or communities, and at the same time connections between words in different modules are relatively weak. This structural feature enhances the efficiency of information processing by facilitating the retrieval of semantically related words. Moreover, highly modular networks are more robust against disruptions because the strong within-module connections contain the spread of perturbations, limiting their impact on the overall network function.

Scale free properties

To explore the degree distributions of lexical networks, we aggregated discourse production data for each task and fitted a group-level degree distribution for each task (Figure 2). To test whether the data follow a power law distribution, the *Kolmogorov–Smirnov* test was used, in which smaller scores indicate a better fit. The

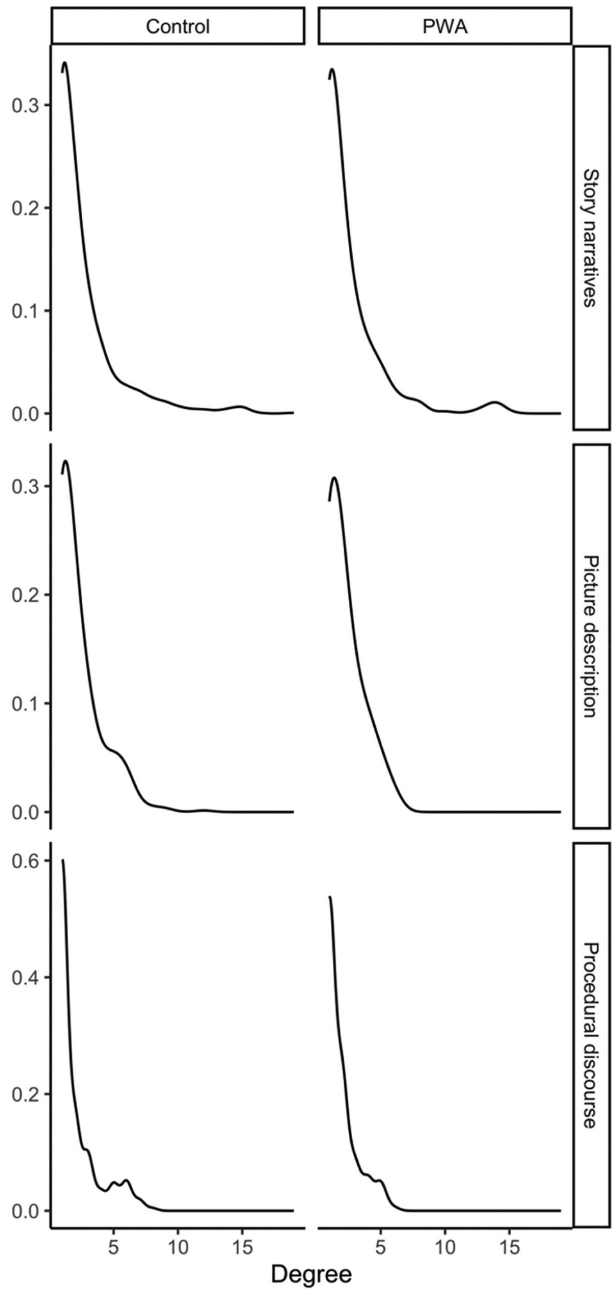

Figure 2. Group-level degree distribution for each task and group.

Table 5. Kolmogorov–Smirnov test results of degree distribution of nodes in lexical networks of PWA and healthy controls.

Tasks	Groups	Kolmogorov–Smirnov scores	*p*
Story narratives	PWA	0.010	1.0
	Control	0.057	0.2
Picture description	PWA	0.078	1.0
	Control	0.073	1.0
Procedural discourse	PWA	0.014	1.0
	Control	0.037	1.0

p-values of the Kolmogorov–Smirnov tests indicated that all group-level networks approximately followed power-law distributions (Table 5).

Discussion

Our analysis revealed that PWA and healthy controls differed in the size, density, diameter, small-worldness, modularity and distribution of node degrees in their lexical networks built on discourse data. Task types also affected these network properties.

Impact of aphasia on lexical networks in discourse production

First, people who suffered from aphasia generally produced fewer words in discourse production, rendering their network smaller in size and shorter in diameter relative to healthy controls. This reflects that impairments caused by aphasia constrain the range of vocabulary PWA can use in discourse production, consistent with previous research on linguistic productivity (Fergadiotis & Wright, 2011) and core lexicons (Chen & Chang, 2024; Dalton & Richardson, 2015). Higher density was observed in the individual PWA networks compared to those of healthy controls. Rather than indicating stronger semantic associations, this pattern may instead be a by-product of constrained production, with PWA relying on a semantically interconnected but limited set of words.

Second, although language impairments affected the size and diameter of lexical networks in discourse production, PWA networks still exhibited small-world characteristics in all tasks. Small-world networks characterised by short path lengths and high clustering coefficients are known for their efficiency in information processing. Small-worldness is considered central to typical language networks in the previous research (Cancho & Solé, 2001; Cong & Liu, 2014; Ferrer i Cancho et al., 2004; Liu & Hu, 2008). From a developmental perspective, language networks of children around the age of two show a transition in syntactic dependency relations characterised by a reduction in the average path length and an increase in clustering coefficient, i.e. the formation of a small-world network (Corominas-Murtra et al., 2010). Thus, we argue that although brain impairments affect the size and diameter of lexical networks, the observed preservation of small-world characteristics reflects compensatory mechanisms or adaptive strategies employed by the brain to overcome the challenges posed by aphasia. In other words, the language system of PWA demonstrates resilience in the face of linguistic deficits and maintains some basic level of efficiency in verbal communication. Nonetheless, PWA networks did display less small-worldness relative to those of the control group in the story narrative task. This is consistent with the finding that children with language delays showed less small-world characteristics in their language networks compared to typically developing children (Beckage et al., 2011).

Furthermore, studies have also showed reduced small-worldness for brain networks of individuals with neurodegenerative diseases, such as Alzheimer's disease (Dai et al., 2019) or acute brain injuries, such as traumatic brain injury (Pandit et al., 2013). The convergence of evidence from both language and brain network studies highlights the intricate relationship between neural architecture and language processing and production. This has broader implications for understanding language-related disorders and cognitive deficits.

Third, the analysis of modularity of networks indicated systematic variations in network organisation across different participant groups. Specifically, healthy controls exhibited larger modularity in their lexical networks compared to PWA. Highly modular networks are characterised by strong within-module connections and weak between-module connections, facilitating efficient information processing, robustness against perturbations and specialisation of function within modules, as observed in networks of the control group. Conversely, low modularity may indicate a more homogeneous network structure with fewer distinct communities, potentially leading to less efficient information flow and reduced resilience to disruptions in the PWA networks.

Fourth, both PWA and controls exhibited scale-free properties in the distribution of node degrees in their lexical networks. In a scale-free network, there exists a small number of highly connected nodes with high degrees, known as hubs, alongside other nodes with low degrees. These hubs play a crucial role in reducing path lengths within the network, facilitating efficient information transfer and communication across the system. Many language networks were also found to be scale-free at lexical (Sigman & Cecchi, 2002) or syntactic levels (Cancho & Solé, 2001; Cong & Liu, 2014; Corominas-Murtra et al., 2010; Ferrer i Cancho et al., 2004; H. Liu & Hu, 2008), though phonological networks (Vitevitch, 2008) did not display this characteristic.

Impact of task types

Lexical networks built on story narrative data were generally larger in size, diameter, higher in modularity scores but lower in density than those of picture descriptions, which in turn exhibited larger size and diameter and higher modularity but lower density compared to procedural discourse, suggesting that linguistic properties of discourse production are affected by task types. According to the Triadic Componential Framework (Robinson & Gilabert, 2007), the task type effect could be attributed to two dimensions of task complexity, i.e. resource-directing and resource-dispersing.

The story narrative task in the present study involves reasoning about causal events and relations, which is the resource-directing dimension. For example, Cinderella was able to attend the ball due to her fairy godmother's intervention. The picture description of the Broken Window also involves some reasoning about causal events. For example, a boy was kicking a ball, and the ball broke a window. Thus, participants need to infer causal relationships between these events. The procedural description of making a sandwich does not involve much reasoning about causal events, although there is a temporal relation among the steps required to assemble a sandwich.

Moreover, the three tasks also differ in the amount of reasoning involved about people's intentions, beliefs and desires. The story narrative task requires extensive reasoning about characters' internal states, as these motivations drive the plot forward. For instance, understanding the stepmother's jealousy motivates her mistreatment of Cinderella. The picture description task involves less reasoning about intentions and beliefs relative to the story narrative task. The procedural discourse task demands primarily an objective description. We speculate that the storytelling task elicited a greater amount of discourse production (as revealed by the larger network size) and allowed speakers to exploit more of their lexical resources (as revealed by the larger network diameter) to express causal relationships and intentionality (forming different modules within the network) relative to picture and procedure description tasks.

In contrast to resource-directing variables, resource-dispersing variables refer to the performative demands of a task, such as planning time or background knowledge. The picture description task directly supplied the background knowledge necessary for the task, whereas storytelling and procedural description tasks offered limited background information. Consequently, speakers may rely more on their existing knowledge of the story and the procedure to fulfil the tasks. In this sense, the storytelling task is more demanding in the resource-dispersing dimension than the picture description task and at the same time requires more information (such as relative time, spatial location and causal relationships) than both picture and procedural description tasks in the resource-directing dimension. This explains why PWA displayed deviations from controls in terms of network efficiency measures, such as small-worldness and modularity scores, only in the story narrative task. In other words, PWA struggle with the storytelling task which has high cognitive and memory demands. Thus, they showed less effective use of their lexical resources, reflected in reduced network efficiency relative to control participants. These findings underscore the importance of considering task complexity in understanding the network dynamics of aphasic discourse production.

Conclusions

Our study provides valuable insights into lexical networks built on discourse data produced by PWA and healthy controls. PWA networks showed smaller size, diameter and lower modularity but higher density compared to healthy controls, suggesting that lexical networks of PWA discourse were impaired, constrained and less efficient. Nonetheless, their remaining vocabulary in the discourse was still closely interconnected. Despite these differences, networks of both groups exhibited small-world characteristics, albeit to a lesser extent in PWA networks and displayed scale-free properties in the distribution of node degrees. The storytelling task is cognitively more demanding than both picture and procedural description tasks and is more revealing in detecting network differences between PWA and healthy controls. Overall, our findings contribute to a deeper understanding of the dynamics between language impairments and task demands in aphasic discourse production.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the Humanities and Social Sciences Fundation of the Ministry of Education of PRC [23YJC740002].

Data availability statement

The data that support the findings of this study are openly available in AphasiaBank English PWA Protocol Data at https://aphasia.talkbank.org/access/English/Protocol/.

References

Abramov, O., & Mehler, A. (2011). Automatic language classification by means of syntactic dependency networks. *Journal of Quantitative Linguistics*, *18*(4), 291–336. https://doi.org/10.1080/09296174.2011.608602

Amancio, D. R. (2015). A complex network approach to Stylometry. *PLOS ONE*, *10*(8), e0136076. https://doi.org/10.1371/journal.pone.0136076

Bassett, D. S., Bullmore, E., Verchinski, B. A., Mattay, V. S., Weinberger, D. R., & Meyer-Lindenberg, A. (2008). Hierarchical organization of human cortical networks in health and schizophrenia. *Journal of Neuroscience*, *28*(37), 9239–9248. https://doi.org/10.1523/JNEUROSCI.1929-08.2008

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

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

Blanco, R., & Lioma, C. (2012). Graph-based term weighting for information retrieval. *Information retrieval*, *15*(1), 54–92. https://doi.org/10.1007/s10791-011-9172-x

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

Bryant, L., Ferguson, A., & Spencer, E. (2016). Linguistic analysis of discourse in aphasia: A review of the literature. *Clinical Linguistics and Phonetics*, *30*(7), 489–518. https://doi.org/10.3109/02699206.2016.1145740

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

Cancho, R. F. I., & Solé, R. V. (2001). The small world of human language. *Proceedings of the Royal Society of London. Series B: Biological Sciences* (Vol. 268, pp. 2261–2265). https://doi.org/10.1098/rspb.2001.1800

Castro, N., Stella, M., & Siew, C. S. Q. (2020). Quantifying the interplay of semantics and phonology during failures of word retrieval by people with aphasia using a multiplex lexical network. *Cognitive Science*, *44*(9), e12881. https://doi.org/10.1111/cogs.12881

Chen, J., & Chang, H. (2024). Core lexicon analysis of spoken discourse production by mandarin Chinese speakers with aphasia. *Aphasiology*, *38*(11), 1766–1785. https://doi.org/10.1080/02687038.2024.2325169

Coelho, R. M., Drummond, C., Mota, N. B., Erthal, P., Bernardes, G., Lima, G., Molina, R., Sudo, F. K., Tannock, R., & Mattos, P. (2021). Network analysis of narrative discourse and attention-deficit hyperactivity symptoms in adults. *PLOS ONE*, *16*(4), e0245113. https://doi.org/10.1371/journal.pone.0245113

Cong, J., & Liu, H. (2014). Approaching human language with complex networks. *Physics of Life Reviews*, *11*(4), 598–618. https://doi.org/10.1016/j.plrev.2014.04.004

Corominas-Murtra, B., Valverde, S., & Solé, R. V. (2010). Emergence of scale-free syntax networks. In S. Nolfi & M. Mirolli (Eds.), *Evolution of communication and language in embodied agents* (pp. 83–101). Springer. https://doi.org/10.1007/978-3-642-01250-1_6

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

Dai, Z., Lin, Q., Li, T., Wang, X., Yuan, H., Yu, X., He, Y., & Wang, H. (2019). Disrupted structural and functional brain networks in Alzheimer's disease. *Neurobiology of Aging*, *75*, 71–82. https://doi.org/10.1016/j.neurobiolaging.2018.11.005

Dalton, S. G., & Richardson, J. D. (2015). Core-lexicon and main-concept production during picture-sequence description in adults without brain damage and adults with aphasia. *American Journal of Speech-Language Pathology*, *24*(4), 923–938. https://doi.org/10.1044/2015_AJSLP-14-0161

Dalton, S. G., Stark, B. C., Fromm, D., Apple, K., MacWhinney, B., Rensch, A., & Rowedder, M. (2022). Validation of an automated procedure for calculating core lexicon from transcripts. *Journal of Speech, Language, & Hearing Research*, *65*(8), 2996–3003. https://doi.org/10.1044/2022_JSLHR-21-00473

Davis, G. F., Yoo, M., & Baker, W. E. (2003). The Small world of the American corporate elite, 1982–2001. *Strategic Organization*, *1*(3), 301–326. https://doi.org/10.1177/14761270030013002

de Arruda, F., Marinho, H. Q., da, V. F., Costa, L., & Amancio, R. D. (2019). Paragraph-based representation of texts: A complex networks approach. *Information Processing & Management*, *56* (3), 479–494. https://doi.org/10.1016/j.ipm.2018.12.008

Dipper, L., Marshall, J., Boyle, M., Botting, N., Hersh, D., Pritchard, M., & Cruice, M. (2021). Treatment for improving discourse in aphasia: A systematic review and synthesis of the evidence base. *Aphasiology*, *35*(9), 1125–1167. https://doi.org/10.1080/02687038.2020.1765305

Dipper, L. T., & Pritchard, M. In Dreux, F. D. M. F., Fernandes, M. (2017). Discourse: Assessment and therapy (Ed.), *Advances in speech-language pathology*. IntechOpen. https://doi.org/10.5772/intechopen.69894

Fan, L., & Jiang, Y. (2021). A syntactic dependency network approach to the study of translational language. *Digital Scholarship in the Humanities*, *36*(3), 595–606. https://doi.org/10.1093/llc/fqaa030

Fergadiotis, G., & Wright, H. H. (2011). Lexical diversity for adults with and without aphasia across discourse elicitation tasks. *Aphasiology*, *25*(11), 1414–1430. https://doi.org/10.1080/02687038.2011.603898

Fergadiotis, G., Wright, H. H., & West, T. M. (2013). Measuring lexical diversity in narrative discourse of people with aphasia. *American Journal of Speech-Language Pathology*, *22*(2). https://doi.org/10.1044/1058-0360(2013/12-0083)

Ferrer i Cancho, R., Solé, R. V., & Köhler, R. (2004). Patterns in syntactic dependency networks. *Physical Review E*, *69*(5), 051915. https://doi.org/10.1103/PhysRevE.69.051915

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, *486*(3–5), 75–174. https://doi.org/10.1016/j.physrep.2009.11.002

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Grave, E., Bojanowski, P., Gupta, P., Joulin, A., & Mikolov, T. (2018). Learning word vectors for 157 languages. In *Proceedings of the International Conference on Language Resources and Evaluation (LREC 2018)* (pp. 3483–3487). European Language Resources Association (ELRA).

Halekoh, U., & Hojsgaard, S. (2014). A kenward-roger approximation and parametric bootstrap methods for tests in linear mixed models-the R package pbkrtest. *Journal of Statistical Software*, *59* (9), 1–30. https://doi.org/10.18637/jss.v059.i09

Humphries, M. D., Gurney, K., & Sporns, O. (2008). Network 'small-world-ness': A quantitative method for determining canonical network equivalence. *PLOS ONE*, *3*(4), e0002051. https://doi.org/10.1371/journal.pone.0002051

Kertesz, A. (2006). *Western aphasia battery-revised (revised)*. Pearson.

Kim, H., & Wright, H. H. (2020). Concurrent validity and reliability of the core lexicon measure as a measure of word retrieval ability in aphasia narratives. *American Journal of Speech-Language Pathology*, *29*(1), 101–110. https://doi.org/10.1044/2019_AJSLP-19-0063

Liu, H. (2008). The complexity of Chinese syntactic dependency networks. *Physica A: Statistical Mechanics and Its Applications*, *387*(12), 3048–3058. https://doi.org/10.1016/j.physa.2008.01.069

Liu, H., & Hu, F. (2008). What role does syntax play in a language network? *EPL (Europhysics Letters)*, *83*(1), 18002. https://doi.org/10.1209/0295-5075/83/18002

Liu, J., Zhao, J., & Bai, X. (2021). Syntactic impairments of Chinese Alzheimer's disease patients from a language dependency network perspective. *Journal of Quantitative Linguistics*, *28*(3), 253–281. https://doi.org/10.1080/09296174.2019.1703485

MacWhinney, B. (2016). The CHILDES project: Tools for analyzing talk. *Child Language Teaching & Therapy*. https://doi.org/10.1177/026565909200800211

MacWhinney, B., & Fromm, D. (2016). AphasiaBank as BigData. *Seminars in Speech and Language*, *37*(1), 010–022. https://doi.org/10.1055/s-0036-1571357

MacWhinney, B., Fromm, D., Forbes, M., & Holland, A. (2011). AphasiaBank: Methods for studying discourse. *Aphasiology*, *25*(11), 1286–1307. https://doi.org/10.1080/02687038.2011.589893

Malcorra, B. L. C., Mota, N. B., Weissheimer, J., Schilling, L. P., Wilson, M. A., & Hübner, L. C. (2021). Low speech connectedness in Alzheimer's disease is associated with poorer semantic memory performance. *Journal of Alzheimer's Disease*, *82*(3), 905–912. https://doi.org/10.3233/JAD-210134

Meunier, D., Lambiotte, R., & Bullmore, E. T. (2010). Modular and hierarchically modular organization of brain networks. *Frontiers in Neuroscience*, *4*. https://doi.org/10.3389/fnins.2010.00200

Mota, N. B., Weissheimer, J., Finger, I., Ribeiro, M., Malcorra, B., & Hübner, L. (2023). Speech as a graph: Developmental perspectives on the organization of spoken language. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *8*(10), 985–993. https://doi.org/10.1016/j.bpsc.2023.04.004

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, *103*(23), 8577–8582. https://doi.org/10.1073/pnas.0601602103

Pandit, A. S., Expert, P., Lambiotte, R., Bonnelle, V., Leech, R., Turkheimer, F. E., & Sharp, D. J. (2013). Traumatic brain injury impairs small-world topology. *Neurology*, *80*(20), 1826–1833. https://doi.org/10.1212/WNL.0b013e3182929f38

Pritchard, M., Hilari, K., Cocks, N., & Dipper, L. (2017). Reviewing the quality of discourse information measures in aphasia. *International Journal of Language & Communication Disorders*, *52*(6), 689–732. https://doi.org/10.1111/1460-6984.12318

Pritchard, M., Hilari, K., Cocks, N., & Dipper, L. (2018). Psychometric properties of discourse measures in aphasia: Acceptability, reliability, and validity. *International Journal of Language & Communication Disorders*, *53*(6), 1078–1093. https://doi.org/10.1111/1460-6984.12420

R Core Team. (2018). *R: A language and environment for statistical computing* [Manual]. https://www.R-project.org/

Robinson, P., & Gilabert, R. (2007). Task complexity, the cognition hypothesis and second language learning and performance. *IRAL*, *45*(3), 161–176. https://doi.org/10.1515/iral.2007.007

Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, *105*(4), 1118–1123. https://doi.org/10.1073/pnas.0706851105

Saussure, F. D. (1959). *Course in General linguistics* (W. Baskin, Trans.). Philosophical Library.

Siew, C. S. Q., Wulff, D. U., Beckage, N. M., Kenett, Y. N., Meštrović, A., & Meštrović, A. (2019). Cognitive network science: A review of research on cognition through the lens of network representations, processes, and dynamics. *Complexity*, *2019*(1), 1–24. https://doi.org/10.1155/2019/2108423

Sigman, M., & Cecchi, G. A. (2002). Global organization of the Wordnet lexicon. *Proceedings of the National Academy of Sciences*, *99*(3), 1742–1747. https://doi.org/10.1073/pnas.022341799

Silge, J., & Robinson, D. (2016). Tidytext: Text mining and analysis using tidy data principles in r. *JOSS*, *1*(3). https://doi.org/10.21105/joss.00037

Stam, C. J. (2014). Modern network science of neurological disorders. *Nature Reviews Neuroscience*, *15*(10), 683–695. https://doi.org/10.1038/nrn3801

Stark, B. C. (2019). A comparison of three discourse elicitation methods in aphasia and age-matched adults: Implications for language assessment and outcome. *American Journal of Speech-Language Pathology*, *28*(3), 1067–1083. https://doi.org/10.1044/2019_AJSLP-18-0265

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

Uzzi, B., & Spiro, J. (2005). Collaboration and creativity: The Small world problem. *The American Journal of Sociology*, *111*(2), 447–504. https://doi.org/10.1086/432782

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

Vitevitch, M. S., & Castro, N. (2015). Using network science in the language sciences and clinic. *International Journal of Speech-Language Pathology*, *17*(1), 13–25. https://doi.org/10.3109/17549507.2014.987819

Wießner, I., Falchi, M., Daldegan-Bueno, D., Palhano-Fontes, F., Olivieri, R., Feilding, A., Araujo, B., Ribeiro, D., Bezerra Mota, S. N., & Tófoli, L. F. (2023). LSD and language: Decreased structural connectivity, increased semantic similarity, changed vocabulary in healthy individuals. *European Neuropsychopharmacology*, *68*, 89–104. https://doi.org/10.1016/j.euroneuro.2022.12.013

Wulff, D. U., Hills, T. T., & Mata, R. (2022). Structural differences in the semantic networks of younger and older adults. *Scientific Reports*, *12*(1), 21459. https://doi.org/10.1038/s41598-022-11698-4

Xu, Q., Markowska, M., Chodorow, M., & Li, P. (2021). Modeling bilingual lexical processing through code-switching speech: A network science approach. *Frontiers in Psychology*, *12*, 662409. https://doi.org/10.3389/fpsyg.2021.662409

Yang, L., Li, K., & Huang, H. (2018). A new network model for extracting text keywords. *Scientometrics*, *116*(1), 339–361. https://doi.org/10.1007/s11192-018-2743-5