Neuropsychologia 170 (2022) 108235

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Semantic flow and its relation to controlled semantic retrieval deficits in the narrative production of people with aphasia☆

Celia P. Litovsky [a,b,*], Ann Marie Finley [a,b], Bonnie Zuckerman [a,b], Matthew Sayers [a,b], Julie A. Schoenhard [a,b], Yoed N. Kenett [c], Jamie Reilly [a,b]

[a] Eleanor M. Saffran Center for Cognitive Neuroscience, Temple University, Philadelphia, PA, USA
[b] Department of Communication Sciences and Disorders, Temple University, Philadelphia, PA, USA
[c] Faculty of Industrial Engineering & Management, Technion – Israel Institute of Technology, Haifa, Israel

ARTICLE INFO

Keywords:
Aphasia
Semantic distance
Thematic relations
Taxonomic relations

ABSTRACT

Aphasia has had a profound influence on our understanding of how language is instantiated within the human brain. Historically, aphasia has yielded an in vivo model for elucidating the effects of impaired lexical-semantic access on language comprehension and production. Aphasiology has focused intensively on single word dissociations. Yet, less is known about the integrity of combinatorial semantic processes required to construct well-formed narratives. Here we addressed the question of how controlled lexical-semantic retrieval deficits (a hallmark of aphasia) might compound over the course of longer narratives. We specifically examined word-by-word flow of taxonomic vs. thematic semantic distance in the storytelling narratives of individuals with chronic post-stroke aphasia (n = 259) relative to age-matched controls (n = 203). We first parsed raw transcribed narratives into content words and computed inter-word semantic distances for every running pair of words in each narrative (N = 232,490 word transitions). The narratives of people with aphasia showed significant reductions in taxonomic and thematic semantic distance relative to controls. Both distance metrics were strongly predictive of offline measures of semantic impairment and aphasia severity. Since individuals with aphasia often exhibit perseverative language output (i.e., repetitions), we performed additional analyses with repetitions excluded. When repetitions were excluded, group differences in semantic distances persisted and thematic distance was still predictive of semantic impairment, although some findings changed. These results demonstrate the cumulative impact of deficits in controlled word retrieval over the course of narrative production. We discuss the nature of semantic flow between words as a novel metric of characterizing discourse and elucidating the nature of lexical-semantic access impairment in aphasia at multiword levels.

Credit author statement

Celia P. Litovsky: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization, Project administration. **Ann Marie Finley:** Conceptualization, Methodology, Software, Investigation, Writing – Original Draft, Writing – Review & Editing. **Bonnie Zuckerman:** Conceptualization, Methodology, Software, Data Curation, Writing – Original Draft, Writing – Review & Editing. **Matthew Sayers:** Methodology, Software, Investigation. **Julie A. Schoenhard:** Conceptualization, Writing – Review & Editing. **Yoed N. Kenett:** Conceptualization, Writing – Original Draft, Writing – Review & Editing. **Jamie Reilly:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Supervision, Project administration, Funding acquisition.

1. Introduction

The pioneering efforts of aphasiology in the nineteenth century forged much of the foundation upon which modern cognitive neuroscience was built. Broca (1861), Wernicke (1874), Lichtheim (1885), Hughlings Jackson (1874) and others used post-mortem lesion methods to first link specific language functions to circumscribed brain regions and later to refine network models of the flow of information through

☆ *This work was funded by US Public Health Service Grant R01 DC013063.
* Corresponding author. Eleanor M. Saffran Center for Cognitive Neuroscience, Temple University, Philadelphia, PA, USA.
E-mail address: celia.litovsky@temple.edu (C.P. Litovsky).

https://doi.org/10.1016/j.neuropsychologia.2022.108235
Received 29 November 2021; Received in revised form 4 April 2022; Accepted 5 April 2022
Available online 14 April 2022

the human brain. Many of these efforts involved characterizing dissociations in the production and/or comprehension of single words. For example, the classical taxonomy of aphasia was premised upon the binary classification of impairments in language fluency, word repetition, and auditory comprehension (e.g., Albert et al., 1981;Geschwind, 1965; Benson, 1979). As aphasiology advanced throughout the twentieth century, the burgeoning disciplines of cognitive neuropsychology and psycholinguistics maintained an intense focus on speech errors and the nature of verbal paraphasias (e.g., semantic errors, neologisms). Much of this attention was fueled by the promise of aphasia as a lesion model for language in the human brain informed by dissociations for single words.

Aphasiology evolved in complexity and sophistication over the twentieth century, interrupted by wars and punctuated by changing computational trends in neuroscience (Levelt, 2012). Much of the evolution of aphasiology during this period was untethered from crucial advances in linguistics (e.g., syntax, compositionality) that have firmly established language as an emergentist system where outcomes (e.g., message-level semantics) are only opaquely related to the individual lexical constituents. As such, an extensive body of research at the single word level in aphasia yields a limited picture of how language operates in real world ecological conditions that involve the coordination of semantics, syntax, prosody, morphology, speech motor control, and other supportive functions (e.g., theory of mind; Bates and Goodman, 1997; Friederici and Alter, 2004). As such, there exists a compelling need to bridge our understanding of single word deficit profiles in aphasia with multiword utterances and narratives. Here we addressed the question of how impaired lexical-semantic retrieval processes cumulatively impact word-by-word language production in the context of narratives.

1.1. The nature of aphasia and semantic deficits

Many patient advocacy organizations for post-stroke aphasia emphasize the point that aphasia is a disorder of language, not of intelligence (National Aphasia Association, 2021). That is, a person with aphasia might have difficulty understanding the spoken, signed, or written word for ‘dog’ but would readily demonstrate preserved knowledge of dogs through alternate response modalities. Language rehabilitation strategies commonly leverage preserved pathways (e.g., gesture, drawing, circumlocution) as a means for compensating for profound anomia.

Tulving (1972) is generally credited with delineating semantic memory as a subsystem of human memory dedicated to word meaning, object knowledge, and encyclopedic facts that allow us to negotiate our most fundamental interactions with the world. Warrington and Shallice (1979) soon thereafter argued that semantic deficits fractionate into disorders of storage vs. access. Under this model, a patient with a semantic storage deficit (also known as a representational deficit) might be unable to name a dog because they have ‘lost’ the concept of dogs. In contrast, patients with lexical-semantic access deficits (i.e., aphasia) would show patterns of anomia in the context of preserved non-verbal dog knowledge.

Prior research on lexical-semantic access impairments primarily comes from single-word production errors. For example, Buckingham and Rekart (1979) reported on a patient with post-stroke Wernicke’s aphasia who exhibited a characteristic lexical-semantic access deficit with maintained semantic knowledge. When discussing hearing aids, for example, the patient called her ear “eye” but pointed to her left ear, indicating that she understood that hearing aids are associated with ears rather than eyes. Across many tasks, the patient’s paraphasias shared a semantic category with the target, and she often indicated through gestures or through self-correction that she knew she had made an error, suggesting preserved semantic knowledge despite impaired semantic production. Individuals with lexical-semantic access impairments typically exhibit better naming performance when given phonemic cues (i.e., when provided with the first phonemes of the target), indicating a preserved concept but difficulty with its retrieval (Jefferies and Lambon Ralph, 2006). Another hallmark of lexical-semantic access impairment is inconsistency on lexical semantics tasks across repeated testing sessions, with patients successfully retrieving particular words on one occasion but not on another (Warrington and Shallice, 1979). Had the semantic representations themselves been damaged, the same damaged concepts would be affected at each testing session (van Scherpenberg et al., 2019). Further, in line with the notion that lexical-semantic access impairments result from executive dysfunction in retrieval of semantic representations, other studies have shown strong correlations of lexical semantics task performance with executive functioning ability (e.g., Corbett et al., 2009). These studies, among others (see Mirman and Britt, 2014, for a review), highlight that the semantic paraphasias in post-stroke aphasia are thought to result from difficulty accessing the lexical item, rather than damage to the semantic representations themselves.[1]

Additionally, individuals with lexical-semantic access impairments typically exhibit an access dissociation in taxonomic vs. thematic semantic relationships (e.g., Jefferies and Lambon Ralph, 2006). Taxonomic relationships are feature-based – DOG and WOLF are related concepts based on shared semantic features (e.g., canine, four-legged). Thematic relationships are predicated upon co-occurrence – DOG and LEASH are related not because of intrinsic shared features (i.e., taxonomically related), but because dogs and leashes frequently co-occur in the world. Some concepts have similar degrees of relatedness in both the thematic and the taxonomic system (e.g., FORK-SPOON); whereas others are loosely related regardless of relational system (e.g., ROBOT-DAISY).

In one study of 25 individuals with post-stroke aphasia, Vivas et al. (2015) found that participants consistently performed worse on tasks that required retrieval of thematic vs. taxonomic concept relationships. However, other studies have reported the opposite phenomenon, such that participants with aphasia made more taxonomic errors than thematic errors (e.g., Schwartz et al., 2011). Further, multiple studies have reported inverse relationships of taxonomic and thematic retrieval impairments, with participants tending to exhibit considerably more deficits in retrieval of one type of semantic relationship than the other (e.g., Kalénine et al., 2012). These dissociations are thought to result from differences in lesion locations across participants, with lesions in the angular gyrus/temporoparietal junction associated with worse performance on thematic relationships, and lesions in the anterior temporal lobe associated with worse performance on taxonomic relationships (Mirman and Graziano, 2012; Schwartz et al., 2011). Given that aphasia subsequent to stroke more often results from lesions in the angular gyrus/temporoparietal junction territory than from anterior temporal lesions (Khedr et al., 2020; Hayward et al., 1977), it is unsurprising that more studies indicate worse thematic than taxonomic performance in this population.

Studies of single-word semantics have considerably added to our knowledge of the semantic system. Yet, language expression is a continuous phenomenon where message-level meaning cannot be derived simply from a concatenation of single words. Relatively fewer studies have investigated lexical-semantic processing in multi-word utterances (e.g., Andreetta et al., 2012). As a consequence, much remains to be learned about how lexical-semantic access impairments manifest in the context of discourse or narrative production. For example, although many individuals with post-stroke aphasia exhibit deficits in retrieval of particular words, the nature of compounding impairment in multiword utterances is unclear.

1.2. Narrative production deficits in aphasia

Successful narrative language production requires coordination of

[1] Also see Malt (2020) for a discussion on the distinction between words and concepts.

many cognitive-linguistic processes (e.g., morphology, syntax, theory of mind) to structure an overarching message from single words. Disruption to any one of these components may lead to impairment in narrative language. Relative to controls, people with aphasia (PWA) produce narratives that are shorter, less syntactically complex, and convey less information (Andreetta et al., 2012; Bryant et al., 2016; Cunningham and Haley, 2020; Stark, 2019; Ulatowska et al., 1981). In turn, listeners perceive PWA's narratives as less interesting, less "good," and lacking in clarity, vocabulary, and content relative to control group narratives (Behrns et al., 2009; Ulatowska et al., 1981). Notably, narrative language has traditionally been studied as the sum of its parts, with a primary focus on macrolevel (e.g., story grammar, Lakoff, 1972; main concepts, Richardson and Dalton, 2016) and phrasal/sentential levels rather than word-level features (Bryant et al., 2016). This stands in contrast to the large body of knowledge that cognitive neuroscientists and clinical researchers have amassed in single-word processing. Although there is increasing interest in standardizing narrative analysis as a treatment outcome in clinical populations (Duff, 2021), to date, relatively few studies have examined the compounding effect of lexical-semantic access impairment in word-by-word narrative production in PWA.

Specifically, one fundamental aspect of semantic processing that has not yet been investigated is how semantic information flows over the course of narrative production. For example, when an individual is producing the Cinderella story, how does semantic information transition from describing the fairy godmother transforming Cinderella's gown and carriage to describing meeting the prince at the ball? In this study, we investigate the flow of semantic information in the narratives of individuals with and without post-stroke aphasia via inter-word semantic distance, which estimates semantic cohesion between adjacent words.

1.3. Semantic distance

Semantic distance is a fundamental issue in psycholinguistic research (Kenett, 2018). While it has been traditionally examined through subjective assessment, the rapid advances in computational linguistic models are now offering quantitative operationalization of semantic distance (Kenett, 2018, 2019). Such computational models are either network-based, corpus-based, or taxonomy-based. Network-based models are graph theoretical representations of semantic memory structure as a network, an approach that has been applied to study various cognitive phenomena related to memory and language (Hills and Kenett, 2022; Siew et al., 2019). For example, cognitive network research in aphasia has shown how semantic distance predicts the performance of PWA on a picture naming task (Castro and Stella, 2019). Corpus-based models are based on analyzing large bodies of text and representing concepts as multidimensional vectors of features, based on co-occurrence statistics (Günther et al., 2019; Mandera et al., 2015, 2017). The application of such corpus-based methods to study different phenomena of semantic memory are becoming increasingly popular (Kumar, 2021). For example, in creativity research, corpus-based measures of semantic distance—the inverse of semantic similarity between concepts—has been strongly linked to subjective measures of idea originality and free association (Beaty and Johnson, 2021; Kenett, 2019), indicating that more creative individuals are better able to search wider and deeper in semantic memory. Taxonomy-based models also represent concepts as multidimensional vectors of features, such that vector components each represent particular taxonomic features (e.g., animal, vertebrate, or mammal; Cree and McRae, 2003; Garrard et al., 2001). Computing the distance between vectors (i.e., Euclidean, cosine) allows for comparison of concepts based on taxonomic similarity. Overall, quantitative measures of semantic distance as incorporated in empirical research are rapidly advancing research in varied areas related to semantic cognition. Such measures have not previously been investigated in individuals with lexical-semantic access impairments, and characterization of semantics over multiword utterances is essential to characterizing the full extent of the deficit.

In this pre-registered study, we examined semantic distance (flow) over the course of narrative production, a measure that forms a bridge between analyses of single words and full narratives, to examine how lexical-semantic access errors compound when aggregating multiple words together into ideas. Evaluating the evolution of word meaning in the context of narrative production is also a more naturalistic, ecologically valid approach than with most tasks that evaluate single words, and doing so may more fully capture the extent of communicative function and impairment in PWA.

Specifically, we investigated semantic distance during spoken narrative production in AphasiaBank, an extensive database of people with chronic aphasia and age-matched controls (MacWhinney et al., 2011). We investigated four main research questions: (1) We first investigated the relationship between semantic impairment and overall severity of language deficits in aphasia, with greater semantic impairment expected to be associated with greater overall aphasia severity. (2) We then examined whether PWA and healthy controls exhibited overall differences in semantic distance, with PWA expected to show lower mean semantic distances than controls. (3) We also investigated the relationship of semantic distance to the degree of semantic impairment in PWA, with greater semantic distances expected to be associated with worse lexical-semantic access impairment. Such a relationship would be non-linear, with greater semantic distances associated with severe impairment and with minimal impairment, and lower semantic distances associated with moderate impairment. (4) Finally, we investigated group differences in the relationship between thematic and taxonomic semantic distances, with controls expected to have stronger correlations in thematic and taxonomic distances than PWA, given previous findings that aphasic participants' thematic and taxonomic abilities tend to be inversely correlated (Kalénine et al., 2012).

2. Methods

2.1. Overview

We analyzed word-to-word semantic relationships in spoken narratives of PWA and age-matched healthy controls, as indexed within AphasiaBank (MacWhinney et al., 2011). AphasiaBank is a database populated by multiple clinical aphasiology centers with standardized protocols characterizing language, neuropsychological functioning, and demographic information. AphasiaBank is unique in that it includes transcription of spoken narrative data (e.g., telling of the Cinderella story) from an unprecedented number of PWA and controls. We crafted a text processing pipeline to analyze word-to-word semantic relationships in the narratives of over 400 individuals. The research questions, hypotheses, and planned analyses were pre-registered on August 6, 2021.[2] Although one of the authors (M.S.) had access to the AphasiaBank data prior to this date for use in studies unrelated to semantic processing, no semantic distance computations were performed on the dataset prior to August 27, 2021.

2.2. Participants

Participants included PWA (N = 259, 107 female, mean age = 61.6 years, SD = 11.9 years) and age-matched controls (N = 203, 113 female, mean age 64.5 years, SD = 16.6 years). We included participants delineated by AphasiaBank as native English speakers. The PWA group included only participants with an aphasia etiology secondary to left hemisphere stroke (i.e., no right hemisphere strokes, tumors, or traumatic brain injury). Table 1 lists participant demographics.

[2] The pre-registration is available at osf.io/93jbx.

Table 1
Participant demographics.

	Aphasia (N = 259)	Control (N = 203)
Gender		
Female	107 (41.3%)	113 (55.7%)
Male	152 (58.7%)	90 (44.3%)
Age		
Mean (SD)	61.6 (11.9)	64.5 (16.6)
Race		
African American	34 (13.1%)	4 (2.0%)
Asian	3 (1.2%)	0 (0%)
Hispanic/Latino	4 (1.5%)	4 (2.0%)
Mixed	1 (0.4%)	0 (0%)
Native Hawaiian/Pacific Islander	2 (0.8%)	0 (0%)
Unavailable	3 (1.2%)	4 (2.0%)
White	212 (81.9%)	191 (94.1%)
Education (years)		
Mean (SD)	15.3 (2.7)	15.7 (2.52)
Aphasia Duration (years)		
Mean (SD)	5.07 (4.79)	NA (NA)

3. Data processing and analysis

3.1. Text cleaning methods

We applied a customized text cleaning algorithm to all narratives with the aim of transforming the original text into vectors of open class (i.e., content) words. We imported raw string data into RStudio (R version 4.0.3), yoking all associated metadata (e.g., age, aphasia severity, sex) with each narrative. We omitted utterances from individuals other than the participants and retained only data corresponding to the standardized narrative story prompts (picture description, procedural narrative, and spontaneous speech prompts). Procedures were standardized such that all participants received the same instructions for each prompt. Feedback for each narrative was minimized and only given when participants produced very limited speech. Detailed descriptions of each prompt and the procedures used to elicit the narratives may be found at the AphasiaBank website.[3] The cleaning script then automatically executed many sequential commands (regular expressions) transforming the original transcriptions (e.g., "The dog is drinking the water.") into a temporally ordered vector of lemmatized content words (e.g., "Dog Drink Water"). Although space limitations prohibit an exhaustive description of this text cleaning pipeline, we describe some of the major methods to follow.[4]

Cleaning involved a series of omissions and substitutions followed by lemmatization. We omitted all non-alphabetic characters (punctuation, numbers, symbols) after replacing contractions (e.g., 'it's' → 'it is'). We converted all characters to lowercase and omitted numerous closed class words (e.g., is, any, that, for) using a customized stopword list. We then executed a variety of replacements including singular pronouns (e.g., 'she' → 'woman'), plural pronouns (e.g., 'they' → 'people'), and common abbreviations (e.g., 'M.D.' → 'doctor'). We replaced proper nouns corresponding to common first and second names of people (e.g., 'Smith' → 'person') by indexing a list of the 300 most common surnames from around in the world. We replaced proper noun place names (e.g., 'New York' → 'place') by matching the 100 most populated countries and 100 most visited cities from around the world. After these replacement and omission procedures were completed, we lemmatized the entire vector using UDpipe (Straka and Straková, 2017) trained on the GUM English corpus (Zeldes, 2017).[5] These procedures yielded an ordered vector of content words (e.g., DOG DRINK MILK) across probes and participants from which we derived two running metrics of inter-word (i.e., word-to-word) semantic distance for each unfolding narrative (e.g., DOG-DRINK, DRINK-MILK, etc.).

3.2. Text-based language measures

To characterize group differences in language output across prompts, we calculated participants' (1) mean number of words produced per prompt, (2) mean length of utterance, and (3) type-token ratio for both the raw and the cleaned texts. First, we applied a minimal cleaning pipeline to the raw data to prepare it for processing (e.g., all text to lowercase, replace common abbreviations). Then we used the tm. plugin.koRpus package in R (Michalke, 2021) to create corpora for both the raw and the cleaned texts, from which we derived our language measures. We additionally calculated the ratio of lemmatized content words to total words produced, a measure we called the "content ratio." A linear mixed effects model was performed to evaluate group-level differences in each language measure, before and after the text cleaning procedure[6] (note that the content ratio is a ratio of post-cleaning to raw word count, so it only has one model).

3.3. Outlier criteria

To determine outlier criteria, we examined the distribution of participant performance across narrative prompts. We eliminated outliers with attention to two specific criteria. We first excluded any narrative prompts for which the participant produced fewer than five content words. Additionally, we eliminated narratives with type-token ratios corresponding to $z < -1.96$ below the mean for that particular prompt, separately for the PWA and control groups. This outlier criterion served to remove prompts with lexical diversity values far below what is typical for the group.

3.4. Semantic distance derivation

We derived two measures of semantic distance (thematic vs. taxonomic) for each running pair of words within narratives. Thematic semantic relationships tend to involve contextual associations (e.g., fork-blender; Belacchi and Artuso, 2018; Mirman et al., 2017). Word-embedding techniques (e.g., latent semantic analysis, GloVe; Pennington et al., 2014) are uniquely suited for quantifying thematic semantic relationships because of their emphasis on lexical co-occurrence in large corpora (Landauer et al., 1998).

No universal standard exists for quantifying taxonomic semantic distance. In general, taxonomic semantic distance comprises feature-based similarity (e.g., DOG vs. WOLF) relative to contextual similarity (e.g., DOG vs. LEASH) for thematic relationships. Human experiential ratings (e.g., color, sound, emotion) can potentially yield such an index (Binder et al., 2016; Reilly et al., 2016). We derived taxonomic semantic distance between each running word pair by first creating a lookup database composed of seven dimensions from the Lancaster Sensorimotor Norms (Lynott et al., 2019) where each English word was characterized by Likert scale ratings for the following dimensions: auditory, gustatory, haptic, interoceptive, olfactory, visual, and hand-arm. Of the

[3] https://aphasia.talkbank.org/protocol/english/materials-aphasia/description.pdf.

[4] Scripts are freely available for inspection and use at osf.io/93jbx.

[5] Lemmatization promotes term aggregation by transforming morphological derivatives of words into their respective dictionary entries (e.g., independently → independence) based on natural language statistics gleaned from pretrained language corpora.

[6] Each of these models served to predict one language measure (word count per prompt, mean length of utterance, type-token ratio, or content ratio) from a fixed effect of Group (aphasia vs. control) and random effects of participant, prompt, and gender.

motor dimensions, we selected only the hand-arm dimension for the following reasons: (1) although the head motor dimension was identified as the dominant motor dimension most frequently (Lynott et al., 2019), it was often selected as the dominant dimension for words such as "daydream" or "stare," and was thus often selected for non-motoric reasons. (2) The hand-arm motor dimension was selected as the next most frequently dominant motor dimension and appears to have been selected for truly motoric reasons such as reaching and grasping. Thus, we excluded the other motor dimensions so that the remaining perceptual and motor dimensions are weighted equally. We then merged the Lancaster ratings with eight additional dimensions specifically related to emotion from the AffectVec Emotion Database (Raji and de Melo, 2020). These included: excitement, surprise, happiness, fear, anger, contempt, disgust, and sadness. We retained all entries that had ratings from both databases ($N = 24{,}771$ words) and scaled ratings to z-scores based on each semantic dimension's distribution. These procedures resulted in a 15-dimension semantic space spanning 24,771 English words.

We derived thematic distance between each word pair in the AphasiaBank narratives using the cosine distance between each word's GloVe word embedding vectors (Pennington et al., 2014). GloVe word embeddings were chosen as the thematic semantic representations since recent research identified GloVe as better capturing thematic relationships than other vector representations (e.g., word2vec, Kacmajor and Kelleher, 2020). We derived taxonomic distance for the same word pairs by computing cosine distance between our proposed 15-dimensional space. We then transformed each distance by taking the inverse and adding one. This transformation constrained each distance to the range 0–2, such that more dissimilar words were associated with greater distances, and more similar words were associated with smaller distances.

3.5. Calculation of semantic impairment

To measure the degree of semantic impairment in the aphasia group, we extracted each participant's scores from all tasks that are thought to require significant semantic processing [Boston Naming Test (Kaplan et al., 2001); Verb Naming Test (Thompson, 2012); and the Auditory Word Recognition subtest, Sentence Completion subtest, and Responsive Speech subtest from the Western Aphasia Battery-Revised (WAB; Kertesz, 2006)]. For each test, we then z-scored each participant's score relative to the aphasia group. The integrity of each participant's semantic system was calculated as the average of his/her z-scores on all semantic tasks, such that greater semantic impairment was associated with lower composite scores.

3.6. Language measures

To characterize differences in participants' language output across prompts, we calculated participants' (1) mean number of words produced per prompt, (2) mean utterance length, and (3) type token ratio, both before and after the text cleaning procedure. We additionally calculated the ratio of lemmatized content words to total words produced, a measure we called the "content ratio." A linear mixed effects model was performed to evaluate group-level differences in each language measure, before and after the text cleaning procedure[7] (note that the content ratio is a ratio of post-cleaning to raw word count, so it only has one model).

3.7. Analyses

To evaluate the relationship between PWA semantic impairment and their overall aphasia severity (Question 1), we calculated a Pearson correlation of aphasic participants' semantic composite score (a measure of their overall semantic deficit) with their WAB AQ (a measure of their overall aphasia severity).

To evaluate potential differences in thematic and taxonomic semantic distance between individuals with aphasia and neurotypical controls (Question 2), we performed two linear mixed effects models (LMEMs) using the lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017) packages in R. The first model predicted all inter-word thematic semantic distances across all participants and all prompts; the second model predicted all inter-word taxonomic semantic distances. Both models included fixed effects of group (aphasia vs. control) and word count. For each model, a random intercept was included by-individuals and by-prompt. Any significant differences between aphasia and control groups were determined by evaluating whether the fixed effect of group was significant.

To evaluate the relationship of integrity of the semantic system and semantic distance in aphasia (Question 3), we performed an additional set of two LMEMs (one each to predict thematic or taxonomic distance). These models included only PWA, as the neurotypical controls are expected to have intact semantic systems. Both models included fixed effects of semantic impairment (composite score), speech fluency (as measured by the WAB Spontaneous Speech Fluency subscore), and word count. For each model, a random effect was included by-individuals (with a random intercept and slope for prompt). The WAB Spontaneous Speech Fluency subscore is a measure of the ease of speech, assessing word finding difficulty, frequency of paraphasias, hesitations, and syntactic ability. Including speech fluency as a fixed effect served to ensure that a significant relationship of semantic impairment with semantic distance was not explained by reductions in the fluency of production. A significant relationship of semantic system impairment with semantic distance was determined by evaluating whether the fixed effect of semantic impairment was significant. To examine whether a non-linear relationship existed between semantic impairment and semantic distances, we performed two polynomial mixed effects models (one each for thematic and taxonomic semantic distance). These models were identical to those described above, with a quadratic term of the square of the semantic composite score as an additional fixed effect. A non-linear relationship in each model was determined by evaluating whether the quadratic term of semantic impairment was significant.

Finally, to evaluate the relationship between participants' mean thematic and taxonomic inter-word semantic distances (Question 4), we performed two Pearson correlations (one per group). For each participant, we calculated his/her median thematic and median taxonomic inter-word semantic distances across all prompts, and then correlated these median distances across all participants (median, rather than mean, values were used since participants' inter-word semantic distances were not normally distributed). To evaluate whether or not the strength of this relationship differed between groups, we implemented the Fisher's z-test in the R cocor package (Diedenhofen and Musch, 2015). Additionally, to evaluate whether the slope of the correlation differed between groups, we performed a multiple linear regression model to predict participants' mean taxonomic semantic distance from their mean thematic semantic distance, group, and an interaction of mean thematic distance and group. If the slope of the correlation differed between groups, then this model would identify a significant effect of the interaction of group and mean thematic distance. Given the directionality of multiple regressions, we ran a second multiple linear regression to predict participants' mean thematic semantic distance from their mean taxonomic semantic distance, group, and an interaction

[7] Each of these models served to predict one language measure (word count per prompt, mean length of utterance, type-token ratio, or content ratio) from a fixed effect of Group (aphasia vs. control) and random effects of participant, prompt, and gender.

of mean taxonomic distance and group.[8] As with the previous regression, a significant effect of the interaction of group and mean taxonomic distance would indicate significant differences in slope between groups.

4. Results

4.1. Language measures

Across all language measures (total words produced per prompt, mean length of utterance, type-token ratio, and content ratio), PWA exhibited significantly different values than control participants, as expected (see Table 2, Supplementary Table S1). Although the total words produced, mean length of utterance, and content ratio were significantly greater for the control group, the type-token ratio was significantly greater for the aphasia group. A correlation matrix depicting Pearson's correlation values for these language measures, WAB AQ, WAB Spontaneous Speech Fluency sub-score, the semantic composite score, and the mean semantic distance measures in the aphasia group (Fig. 1) indicates that the semantic distance measures are moderately correlated with many of the language measures (see Fig. 2).

4.2. Relationship of semantic impairment to aphasia severity

Regarding the relationship of semantic impairment to overall aphasia severity, a Pearson correlation indicated that participants' semantic composite score was strongly positively correlated with their overall aphasia severity as measured by their WAB AQ [r (250) = 0.903 $p < .001$], as predicted.

4.3. Group-level differences in semantic distance

Approximately 10% of words that participants produced were replaced (e.g., she → WOMAN) in the lemmatization pipeline (10.4% of words by PWA; 9.5% of words by control participants). Additionally, approximately 60% of words produced by participants in both groups were identified as stop words (e.g., THE, AND) and were removed from semantic distance computations (62.3% of words by PWA; 58.9% of words by control participants).

We observed significant differences in mean thematic and taxonomic semantic distances between PWA and control participants. As hypothesized, PWA exhibited lower thematic distances (0.723 ± 0.079) than control participants (0.791 ± 0.025; Table 3; Fig. 2). Contrary to our hypothesis, we observed that PWA also exhibited lower taxonomic distances (0.697 ± 0.097) than control participants (0.779 ± 0.029; Table 3; Fig. 2). In LMEMs developed to identify potential group-level differences in semantic distance (see Methods), the fixed effect of group (aphasia vs. control) significantly predicted participants' inter-word taxonomic ($p < .001$) and thematic ($p < .001$) semantic distance, indicating that control participants have greater overall taxonomic and thematic semantic distances than PWA (see Tables 4–5).

Regarding the relationship between semantic impairment and thematic and taxonomic semantic distance, the fixed effect of semantic impairment (as measured by the semantic composite score) significantly predicted participants' thematic ($p < .001$) and taxonomic ($p < .001$) semantic distances, indicating that better semantic ability in aphasia was associated with greater semantic distances (see Tables 6–7, Fig. 3). Participants' speech fluency additionally predicted their taxonomic semantic distances, with higher speech fluency scores associated with greater taxonomic distances ($p < .01$). Additional polynomial mixed effects models, which include an additional quadratic semantic composite term, show no significant relationship of the quadratic term to thematic or taxonomic semantic distances (both $ps > .05$; Supplementary Tables S2-S3). Thus, the linear relationship of semantic integrity positively correlating with semantic distances holds across the spectrum of semantic impairment.

Finally, regarding the relationship between thematic and taxonomic semantic distances in each participant group, Pearson correlations indicated that median taxonomic and thematic semantic distances were moderately positively correlated in the aphasia group ($r = .476$) and more modestly correlated in the control group ($r = 0.165$; see Fig. 4). A Fisher's z-transformation indicated that the aphasia group's semantic distance correlation was significantly stronger than that of the control group ($z = 3.72$, $p < .001$), the opposite direction as we predicted. Additionally, the presence of a significant interaction effect in the multiple linear regressions indicates that the slopes of the two correlations are significantly different (see Tables 8–9 and Supplementary Tables S4-S5).

Table 2
Language measures by group.

Total words produced per prompt	Aphasia (N = 259)	Control (N = 203)
Raw, Mean (SD)	118 (84.9)	192 (78.0)
Post-cleaning, Mean (SD)	37.8 (27.6)	73.2 (29.7)
Mean length of utterance (words)		
Raw, Mean (SD)	5.73 (2.39)	10.40 (2.18)
Post-cleaning, Mean (SD)	2.49 (0.80)	4.25 (0.99)
Type-token ratio		
Raw, Mean (SD)	0.56 (0.09)	0.55 (0.06)
Post-cleaning, Mean (SD)	0.68 (0.09)	0.68 (0.06)
Content ratio		
Mean (SD)	0.34 (0.08)	0.39 (0.03)

Table 2. Mean and standard deviation of four language measures in each participant group. "Raw" refers to the words spoken by the participant, and "Post-cleaning" refers to the same measures after the text cleaning procedure was applied. Means were first calculated across all prompts within each participant, and then across all participants. Standard deviations were calculated across participants within-group.

5. General discussion

The grand promise of language science is that parallel advances in computational linguistics, psycholinguistics, and neurolinguistics might one day yield the data needed to converge upon a unified theory of language. Such a synthesis requires integration across multiple linguistic processes, each unfolding at different scales. This is an especially important consideration for theories of lexical-semantic processing, which must account for compositionality in the construction of message-level meaning. These efforts must ultimately reconcile an extensive body of research elicited by single word paradigms with a sparser literature focused on meaning in connected language. Here we evaluated an intermediary step between these domains by linking semantic characteristics of single words to their neighbors (i.e., inter-word semantic distance) in the context of running narratives. This method yielded tens of thousands of temporally ordered data points, reflecting the relatedness of one word to the next.

Our research aims followed a sequential logic. First, we evaluated the relationship between lexical-semantic impairment and aphasia severity. These factors were strongly positively correlated (Pearson's $r = .903$), indicating that the integrity of lexical-semantic processing is a robust predictor of language processing in aphasia. This is in line with prior literature indicating tight relationships between semantics task performance and overall language impairment (e.g., Carpenter et al., 2020).

In our second aim, we contrasted inter-word semantic distances in aphasia relative to age-matched controls. The aphasia group exhibited significantly decreased thematic and taxonomic semantic distance relative to the control group, indicating compression of retrieval processes in aphasia.

[8] Supplementary analyses were also performed to predict participants' median, rather than mean, semantic distances (see Supplementary Tables S5-S6).

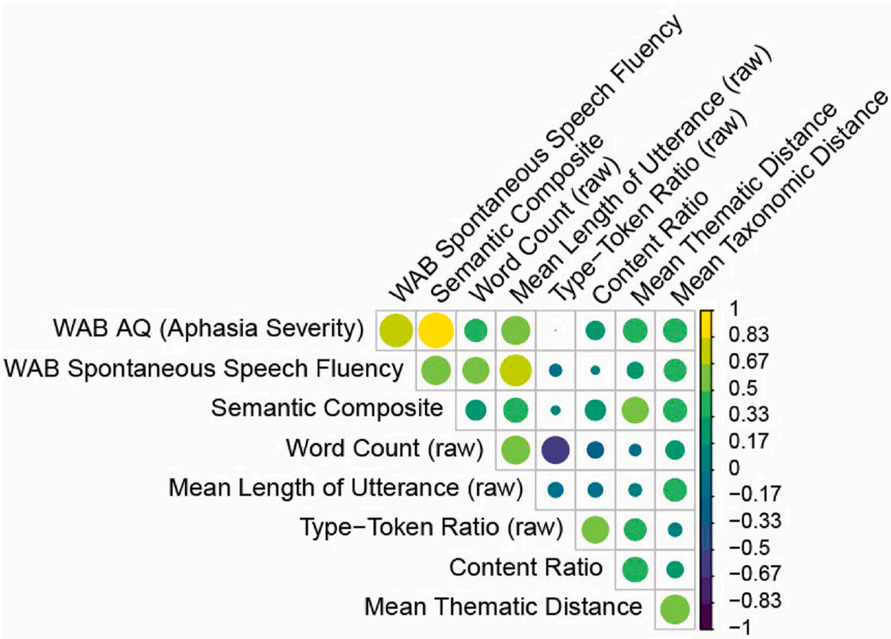

Fig. 1. Correlation matrix indicating Pearson's correlations for each pair of language measures in the aphasia group. Circle area shows the absolute value of the correlation coefficient. More yellow indicates more positive correlation coefficients; more purple indicates more negative correlation coefficients. "Raw" indicates that the measure was calculated before implementation of the text cleaning algorithm. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 3
Mean taxonomic and thematic semantic distance by group.

Thematic Distance	Aphasia (n = 259)	Control (n = 203)
Thematic Distance		
Mean (SD)	0.723 (0.079)	0.791 (0.025)
Taxonomic Distance		
Mean (SD)	0.697 (0.097)	0.779 (0.029)

Table 4
Results of LMEM predicting thematic distance from fixed effect of participant group.

Fixed effects	Estimate	Std. Error	t-value	P-value
Intercept	−0.1785	0.0349	−5.112	**< 0.001**
Group (control)	0.2701	0.0178	15.143	**< 0.001**
Word Count	−0.0086	0.0039	−2.193	**< 0.05**

Table 5
Results of LMEM predicting taxonomic distance from fixed effect of participant group.

Fixed effects	Estimate	Std. Error	t-value	P-value
Intercept	−0.1508	0.0342	−4.411	**< 0.001**
Group (control)	0.1785	0.0118	15.079	**< 0.001**
Word Count	−0.0162	0.0039	−4.131	**< 0.001**

Table 6
Results of LMEM predicting thematic distance from fixed effects of semantic composite score, speech fluency, and word count.

Fixed effects	Estimate	Std. Error	t-value	P-value
Intercept	−0.1701	0.0126	−13.486	**< 0.001**
Semantic composite	0.1895	0.0134	14.159	**< 0.001**
Speech fluency	0.00081	0.0132	0.061	0.95
Word Count	−0.0471	0.0011	−4.423	**< 0.001**

Given the significant group-level differences in thematic and taxonomic distance, we then evaluated the relationship between inter-word semantic distance and lexical-semantic processing in aphasia. These analyses revealed moderate positive correlations between both semantic distance measures and semantic impairment. These relationships were such that worsening lexical-semantic deficits were associated with more compression of inter-word semantic distances. That is, greater inter-word semantic distances (i.e., more control-like) were associated with better lexical-semantic processing, even when controlling for participants' speech fluency.

Table 7
Results of LMEM predicting taxonomic distance from fixed effects of semantic composite score, speech fluency, and word count.

Fixed effects	Estimate	Std. Error	t-value	P-value
Intercept	−0.1185	0.0099	−11.968	**< 0.001**
Semantic composite	0.0577	0.0121	4.752	**< 0.001**
Speech fluency	0.0378	0.0121	3.134	**< 0.01**
Word Count	0.0121	0.0085	1.418	0.16

In our final aim, we investigated the relationship between participants' thematic and taxonomic inter-word semantic distances. In the aphasia group, taxonomic and thematic distances were moderately positively correlated (Pearson's $r = .476$) and more modestly positively correlated for the control group (Pearson's $r = 0.165$). These positive correlations indicate that thematic and taxonomic semantics (1) are not fully dissociable, and (2) are more closely related in individuals with lexical-semantic retrieval impairments. Below we discuss these findings and their implications in more detail.

5.1. Inter-word taxonomic vs. thematic distance: differences and similarities

Patients with circumscribed brain damage have manifested a wide variety of category-specific impairments. Such deficits, although often controversial, have spanned both broad dichotomies such as natural kinds vs. manufactured artifacts (Warrington and McCarthy, 1983, 1987; Warrington and Shallice, 1984), nouns vs. verbs (Laiacona and Caramazza, 2004), and abstract vs. concrete words (Warrington and Shallice, 1984), as well as more granular semantic distinctions such as fruits and vegetables (Hillis and Caramazza, 1991). Recent studies have focused on the major distinction between taxonomic vs. thematic semantic knowledge and the neural substrates of these systems. Current

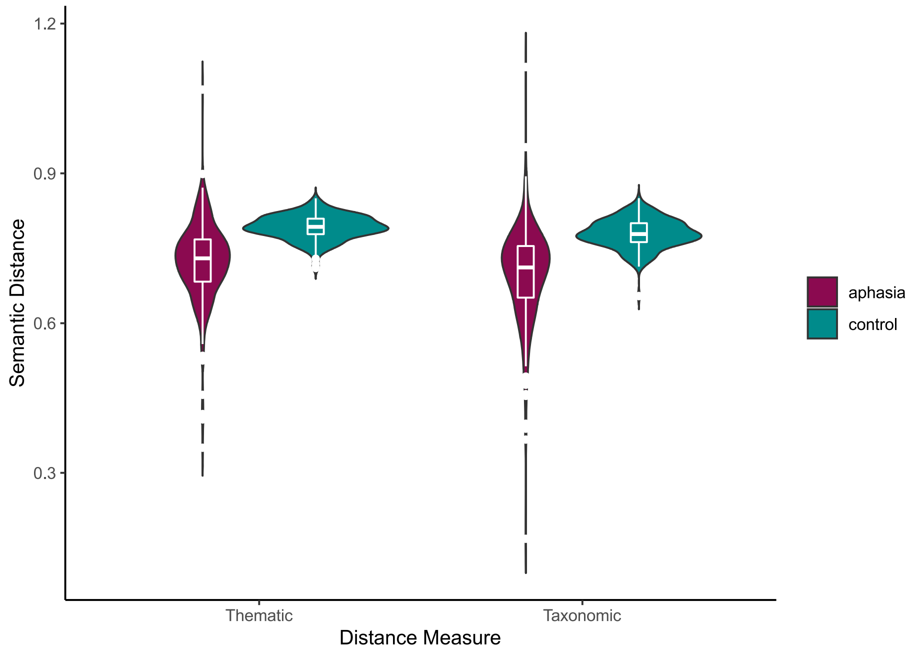

Fig. 2. Distribution of mean thematic and taxonomic semantic distance for the aphasia and control groups in AphasiaBank. The box plots indicate the median, interquartile range, and variability for each group, and the violin plots additionally illustrate each group's kernel probability density.

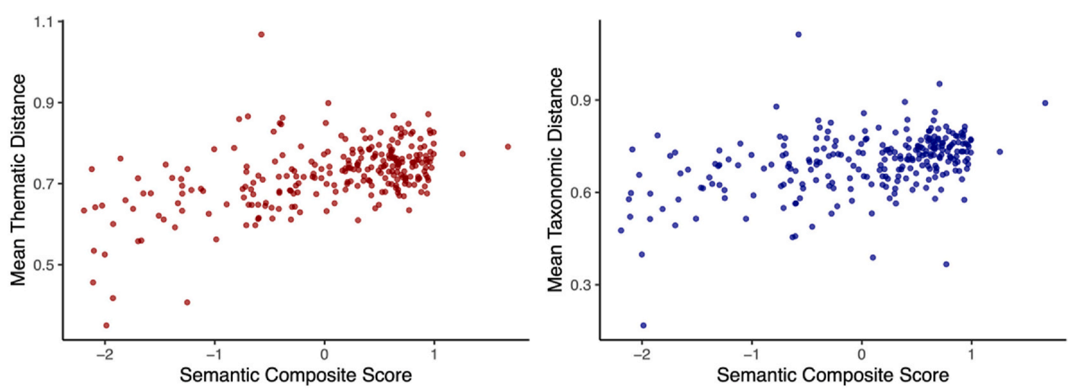

Fig. 3. Relationship of the semantic composite score to thematic semantic distance (left, red) and taxonomic (right, blue) in the aphasia group. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

thinking is that taxonomic or feature-based knowledge (e.g., visual feature overlap) is mediated by anterior and inferior regions of the temporal lobes. For example, dogs and wolves share considerable feature overlap and may be represented within a phylogenetic tree structure that is susceptible to selective pruning or more widespread damage. In contrast, dogs and leashes are primarily linked through contextual associations and/or common event schemas (Mirman et al., 2017). This type of thematic semantic knowledge is thought to rely more heavily upon associative processes within the temporoparietal junction.

The canonical distribution of brain damage in aphasia yields principled predictions about the nature of how semantic distance might be perturbed. The middle cerebral artery (MCA) territory is the most vulnerable region of the brain to cerebral ischemia (Nogles and Galuska, 2021). This swath of cortex includes much of perisylvian language network extending from anterior regions such as Brodmann areas 44/45 to posterior temporal and inferior parietal structures encompassing the temporoparietal junction. Unilateral anterior temporal lobe damage is not unheard of, however, particularly in posterior communicating artery (PCA) strokes. Yet, the relative impact of aphasia is more commonly seen in regions such as the angular gyrus that are thought to support thematic or association-based semantic knowledge.

Further, previous research indicates that aphasia often results in topic perseveration, with individuals exhibiting decreased ability to shift from one topic to another (Chiou and Kennedy, 2009). Given the perseverative nature of speech in aphasia, we hypothesized that the words that PWA produce would be more thematically related than those produced by controls. Our finding of compressed thematic semantic distances in aphasia supports this hypothesis. Further, given other evidence that individuals with other conditions (e.g., schizophrenia, Elvevåg et al., 2007) tend to exhibit greater semantic distances than controls, these findings suggest that there is a "sweet spot" of semantic distance in narrative production. Although abnormally large semantic distances (as in the speech of individuals with schizophrenia) may be perceived as disjointed or inharmonious (Elvevåg et al., 2007), speech in aphasia, with its abnormally low semantic distances, may be perceived as thematically narrow and repetitive (Behrns et al., 2009; Ulatowska et al., 1981). Future research should evaluate the prediction that thematic semantic distance impacts perceived cohesiveness of narratives.

The anterior temporal lobe is thought to be the hub of taxonomic relationships (Lewis et al., 2015) and is relatively spared in individuals with post-stroke aphasia (Jefferies et al., 2007). As such, we hypothesized that taxonomic distances in aphasia would not be as compressed as

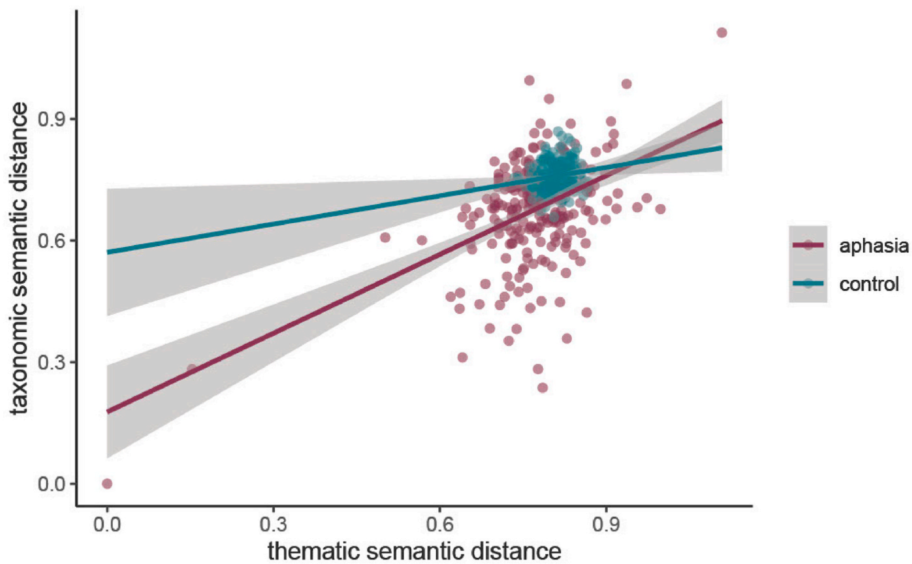

Fig. 4. Correlation of median thematic and taxonomic semantic distance by group. The aphasia group's median semantic distance measures are more highly correlated and have a greater slope than that of the control group. Shaded areas represent 95% confidence intervals.

Table 8
Results of multiple linear regression predicting mean taxonomic distance.

Coefficient	Estimates	Std. Error	t-value	P-value
Intercept	0.152	0.035	4.395	**< 0.001**
Thematic distance	0.755	0.048	15.880	**< 0.001**
Group (control)	0.425	0.139	3.047	**< 0.01**
Interaction: Thematic distance x Group (control)	−0.499	0.177	−2.818	**< 0.01**
Observations: 462 Adjusted $R^2 = 0.500$				

Table 8. Results of multiple linear regression predicting mean taxonomic distance from independent predictors of (1) mean thematic distance, (2) group (aphasia vs. control), and (3) interaction of mean thematic distance and group. Results indicate a significant correlation of mean thematic distance and mean taxonomic distance ($p < .001$) and a significant difference in the slopes of the correlations between the two groups ($p < .01$).

Table 9
Results of multiple linear regression predicting mean thematic distance.

Coefficient	Estimates	Std. Error	t-value	P-value
Intercept	0.373	0.022	16.667	**< 0.001**
Taxonomic distance	0.502	0.032	15.798	**< 0.001**
Group (control)	0.272	0.096	2.828	**< 0.01**
Interaction: Thematic distance x Group (control)	−0.314	0.124	−2.528	**< 0.05**
Observations: 462 Adjusted $R^2 = 0.504$				

Table 9. Results of multiple linear regression predicting mean thematic distance from independent predictors of (1) mean taxonomic distance, (2) group (aphasia vs. control), and (3) interaction of mean taxonomic distance and group. Results indicate a significant correlation of mean thematic distance and mean taxonomic distance ($p < .001$) and a significant difference in the slopes of the correlations between the two groups ($p < .05$).

thematic distances are. Additionally, PWA often produce semantic paraphasias and circumlocutory speech in attempts to access lexical items (Andreetta et al., 2012; Azhar et al., 2016). Occasionally, paraphasias may represent concepts that are distantly related to the intended target (Ardila and Rosselli, 1993). Therefore, we predicted that the aphasia group would exhibit greater taxonomic distances than the control group. However, we found that, similar to thematic distances, the aphasia group exhibited compressed taxonomic distances relative to the control group. The compressed taxonomic distances in aphasia may be due to a few factors. Previous research has shown that semantic processing of more closely related concepts occurs more quickly and requires less cognitive control than more distally related concepts (Geller et al., 2019; Lewis et al., 2015b). As such, the lexical-semantic retrieval deficit in aphasia may make selection of less taxonomically related items more challenging than selection of more closely related items. In narrative discourse, this would manifest behaviorally as lower inter-word distances and increased overall taxonomic relatedness. This is supported by our finding of tight thematic relationships in PWA. Specifically, difficulty in topic-shifting (e.g., thematic jumps) may also constrain inter-word taxonomic relationships to more limited categories. For example, one participant with aphasia had difficulty transitioning from a description of Cinderella's family to her work as their servant, saying "She lived with her stepmother and two stepsisters or, not nephews, but, well, stepsisters, I guess." This perseveration of tightly taxonomically linked relationships (i.e., STEPMOTHER – STEPSISTER – NEPHEW – STEPSISTER) results in compressed taxonomic inter-word distances. In summary, in both relational systems, PWA tended to select items more closely related to the previous word, resulting in more compressed thematic and taxonomic distances.

This compression across both taxonomic and thematic systems was also reflected in the relationship between the two systems. Specifically, we found that thematic and taxonomic inter-word distances were positively correlated for both the aphasia and control groups. These positive correlations indicate that individuals with lower mean thematic distances also exhibited lower mean taxonomic distances (i.e., compression operated similarly across systems). This finding may be surprising given previous evidence that aphasia typically results in much greater impairment to either thematic or taxonomic semantics rather than significant impairment to both systems (Kalénine et al., 2012). However, the positive associations of thematic and taxonomic distances are in line with previous research indicating that taxonomic and thematic semantics are not fully dissociable. Although there is general agreement that taxonomic and thematic semantics partially dissociate behaviorally (e.g., Seckin et al., 2016; Thompson et al., 2017) and neurally (e.g., Carota et al., 2021; Savic et al., 2017; Zhang et al., 2021), other studies have identified strong correlations in thematic and taxonomic semantics (e.g., Wang et al., 2018). This study provides additional evidence that thematic and taxonomic semantics are not fully dissociable and may be similarly impacted in narrative discourse in aphasia.

6. Limitations

There are a few limitations of the current study. First, although aphasia is often discussed as a lesion model of lexical-semantic access, we have since come to understand that the distinction between semantic representation and controlled lexical-semantic access is nuanced. Lambon Ralph, Jefferies, Rogers, Patterson and their many colleagues have argued that lexical-semantic control and representational knowledge are highly interactive components of the semantic system (e.g., Lambon Ralph et al., 2017). Moreover, recent studies have also shown that individuals with aphasia do sometimes exhibit deficits in semantic knowledge in addition to deficits in lexical-semantic retrieval, thus complicating the characterization of aphasia as a deficit specifically in retrieval (Chapman et al., 2020).

Additionally, although Kacmajor and Kelleher (2020) recently showed that GloVe represents thematic relationships better than other vector representations (e.g., word2vec), GloVe does also capture taxonomic relationships to a lesser degree. As such, future research should investigate whether other co-occurrence models (e.g., n-gram frequency models, as in Banks et al., 2021) better capture purely thematic semantic distances. Further, in this study, we averaged semantic distances across all inter-word pairs, collapsing by narrative prompt, rather than examining the time series of semantic flow over the course of narrative production. By collapsing semantic distances across all narrative prompts, we have examined semantic flow in a "brute force" manner, under the assumption that narrative production is taxing in aphasia (Ulatowska et al., 1981) and typically produces semantic paraphasias (Andreetta et al., 2012). Future studies should examine how semantic flow changes in the words directly following a paraphasia. Additionally, although participants, especially individuals with aphasia, often repeat words or phrases, our semantic distance computations did not remove immediate repetitions of the same word. Given that our pipeline removes punctuation, it does not distinguish between a true repetition such as "The cat chased the dog, was it a dog?" and an utterance such as "The cat chased the dog. The dog ran away." In both cases, the utterances would include the repeated lemma DOG, resulting in inter-word semantic distances of zero. Since apparent repetitions of words may be true repetitions or an effect of the lemmatization processing, the analyses presented here do not remove repeated words. To address the concern that repetitions might bias the results reported here, we also performed all analyses with the repetitions removed, although we note that the removed repetitions include both true repetitions and false repetitions due to our lemmatization processing (see Supplementary Information). When repetitions were removed, we found that PWA still exhibited significantly lower thematic and taxonomic distances than controls (Supplementary Tables S6-S8; Supplementary Figure S2) and that better semantic ability was associated with greater thematic distances (Supplementary Table S9; Supplementary Figure S3), as we found when repetitions were included. Although the significant correlations of thematic and taxonomic distance (Supplementary Figure S4; Supplementary Tables S11-S12) and effect of semantic ability on taxonomic distances (Supplementary Table S10) do not hold when repetitions were removed, future research should assess whether these relationships exist when only true repetitions are excluded. Finally, given that this retrospective study included only participants in the AphasiaBank database, the analyses were constrained to the predominantly White, monolingual English-speaking, highly educated participants in AphasiaBank. Future studies should include participants of more diverse backgrounds to confirm that these findings are reliable across diverse groups.

7. Future directions & conclusion

Although individuals with lexical-semantic retrieval deficits (aphasia) exhibit deficits in discourse and narrative production, most research on their deficits examines impairment at the single-word level. However, far less is known about how lexical-semantic deficits compound over the course of more naturalistic language. In this study, we build upon previous research on lexical-semantic access impairments by investigating semantic flow (semantic distance) during narrative production in individuals with aphasia and age-matched controls. We found that lexical-semantic access deficits compound during narrative production in aphasia, such that semantic distances were decreased relative to controls. Further, semantic distance appears to be a reliable predictor of semantic impairment in aphasia, with more compression associated with worse impairment. This semantic compression stands in contrast to other disordered populations (e.g., schizophrenia), who tend to exhibit more disjointed, expanded semantic distances. Future work may examine the role of semantic distance in listeners' judgments of narrative quality. The findings we report indicate that individuals with aphasia exhibit marked topic perseveration and discourse-level semantic impairments, extending our knowledge about lexical-semantic retrieval processes during narrative production. Future research should investigate the time course of semantic distance over narrative production. Examination of semantic distance over the course of a narrative may uncover important information on the compounding of lexical-semantic deficits over multiword utterances. More broadly, semantic distance forms a bridge between analyses of single words and narratives, making it a promising measure for future work examining how language operates at the multiword scale.

Acknowledgements

We thank the many participants and researchers who contributed to AphasiaBank. This work was funded by US Public Health Service Grant R01 DC013063.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuropsychologia.2022.108235.

References

Albert, M.L., Goodglass, H., Helm, N.A., Rubers, A.B., Alexander, M.P., 1981. Clinical Aspects of Dysphasia. Springer-Verlag, New York.

Andreetta, S., Cantagallo, A., Marini, A., 2012. Narrative discourse in anomic aphasia. Neuropsychologia 50 (8), 1787–1793 https://doi.org/10/f33nfh.

Ardila, A., Rosselli, M., 1993. Language deviations in aphasia: a frequency analysis. Brain Lang. 44, 165–180.

Azhar, A., Maqbool, S., Awais Butt, G., iftikhar, S., Iftikhar, G., 2016. Frequency of aphasia and its symptoms in stroke patients. J. Speech Pathol. Ther. 2 (1) https://doi.org/10/gngtrd.

Banks, B., Wingfield, C., Connell, L., 2021. Linguistic distributional knowledge and sensorimotor grounding both contribute to semantic category production. Cognit. Sci. 45, e13055.
Bates, E., Goodman, J.C., 1997. On the inseparability of grammar and the lexicon: evidence from acquisition, aphasia, and real-time processing. Lang. Cognit. Process. 12 (5–6), 507–584 https://doi.org/10/d8qwgt.
Bates, Douglas, Mächler, Martin, Bolker, Ben, Walker, Steve, 2015. Fitting linear mixed-effects models using lme4. J. Statist. Software 67 (1), 1–48. https://doi.org/10.18637/jss.v067.i01.
Beaty, R.E., Johnson, D.R., 2021. Automating creativity assessment with SemDis: an open platform for computing semantic distance. Behav. Res. Methods 53 (2), 757–780. https://doi.org/10.3758/s13428-020-01453-w.
Behrns, I., Wengelin, Å., Broberg, M., Hartelius, L., 2009. A comparison between written and spoken narratives in aphasia. Clin. Linguist. Phon. 23 (7), 507–528. https://doi.org/10.1080/02699200902916129.
Belacchi, C., Artuso, C., 2018. How taxonomic and thematic associations in semantic memory modulate recall in young through old-old adults. Psychol. Aging 33 (7), 1060–1069 https://doi.org/10/gngszm.
Benson, D.F., 1979. Aphasia, Alexia and Agraphia. Churchill Livingstone, New York.
Binder, J.R., Conant, L.L., Humphries, C.J., Fernandino, L., Simons, S.B., Aguilar, M., Desai, R.H., 2016. Toward a brain-based componential semantic representation. Cogn. Neuropsychol. 1–45. https://doi.org/10.1080/02643294.2016.1147426.
Broca, M.P., 1861. Remarques sur le siege de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole). Bullet. Soc. Anat. 6, 330–357.
Bryant, L., Ferguson, A., Spencer, E., 2016. Linguistic analysis of discourse in aphasia: a review of the literature. Clin. Linguist. Phon. 30 (7), 489–518 https://doi.org/10/gf2tdh.
Buckingham, H.W., Rekart, D.M., 1979. Semantic paraphasia. J. Commun. Disord. 12, 197–209.
Carota, F., Nili, H., Pulvermüller, F., Kriegeskorte, N., 2021. Distinct fronto-temporal substrates of distributional and taxonomic similarity among words: evidence from RSA of BOLD signals. Neuroimage 224, 117408 https://doi.org/10/gngs3r.
Carpenter, E., Rao, L., Peñaloza, C., Kiran, S., 2020. Verbal fluency as a measure of lexical access and cognitive control in bilingual persons with aphasia. Aphasiology 34 (11), 1341–1362. https://doi.org/10.1080/02687038.2020.1759774.
Castro, N., Stella, M., 2019. The multiplex structure of the mental lexicon influences picture naming in people with aphasia. J. Complex Netw. 7 (6), 913–931.
Chapman, C.A., Hasan, O., Schulz, P.E., Martin, R.C., 2020. Evaluating the distinction between semantic knowledge and semantic access: evidence from semantic dementia and comprehension-impaired stroke aphasia. Psychonomic Bullet. Rev. 27 (4), 607–639 https://doi.org/10/ghxbjm.
Chiou, H.S., Kennedy, M.R.T., 2009. Switching in adults with aphasia. Aphasiology 23 (7–8), 1065–1075 https://doi.org/10/d23sqs.
Corbett, F., Jefferies, E., Ehsan, S., Lambon Ralph, M.A., 2009. Different impairments of semantic cognition in semantic dementia and semantic aphasia: evidence from the non-verbal domain. Brain 132 (9), 2593–2608 https://doi.org/10/cmzjpm.
Cree, G.S., McRae, K., 2003. Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). J. Exp. Psychol. Gen. 132 (2), 163–201 https://doi.org/10/b83rwd.
Cunningham, K.T., Haley, K.L., 2020. Measuring lexical diversity for discourse analysis in aphasia: moving-average type–token ratio and word information measure. J. Speech Lang. Hear. Res. 63 (3), 710–721. https://doi.org/10.1044/2019_JSLHR-19-00226.
Diedenhofen, B., Musch, J., 2015. Cocor: a comprehensive solution for the statistical comparison of correlations. PLoS One 10 (4), e0121945.
Duff, M.C., 2021. Introduction to the special issue for the 49th clinical aphasiology conference. Am. J. Speech Lang. Pathol 30 (1S), 325–325.https://doi.org/10/gnc9fp.
Elvevåg, B., Foltz, P.W., Weinberger, D.R., Goldberg, T.E., 2007. Quantifying incoherence in speech: an automated methodology and novel application to schizophrenia. Schizophr. Res. 93 (1–3), 304–316 https://doi.org/10/cjrx9v.
Friederici, A.D., Alter, K., 2004. Lateralization of auditory language functions: a dynamic dual pathway model. Brain Lang. 89 (2), 267–276 https://doi.org/10/bwmp6k.
Garrard, P., Lambon Ralph, M.A., Hodges, J.R., Patterson, K., 2001. Prototypicality, distinctiveness, and intercorrelation: analyses of the semantic attributes of living and nonliving concepts. Cogn. Neuropsychol. 18 (2), 125–174 https://doi.org/10/fp7jjv.
Geller, J., Thye, M., Mirman, D., 2019. Estimating effects of graded white matter damage and binary tract disconnection on post-stroke language impairment. Neuroimage 189, 248–257. https://doi.org/10.1016/j.neuroimage.2019.01.020. January.
Geschwind, N., 1965. Disconnection syndromes in animals and man. Brain 88, 237–294.
Günther, F., Rinaldi, L., Marelli, M., 2019. Vector-space models of semantic representation from a cognitive perspective: a discussion of common misconceptions. Perspect. Psychol. Sci. 14 (6), 1006–1033.
Hayward, R.W., Naeser, M.A., Zatz, L.M., 1977. Cranial computed tomography in aphasia. Radiology 123, 653–660.
Hillis, A.E., Caramazza, A., 1991. Category-specific naming and comprehension impairment: a double dissociation. Brain 114, 2081–2094.
Hills, T.T., Kenett, Y.N., 2022. Is the mind a network? Maps, vehicles, and skyhooks in cognitive network science. Top. Cognit. Sci. 14 (1), 189–208 https://doi.org/https://doi.org/10.1111/tops.12570.
Hughlings Jackson, J., 1874/1932. On affections of speech from disease of the brain. In: Taylor, J. (Ed.), Selected Writings of John Hughlings Jackson, vol. 2. Hodder & Stoughton, London, United Kingdom, pp. 155–204.
Jefferies, E., Baker, S.S., Doran, M., Ralph, M.A.L., 2007. Refractory effects in stroke aphasia: a consequence of poor semantic control. Neuropsychologia 45 (5), 1065–1079 https://doi.org/10/bqbfkh.
Jefferies, E., Lambon Ralph, M.A., 2006. Semantic impairment in stroke aphasia versus semantic dementia: a case-series comparison. Brain 129 (8), 2132–2147 https://doi.org/10/c4m7tn.
Kacmajor, M., Kelleher, J.D., 2020. Capturing and measuring thematic relatedness. Comput. Humanit. 54, 645–682. https://doi.org/10.1007/s10579-019-09452-w.
Kalénine, S., Mirman, D., Buxbaum, L.J., 2012. A combination of thematic and similarity-based semantic processes confers resistance to deficit following left hemisphere stroke. Front. Hum. Neurosci. 6 https://doi.org/10.3389/fnhum.2012.00106.
Kaplan, E., Goodglass, H., Weintraub, S., 2001. In: Boston Naming Test. Austin, TX: Pro-.
Kenett, Y.N., 2018. Going the extra creative mile: the role of semantic distance in creativity – theory, research, and measurement. In: Jung, R.E., Vartanian, O. (Eds.), The Cambridge Handbook of the Neuroscience of Creativity. Cambridge University Press, pp. 233–248.
Kenett, Y.N., 2019. What can quantitative measures of semantic distance tell us about creativity? Curr. Opin. Behav. Sci. 27, 11–16 https://doi.org/https://doi.org/10.1016/j.cobeha.2018.08.010.
Khedr, E.M., Abbass, M.A., Soliman, R.K., Zaki, A.F., Gamea, A., Abo El-Fetoh, N., Abdel-Aaal, M.A., 2020. A hospital-based study of post-stroke aphasia: frequency, risk factors, and topographic representation. Egypt. J. Neurol. Psychiatr. Neurosurg. 56, 2.
Kumar, A.A., 2021. Semantic memory: a review of methods, models, and current challenges. Psychonomic Bull. Rev. 28 (1), 40–80. https://doi.org/10.3758/s13423-020-01792-x.
Kuznetsova, Alexandra, Brockhoff, Per B., Christensen, Rune H.B., 2017. lmerTest package: Tests in linear mixed effects models. J. Statist. Software 82 (13), 1–26. https://doi.org/10.18637/jss.v082.i13.
Laiacona, M., Caramazza, A., 2004. The NOUN/verb dissociation IN language production: varieties OF causes. Cogn. Neuropsychol. 21 (2–4), 103–123 https://doi.org/10/csd22s.
Lakoff, G., 1972. Structural complexity in fairy tales. Stud. Man 1, 128–150.
Lambon Ralph, M.A., Jefferies, E., Patterson, K., Rogers, T.T., 2017. The neural and computational bases of semantic cognition. Nat. Rev. Neurosci. 18, 42–55.
Landauer, T.K., Foltz, P.W., Laham, D., 1998. An introduction to latent semantic analysis. Discourse Process 25 (2–3), 259–284 https://doi.org/10/fsc78g.
Levelt, W., 2012. A History of Psycholinguistics: the Pre-chomskyan Era. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199653669.001.0001.
Lewis, G.A., Poeppel, D., Murphy, G.L., 2015. The neural bases of taxonomic and thematic conceptual relations: an MEG study. Neuropsychologia 68, 176–189 https://doi.org/10/f6269q.
Lichtheim, L., 1885. On aphasia. Brain 7, 433–484.
Lynott, D., Connell, L., Brysbaert, M., Brand, J., Carney, J., 2019. The Lancaster Sensorimotor Norms: multidimensional measures of perceptual and action strength for 40,000 English words. Behav. Res. Methods 1–21.
MacWhinney, B., Fromm, D., Forbes, M., Holland, A., 2011. AphasiaBank: methods for studying discourse. Aphasiology 25 (11), 1286–1307. PubMed.https://doi.org/10/cd6vc3.
Malt, B.C., 2020. Words, thoughts, and brains. Cogn. Neuropsychol. 37 (5–6), 241–253.
Mandera, P., Keuleers, E., Brysbaert, M., 2015. How useful are corpus-based methods for extrapolating psycholinguistic variables? Q. J. Exp. Psychol. 8, 1628–1642 https://doi.org/https://doi.org/10.1080/17470218.2014.988735.
Mandera, P., Keuleers, E., Brysbaert, M., 2017. Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: a review and empirical validation. J. Mem. Lang. 92, 57–78 https://doi.org/https://doi.org/10.1016/j.jml.2016.04.001.
Michalke, M., 2021. tm.plugin.koRpus: Full Corpus Support for the 'koRpus' Package. Version 0.4-2. https://reaktanz.de/?c=hacking&s=koRpus.
Mirman, D., Britt, A.E., 2014. What we talk about when we talk about access deficits. Phil. Trans. Biol. Sci. 369 (1634), 20120388 https://doi.org/10/gngtfp.
Mirman, D., Graziano, K.M., 2012. J. Exp. Psychol. Gen. 141 (4), 601–609.
Mirman, D., Landrigan, J.-F., Britt, A.E., 2017. Taxonomic and thematic semantic systems. Psychol. Bull. 143 (5), 499–520 https://doi.org/10/gkbqjp.
National Aphasia Association, 2021. Aphasia FAQs. Retrieved November 3, 2021, from. www.aphasia.org/aphasia-faqs/.
Nogles, T.E., Galuska, M.A., 2021. Middle cerebral artery stroke. In: StatPearls. StatPearls Publishing.
Pennington, J., Socher, R., Manning, C., 2014. Glove: Global Vectors for Word Representation. *Proceedings Of the 2014 Conference On Empirical Methods In Natural Language Processing (EMNLP)*, pp. 1532–1543 https://doi.org/10/gfshwg.
Raji, S., de Melo, G., 2020. What sparks joy: the AffectVec emotion database, 2020 Proc. Web Conf. 2991–2997. https://doi.org/10/gngs2g.
Reilly, J., Peelle, J.E., Garcia, A., Crutch, S.J., 2016. Linking somatic and symbolic representation in semantic memory: the dynamic multilevel reactivation framework. Psychonomic Bullet. Rev. 23 (4), 1002–1014. https://doi.org/10.3758/s13423-015-0824-5.
Richardson, J.D., Dalton, S.G., 2016. Main concepts for three different discourse tasks in a large non-clinical sample. Aphasiology 30 (1), 45–73 https://doi.org/10/ggqdj7.
Savic, O., Savic, A.M., Kovic, V., 2017. Comparing the temporal dynamics of thematic and taxonomic processing using event-related potentials. PLoS One 12 (12), e0189362 https://doi.org/10/gcn7q8.
Schwartz, M.F., Kimberg, D.Y., Walker, G.M., Brecher, A., Faseyitan, O.K., Dell, G.S., Mirman, D., Coslett, H.B., 2011. Neuroanatomical dissociation for taxonomic and thematic knowledge in the human brain. Proc. Natl. Acad. Sci. Unit. States Am. 108 (20), 8520–8524.

Seckin, M., Mesulam, M.-M., Voss, J.L., Huang, W., Rogalski, E.J., Hurley, R.S., 2016. Am I looking at a cat or a dog? Gaze in the semantic variant of primary progressive aphasia is subject to excessive taxonomic capture. J. Neurolinguistics 37, 68–81 https://doi.org/10/ghh3xv.

Siew, C.S.Q., Wulff, D.U., Beckage, N.M., Kenett, Y.N., 2019. Cognitive network science: a review of research on cognition through the lens of network representations, processes, and dynamics. Complexity 2019, 2108423. https://doi.org/10.1155/2019/2108423.

Stark, B.C., 2019. A comparison of three discourse elicitation methods in aphasia and age-matched adults: implications for language assessment and outcome. Am. J. Speech Lang. Pathol 28 (3), 1067–1083. https://doi.org/10.1044/2019_AJSLP-18-0265.

Straka, M., Straková, J., 2017. Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe. Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies. Association of Computational Linguistics, Vancouver, Canada, pp. 88–99.

Thompson, C.K., 2012. Northwestern Assessment of Verbs and Sentences (NAVS). Evanston, IL: 2011.

Thompson, H., Davey, J., Hoffman, P., Hallam, G., Kosinski, R., Howkins, S., Wooffindin, E., Gabbitas, R., Jefferies, E., 2017. Semantic control deficits impair understanding of thematic relationships more than object identity. Neuropsychologia 104, 113–125 https://doi.org/10/gch7nd.

Tulving, E., 1972. Episodic and semantic memory. In: Organization of Memory. Academic Press (pp. xiii, 423–xiii, 423.

Ulatowska, H.K., North, A.J., Macaluso-Haynes, S., 1981. Production of narrative and procedural discourse in aphasia. Brain Lang. 13 (2), 345–371 https://doi.org/10/dhdbxd.

van Scherpenberg, C., Fieder, N., Savage, S., Nickels, L., 2019. The relationship between response consistency in picture naming and storage impairment in people with semantic variant primary progressive aphasia. Neuropsychology 33 (1), 13–34.

Vivas, L., García García, R., Perea Bartolomé, M., Leite D'almeida, A., Ladera Fernández, V., 2015. Recognition of thematic and taxonomic conceptual relations in patients with aphasia. Aphasiology 30 (6), 657–677.

Wang, X., Xu, Y., Wang, Y., Zeng, Y., Zhang, J., Ling, Z., Bi, Y., 2018. Representational similarity analysis reveals task-dependent semantic influence of the visual word form area. Sci. Rep. 8 (1), 3047 https://doi.org/10/gc4gz4.

Warrington, E., McCarthy, R., 1983. Category specific access dysphasia. Brain 106, 859–878.

Warrington, E., McCarthy, R., 1987. Categories of knowledge: further fractionations and an attempted integration. Brain 110, 1273–1296.

Warrington, E.K., Shallice, T., 1979. Semantic access dyslexia. Brain 102 (1), 43–63.

Warrington, E.K., Shallice, T., 1984. Category specific semantic impairments. Brain 107, 829–953.

Wernicke, C., 1874. Der aphasische Symptomencomplex; eine psychologische Studie auf anatomischer Basis (Cohn).

Zeldes, A., 2017. The GUM corpus: creating multilayer resources in the classroom. Comput. Humanit. 51 (3), 581–612.

Zhang, M., Varga, D., Wang, X., Krieger-Redwood, K., Gouws, A., Smallwood, J., Jefferies, E., 2021. Knowing what you need to know in advance: the neural processes underpinning flexible semantic retrieval of thematic and taxonomic relations. Neuroimage 224, 117405 https://doi.org/10/gngssw.