Spring 2026

# Main Concept Analysis as a Predictor of Aphasia Severity: Semantic, Phonological, and Executive Influences on Main Concept Performance

Chathurvedi Ganguru
*University of South Carolina - Columbia*

**Director of Thesis:** William Matchin
**Second Reader:** Sigfus Kristinsson

**Acknowledgements**

I am deeply grateful for the support and mentorship I received throughout the completion of this honors thesis. I would first like to thank my director, William Matchin, for welcoming me into his lab, for his constant guidance and encouragement, and for helping me gain full access to the data that made this project possible. I am also sincerely thankful to my second reader, Sigfus Kristinsson, for his thoughtful support in shaping my research question, organizing my thesis, and for his extensive knowledge and involvement in the POLAR study, especially following Dr. Dirk den Ouden's transition to Chapman University. I am especially appreciative of Dr. Dirk den Ouden for introducing me to the main concept task and for providing the training necessary to complete the discourse analysis, which was foundational to this work. I would also like to thank Zayd Kidwai for collaborating with me on the MCA analysis for inter-rater variability and for his time and insight throughout that process. Additionally, I am grateful to the members of the NeuroSyntax Lab for their feedback and support. Finally, I want to extend my heartfelt appreciation to the staff in the Center for Aphasia Study and Recovery and the Neurolinguistics Lab for welcoming me as a volunteer during my sophomore year and, along with Dr. Matchin, helping me gain full access to the data and resources necessary to complete this thesis. Their support has been invaluable to both this project and my development as a researcher.

**Table of Contents**

**Abstract**

**Background:** Chronic post-stroke aphasia (PSA) is a language disorder resulting from left-hemisphere stroke that leads to persistent impairments in language production and comprehension. These deficits often extend beyond single-word processing to discourse-level communication, limiting an individual's ability to convey meaningful information in everyday contexts. While many standardized assessments focus on word- and sentence-level abilities, they do not fully capture functional communication. Main Concept Analysis (MCA) is a discourse-based measure that evaluates the accuracy and completeness of essential information conveyed during narrative speech and has emerged as a clinically feasible and ecologically valid tool for assessing discourse production. Although MCA has demonstrated sensitivity to language impairment, the extent to which it reflects overall aphasia severity and the underlying linguistic and cognitive processes that support performance remains unclear.

**Aims:** This study aims to (1) determine whether Main Concept composite scores derived from the Cinderella narrative task are associated with aphasia severity and (2) identify the relative contributions of semantic, phonological, and executive functioning to Main Concept performance in individuals with chronic PSA.

**Methods:** Approximately 100 individuals with chronic post-stroke aphasia were selected from the baseline cohort of the POLAR (Predicting Outcomes of Language Rehabilitation in Aphasia) study. Discourse samples were collected using the Cinderella narrative task and analyzed using Main Concept Analysis to derive composite scores. Baseline measures included the Western Aphasia Battery Aphasia Quotient (WAB-AQ) to assess aphasia severity, the Pyramids and Palm Trees Test to evaluate semantic processing, the Philadelphia Repetition Test to assess phonological ability, and the Wechsler Adult Intelligence Scale (WAIS) Matrix Reasoning subtest as a measure of executive functioning. Correlational analyses were conducted to examine relationships between Main Concept performance and aphasia severity, and multiple regression analyses were used to determine the relative contributions of semantic, phonological, and cognitive factors to Main Concept performance.

**Results:** Main Concept scores were strongly correlated with aphasia severity (WAB-AQ; $r = .88$, $p < .001$) and accounted for 78% of its variance ($R^2 = .78$). Descriptive statistics showed wide variability in WAB-AQ and MCA, while semantic measures were tightly clustered, suggesting possible ceiling effects. In regression analyses, phonological processing was the only significant predictor of Main Concept performance ($\beta = .595$, $p < .001$).

**Conclusion:** Main Concept Analysis closely reflects aphasia severity and provides a practical, discourse-based measure of functional communication. Phonological processing appears to play a central role in supporting discourse performance. However, limited variability in semantic and cognitive measures and a predominantly mild-to-moderate sample may have constrained findings. Future work should include a broader range of severity levels and more sensitive measures to better capture the contributions of multiple cognitive-linguistic processes

## Introduction

### *Background on Aphasia*

Stroke is the leading cause of long-term disability in Western societies and affects approximately 800,000 individuals in the United States each year (Adamson et al., 2004; Virani et al., 2020). A common consequence of stroke is aphasia, an acquired language disorder resulting from damage to left-hemisphere brain regions that support language processing (Hillis, 2007). These regions, particularly within the frontal and temporal lobes, are critical for functions such as lexical retrieval, syntactic processing, and speech production. Disruption to these networks due to reduced blood flow can lead to widespread impairments in language, often accompanied by damage to underlying white matter pathways (Marebwa et al., 2017). While many individuals experience some degree of spontaneous recovery in the months following stroke (Lazar et al., 2010; Pedersen et al., 1995), approximately one-third of stroke survivors develop chronic post-stroke aphasia (PSA), defined as persistent language impairment beyond six months post-onset (Flowers et al., 2016).

The impact of aphasia extends far beyond language impairment alone. Individuals with aphasia often experience significant challenges in everyday communication, which can negatively affect social participation, employment, and overall quality of life (Graham et al., 2011; Kauhanen et al., 2000). Aphasia has also been associated with poorer health outcomes and has been reported to have a greater impact on quality of life than many other neurological conditions (Lazar & Boehme, 2017; Lam & Wodchis, 2010). With increasing stroke incidence among younger populations and improved life expectancy, the prevalence of chronic aphasia continues to rise, further emphasizing the need for effective, evidence-based approaches to assessment and rehabilitation (Joundi et al., 2021; Li et al., 2020; Wafa et al., 2020; Simmons-Mackie, 2018).

Aphasia presents a range of impairments affecting fluency, auditory comprehension, repetition, and naming, all of which contribute to an individual's overall communicative ability. These domains are commonly assessed using standardized tools such as the Western Aphasia Battery–Revised (WAB-R), which provides a comprehensive evaluation of language function and yields an Aphasia Quotient (AQ) as a measure of overall severity (Kertesz, 2007). The WAB-R also classifies individuals into aphasia syndromes based on patterns of impairment, including Broca's, Wernicke's, conduction, anomic, global, and transcortical subtypes (Barford, 2013). These syndromes are often broadly categorized into fluent and non-fluent types, with fluent aphasia characterized by relatively preserved speech rate but reduced content accuracy, and non-fluent aphasia marked by effortful, reduced output and simplified grammatical structure (Le & Lui, 2023). Although aphasia has traditionally been viewed as relatively stable in the chronic phase, emerging evidence suggests that language recovery can continue well beyond the initial period of spontaneous recovery. For example, Johnson et al. (2019) demonstrated that over half of individuals with chronic left-hemisphere stroke showed improvements in language comprehension and production years after onset, as measured by the WAB-R.

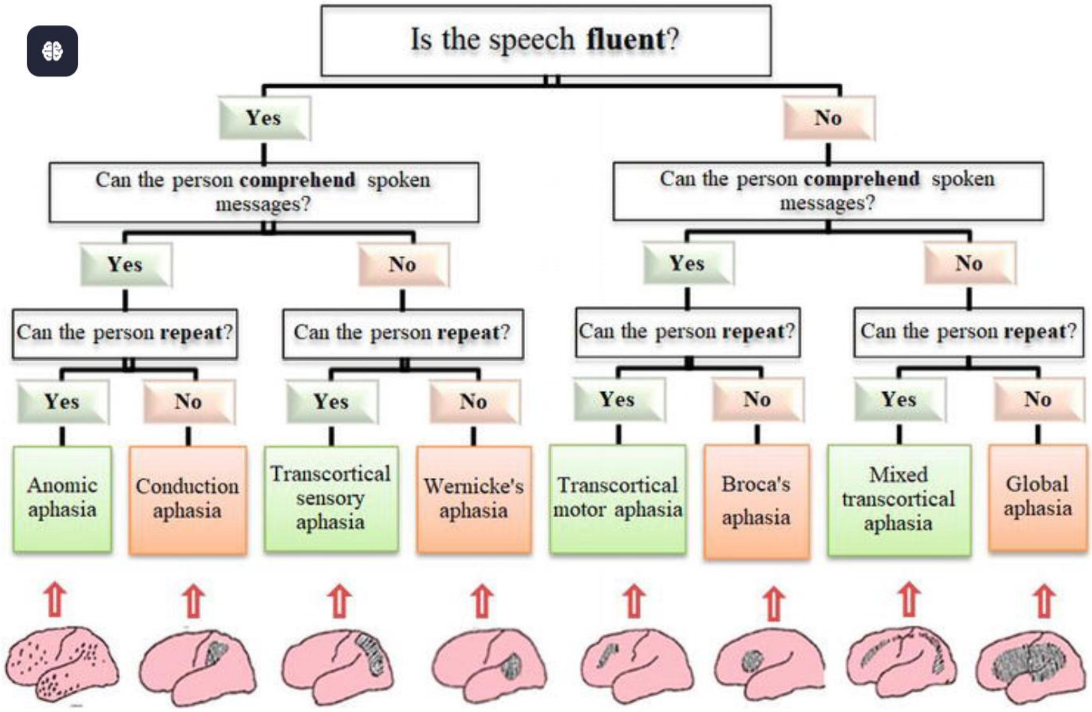


*Figure 1: Classification of aphasia subtypes based on fluency, comprehension, and repetition, distinguishing fluent and non-fluent aphasia types.*

Aphasia syndromes are commonly categorized into fluent and non-fluent types based on patterns of speech production, comprehension, and repetition (see Figure 1). While these classifications are useful for describing language impairment, they do not fully capture the complexity of functional communication, particularly at the level of connected speech. Many individuals with aphasia experience difficulties that extend beyond single-word or sentence-level production, affecting their ability to convey meaningful and coherent information in discourse. As a result, there is a growing need for assessment approaches that capture these higher-level communication abilities.

***Discourse Impairment in Aphasia***

Although aphasia is commonly assessed using single word and sentence level tasks, many of the most meaningful communication deficits emerge at the level of discourse. Discourse production requires the integration of multiple linguistic processes such as lexical retrieval, semantic access, and syntactic organization, along with higher level cognitive functions including attention, working memory, and planning. Because of these combined demands, individuals with aphasia often produce narratives that are less informative, less coherent, and incomplete, limiting their ability to convey meaningful messages in everyday communication. These impairments reflect the distributed nature of the neural systems that support language. Left hemisphere stroke can disrupt not only core language regions in the frontal and temporal lobes, but also the white matter pathways that support communication between these regions. In addition to impairing word retrieval and phonological encoding, this damage can affect cognitive systems responsible for organizing and maintaining a structured narrative. As a result, discourse production is particularly sensitive to both linguistic and executive deficits, often revealing impairments that are not captured by more constrained tasks.

Importantly, performance on traditional aphasia assessments does not always generalize to real world communication. Individuals may perform relatively well on structured naming or repetition tasks while still struggling to produce clear and complete discourse. This disconnect highlights a key limitation of standard assessment approaches and underscores the need for measures that capture functional communication. Discourse based analyses provide a more ecologically valid perspective on language ability by evaluating how effectively individuals convey information in connected speech, making them critical for both assessment and treatment planning (Nicholas & Brookshire, 1995).

***Main Concept Analysis (MCA)***

Main Concept Analysis (MCA) is a structured, discourse-level assessment that evaluates the extent to which individuals convey essential information during narrative speech. Rather than focusing on isolated word or sentence production, MCA captures the accuracy and completeness of key story elements, providing a more functional measure of communicative effectiveness. This approach is particularly valuable in aphasia research, as it reflects an individual's ability to produce meaningful, goal-directed language in contexts that more closely resemble everyday communication. MCA has been widely supported as an informative and reliable method for assessing discourse adequacy, demonstrating sensitivity to differences in information content as well as strong inter-rater reliability (Nicholas & Brookshire, 1995; Kong, 2009).

MCA is commonly applied to standardized narrative tasks, such as the Cinderella story, in which a predefined set of "main concepts" represents the critical informational units required for an adequate retelling. These concept lists are typically derived from control narratives and used as a reference for scoring performance across individuals. During analysis, transcripts are systematically coded by identifying utterances that correspond to each target concept and evaluating whether each component is produced accurately, inaccurately, or is absent. These scores are then combined to generate a composite measure of Main Concept performance, reflecting overall discourse informativeness. In the present study, MCA was conducted using a standardized scoring protocol and analysis tool developed for the POLAR dataset, ensuring consistency across participants and raters.

MCA performance has been shown to relate to functional communication outcomes. Increases in the number of main concepts produced have been significantly associated with listener-rated improvements in communicative effectiveness (Ross & Wertz, 1999), suggesting that MCA captures meaningful aspects of real-world communication. Additionally, MCA has been successfully applied across multiple aphasia subtypes, including anomic, Broca's, conduction, and Wernicke's aphasia, further supporting its utility as a broadly applicable assessment tool.

Recent research has also highlighted the role of cognitive and linguistic processes in supporting Main Concept performance. In a large-scale analysis of Cinderella narratives from Aphasia Bank, Richardson et al. (2021) found that reduced Main Concept production and narrative coherence were strongly associated with impairments in processing speed, executive functioning, and semantic retrieval. These findings suggest that successful discourse production in aphasia depends not only on lexical access, but also on broader cognitive mechanisms that support the organization and integration of information. Similarly, Lee et al. (2009) demonstrated that individuals with non-fluent aphasia showed greater improvements in conversational script performance when treatment intensity was adjusted to account for executive and attentional limitations. Despite its clinical promise, the extent to which MCA performance reflects overall aphasia severity and the specific contributions of semantic, phonological, and executive processes remains insufficiently understood. Addressing this gap is critical for determining how MCA can be used not only as an assessment tool, but also as a predictor of language impairment and a guide for individualized intervention.

***Cognitive and Linguistic Contributors to Discourse***

Discourse production in aphasia depends on the interaction of multiple linguistic and cognitive systems, including semantic access, phonological processing, and executive functioning. Disruptions in these domains can affect not only the accuracy of individual words but also the ability to produce complete, specific, and well-organized narratives. As a result, reduced discourse informativeness may reflect breakdowns at multiple levels of processing rather than a single underlying deficit. The present study aims to directly compare the contributions of semantic processing, phonological processing, and general cognitive functioning to Main Concept performance using baseline measures from the POLAR study.

Semantic processing is assessed using two complementary measures. The Pyramids and Palm Trees Test (PPTT) is a 52-item forced-choice task in which participants match a target item to one of two semantically related options, providing a measure of amodal semantic knowledge and conceptual associations. The Kissing and Dancing Test (KDT) assesses relational and event-based semantics by requiring participants to identify the appropriate relationship between agents in pictured scenes, capturing sensitivity to thematic roles and event structure. Together, these measures provide a broader assessment of semantic processing by distinguishing between object-based conceptual knowledge and relational semantic understanding, both of which support discourse production.

Phonological processing is measured using the Philadelphia Repetition Test (PRT), which consists of spoken versions of the 175 Philadelphia Naming Test (PNT) picture names. Participants are required to repeat each word aloud following auditory presentation, and responses are scored for accuracy. Performance on the PRT reflects phonological processing, including auditory-phonological input, short-term maintenance, and phonological output processes. Because all items are real words, performance may also be supported by lexical-semantic knowledge, and thus reflects the interaction of phonological and lexical systems rather than phonological processing in isolation.

General cognitive functioning is assessed using the Matrix Reasoning subtest of the Wechsler Adult Intelligence Scale (WAIS). This nonverbal task requires participants to select the missing element in a visual pattern and is widely interpreted as a measure of fluid reasoning, pattern recognition, and aspects of executive organization. All measures were administered at the baseline visit in the POLAR study, prior to any intervention, allowing for examination of how pre-treatment linguistic and cognitive abilities relate to discourse performance

***Gaps in Literature***

Although discourse analysis has been widely recognized as a valuable approach for assessing functional communication in aphasia, several important gaps remain. Measures such as Main Concept Analysis have demonstrated strong reliability and sensitivity to differences between individuals with and without aphasia, particularly in capturing the accuracy and completeness of essential information in connected speech (Nicholas & Brookshire, 1995) . However, it remains unclear to what extent Main Concept performance reflects overall aphasia severity as measured by standardized assessments such as the Western Aphasia Battery Aphasia Quotient. Establishing this relationship is critical for determining whether discourse based measures can serve as efficient and clinically meaningful indicators of language impairment.

In addition, although discourse production is known to rely on multiple interacting linguistic and cognitive systems, there is limited research directly examining which of these processes most strongly influence functional communication. Prior work has linked discourse performance to impairments in semantic processing, phonological processing, and executive functioning, but these domains are often examined in isolation rather than within a unified framework (Richardson et al., 2021) . Similarly, large scale clinical studies such as the POLAR trial highlight that language outcomes are influenced by multiple interacting factors, yet do not isolate the relative contributions of specific processes to discourse level performance.

Importantly, commonly used measures of semantic processing, phonological processing, and general cognition each capture specific components of language function, but do not independently reflect an individual's ability to communicate effectively in real world contexts. Discourse production inherently requires the integration of these systems, combining lexical, phonological, and cognitive processes to support the accurate and complete expression of meaningful information. As a result, there is a need for research that not only examines these components individually but also evaluates how they collectively contribute to discourse level communication. Existing literature has focused on word level or impairment-based outcomes, particularly in the context of treatment studies, rather than linking these underlying processes to everyday communication. Because discourse more closely reflects real world language use, identifying the mechanisms that support discourse performance is essential for improving both assessment and intervention.

### ***Goals and Significance***

The purpose of this study is to evaluate Main Concept Analysis as a functional measure of discourse production in individuals with chronic post stroke aphasia and to examine the underlying linguistic and cognitive processes that support discourse performance. Although traditional aphasia assessments provide valuable information regarding specific language domains, they do not fully capture an individual's ability to communicate effectively in connected speech. As a result, there is a need to determine whether discourse based measures more accurately reflect functional communication and to identify the mechanisms that contribute to successful discourse production. Specifically, this study addresses the following research questions:

Aim 1: Relationship Between Discourse and Aphasia Severity

- To determine whether Main Concept composite scores are associated with overall aphasia severity as measured by the Western Aphasia Battery Aphasia Quotient
- To evaluate whether discourse performance reflects the global severity of language impairment

Aim 2: Cognitive and Linguistic Predictors of Discourse Performance

- To examine the relative contributions of:
    - Semantic processing (Pyramids and Palm Trees Test, Kissing and Dancing Test)
    - Phonological processing (Philadelphia Repetition Test)
    - General cognitive functioning (WAIS Matrix Reasoning)
- To identify which of these processes most strongly predicts Main Concept performance

This study aims to advance the understanding of functional communication in aphasia by conceptualizing discourse as the product of multiple interacting linguistic and cognitive systems rather than as a reflection of isolated impairments. By directly comparing the contributions of semantic processing, phonological processing, and general cognitive functioning, this work addresses a critical gap in the literature regarding which underlying processes most strongly support the accurate and complete expression of information in connected speech. Clarifying these relationships has important implications for intervention, as it may inform more targeted and efficient treatment approaches that prioritize the processes most relevant to functional communication.

## Materials and Methods

### *Description of the POLAR Study's Participant/Criteria*

Participants for this study were drawn from the baseline cohort of the POLAR (Predicting Outcomes of Language Rehabilitation in Aphasia) study, a large-scale clinical trial investigating predictors of language recovery in individuals with chronic post stroke aphasia. Study procedures were conducted at the University of South Carolina and the Medical University of South Carolina, and all participants provided informed consent in accordance with institutional review board guidelines.

A total of 107 participants were initially enrolled in the POLAR study, of whom 102 completed baseline assessments and were included in primary analyses. To be eligible for participation, individuals were required to be between 21 and 80 years of age and have chronic aphasia resulting from a left hemisphere ischemic or hemorrhagic stroke, defined as at least 12 months post onset. Participants with multiple strokes were included only if all lesions were confined to the left hemisphere. Individuals were excluded if they were presented with severely limited speech output or auditory comprehension that would prevent completion of language tasks, as determined by the Western Aphasia Battery, or if they had contraindications for neuroimaging or evidence of bilateral stroke. Baseline data for discourse and relevant assessment measures were collected from the participants. The sample represented a range of aphasia severities and subtypes, including anomic, Broca's, conduction, Wernicke's, and other less frequent classifications. Aphasia severity ranged widely (WAB AQ approximately 14.5 to 93), capturing individuals with mild, moderate, and severe impairment. Demographically, participants had a mean age of approximately 60.5 years (SD = 10.9, range 29 to 80) and an average of 15.4 years of education (SD = 2.3) . Time post stroke also varied substantially, with a mean of 46.2 months (SD = 48.1, range 12 to 241). The sample included both male and female participants and represented a range of clinical profiles and recovery trajectories.

| **Variable** | **Mean (SD)** | **Range / N (%)** |
|---|---|---|
| **Age (years)** | 60.5 (10.9) | 29 – 80 |
| **Education (years)** | 15.4 (2.3) | 12 – 20 |
| **Time Post Stroke (months)** | 46.2 (48.1) | 12 – 241 |
| **WAB Aphasia Quotient** | 60.0 (22.7) | 14.5 – 93.1 |
| **Sex** | — | Female: 42 (41.2%); Male: 60 (58.8%) |
| **Handedness** | — | Right-handed: 90 (88.2%) |
| **Aphasia Type** | — | Broca's: 45 (44.1%);<br>Anomic: 29 (28.4%); Conduction: 16 (15.7%); Wernicke's: 7 (6.9%);<br>Global: 4 (3.9%);<br>Transcortical Motor: 1 (1.0%) |

***Table 1. Participant Demographic and Clinical Characteristics***
*Table 1 presents the demographic and clinical characteristics of participants drawn from the POLAR study sample, including age, education, time post stroke, aphasia severity, and distribution of aphasia subtypes.*

***Discourse Task: Cinderella Narrative***

Discourse samples were obtained using the Cinderella narrative task, a standardized story retelling paradigm commonly used in aphasia research. In this task, participants are presented with a wordless picture book depicting the story of Cinderella and are instructed to describe the story in as much detail as possible. The use of a picture-supported narrative minimizes memory demands while encouraging the production of extended, connected speech.

The Cinderella narrative task was selected due to its widespread use in aphasia research and its ability to elicit comparable discourse samples across individuals. Because all participants describe the same story, variability in content is reduced, allowing for systematic evaluation of discourse performance. At the same time, the task provides sufficient flexibility for individuals to generate spontaneous language, making it well suited for assessing functional communication. Narrative samples collected as part of the POLAR baseline assessment were audio recorded and subsequently transcribed by trained speech language pathologists. Transcriptions included detailed representation of participant responses, including phonological and speech production errors where relevant, to ensure accurate capture of language output.

Main concepts for the Cinderella narrative were determined based on established procedures in the discourse analysis literature. Following the approach outlined by Nicholas and Brookshire (1995), main concepts represent the essential informational units that convey the core meaning or “gist” of a narrative . These concepts are typically derived from narratives produced by neurologically healthy speakers and reflect the most consistently expressed events across individuals. More recent work has expanded this approach using large discourse databases such as AphasiaBank, which provide standardized transcripts and normative data for commonly used narrative tasks (e.g., Richardson et al., 2021) . Based on these resources, predefined main concept checklists have been developed for tasks such as the Cinderella story to ensure consistency and reliability in scoring.


***Figure 2. Main Concepts in the Cinderella Narrative Task***
*This figure presents a subset of Main Concepts from the Cinderella story, adapted from the Aphasia Bank Main Concept framework that includes 34 main concepts.The selected concepts reflect the types of information captured in Main Concept analysis and are presented in chronological order to demonstrate the overall narrative structure.*

### *Main Concept Analysis Scoring Procedure*

Discourse samples were analyzed using Main Concept Analysis, a structured and widely used method for evaluating the informativeness of connected speech in individuals with aphasia. Main concepts represent the essential informational units required to convey the core meaning or “gist” of a narrative (Nicholas & Brookshire, 1995) . This approach has been further developed and standardized through subsequent work using large discourse datasets, including AphasiaBank, and refined scoring systems such as those described by Richardson and Dalton (2021) .Training in Main Concept Analysis for this study was completed using materials and procedures provided by Dr. Dirk den Ouden as part of the POLAR study protocol . This training included review of foundational literature, exposure to standardized scoring guidelines, and completion of practice coding tasks using example transcripts (e.g., Broken Window and Cat Rescue tasks) prior to scoring Cinderella narratives. This process ensured familiarity with both the conceptual framework and application of scoring criteria.

For the Cinderella narrative, a predefined checklist of 34 main concepts was used, representing the key events, characters, and outcomes necessary for an adequate retelling of the story. Each participant’s transcript was analyzed using Rob Cavanaugh’s Main Concept tool in AphasiaBank, an online platform designed to standardize discourse scoring. Within this tool, the full transcript was loaded and segmented into individual utterances. The coder then reviewed the transcript and identified utterances corresponding to each target main concept, noting that concepts may appear in any order within the narrative.

Each main concept consists of multiple components, typically including an agent, an action, and an object or outcome. For each component, the coder determined whether it was produced as accurate, inaccurate, or absent. These component level judgments were then used to assign an overall classification to each concept. Concepts were categorized as accurate and complete (AC) when all required elements were present and correct, accurate but incomplete (AI) when the concept was correct but missing one or more elements, inaccurate but complete (IC) when all elements were present but contained incorrect information, or inaccurate and incomplete (II) when the concept was both incorrect and missing information. Concepts that were not expressed were coded as absent. To quantify discourse performance, a composite Main Concept score was calculated using a weighted scoring system in which accurate and complete concepts received the highest value. Specifically, the composite score was calculated using the formula:

$$\textit{MCA Composite Score} = (3 \times AC) + (2 \times AI) + (1 \times IC) + (0 \times II)$$

Absent concepts also did not impact the Composite Scores. Higher composite scores reflect greater discourse informativeness, capturing both the accuracy and completeness of information conveyed. Scoring summary reports and data spreadsheets were downloaded to retrieve information later if scores were not recorded correctly.

An example of scoring is provided for illustration. For the target concept “Cinderella is forced to do the housework,” a participant response such as “Cinderella had to clean the house for her stepmother” would be scored as accurate and complete (AC), as all essential components are present and correct. A response such as “Cinderella had to clean” would be classified as accurate but incomplete (AI), as the action is correct but lacks full contextual detail. A response such as “The stepmother cleaned the house” would be scored as inaccurate but complete (IC), as all elements are present but incorrectly assigned. A vague response such as “She did stuff” would be categorized as inaccurate and incomplete (II), as it lacks both accuracy and necessary detail. If the concept was not mentioned at all, it would be coded as absent.

All transcripts were scored using this standardized procedure. To ensure reliability and consistency in coding, two trained reviewers independently scored the discourse samples. In cases where discrepancies or subjective judgments arose, scores were reviewed and discussed collaboratively until consensus was reached. If the concept was absent then it was marked Absent.

***Challenges of MCA Scoring***

Although Main Concept Analysis provides a structured and standardized approach to evaluating discourse, several challenges are inherent in the scoring process. Determining the boundaries of utterances and mapping them to specific concepts can be difficult when narratives are disorganized, produced out of sequence, or include tangential content. In some cases, participants deviated from the main storyline, requiring the coder to distinguish between relevant and irrelevant information. Familiarity with the full set of main concepts and the overall narrative structure was therefore essential for identifying intended concepts, particularly when participants expressed ideas indirectly or out of order. Additionally, distinguishing between accurate and inaccurate productions often required subjective judgment, especially in the presence of semantic substitutions, vague language, or phonological errors that partially obscured meaning. These difficulties were further amplified in cases where participants produced minimal speech, as individuals with more severe aphasia sometimes generated limited output. Despite these challenges, adherence to standardized guidelines, training procedures, and consensus scoring supported consistent and reliable analysis across participants.

**Results/Data Analysis**

Following completion of Main Concept Analysis, a formal data request was submitted to the Center for the Study of Aphasia and Recovery (C-STAR) laboratory to obtain participant-level demographic and cognitive-linguistic data. These data included demographic variables as well as performance on the Kissing and Dancing Test, Pyramids and Palm Trees Test, Philadelphia Repetition Test, and Wechsler Adult Intelligence Scale Matrix Reasoning subtest. These variables were combined with Main Concept composite scores to create a dataset for subsequent statistical analyses.

Descriptive statistics for all measures, including Western Aphasia Battery Aphasia Quotient scores, were calculated using Microsoft Excel to summarize central tendencies and variability within the dataset. Means, standard deviations, and ranges were examined to characterize the distribution of participant performance across variables. All inferential statistical analyses were conducted using JASP, a free and open-source statistical software (https://jasp-stats.org/). For Aim 1, correlational analysis was performed to examine the relationship between Main Concept composite scores and aphasia severity as measured by the Western Aphasia Battery Aphasia Quotient. Pearson's correlation coefficient (r) was calculated to assess the strength and direction of the relationship, and the coefficient of determination ($R^2$) was used to quantify the proportion of variance in aphasia severity explained by Main Concept performance. For Aim 2, multiple regression analysis was conducted to evaluate the relative contributions of semantic processing, phonological processing, and general cognitive functioning to Main Concept performance. Main Concept composite score served as the dependent variable, while performance on the Pyramids and Palm Trees Test and Kissing and Dancing Test (semantic processing), Philadelphia Repetition Test (phonological processing), and WAIS Matrix Reasoning (general cognition) were included as predictor/covariate variables. Standardized regression coefficients (β) were used to assess the relative strength of each predictor, and overall model fit was evaluated using $R^2$. Statistical significance of individual predictors and the overall model was assessed using p-values. Confidence intervals were set at 95%, corresponding to a significance level of $\alpha = .05$. In addition, a Pearson correlation heatmap was generated to descriptively examine relationships among predictor variables, Main Concept performance, and aphasia severity (WAB_AQ). This analysis was included to provide context for the regression by illustrating the pattern of associations among variables, including how predictors relate to both MCA and WAB_AQ.

***Descriptive Statistics***

*Descriptives*

| | N | Mean | SD | SE |
|---|---|---|---|---|
| MCA Composite | 101 | 36.10 | 18.11 | 1.802 |
| KDT | 101 | 48.18 | 3.156 | 0.314 |
| PPT | 101 | 49.52 | 2.248 | 0.224 |
| WAIS | 101 | 8.911 | 1.738 | 0.173 |
| PRT Scores | 101 | 109.2 | 54.73 | 5.446 |

***Figure 3. Descriptive Statistics for Primary Measures***
*The distribution of scores for Main Concept composite performance and cognitive and linguistic measures, including the Philadelphia Repetition Test, Pyramids and Palm Trees Test, Kissing and Dancing Test, and WAIS Matrix Reasoning.*

Descriptive statistics revealed variability across measures, with notable differences in score distributions relative to each test's maximum possible score. Main Concept composite scores (M = 36.10, SD = 18.11) showed a wide range, indicating substantial variability in discourse performance across participants. Aphasia severity, as measured by the WAB Aphasia Quotient, also demonstrated a broad distribution, reflecting the inclusion of individuals with varying levels of impairment. Semantic measures, including the Pyramids and Palm Trees Test (maximum score = 52) and the Kissing and Dancing Test (maximum score = 52), showed relatively high mean scores (PPT: M = 49.52, SD = 2.25; KDT: M = 48.18, SD = 3.16), suggesting that many participants performed near the upper limit of these tasks. This pattern indicates reduced variability and potential ceiling effects. In contrast, phonological processing, as measured by the Philadelphia Repetition Test (maximum score = 175), demonstrated substantially greater variability (M = 109.20, SD = 54.73), indicating a wider spread of performance across participants. General cognitive functioning, as measured by WAIS Matrix Reasoning (maximum scaled score = 19), showed relatively limited variability (M = 8.91, SD = 1.74).

***Aim 1: Relationship Between Main Concept Performance and Aphasia Severity***

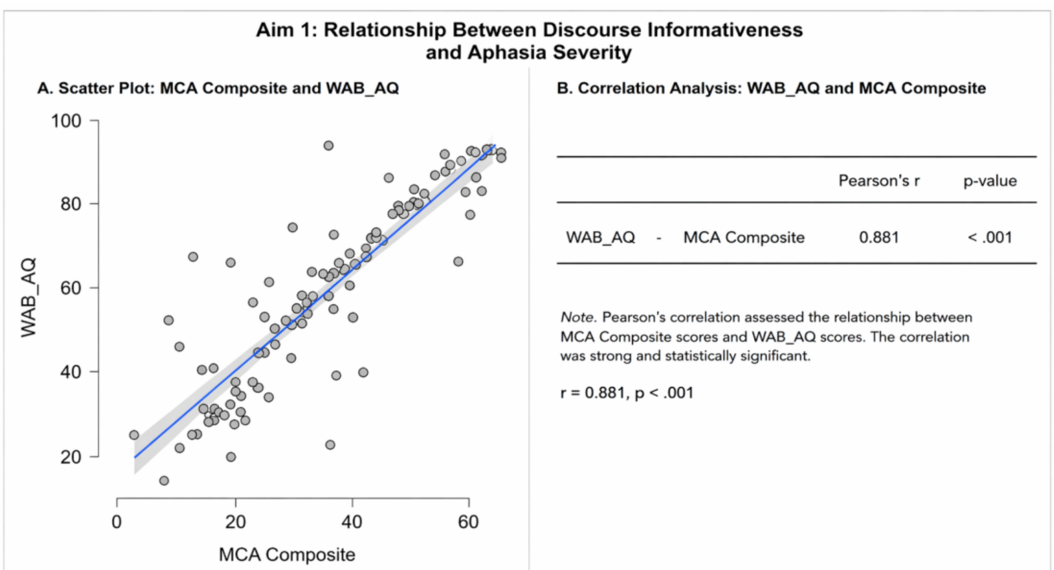


***Figure 4. Relationship Between Discourse Informativeness and Aphasia Severity***
*Panel A displays a scatter plot with a fitted regression line illustrating the association between MCA and WAB_AQ. Panel B summarizes the results of the Pearson correlation analysis, including the correlation coefficient and corresponding p-value.*

Pearson's correlation analysis revealed a strong and statistically significant positive relationship between Main Concept (MCA) composite scores and aphasia severity as measured by WAB_AQ, r = 0.881, p < .001. The scatterplot demonstrates a clear positive linear trend, with higher MCA scores associated with higher WAB_AQ scores. However, variability is present, with several data points deviating from the overall trend. These observations indicate that individuals with similar aphasia severity may differ in discourse performance, and vice versa.

***Aim 2: Cognitive and Linguistic Predictors of Main Concept Performance***

*Pearson's Correlations*

| | | | Pearson's r | p |
|---|---|---|---|---|
| MCA Composite | - | WAB_AQ | 0.881 | < .001 |
| MCA Composite | - | WAIS | 0.163 | .102 |
| MCA Composite | - | KDT | 0.144 | .149 |
| MCA Composite | - | PPT | 0.308 | .001 |
| MCA Composite | - | PRT Scores | 0.611 | < .001 |
| WAB_AQ | - | WAIS | 0.133 | .184 |
| WAB_AQ | - | KDT | 0.170 | .090 |
| WAB_AQ | - | PPT | 0.373 | < .001 |
| WAB_AQ | - | PRT Scores | 0.759 | < .001 |
| WAIS | - | KDT | 0.227 | .022 |
| WAIS | - | PPT | 0.445 | < .001 |
| WAIS | - | PRT Scores | 0.014 | .886 |
| KDT | - | PPT | 0.607 | < .001 |
| KDT | - | PRT Scores | 0.145 | .147 |
| PPT | - | PRT Scores | 0.328 | < .001 |

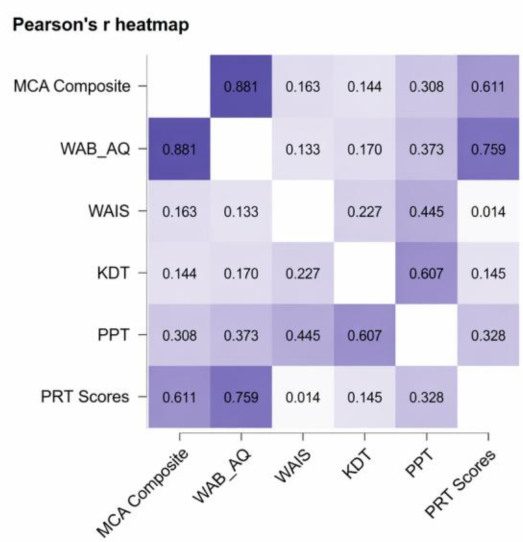


***Figure 5: Pearson Correlation Heatmap and Table of Study Variable***

*This figure displays Pearson's correlation coefficients (r) among Main Concept (MCA) composite scores, Western Aphasia Battery Aphasia Quotient (WAB_AQ), and the various predictors. The heatmap provides a visual representation of correlation strength and direction, with darker shading indicating higher correlation values.*

The Pearson correlation heatmap (Figure 5) provides a descriptive overview of relationships among Main Concept (MCA) composite scores, aphasia severity (WAB_AQ), and the cognitive-linguistic predictor variables. Overall, variables that showed stronger positive associations with MCA tended to show similarly strong associations with WAB_AQ. For example, the Philadelphia Repetition Test (PRT) demonstrated relatively strong correlations with both MCA ($r$ = .611) and WAB_AQ ($r$ = .759), whereas weaker associations were observed for WAIS Matrix Reasoning and the Kissing and Dancing Test (KDT) with both measures. This pattern suggests that performance across these tasks may relate in similar ways to both discourse informativeness and overall aphasia severity, consistent with shared variance across measures. As expected, the semantic measures (Pyramids and Palm Trees Test and KDT) were moderately correlated ($r$ = .607), reflecting their shared reliance on semantic processing. While these correlations indicate that the variables are interrelated, they do not distinguish the extent to which each measure uniquely contributes to Main Concept performance. Therefore, a multiple regression analysis was conducted to examine the relative contributions of semantic processing (PPT, KDT), phonological processing (PRT), and general cognitive functioning (WAIS) to MCA composite scores.

**Linear Regression**

*Model Summary - MCA Composite*

| Model | R | R² | Adjusted R² | RMSE | R² Change | df1 | df2 | p |
|---|---|---|---|---|---|---|---|---|
| $M_0$ | 0.000 | 0.000 | 0.000 | 18.11 | 0.000 | 0 | 100 | |
| $M_1$ | 0.635 | 0.403 | 0.378 | 14.28 | 0.403 | 4 | 96 | < .001 |

*Note.* $M_1$ includes KDT, PPT, PRT Scores, WAIS

*ANOVA*

| Model | | Sum of Squares | df | Mean Square | F | p |
|---|---|---|---|---|---|---|
| $M_1$ | Regression | 13210 | 4 | 3303 | 16.20 | < .001 |
| | Residual | 19580 | 96 | 203.9 | | |
| | Total | 32790 | 100 | | | |

*Note.* $M_1$ includes KDT, PPT, PRT Scores, WAIS
*Note.* The intercept model is omitted, as no meaningful information can be shown.

*Coefficients*

| Model | | Unstandardized | Standard Error | Standardized | t | p |
|---|---|---|---|---|---|---|
| $M_0$ | (Intercept) | 36.10 | 1.802 | | 20.04 | < .001 |
| $M_1$ | (Intercept) | −17.52 | 33.87 | | −0.517 | .606 |
| | KDT | −0.022 | 0.573 | −0.004 | −0.038 | .970 |
| | PPT | 0.422 | 0.934 | 0.052 | 0.452 | .653 |
| | PRT Scores | 0.197 | 0.028 | 0.595 | 6.962 | < .001 |
| | WAIS | 1.378 | 0.933 | 0.132 | 1.477 | .143 |

***Figure 6. Multiple regression analysis illustrating the relationship between Main Concept composite scores and predictor variables***

A multiple regression analysis was conducted to examine the relative contributions of semantic processing (PPT, KDT), phonological processing (PRT), and general cognitive functioning (WAIS) to Main Concept (MCA) performance (Figure 6). The overall model was statistically significant, $F(4, 96) = 16.20$, $p < .001$, and accounted for 40.3% of the variance in MCA scores ($R^2 = .403$; adjusted $R^2 = .378$), indicating a moderate model fit. Among the predictors, performance on the Philadelphia Repetition Test (PRT) demonstrated the strongest association with MCA scores ($\beta = .595$) and was the only statistically significant predictor ($p < .001$). In contrast, the semantic measures (PPT: $\beta = .052$, $p = .653$; KDT: $\beta = -.004$, $p = .970$) and general cognitive functioning (WAIS: $\beta = .132$, $p = .143$) did not significantly contribute to the model.

## Discussion

### *Overview of Findings and Clinical Relevance of MCA*

This study examined whether Main Concept Analysis (MCA) reflects aphasia severity and evaluated how semantic, phonological, and general cognitive measures relate to discourse performance in individuals with chronic post-stroke aphasia. The findings indicate that MCA performance is closely aligned with overall aphasia severity, suggesting that the ability to produce complete and accurate main concepts corresponds with broader language functioning. In addition to this strong relationship, variability observed in discourse performance indicates that individuals with similar levels of aphasia severity may differ in how effectively they communicate information in connected speech.

MCA captures performance within a discourse context and reflects how individuals organize and express meaningful content, rather than how they perform on isolated language tasks. Standardized assessments such as the Western Aphasia Battery Aphasia Quotient (WAB-AQ) evaluate multiple language domains under controlled conditions, whereas MCA reflects how effectively individuals convey essential information during connected speech. For this reason, MCA is best interpreted as a measure of functional communication that complements, rather than replaces, standardized severity metrics.

### *Complementary Roles of MCA and WAB-AQ*

The relationship between MCA and WAB-AQ indicates that these measures capture related but distinct aspects of language performance. The WAB-AQ provides a structured, norm-referenced index of impairment across fluency, comprehension, repetition, and naming, allowing for standardized classification of aphasia severity. However, these tasks are administered in constrained formats and may not fully reflect how language is used in everyday communication.

In contrast, MCA evaluates the extent to which individuals successfully communicate key narrative information, requiring the integration of content selection, lexical retrieval, and speech production in real time. As a result, individuals with similar WAB-AQ scores may differ in their ability to convey meaningful information during discourse. MCA therefore adds clinically relevant information by capturing communicative effectiveness rather than isolated task performance.

At the same time, MCA has limitations that restrict its use as a standalone severity measure. Scoring requires judgment regarding whether a concept is sufficiently expressed, particularly when responses are incomplete or disorganized. Performance may also vary depending on communication style, verbosity, and task engagement. In contrast, the WAB-AQ minimizes these sources of variability through standardized administration and scoring but may underestimate functional communication ability. Together, these measures provide complementary perspectives by combining standardized assessment with evaluation of real-world language use.

***Clinical Feasibility and Implementation of MCA***

In addition to conceptual differences, MCA and standardized assessments differ in feasibility and implementation. The WAB-AQ requires administration by trained clinicians and involves structured testing procedures that can be time-intensive. MCA, in contrast, can be applied to naturalistic speech samples and may be learned by researchers or clinicians with appropriate training, making it a flexible tool for both clinical and research contexts.

However, MCA also requires time for transcription, scoring, and rater training, which may limit its practicality in fast-paced clinical environments. Inter-rater reliability depends on familiarity with scoring conventions and the narrative framework, and variability may arise without standardized training procedures. As a result, although MCA offers advantages in ecological validity, its implementation in routine clinical practice may require additional resources, including training time and improved scoring standardization.

***Cognitive-Linguistic Contributions to Discourse***

The findings from Aim 2 indicate that the predictor variables are interrelated and reflect overlapping aspects of overall language impairment. The correlation patterns show that variables strongly associated with MCA were also more strongly associated with WAB-AQ, suggesting that overall aphasia severity influences performance across measures.

Although performance on the Philadelphia Repetition Test (PRT) showed the strongest independent association with MCA scores when the other measures were considered, this finding should not be interpreted as evidence that phonological processing independently determines discourse performance. Rather, it suggests that tasks requiring accurate and efficient speech production are more sensitive to individual differences within this sample

Discourse production requires both the representation of meaning and the ability to express that meaning in spoken form. Semantic processing supports the selection and organization of relevant content, while phonological and articulatory processes support its production. The present findings indicate that, in individuals with relatively preserved conceptual knowledge, limitations in speech production processes may constrain the extent to which information is expressed. This distinction is critical: intact semantic knowledge does not ensure that meaningful information can be successfully conveyed in connected speech if the mechanisms required to produce language are disrupted.

General cognitive functioning, as measured by WAIS Matrix Reasoning, did not show a measurable relationship with discourse performance. This likely reflects a mismatch between the task and the cognitive demands of discourse. Matrix Reasoning assesses abstract nonverbal reasoning, whereas discourse production depends on language-specific processes such as verbal working memory, sequencing, and planning. Because these functions were not directly measured, the contribution of language-relevant cognitive control processes may not be fully captured in the present study.

Taken together, these findings suggest that discourse performance in this sample reflects the interaction of processes involved in selecting relevant content, retrieving lexical items, and producing that information in real time, all of which are shaped by overall language impairment. Within this context, the apparent strength of the PRT likely reflects both its alignment with speech production demands and its sensitivity to variability in the sample, rather than an isolated contribution of phonological processing.

***Ceiling Effects and Sample Characteristics***

The contribution of semantic measures was limited by restricted variability in the sample. Most participants presented with anomic aphasia and mild to moderate impairment, a profile associated with relatively preserved semantic representations. This was reflected in the clustering of scores on the Pyramids and Palm Trees Test and Kissing and Dancing Test, where most participants performed near ceiling. When performance is concentrated within a narrow range, statistical models cannot effectively distinguish between individuals, even when underlying differences exist.

This pattern aligns with neurocognitive models of aphasia. Semantic knowledge is primarily supported by temporal lobe systems, which are often preserved in anomic aphasia. In contrast, phonological encoding and speech production rely on posterior perisylvian and frontal networks, which are more frequently affected. As a result, variability in discourse performance within this sample is more likely to reflect differences in speech production demands than differences in conceptual knowledge. These findings are specific to the composition of the current sample. In individuals with more severe aphasia or greater temporal lobe involvement, semantic deficits would be expected to contribute more directly to discourse breakdown.

## Future Directions

Future research should include a broader range of aphasia severity and subtype representation to better capture variability in semantic, phonological, and discourse abilities. Expanding the sample to include individuals with more severe aphasia, as well as a wider distribution of fluent and nonfluent profiles, would allow for clearer evaluation of how different processes contribute to discourse performance.

Future studies should also incorporate assessment tools that more closely match the demands of connected speech. For semantic processing, tasks that require active generation of language and integration of meaning across utterances may provide a more accurate reflection of discourse ability than recognition-based tests. Examples include narrative production tasks, picture description tasks such as the Cookie Theft or Picnic Scene, verb generation tasks, and measures of event knowledge. Tasks that assess semantic control, such as the Camel and Cactus Test, may capture variability in semantic selection that is not detected by association-based measures. For phonological processing, measures that place greater demands on phonological processing with reduced support from lexical knowledge may be useful. Nonword repetition tasks and phoneme discrimination tasks can assess phonological processing with minimal reliance on stored word representations. Analysis of phonological errors in connected speech may also provide a more ecologically valid measure of how these deficits affect discourse. For general cognition, future work should include measures that assess language-relevant executive functions. Tasks that measure verbal working memory, such as digit span or sentence repetition, as well as tasks that assess planning and sequencing, may better capture the cognitive demands required to maintain coherence and structure during connected speech.

Further development of MCA is necessary to support broader clinical use. Standardized scoring procedures should be refined to reduce variability across raters, including clearer definitions of what constitutes an accurate, complete, or sufficient expression of a concept. Formal training protocols and reliability benchmarks would improve scoring consistency. In addition, the development of automated or semi-automated tools for transcription and scoring could improve efficiency and feasibility in clinical settings.

Finally, longitudinal research is needed to examine how changes in semantic access, phonological production, and language-related cognitive control relate to changes in discourse performance over time. This approach would support the development of models that link specific mechanisms to functional communication outcomes and inform more targeted intervention strategies.

## Conclusion

The present study examined the relationship between discourse performance and aphasia severity and evaluated how semantic, phonological, and cognitive processes relate to connected speech in individuals with chronic post-stroke aphasia. The findings indicate that Main Concept Analysis aligns closely with overall aphasia severity and captures how individuals organize and express essential information in connected speech, which is not fully reflected in structured assessment tasks.

At the same time, the results demonstrate that discourse performance reflects the interaction of multiple cognitive and linguistic processes. Although performance on the Philadelphia Repetition Test showed the strongest independent association with discourse performance after accounting for the other measures, this pattern likely reflects the sensitivity of production-based tasks and the characteristics of the participant sample rather than an isolated influence of phonological processing. The limited contribution of semantic and general cognitive measures is best understood in the context of restricted variability and differences in task demands.

These findings reinforce the importance of analyzing discourse to understand how language impairments affect functional communication. While standardized assessments provide essential information about language impairment, they do not fully capture how individuals express meaningful content in real-world contexts. MCA offers a structured method for evaluating this level of performance, although considerations related to scoring consistency, efficiency, and standardization must be addressed before broader clinical implementation. Overall, this work supports the integration of discourse-based measures into aphasia assessment to provide a more complete understanding of language function and to better reflect functional communication outcomes.

**References**

Dignam, J., Copland, D., O'Brien, K., Burfein, P., Khan, A., & Rodriguez, A. D. (2017). Influence of cognitive ability on therapy outcomes for anomia in adults with chronic poststroke aphasia. *Journal of Speech, Language, and Hearing Research, 60*(2), 406–421. https://doi.org/10.1044/2016_JSLHR-L-15-0384

Flowers, H. L., Skoretz, S. A., Silver, F. L., Rochon, E., Fang, J., Flamand-Roze, C., & Martino, R. (2016). Poststroke aphasia frequency, recovery, and outcomes: A systematic review and meta-analysis. *Archives of Physical Medicine and Rehabilitation, 97*(12), 2188–2201.

Forbes, M. M., Fromm, D., & MacWhinney, B. (2012). AphasiaBank: A resource for clinicians. *Seminars in Speech and Language, 33*(3), 217–222. https://doi.org/10.1055/s-0032-1320041

Gilmore, N., Meier, E. L., Johnson, J. P., & Kiran, S. (2019). Nonlinguistic cognitive factors predict treatment-induced recovery in chronic poststroke aphasia. *Archives of Physical Medicine and Rehabilitation, 100*(7), 1251–1258. https://doi.org/10.1016/j.apmr.2018.12.024

Graham, N. L., Hickey, E., & Freeman, J. (2011). Post-stroke aphasia: Long-term recovery and management. *Clinical Rehabilitation, 25*(8), 685–695.

Hillis, A. E. (2007). Aphasia: Progress in the last quarter of a century. *Neurology, 69*(2), 200–213.

Johnson, L., Basilakos, A., Yourganov, G., Cai, B., Bonilha, L., Rorden, C., & Fridriksson, J. (2019). Progression of aphasia severity in the chronic stages of stroke. *Journal of Speech, Language, and Hearing Research, 62*, 1914–1920.

Kauhanen, M. L., Korpelainen, J. T., Hiltunen, P., Määttä, R., Mononen, H., Brusin, E., & Myllylä, V. V. (2000). Aphasia, depression, and quality of life after stroke. *Stroke, 31*(10), 2426–2430.

Kertesz, A. (2007). *Western Aphasia Battery–Revised (WAB-R)*. Pearson.

Lam, J. M., & Wodchis, W. P. (2010). The impact of aphasia on quality of life. *Stroke, 41*(7), 145–151.

Lazar, R. M., & Boehme, A. K. (2017). Aphasia as a predictor of stroke outcome. *Stroke, 48*(5), 1352–1358.

Lazar, R. M., Minzer, B., Antoniello, D., Festa, J. R., Krakauer, J. W., & Marshall, R. S. (2010). Improvement in aphasia in the first year after stroke. *Stroke, 41*(7), 148–153.

Lee, J. B., Kaye, R. C., & Cherney, L. R. (2009). Conversational script performance in adults with non-fluent aphasia: Treatment intensity and aphasia severity. *Aphasiology, 23*(7–8), 885–897. https://doi.org/10.1080/02687030802669534

Le, Y., & Lui, C. (2023). Fluent and nonfluent aphasia revisited. *Neuropsychology Review, 33*(1), 112–130.

Nardo, D., Holland, R., Leff, A. P., Price, C. J., & Crinion, J. T. (2017). Less is more: Neural mechanisms underlying anomia treatment in chronic aphasic patients. *Brain, 140*(11), 3039–3054. https://doi.org/10.1093/brain/awx234

Nicholas, L. E., & Brookshire, R. H. (1995). Presence, completeness, and accuracy of main concepts in the connected speech of non-brain-damaged adults and adults with aphasia. *Journal of Speech and Hearing Research, 38*(1), 145–156. https://doi.org/10.1044/jshr.3801.145

Richardson, J. D., Dalton, S. G., Greenslade, K. J., Jacks, A., Haley, K. L., & Adams, J. (2021). Main concept, sequencing, and story grammar analyses of Cinderella narratives in a large sample of persons with aphasia. *Brain Sciences, 11*(1), 110. https://doi.org/10.3390/brainsci11010110

Ross, K. B., & Wertz, R. T. (1999). Quality of life and communication effectiveness in aphasia. *Aphasiology, 13*(2), 105–121.

Simmons-Mackie, N. (2018). Aphasia in the context of aging populations. *Seminars in Speech and Language, 39*(1), 1–14.

Kristinsson, S., Basilakos, A., den Ouden, D. B., Cassarly, C., Spell, L. A., Bonilha, L., Rorden, C., Hillis, A. E., Hickok, G., Johnson, L., Busby, N., Walker, G. M., McLain, A., & Fridriksson, J. (2023). Predicting outcomes of language rehabilitation: Prognostic factors for immediate and long-term outcomes after aphasia therapy. *Journal of Speech, Language, and Hearing Research, 66*(3), 1068–1084. https://doi.org/10.1044/2022_JSLHR-22-00347