Aphasiology

ISSN: 0268-7038 (Print) 1464-5041 (Online) Journal homepage: www.tandfonline.com/journals/paph20

Characterizing narrative discourse in individuals with latent aphasia: An analysis of microlinguistic and macrolinguistic abilities

Manaswita Dutta & Melissa Ugianskis

To cite this article: Manaswita Dutta & Melissa Ugianskis (29 May 2026): Characterizing narrative discourse in individuals with latent aphasia: An analysis of microlinguistic and macrolinguistic abilities, Aphasiology, DOI: 10.1080/02687038.2026.2680447

To link to this article: https://doi.org/10.1080/02687038.2026.2680447

APHASIOLOGY
https://doi.org/10.1080/02687038.2026.2680447

Routledge
Taylor & Francis Group

Characterizing narrative discourse in individuals with latent aphasia: An analysis of microlinguistic and macrolinguistic abilities

Manaswita Dutta[a] and Melissa Ugianskis[b]

[a]Department of Speech and Hearing Sciences, Portland State University, Portland, USA; [b]Audiology and Speech-Language Pathology, VA Portland Health Care System, Portland, USA

ABSTRACT

Background: Individuals with latent aphasia frequently experience persistent communication challenges despite scoring above the cut-off ontraditional aphasia assessments, often resulting in reduced life participation and lower self-confidence (Armstrong etal. 2013; Cavanaugh & Haley, 2020). Due to their high-level language deficits, people with latent aphasia often do not meet the criteria for aphasia services, resultingin a lack of essential treatment (Richardson et al. 2021). Effectively identifying language impairments in latent aphasia is crucial and discourse analysis has emerged as a key tool, with recent research highlighting a range of linguistic deficits in this population (Fromm et al. 2017; Stark et al. 2025). Continued research should prioritize identification of more sensitive discourse analysis tools and metrics to improve the detection and accurate characterization of language abilities in latent aphasia.

Aim: This study investigated narrative abilities of individualswith latent aphasia who score within normal limits on the Western Aphasia Battery-Revised (WAB-R; Kertesz, 2007) by comparing their performance to non-brain-injured control participants without aphasia. Specifically, we examined narrative coherence – a macrostructural feature of language – in relation to microlinguistic abilities.

Method: The study evaluated data from 118 age-, education-, andgender-matched participants from AphasiaBank: 59 with latent aphasia and 59 non-brain-injured control participants without aphasia. Cinderella story narratives of all participants were analyzed for microlinguistic features including productivity,lexical diversity, and verb production. Coherence was assessed using an adapted transcription-free version of the rubric developed by Linnik and colleagues. Group differences in discourse scores were examined using Mann-WhitneyU tests. Additionally, Spearman's rank correlations were conducted to explore relationships among narrative variables.

Results: Significant group differences were noted on all microlinguistic measures. Importantly, both participant groups showed

ARTICLE HISTORY
Received 10 March 2025
Accepted 21 May 2026

KEYWORDS
Latent aphasia; aphasia; coherence; discourse analysis

CONTACT Manaswita Dutta mdutta@pdx.edu Department of Speech and Hearing Sciences, Portland State University, 527 SW Hall St UCB 430F, Portland, OR 97201, USA

Supplemental data for this article can be accessed online at https://doi.org/10.1080/02687038.2026.2680447

marked differences across all four coherence domains, with the latent aphasia group scoring consistently lower than controls. Correlational analyses revealed few, weak associations between microlinguistic and coherence variables in both the latent aphasia and control groups.
Conclusions: Discourse analysis is a valuable tool for examining language characteristics in latent aphasia. Microlinguistic measures capture structural language deficits; however, they may have limited reliability in distinguishing the subtle impairments associated with latent aphasia. In contrast, coherencevariables more consistently and effectively capture these subtle deficits. In individuals with latent aphasia, breakdowns in coherence are less likely to be explained solely by microlinguistic features. Therefore, a comprehensive approach that integrates both micro- and macrolinguistic measures is necessary to enhance the diagnostic sensitivity of language assessments for latent aphasia.

Background

Language is an integral part of our lives. Personal, academic, and professional success is reliant upon language. As such, when there is a breakdown in the linguistic system, it can have a significant impact on a person's life participation (Wallace, 2010). This is commonly seen in aphasia, a communication disorder commonly caused by an acquired brain injury that can affect language comprehension and expression in terms of speaking, listening, writing, and reading (Murray & Clark, 2015). Moreover, the ability to engage in *discourse* - language use that extends beyond individual words and sentences to convey information for specific purposes – is crucial for everyday communication and maintaining one's societal role (Dipper et al., 2021). Limitations in discourse can have a negative cascading effect on individuals' life participation and rehabilitation outcomes and are linked to increased depression, social isolation, and a shrinking social network (Code & Herrmann, 2003; Hilari & Northcott, 2017).

Latent aphasia and associated cognitive-linguistic deficits

Latent aphasia is characterized by very mild language difficulties that can significantly impact communication and quality of life. Unlike more severe forms of aphasia, where impairments in speaking, understanding, reading, or writing are noticeable, latent aphasia presents more subtly. Individuals with latent aphasia may not always exhibit marked deficits on standardized language assessments like the Western Aphasia Battery-Revised (WAB-R) or discrete tests (e.g., object naming, verbal fluency) that are designed to evaluate basic language functions and estimate aphasia severity (Martzoukou et al., 2023). Despite scoring above cut-offs on standardized language tests, this group – often labeled "*subclinical aphasia*" (Olsen et al., 1986), or "*not aphasic by the WAB*" [NABW] (Fromm et al., 2017; Richardson et al., 2021) – still experiences persisting language and cognitive-linguistic difficulties. Documented impairments include challenges in auditory comprehension, word finding, spoken discourse, and conversation (Armstrong et al., 2013; Boller, 1968; Fox et al., 2009; Salis & De De, 2022), as well as deficits in attention and

executive functioning (Frankel et al., 2007; Laks et al., 2025; Silkes et al., 2021). These cognitive-linguistic impairments interfere with reengagement in pre-injury activities and community reintegration, often leading to frustration (Cavanaugh & Haley, 2020; Richardson et al., 2021). For example, they may encounter challenges in navigating communication demands in the workplace, participating in group conversations with friends and family, meeting the expectations of communication partners, and forming new relationships (Armstrong et al., 2013; Cruice et al., 2006; Hickin et al., 2015; Marshall, 1993; Morris et al., 2011). Commonly used aphasia assessments fail to capture these daily challenges experienced by individuals with latent aphasia, yet they are frequently used as the primary criteria for determining eligibility for speech and language services. As a consequence, these individuals may lack needed access to treatment or be prematurely discharged from treatment if they do not receive a formal diagnosis. Therefore, there is a critical need to move beyond traditional assessment approaches and implement efficient, accurate, and ecologically valid methods to effectively identify language challenges in individuals with latent aphasia.

Discourse analysis as a tool to identify language impairments in latent aphasia

Discourse assessments are increasingly employed in research and clinical settings as a way to move beyond impairment-focused evaluations and comprehensively characterize language impairments as well as assess treatment related effects in aphasia (Dipper et al., 2021; Mayer & Murray, 2003; Prins & Bastiaanse, 2004). Spoken narratives, a discourse genre involving monologue-based tasks that incorporate spatial and temporal information, are frequently utilized as clinically useful opportunities for linguistic examination in aphasia (De De & Salis, 2020; Dutta et al., 2025; Leaman & Archer, 2023; Salis & De De, 2022). Story retelling can be used to elicit narrative discourse by using wordless picture books (Kintz, 2024). The production of narrative discourse is a cognitively and linguistically demanding task as it relies on the integrity of the language and extralinguistic cognitive systems. It involves processing at the microlinguistic (i.e., phonological, lexical, semantic, morphological, and syntactic) as well as macrolinguistic (i.e., information content, cohesion, and coherence) levels. Cognitive functions such as processing speed, attention, working memory, and executive functioning are closely linked to the successful production of narrative content. Impairments in any of these domains can significantly impact the quality of the language output of individuals with aphasia (Cahana-Amitay & Jenkins, 2018; Dipper et al., 2021; Dutta et al., 2024).

Recent advancements in spoken discourse analysis have demonstrated significant effectiveness in assessing and documenting language impairments among individuals with latent aphasia. For example, Dalton and Richardson (2015) used main concepts and core lexicon analysis to assess discourse informativeness in 235 AphasiaBank participants (MacWhinney et al., 2011), including 25 with latent aphasia (i.e., classified as NABW based on WAB-AQ scores above the cutoff). Transcripts from the Broken Window picture sequence task were extracted and analyzed. They found that the NABW group performed significantly differently compared to neurotypical adults without aphasia and from participants with other types of aphasia on CoreLex (i.e., the number of core lexicon words used based on normed core lexicon lists; Dalton et al., 2019) and main concept (i.e., the number of key ideas relevant to the discourse prompt) measures. Unsurprisingly, the

discourse scores of those in the NABW group showed no significant correlations with their WAB-AQ scores. Fromm et al. (2017) compared narrative performances of individuals classified as NABW to anomic aphasia and control groups using Cinderella narrative transcripts from AphasiaBank. Findings revealed that the NABW group participants showed higher speech rate, lexical diversity, and propositional density, fewer word errors, more main concepts, and longer MLU than those with anomic aphasia, but still underperformed compared to controls.

De De and Salis (2020) compared Cinderella story retellings among individuals with latent aphasia, anomic aphasia, and no aphasia. Those with latent aphasia showed worse performance than individuals with no aphasia on word count, pause duration, speech rate, and formulation time. Their narratives were slower and less detailed, with deficits in articulation rate, and episode continuation. The findings highlight impaired planning, lexical retrieval, and processing speed, supporting the use of temporal metrics in assessing latent aphasia. Recently, Stark et al. (2025) examined discourse in individuals with latent aphasia and mild cognitive impairment (MCI) by analyzing core lexicon use across three tasks (Sandwich Procedure, Cat Rescue, and Cinderella) using TalkBank data. Both clinical groups produced fewer core lexical items than cognitively healthy controls in the Cinderella task, with the latent aphasia group showing slower production than the MCI group across all tasks. In the Sandwich task, the latent aphasia group used significantly fewer core items than both other groups. Sensitivity analyses showed the Sandwich task best distinguished latent aphasia from controls, while the Cinderella task best differentiated latent aphasia from MCI, supporting core lexicon as a sensitive metric for detecting subtle lexical-semantic deficits.

Recent research has substantially advanced our understanding of language impairments in latent aphasia, revealing distinct spoken discourse deficits compared to both non-aphasic adults and those with other aphasia types (De De & Hoover, 2021; Dutta et al., 2024). While most studies have centered on microlinguistic and propositional aspects – such as syntax, lexical access, semantics, and phonology – fewer have examined macrostructural features like main concepts and story grammar (e.g., Main Story Sequence Grammar [MSSG; Richardson et al., 2021]). Notably, to date, no studies have specifically examined narrative coherence in individuals with latent aphasia.

Discourse coherence

Coherence is a macrolinguistic discourse feature that refers to the ability to logically organize verbal content, enabling listeners to easily follow and interpret the overall message or meaning conveyed by the speaker (Kintz, 2024). This measure reflects semantic connectedness across multiple propositions and the overall discourse. Coherence is often studied at two levels – local and global coherence – both of which individuals with aphasia frequently experience impairments in (Christiansen, 1995; Halliday & Hasan, 1976; Wright et al., 2013). Local coherence refers to how well each adjacent proposition or utterance produced during discourse are connected. Global coherence focuses on the overall organization of the discourse, such as the clarity of the main topic or theme and how well the speaker maintains a coherent storyline or argument (Glosser & Deser, 1992). Most commonly, rating scales have been used to judge the overall coherence of

a speaker's discourse, assessing whether the message is clear and logically organized (Coelho & Flewellyn, 2003; Pallickal & Hema, 2020). For instance, Wright et al. (2010) developed a 4-point scale to assess global coherence in adults with and without aphasia. The scores can range from 1 through 4 wherein a score of "4" indicates clear and specific mention of elements of the stimulus (e.g., actors, actions, objects) related to the main details of the stimulus and thus requires no inferencing by the listener; whereas a score of "1" indicates the utterance is entirely unrelated to the stimulus and may be a comment on the stimulus item. Participants with aphasia and controls without aphasia completed two storytelling tasks, with individuals with aphasia scoring significantly lower in global coherence than the control group. A few other studies have evaluated coherence using error analysis to look for markers of incoherence, such as tangential speech, illogical connections between ideas, idea repetitions, or irrelevant utterances (Andreetta et al., 2012; Dutta et al., 2024; Hazamy & Obermeyer, 2020; Leaman & Edmonds, 2021). Andreetta and Marini (2015) found that individuals with fluent aphasia produced significantly more local and global coherence errors in narratives than adults without aphasia. Local coherence errors involved ambiguous or omitted referents or topic shifts, while global coherence errors included fillers, repetitions, semantically incongruent or off-topic remarks.

Individuals with latent aphasia have not been included or minimally represented in discourse coherence studies, limiting insight into their discourse abilities and highlighting the need for systematic investigation in a larger sample. Further, as noted above, traditional definitions and assessments of coherence have largely emphasized thematic relatedness and topic maintenance. However, these definitions and rating scales do not fully capture the multidimensional nature of discourse coherence and overlook other crucial linguistic factors that contribute to it. In response, Linnik et al. (2022) examined Russian speakers with and without aphasia to investigate how language impairments relate to various aspects of discourse coherence and to identify the linguistic mechanisms that support its establishment and maintenance. They operationalized coherence in terms of four key factors: *informativeness, clarity, understandability*, and *connectedness*. Informativeness pertains to the amount of relevant content within the discourse. Clarity refers to the semantic cohesion of the discourse, achieved through the organization of its components in relation to the overall purpose, goal, or topic. Understandability refers to the subjective evaluation of communicative competence, encompassing the interpretation of nonverbal behaviors, whereas connectedness captures the local semantic and pragmatic relationships between elements of the discourse output. Linnik and colleagues also developed a rating scale based on this updated definition to rate transcribed narrative productions from a video retelling task and demonstrated that the narrative retellings produced by speakers with aphasia were rated lower than those without aphasia across all aspects of coherence, providing further evidence of coherence impairments in aphasia. Thus, whereas existing metrics such as Wright and colleagues' rating scale offer a fine-grained linguistic measure of topic maintenance and a single overall rating of how well a narrative relates to the topic, Linnik's system provides a broader assessment of narrative coherence across multiple dimensions of discourse quality that can be evaluated simultaneously.

The current study

The present study aimed to investigate narrative coherence in individuals with latent aphasia by applying the rubric developed by Linnik et al. (2022) to capture the multifaceted nature of coherence impairments in this population. Specifically, we examined whether narrative retellings from individuals with latent aphasia differ in coherence across four dimensions – informativeness, clarity, understandability, and connectedness – compared to those produced by individuals without aphasia. We were particularly interested in how individuals with latent aphasia scored on the narrative coherence metric relative to microlinguistic variables compared to the control group. For the purposes of this study, we used an adapted version of the measure developed by Linnik et al. (2022). Our adaptation of the original framework excludes multimodal communicative cues, does not rely on transcribed samples, is applied to English-speaking participants, and is used with the Cinderella narrative task.

Beyond theoretical inquiry, our study aimed to address a critical gap in clinical practice: existing discourse analysis methods frequently rely on transcription and require substantial time and specialized training, limiting their feasibility in everyday clinical settings (Stark, Dutta et al., 2021). These demands create significant barriers to the widespread assessment of discourse-level impairments. Transcription-less tools offer several advantages – such as enabling clinicians to score discourse in real time from audio or video recordings and increasing the likelihood of broader clinical adoption (Stark & Dalton, 2024). Yet few have been optimized for clinical use (e.g., Casilio, Dutta et al., 2025). By testing the clinical applicability of Linnik et al.'s rubric, our goal was to evaluate its potential as a practical and efficient tool for identifying and documenting language impairments, particularly in milder aphasia, in a more accessible and clinically workable manner. Based on previous research that highlights distinct discourse patterns in latent aphasia (e.g., Fromm et al., 2017; Salis & De De, 2022), we hypothesized that this group would exhibit differences in their narrative abilities, particularly in terms of narrative coherence, given their high-level language challenges.

Method

Participants

Spoken discourse samples of 118 individuals were obtained from AphasiaBank, a shared database that provides a repository for discourse data and standardized testing results from people with and without aphasia who have been enrolled in research studies (Forbes et al., 2012; MacWhinney et al., 2011). We included a latent aphasia group and a control group of non-brain-injured individuals without aphasia, each comprising 59 participants. All participants were fluent English speakers with adequate vision and hearing abilities and matched for age, gender, and years of education.

All participants in the latent aphasia groups had sustained a stroke and completed the WAB-R. An aphasia quotient (AQ) of greater than 93.8 on the WAB-R was used to classify individuals with latent aphasia. AphasiaBank categorizes this group as "not aphasic" based on their WAB-R performance.

Clinical impressions regarding the presence or absence of apraxia of speech (AOS) and dysarthria were available for 27 of the 59 participants in the latent aphasia group in

Table 1. Demographic information for the latent aphasia and healthy controls groups.

Variable	Latent Aphasia ($n = 59$)	Controls Without Aphasia ($n = 59$)	Group Comparison Statistic
	Mean ± SD (range)	***Mean ± SD (range)***	
Age	59.33 ± 13.98 (25.00–79.80)	59.50 ± 14.10 (25.00–80.07)	$U = 1707.50$, $z = -.178$, $p = .859$
Education	16.41 ± 3.34 (10.00–24.00)	16.72 ± 2.93 (10.00–25.00)	$U = 1394.50$, $z = -.688$, $p = .491$
Gender	25 males, 34 females	26 males, 33 females	$\chi 2$ (2, 118) = 1.015, $p = .602$
WAB-R AQ*	96.77 ± 1.77 (93.80–100)	–	–

Note. WAB-R = Western Aphasia Battery-Revised; AQ = Aphasia Quotient; SD = Standard Deviation.

AphasiaBank and were independently confirmed by author MD, a certified and licensed speech-language pathologist. In addition, all participants in the aphasia group were rated independently by the author MD on the 5-point motor speech scale (0 = severe, 1 = marked, 2 = moderate, 3 = mild, and 4 = normal) from the Quick Aphasia Battery (Wilson et al., 2018) that is designed to assess connected speech features. Across the full aphasia sample, AOS was identified in five participants, dysarthria in four participants, and co-occurring AOS and dysarthria in one participant; all were rated as mild (i.e., a rating of 3). No participants were excluded on this basis, as these mild motor speech features were not clinically judged to significantly compromise speech intelligibility or task performance.

All individuals in the control group reported a negative history of stroke, head injury, or neurological condition, and no diagnosis of communication disorders. See Table 1 for summary of participants' demographic, clinical assessment information, and statistical comparisons. Individual demographic and clinical information about all study participants are provided in Supplementary Material 1.

Discourse samples

Discourse samples of the first narration of the Cinderella storytelling were downloaded from AphasiaBank. As per the elicitation procedures outlined on AphasiaBank, participants were instructed to narrate a story after viewing a wordless picture book. As the book was presented, the examiner prompted, "I'm going to ask you to tell a story. Have you ever heard the story of Cinderella? . . . Do you remember much about it? These pictures might remind you of how it goes. Participants were then given time to review the pictures at their own pace. Once the participant indicated they had finished, the clinician provided the following instruction: "Now tell me as much of the story of Cinderella as you can in your own words. You can use any details you know about the story, as well as the pictures you just looked at". Author MD verified that all downloaded transcripts met the document formatting guidelines (i.e., they included the accurate data headers, punctuation marks, and the label of various error words), and that the transcribed text included the %mor tagging to mark the parts of speech.

Discourse data coding and analysis

Transcripts coded by trained AphasiaBank raters using CHAT software were extracted for microlinguistic analyses. Transcripts captured many aspects of spoken language

production, including utterances, dysfluencies, and word- and utterance-level errors. All samples were evaluated for coherence using Linnik et al.'s (2022) analysis rubric. We selected this framework because it comprehensively addresses the intricate aspects of coherence. The four aforementioned aspects align with established definitions of coherence found in prior research (e.g., Ulatowska et al., 2004; Wright & Capilouto, 2016). As per the rubric, coherence is operationalized in terms of *informativeness, clarity, understandability*, and *connectedness* (see Figure 1). To further operationalize informativeness (i.e., essential information), we used the key events central to the Cinderella story as defined by Richardson and Dalton's (2016) framework, which identifies 34 main concepts in the Cinderella narrative (see Supplementary Material 3). In our analysis, narratives that conveyed most or all of the core events (approximately ≥ 70% of concepts) were assigned higher informativeness scores (3–4), with scores of 4 reflecting narratives that included

Informativeness. Informativeness refers to the amount of relevant information content in the retold story.	
Informative	(4) All of the essential information is conveyed
	(3) Most of the essential information is conveyed
Uninformative	(2) Some of the essential information is conveyed; there are omissions that affect the understanding of the retold story
	(1) The information provided in the retelling is incomplete, a lot of information is missing
Clarity. The term clarity is used to describe the overall meaningfulness of discourse achieved through the appropriateness of its elements with respect to the overall topic of the story and narrative structure. Clarity is the degree to which the story, as a whole, hangs together or makes sense.	
Clear	(4) The story is well structured, every episode of the story is appropriately placed, the speaker does not deviate from the storyline too much
	(3) The story is structured well; there are occasional comments and deviations from the storyline, but they do not disrupt the sequence of elements/ episodes of the story
Unclear	(2) The storyline is hard to follow because of multiple deviations or omissions from the storyline or confused order of episodes
	(1) The retelling is hard or impossible to follow because the storyline is confused and/ or there were serious omissions
Understandability. Understandability reflects how well the meaning intended by the speaker could be perceived and/ or interpreted by the listener. This is not a measure of information content, although omitted information may influence the understandability of a retelling.	
Understandable	(4) It is easy to understand the plot of the story, I understood everything well
	(3) It is possible to understand the overall plot of the story, although some parts are somewhat fuzzy
Not Understandable	(2) It is hard to understand much of the story from the retelling; often it is not clear what the speaker is talking about
	(1) It is almost impossible to understand anything from the retelling
Connectedness. Connectedness touches upon the relationship between consecutive elements of the retelling and appropriate organization of its parts into a unified whole through the use of semantic and pragmatic relations.	
Connected	(4) Transitions between sentences and larger parts of the story are smooth
	(3) Transitions between parts of the story are mostly smooth, although some transitions between episodes and within them are missing or too abrupt
Disconnected	(2) The flow of the story is often disrupted, many transitions between sentences and larger parts of the story are missing or too abrupt

Figure 1. Rubric used to evaluate coherence of story narratives (adapted from Linnik et al., 2022).

nearly all core events. Narratives that included fewer core events (<70%) received lower informativeness scores (1–2), reflecting increasing omissions of essential information. In particular, narratives containing very few core events (approximately < 40%) were assigned the lowest informativeness score, as they lacked sufficient information to convey a coherent version of the story. These percentage ranges served as guidelines to support rubric-based judgments rather than strict numerical cutoffs. Episodes were operationally defined as discrete, goal-directed events within the narrative that included a clear initiating event (e.g., the king decides the prince must marry) and outcome (e.g., a royal ball is announced to find a suitable wife). Episodes were considered out of order when the temporal sequence of events in the participant's narrative deviated from the canonical chronological order of the stimulus story. Coders determined order by referencing the stimulus event timeline and evaluating whether an episode was recounted earlier or later than its correct position. In Linnik's work, understandability was defined broadly to include nonverbal communication following Holland (1982). However, our study did not include nonverbal communication and operationalized understandability solely on the basis of verbal output. Linnik and colleagues reported moderate inter-rater agreement across the four domains, as measured by Kendall's coefficient of concordance (W): informativeness (W = 0.715), clarity (W = 0.710), understandability (W = 0.648), and connectedness (W = 0.763).

To note, we applied Linnik et al.'s rubric as a transcription-free measure, assigning coherence ratings based on direct review of the video samples. Author MU served as the primary coder. Video recordings of each participant's language sample were first reviewed by the coder, and the instructions encouraged the coder to watch each video at least once before beginning to code. Following this initial viewing, the coder watched the recordings again to score the discourse sample produced by each participant. Coders were encouraged to rewatch the video to assign scores for all four coherence domains. Each narrative was rated holistically, by reviewing the full sample and assigning one numerical score for each of Linnik's rubric components for the entire narrative. Prior to scoring all samples, authors MD and MU collaboratively coded the first two samples from AphasiaBank from the latent aphasia group (ACWT04a and adler03a) for scoring calibration and discussion. Each sample required approximately 5–7 minutes to code on average.

Additionally, the selected narrative samples were analyzed for microlinguistic (i.e., word and sentence level) features using productivity, lexical, and syntactic variables. The transcriptions were extracted and examined using the Computerized Language Analysis (CLAN) program (MacWhinney et al., 2011; v25, downloaded in April 2023), the accompanying analysis program to CHAT. Checking of each downloading transcript and extraction of microlinguistic variables was conducted by author MD. See Table 2 for definitions and reported psychometric properties of the extracted discourse variables.

Reliability

Twenty percent of the samples from each participant group were re-rated independently by a different rater (author MD) to establish inter-rater reliability, and again by the same rater (author MU) to establish intra-rater reliability. Reliability analysis was conducted for the hand-scored coherence ratings. Any disagreements in scoring were resolved through discussion. Inter-rater and intra-rater reliability were determined using Kendall's tau

Table 2. Definitions and published psychometric properties of extracted discourse variables.

Variable	Description	Psychometrics
Speech rate	The number of intelligible words produced in a minute	Test-retest reliability ICC = .97 for PWA, .79 for controls; Stark et al. (2023) Construct validity: Expected group differences $p < .0001$ with controls > PWA, Stark et al. (2023)
Mean Length of Utterance (MLU)	Average number of words per utterance	Test-retest reliability ICC = .94 for PWA, .66 for controls; Stark et al. (2023) Construct validity: Expected group differences $p < .0001$ with controls > PWA, Stark et al. (2023)
Type-token ratio (TTR)	Ratio of unique words in a sample (types) to the number of words produced (tokens)	Test-retest reliability ICC = .91 for PWA, .76 for controls, Stark et al. (2023) Intra- and inter-rater reliability Stark et al. (2023) Minimal detectable change at 90% confidence Stark et al. (2023)
Verbs per utterance	Average number of verbs per utterance (including verbs, copulas, and auxiliaries followed by past or present participles)	Test-retest reliability ICC = .91 for PWA, .56 for controls; Stark et al. (2023)
Informativeness	The amount of relevant content within the discourse	Inter-rater agreement = 0.715 Linnik et al. (2022)
Clarity	The semantic cohesion of the discourse, achieved through the organization of its components in relation to the overall purpose, goal, or topic.	Inter-rater agreement = 0.710 Linnik et al. (2022)
Understandability	The subjective evaluation of communicative competence	Inter-rater agreement = 0.648 Linnik et al. (2022)
Connectedness	The local semantic and pragmatic relationships between elements of the discourse	Inter-rater agreement = 0.763 Linnik et al. (2022)

Note. ICC = Intraclass correlation coefficient.

Table 3. Reliability estimates for inter-rater, intra-rater, and test-retest reliability of coherence rubric scores, reported using Kendall's tau (τ) and Spearman's rho (ρ) correlations with corresponding 95% confidence intervals. Systematic differences were estimated by Wilcoxon's signed-ranks test (z) for paired data.

Inter-and Intra-Rater Reliability			
		Inter-rater τ value	Intra-rater τ value
Informativeness		.815 [.658, .938], $p < .001$	.897 [.716, 1.000], $p < .001$
Clarity		.805 [.660, .913], $p < .001$	.891 [.754, 1.000], $p < .001$
Understandability		.797 [.675, .901], $p < .001$	.866 [.578, 1.000], $p < .001$
Connectedness		.784 [.610, .915], $p < .001$	.807 [.603, 1.000], $p = .001$
Test-Retest Reliability			
	Group	**Spearman rho (ρ)**	**Systematic difference (Z)**
Informativeness	Aphasia	ρ = .580 [.112, .807], $p = .005$	$Z = -2.333$, $p = .068$
	Control	ρ =.826 [.529, 1.000], $p < .001$	$Z = .000$, $p = 1.000$
Clarity	Aphasia	ρ =.693 [.479, .823], $p < .001$	$Z = -2.887$, $p = .050$
	Control	ρ =.700 [.154, 1.000], $p = .012$	$Z = -.447$, $p = .655$
Understandability	Aphasia	ρ =.594 [.298, .801], $p < .001$	$Z = -2.615$, $p = .106$
	Control	ρ =.596 [.000, .913], $p = .034$	$Z = .000$, $p = 1.000$
Connectedness	Aphasia	ρ =.587 [.104, .785], $p = .003$	$Z = -2.244$, $p = .212$
	Control	ρ =.742 [.398, .964], $p = .007$	$Z = .000$, $p = 1.000$

correlations, accounting for our sample size and the non-normal distribution of the coherence ratings with bootstrapped 95% confidence intervals (CIs) (Higgins, 2004). As documented in Table 3, Kendall rank correlations suggested that the inter and intra-rater

reliability across the four coherence domains ranged from good to excellent. Additionally, consistent with prior discourse reliability studies in aphasia (e.g., Boyle, 2014; Capilouto et al., 2006; Stark et al., 2023), we assessed test-retest reliability for the coherence variables for those samples that had the Cinderella narratives from a second testing timepoint (n = 19 in the latent aphasia group; n = 12 in the control group) on AphasiaBank. To measure the strength of association between two variables (i.e., consistency), we computed Spearman rho on measures between test and retest. All coherence variables demonstrated a significant Spearman rho, however the values indicate moderate test-retest reliability for the aphasia group compared to good-to-excellent reliability for the control group across both timepoints. To evaluate if there was a systematic difference between test and retest on the coherence variables (e.g., to identify significant changes between testing timepoints), we employed the Wilcoxon signed-ranks test. Results indicated no significant differences between test and retest for all coherence domain scores across both groups, suggesting stability. The informativeness and clarity domain scores for the latent aphasia group showed borderline differences (see Table 3).

Statistical analysis

Data were analyzed using the Statistical Package for the Social Sciences (SPSS IBM Statistics) Version 31.1. Normality was assessed using Shapiro – Wilk tests and Q – Q plots. Most discourse variables were not normally distributed; therefore, non-parametric tests were used. Group differences were examined using Mann-Whitney U tests. An analysis of coherence z-scores was conducted to evaluate overall discourse performance across both groups. It is expected that the control group would demonstrate above-average performance, indicated by mean z-scores greater than zero. To examine the relationship between microlinguistic discourse variables and coherence scores, Spearman's correlation analyses were conducted for both latent aphasia and control groups. Prior to conducting the correlational analysis, the assumptions of monotonicity and linearity were examined using scatterplots with fitted lines in SPSS. Visual inspection of the scatterplots indicated that the relationships between variables were both monotonic and approximately linear, supporting the use of correlation analysis. The significance threshold used for all statistical tests was $p < .05$, two-tailed. Multiple comparisons were corrected for by using False Discovery Rate (FDR; Benjamini & Hochberg, 1995). Effect sizes were calculated using Cohen's d to quantify the magnitude of group differences across variables extracted from the narrative discourse task, with values of 0.2, 0.5, and 0.8 interpreted as small, medium, and large effects, respectively.

Results

Microlinguistic and coherence analysis results

Results from the analysis of the microlinguistic and coherence variables are outlined in Table 4. In terms of productivity, individuals with latent aphasia demonstrated slower speech rates and produced shorter utterances than controls. Additionally, those in the latent aphasia group demonstrated a lower type-token ratio than controls. There was also

Table 4. Results from the Mann-whitney U tests to compare performances of the latent aphasia ($n = 59$) and control ($n = 59$) groups on spoken discourse variables.

Variable	Latent Aphasia ($n = 59$)	Controls Without Aphasia ($n = 59$)	Group Comparison Statistic	Cohen's *d*
	Mean (SD)	Mean (SD)		
Microlinguistic Variable Scores (CLAN)				
Speech rate (Words/min) *	100.35 (36.88)	160.79 (31.50)	$U = 392.00$, $z = -7.258$, $p < .001$	1.762
Mean Length of Utterance (MLU) *	8.53 (1.91)	10.12 (2.28)	$U = 1048.50$, $z = -3.725$, $p < .001$	0.756
Type-token Ratio (TTR) *	0.37 (0.09)	2.00 (0.81)	$U = 1252.00$, $z = -2.502$, $p = 0.012$	0.819
Verbs per utterance *	1.48 (0.34)	1.73 (0.32)	$U = 1007.00$, $z = -3.838$, $p < .001$	0.757
Coherence Rubric Scores (Linnik et al., 2022)				
Informativeness *	2.35 (0.73)	3.57 (0.62)	$U = 432.00$, $z = -7.413$, $p < .001$	1.801
Clarity *	2.25 (0.65)	3.45 (0.62)	$U = 415.00$, $z = -7.519$, $p < .001$	1.889
Understandability*	2.47 (0.62)	3.67 (0.53)	$U = 364.50$, $z = -7.862$, $p < .001$	2.080
Connectedness*	2.22 (0.72)	3.67 (0.50)	$U = 266.00$, $z = -8.359$, $p < .001$	2.339

Note. * = variables with statistically significant group difference; WAB-R AQ = Western Aphasia Battery – Aphasia Quotient; SD = standard deviation; reported *p* values are corrected for multiple comparisons using false discovery rate.

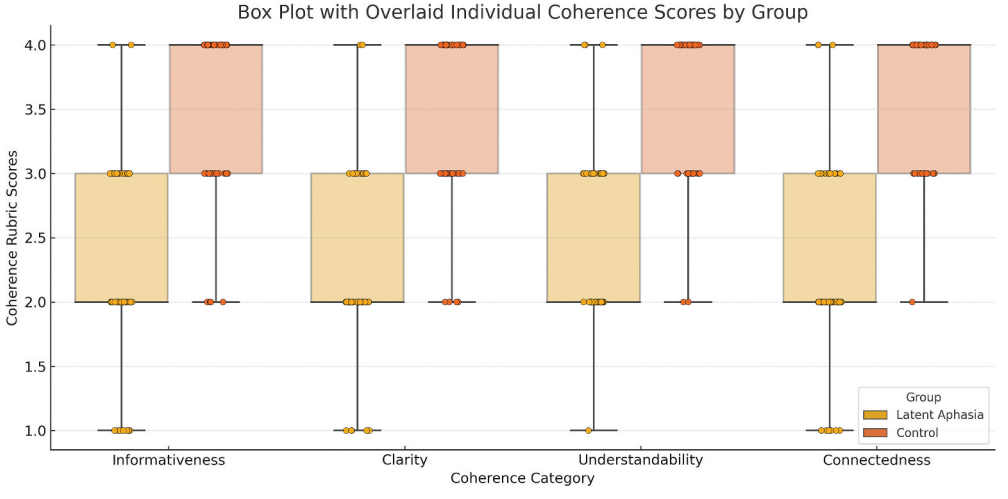

Figure 2. Coherence rubric scores for latent aphasia ($n = 59$) and control ($n = 59$) groups.

a significant group difference noted for verbs per utterance between the latent aphasia and control groups.

Groups differed significantly across all four coherence domains: individuals with latent aphasia scored significantly lower than controls in terms of informativeness, clarity, understandability, and connectedness (Figure 2). The effect sizes for group comparisons of microlinguistic variables ranged from small to large (Cohen, 1988), whereas coherence rubric scores consistently showed large effect sizes.

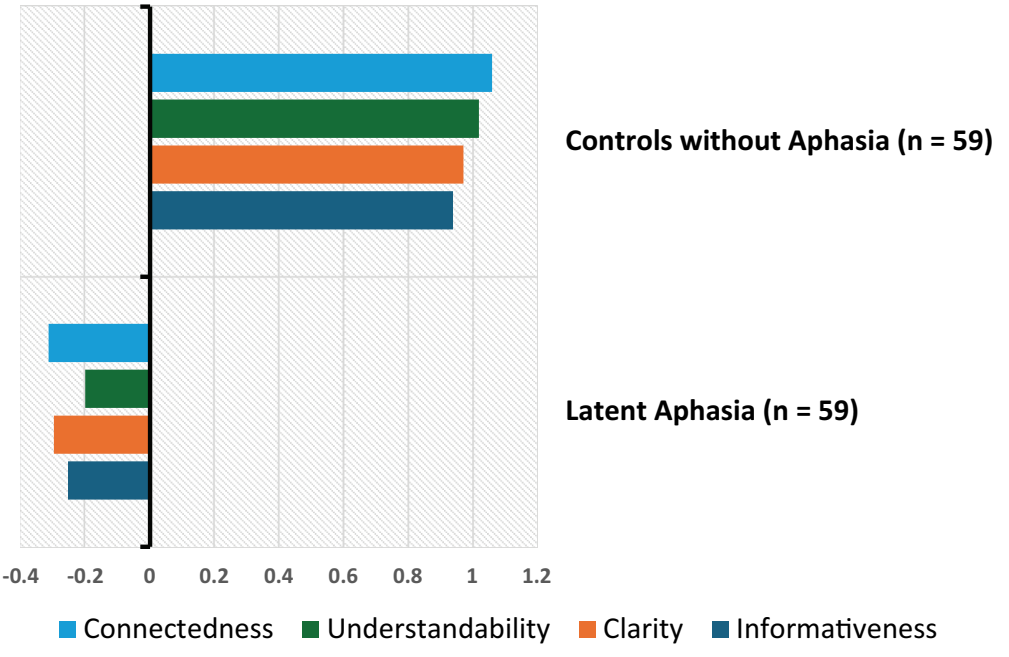

Figure 3. Mean composite z-scores for informativeness, clarity, understandability, and connectedness for the latent aphasia and healthy control groups, presented relative to the overall performance across both groups.

An analysis of the coherence *z*-scores showed that the control group performed above average (i.e., composite score above 0), while the latent aphasia group scored below average although this group remained within one standard deviation below the mean (See Figure 3).

Relationship between microlinguistic variables and coherence scores

In the latent aphasia group, both clarity (ρ =0.379, p = .024) and connectedness (ρ =0.338, p = .036) domains of the coherence rubric were positively correlated with mean length of utterance (MLU). Comparatively, for the control group, the same coherence domains – clarity (ρ =0.332, p = .030) and connectedness (ρ =0.334, p = .030) were significantly positively correlated with speech rate. Overall, the strength of all significant correlations ranged as weak across both groups. No additional significant correlations were observed (refer to Table 5 and Supplementary Material 2 for the full correlation results).

Discussion

The current study aimed to evaluate and characterize spoken discourse abilities of individuals with latent aphasia, with a focus on examining narrative coherence in relation to microlinguistic variables. We compared spoken discourse productions of the latent aphasia group to those without aphasia. Our findings add to the extant literature and demonstrate

Table 5. Significant Spearman rank correlations (ρ) between microlinguistic discourse variables and coherence scores for the latent aphasia and control groups.

Discourse Variable	Informativeness	Clarity	Understandability	Connectedness
Latent Aphasia (n = 59)				
Mean Length of Utterance (MLU)	n.s	**ρ = 0.379,** p = .024	n.s	**ρ = 0.338,** p = .036
Controls Without Aphasia (n = 59)				
Speech Rate	n.s	**ρ = 0.332,** p = .030	n.s	**ρ = 0.334,** p = .030

Note. Correlation is significant at the 0.01 level (2-tailed); reported p values are corrected for multiple comparisons using false discovery rate.
n.s = correlation not significant
Please refer to Supplementary Material 2 for all reported correlations between microlinguistic discourse and coherence variables.

the effectiveness of discourse analysis in identifying language deficits in individuals with latent aphasia. Below, we present a detailed description and discussion of the results.

Discourse-level differences on microlinguistic and coherence variables between latent aphasia and control groups

Individuals with latent aphasia exhibited clear differences in storytelling performance compared to individuals without aphasia, supporting previous findings in this population (De De & Salis, 2020; Fromm et al., 2017). Specifically, for the microlinguistic variables, the latent aphasia group demonstrated significantly reduced discourse productivity – characterized by lower speech rate and MLU – relative to the control group. Significant group differences emerged in terms of lexical diversity (i.e., lower type-token ratio) as well. These findings align with previous research demonstrating lexical impairments in individuals with latent aphasia compared to neurotypical controls (Fromm et al., 2017; Stark et al., 2025), suggesting that reduced lexical diversity are common among these individuals, even when surface-level fluency may appear intact. The number of verbs produced per utterance significantly differed between both groups, suggesting reduced syntactic complexity in individuals with latent aphasia. A more nuanced analysis of verb usage such as examining the proportion of light versus heavy verbs (e.g., Cruice et al., 2014; Park et al., 2023) may offer deeper insights into lexical and syntactic patterns associated with latent aphasia compared to neurotypical controls.

In addition to microlinguistic analysis, this study examined narrative coherence – a critical macrolinguistic feature that had not been previously sufficiently explored in individuals with latent aphasia. In the current study, coherence was evaluated in terms of four distinct domains – informativeness, clarity, understandability, and connectedness, as operationally defined by Linnik et al. (2022). Consistent with prior findings in broader aphasia populations (Kong et al., 2018; Wright et al., 2010), the analysis revealed significant differences across all four coherence domains between both groups. These results not only reinforce existing evidence that coherence deficits are a common feature of aphasic discourse (Linnik et al., 2022; Wright & Capilouto, 2012), but also extend this understanding to latent aphasia, highlighting that even subtle language impairments can disrupt higher-level discourse organization.

Notably, while significant group differences emerged across most microlinguistic and coherence variables, the patterns of effect sizes differed – suggesting that these two levels of discourse analysis may vary in their sensitivity to underlying language impairments. The

variation in effect sizes for microlinguistic variables – ranging from medium to large – suggests that not all aspects of microlinguistic performance equally distinguish between groups. Some variables may reflect subtle differences, while others show more pronounced group-level contrasts. Comparatively, the consistently large effect sizes observed for coherence rubric scores indicate that coherence may be a more robust and sensitive indicator of group differences. Overall, while microlinguistic features varied in their discriminative power, coherence measures reliably capture the underlying discourse-level impairments across groups.

Correlational patterns between microlinguistic measures and narrative coherence in latent aphasia and control groups

When microlinguistic measures were correlated with coherence rubric variables for both latent aphasia and control groups, only a few correlations emerged as significant. Specifically, for the latent aphasia group, in line with Dutta et al. (2024), greater clarity and connectedness of narrative output were associated with longer MLUs. In contrast, for the control group, higher scores on clarity and connectedness domains of the coherence rubric correlated with faster speech rate. Collectively, these findings differ from previous studies on more severe or overt forms of aphasia (Andreetta & Marini, 2015; Wright & Capilouto, 2012), which report stronger and more consistent relationships between microlinguistic and macrolinguistic variables. In the current study, both latent aphasia and control groups showed fewer and less robust correlations between microlinguistic measures and coherence ratings. In fact, both groups showed similar patterns of association, with discourse productivity variables mainly being linked to narrative coherence. These results build upon the findings of Linnik et al. (2022), suggesting that in speakers with latent aphasia, the relationship between microlinguistic features and coherence may be weaker – resembling patterns observed in individuals without aphasia (Dutta et al., 2024) – or that other factors not captured by the measures used here may play a more significant role in shaping their narrative performance. Importantly, these findings provide preliminary evidence that the relationship between microlinguistic and macrolinguistic measures may differ in latent aphasia. These links potentially may also be influenced by the nature of the discourse task. Notably, most studies examining the interplay between these levels have used on narrative or picture description tasks (Andreetta et al., 2025; Dutta et al., 2024; Marini et al., 2011; Wright & Capilouto, 2012). Thus, future research should investigate these relationships across a broader range of discourse genres in latent aphasia.

To the best of our knowledge, this is the first study to examine the relationship between microlinguistic and macrolinguistic discourse levels in latent aphasia. Our study findings imply that while microlinguistic features are important for maintaining narrative coherence (Wright & Capilouto, 2012), their impact may vary between different severities of aphasia and may be influenced by factors beyond the linguistic elements. The limited number of significant correlations in the latent aphasia group further suggests that the dynamics of narrative production in this group may be more complex and not as strongly tied to the specific microlinguistic variables assessed. Future research should explore these microlinguistic-macrolinguistic interactions by incorporating additional linguistic variables. It is also possible that cognitive factors such as processing speed,

attention, executive functioning, working memory may contribute to and better explain the breakdowns at the discourse level in individuals with latent aphasia (Salis & De De, 2022) as has been documented in previous aphasia studies (Cahana-Amitay & Jenkins, 2018; Dutta et al., 2024). For example, individual analyses of some participants in the latent aphasia group revealed difficulties in organizing and recalling elements of the Cinderella story. They frequently exhibited coherence violations, such as tangential or abandoned utterances (Andreetta et al., 2012). While studies have examined discourse characteristics in latent aphasia (e.g., Dalton & Richardson, 2015; Richardson et al., 2021), these impairment patterns are potentially suggestive of challenges with memory for story episodes, inhibition of irrelevant details, and planning of utterances for coherent storytelling. Investigating cognitive assessment scores and their relationship to discourse performance could provide insight into these difficulties.

General discussion and clinical implications

Our findings demonstrate that individuals with latent aphasia exhibit measurable discourse deficits despite scoring above standard aphasia assessment cutoffs. Beyond traditional language testing, analysis of their discourse abilities allowed us to examine the exact nature of their language difficulties and differentiating them from neurologically unimpaired individuals without aphasia. Our analysis revealed that, compared to controls, the latent aphasia group exhibited significantly lower productivity, distinct lexical and syntactic differences, a higher frequency of word errors, and reduced narrative coherence.

Although the microlinguistic analysis revealed performance differences between the two participant groups, the effect sizes associated with these group differences varied. Thus, relying solely on microlinguistic variables may therefore increase chances of missing important discourse-level deficits experienced by individuals with latent aphasia. These results underscore the need to broaden assessment procedures for individuals with aphasia, especially those with latent aphasia, to encompass macrolinguistic aspects of discourse organization. In our study, we focused on evaluating narrative coherence. Although the rating rubric developed by Linnik and colleagues (2022) was originally designed based on research with Russian speakers, it proved to be an effective tool for identifying coherence deficits in English speakers with aphasia, particularly in individuals with subtle language impairments. This scale enabled a comprehensive assessment of linguistic aspects contributing to coherence and revealed that individuals with latent aphasia consistently differed from controls across four key variables – informativeness, clarity, understandability, and connectedness – indicating reduced narrative coherence. Unlike prior studies that examined these components separately, this integrated rating scale provided a unified evaluation of how coherence is affected in aphasia and how its components collectively influence spoken discourse. Different from other coherence scales that rely on broad, less-specific rating definitions, Linnik et al.'s rating rubric offers a more detailed and precise assessment. However, although the Linnik et al.'s rating captured narrative coherence impairments in latent aphasia and differentiated them from controls, it is important to interpret these findings in light of the moderate test – retest reliability reported for these measures. Moderate stability indicates that while the measures are reasonably consistent across administrations, they do not demonstrate the level of robustness typically desired for strong inferential conclusions. As a result, the strength

and precision of the relationships observed in the present data should be viewed with appropriate caution. Overall, a comprehensive discourse analysis approach that integrates both micro- and macro-linguistic measures is essential for improving the diagnostic sensitivity of language assessments in latent aphasia. When assessments are conducted accurately and impairments are properly identified, clinicians can effectively document these challenges and tailor aphasia therapy to address the right areas. In addition to structural language components (e.g., improving productivity, word retrieval), macrolinguistic skills should be a key focus in language therapy, as strengthening these abilities can significantly enhance everyday communication, enabling individuals to better engage in meaningful activities and social participation – aspects that are often challenging for individuals with milder forms of aphasia (Cavanaugh & Haley, 2020).

This study reasserts the fact that the inclusion of spoken discourse tasks in aphasia assessments is critical as standardized aphasia assessments alone are often insufficient and may fail to capture deficits. Furthermore, the findings of this study highlight the value of narrative-based evaluations in assessing language abilities in individuals with latent aphasia – a discourse genre widely used in clinical practice. Consistent with previous research (Richardson et al., 2021; Salis & De De, 2022), the Cinderella storytelling task proved to be sensitive in detecting linguistic deficits in the latent aphasia group compared to controls. Unlike simpler tasks such as picture description, which may overlook the subtle impairments that are characteristic of latent aphasia, narrative tasks elicit richer language output and require individuals to integrate multiple linguistic and cognitive processes, including lexical retrieval, syntactic structuring, and discourse organization. Although different from more functional real-world tasks such as spontaneous conversations, narrative tasks allow for a more structured assessment of monologic language production for individuals aiming to strengthen their narrative skills which are often required for activities that involve organized storytelling such as videoblogging or telling bedtime or personal stories (Dutta et al., 2024; Whitworth et al., 2015). Importantly, evaluating discourse through multiple tasks or genres is recommended to provide a more comprehensive assessment of language abilities and to accurately capture the various types of discourse that individuals with aphasia may wish to improve (Armstrong, 2000; Dipper et al., 2018; Dutta et al., 2025; Leaman & Archer, 2023; Stark, 2019).

Relating our study's findings to contemporary discourse models, such as the Linguistic Underpinnings of Narrative in Aphasia (LUNA) framework (Dipper et al., 2021), offers a theoretically grounded lens for interpreting the linguistic profiles of individuals with latent aphasia. The LUNA framework outlines four core components of discourse production: (1) linguistic, (2) propositional, (3) macrostructure planning, and (4) pragmatic – each of which may be selectively disrupted in aphasia. The impairment patterns identified in our study – reduced lexical access, limited utterance formulation, and diminished narrative coherence – align with disruptions in the linguistic, propositional, and macrostructure planning components of the LUNA model. However, our findings also revealed that not all microlinguistic impairments equally predicted coherence outcomes, providing limited evidence for a straightforward interaction among LUNA components. The weak correlations between micro- and macrolinguistic measures instead suggest a more complex, potentially non-linear relationship between levels of discourse processing in latent aphasia. Collectively, this study reinforces the LUNA framework's emphasis on assessing multiple discourse components and underscores the importance of including narrative-based,

discourse-level analyses in the evaluation of individuals with latent aphasia, who may otherwise be overlooked by standard language assessments.

Limitations and future directions

This study examined a limited set of discourse variables, highlighting the need for more comprehensive assessment of microlinguistic and macrolinguistic abilities and their relationship in latent aphasia. Consistent with previous studies that included individuals with latent aphasia (e.g., Cunningham & Haley, 2020; Fromm et al., 2017; Richardson et al., 2021), our research primarily focused on storytelling, yet we know that discourse outcomes vary depending on the task used (Fergadiotis & Wright, 2011; Stark, 2019). To fully understand the extent of impairments in latent aphasia, future studies should examine diverse discourse genres. Furthermore, emerging evidence suggests a strong link between cognitive processes – such as attention and executive functioning – and discourse production in aphasia (e.g., Cahana-Amitay & Jenkins, 2018; Dutta et al., 2024). Individuals with latent aphasia may experience extralinguistic cognitive deficits (Salis & De De, 2022; Silkes et al., 2021), which, alongside language impairments, can affect discourse coherence and overall communicative effectiveness. However, due to the lack of cognitive testing data on AphasiaBank, our ability to investigate this relationship was limited. Future research should examine how cognitive impairments influence discourse coherence in latent aphasia to inform underlying mechanisms and targeted interventions.

Discourse features such as coherence can be culturally and linguistically influenced, and their manifestation can vary significantly across cultural backgrounds, languages, and dialects. While our study utilized data from AphasiaBank due to its accessibility and established use in aphasia research, we acknowledge its limited demographic diversity. This presents a constraint on the generalizability of our findings and underscores the need for more research to examine coherence across a broader range of languages and cultural contexts to ensure more inclusive and representative discourse assessments. Relatedly, the choice of the Cinderella narrative task – selected for its prevalence in prior aphasia and narrative research – may not be equally appropriate across all linguistic and cultural contexts. It may also introduce biases related to gender and age (Dutta et al., 2025; Fromm et al., 2011; Kevická et al., 2025). Future work should explore the use of more culturally diverse narratives. These limitations highlight the need for future research to examine discourse coherence using more culturally diverse populations and narrative stimuli, ensuring that discourse assessments are both inclusive and representative. Moreover, the instructions used in AphasiaBank to elicit the Cinderella narrative do not explicitly prompt participants to structure their retelling with a clear beginning, middle, and end. The absence of such instructions may have influenced narrative organization and, consequently, performance on the *Clarity* and *Understandability* components of the coherence measure, which rely in part on story grammar elements (Olness, 2006; Wright & Capilouto, 2009). Although participants were not instructed to produce the story in chronological order, deviations from the canonical sequence of events were considered within the clarity dimension of Linnik et al.'s rubric when they affected the ease with which the narrative could be followed. While temporal organization is often assumed in discourse studies using the Cinderella narrative stimulus, narratives may remain understandable even when events are presented nonlinearly. In addition, storytelling

conventions vary across cultures; thus, evaluating narratives against the canonical temporal structure of the stimulus story may disadvantage storytelling styles that are less strictly chronological. Our results were restricted to English speakers in the United States; however, future research should examine how this framework and elicitation procedures apply across diverse cultural and linguistic contexts where narrative organization may differ.

A further limitation is that the coherence metric employed was developed specifically for storytelling narratives. As such, it is not suitable for evaluating coherence in other forms of discourse, such as procedural explanations, expository texts, or spontaneous conversations, which differ substantially in structure and communicative intent. The second author was not blinded to participant group during the coherence scoring process and was aware of the study's goals. We acknowledge that this introduces a potential source of rater bias in the coherence assessment. While efforts were made to ensure consistency through the use of a detailed rubric and interrater reliability checks, the lack of blinding remains a limitation. We recommend that future studies implement blinded scoring procedures to reduce the risk of expectancy effects in discourse evaluations. Although no significant test – retest differences were observed, borderline effects for informativeness and clarity in the latent aphasia group may reflect measurement variability or practice effects; thus, findings should be interpreted cautiously. The current study represents an initial exploration. The patterns identified here warrant further investigation using controlled retest intervals and through multi-session data collection that allows for modeling of stability-related variance of this coherence measure.

As the first application of a transcription-free perceptual method for coherence, its correspondence with traditional transcription-based scoring has not yet been empirically established. Additional work is needed to directly compare perceptual and transcription-based approaches to evaluate convergence and to examine how this time-efficient metric compares with other coherence analysis systems, including its potential for routine clinical use (Dutta et al., 2025; Linnik et al., 2022). Relatedly, the moderate test – retest stability reported for Linnik's rating scale suggests that reliability may differ when applied using transcription-based methods, highlighting an important direction for future work.

Conclusion

The current study adds to the extant literature by providing compelling evidence that individuals with latent aphasia – despite scoring within normal limits on standardized assessments – exhibit measurable deficits in both microlinguistic and macrolinguistic aspects of discourse, particularly in narrative coherence. By applying Linnik et al.'s (2022) multidimensional rubric, we demonstrated that coherence impairments in latent aphasia are not only detectable but are also consistently present across the domains of informativeness, clarity, understandability, and connectedness. These coherence deficits were distinct from, and in some cases more pronounced than, differences in traditional microlinguistic measures, highlighting that macrolinguistic coherence measures may offer greater sensitivity than microlinguistic variables in distinguishing subtle language impairments in this population. The limited and inconsistent correlations between microlinguistic and coherence variables further suggest that these levels of language production may be influenced by different underlying mechanisms that are not fully captured by

linguistic analysis alone. Taken together, this study emphasizes the need for more comprehensive assessments that extend beyond isolated language components to capture the complex, integrative nature of discourse. Selecting appropriate assessments to evaluate discourse-level impairments in latent aphasia can enhance diagnostic precision and enable clinicians and researchers to more effectively support individuals in achieving meaningful communicative outcomes.

Acknowledgements

We gratefully acknowledge and thank the developers, contributors, and research participants of AphasiaBank. We would also like to thank the research assistants in the Adult Language and Cognition Lab at Portland State University for their assistance with data entry, organization, and coding.

Disclosure statement

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

ORCID

Manaswita Dutta http://orcid.org/0000-0001-8371-2246

Data availability statement

The data that support the findings of this study are available from AphasiaBank (http://aphasia.talkbank.org).

Patient consent information

All the participants had signed informed consent.

References

Andreetta, S., Cantagallo, A., & Marini, A. (2012). Narrative discourse in anomic aphasia. *Neuropsychologia*, *50*(8), 1787–1793. https://doi.org/10.1016/j.neuropsychologia.2012.04.003

Andreetta, S., & Marini, A. (2015). The effect of lexical deficits on narrative disturbances in fluent aphasia. *Aphasiology*, *29*(6), 705–723. https://doi.org/10.1080/02687038.2014.979394

Andreetta, S., Marini, A., Menichelli, A., Furlanis, G., Vincis, E., Caruso, P., Naccarato, M., & Manganotti, P. (2025). Language assessment in persons with aphasia early after thrombolysis: The utility of multilevel procedures of discourse analysis. *Aphasiology*, *39*(2), 182–213. https://doi.org/10.1080/02687038.2024.2340797

Armstrong, E. (2000). Aphasic discourse analysis: The story so far. *Aphasiology*, *14*(9), 875–892. https://doi.org/10.1080/02687030050127685

Armstrong, E., Fox, S., & Wilkinson, R. (2013). Mild aphasia: Is this the place for an argument? *American Journal of Speech-Language Pathology*, *22*(2), S268–S278. https://doi.org/10.1044/1058-0360(2012/12-0084)

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57* (1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Boller, F. (1968). Latent aphasia: Right and left "non aphasic" brain-damaged patients compared. *Cortex*, *4*(3), 245–256. https://doi.org/10.1016/S0010-9452(68)80003-6

Boyle, M. (2014). Test-retest stability of word retrieval in aphasic discourse. *Journal of Speech, Language, & Hearing Research*, *57*(3), 966–978. https://doi.org/10.1044/2014_JSLHR-L-13-0171

Cahana-Amitay, D., & Jenkins, T. (2018). Working memory and discourse production in people with aphasia. *Journal of Neurolinguistics*, *48*(April), 90–103. https://doi.org/10.1016/j.jneuroling.2018.04.007

Capilouto, G. J., Wright, H. H., & Wagovich, S. A. (2006). Reliability of main event measurement in the discourse of individuals with aphasia. *Aphasiology*, *20*(2–4), 205–216. https://doi.org/10.1080/02687030500473122

Cavanaugh, R., & Haley, K. L. (2020). Subjective communication difficulties in very mild aphasia. *American Journal of Speech-Language Pathology*, *29*(1S), 437–448. https://doi.org/10.1044/2019_AJSLP-CAC48-18-0222

Christiansen, J. A. (1995). Coherence violations and propositional usage in the narratives of fluent aphasics. *Brain and Language*, *51*(2), 291–317. https://doi.org/10.1006/brln.1995.1062

Code, C., & Herrmann, M. (2003). The relevance of emotional and psychosocial factors in aphasia to rehabilitation. *Neuropsychological Rehabilitation*, *13*(1–2), 109–132. https://doi.org/10.1080/09602010244000291

Coelho, C., & Flewellyn, L. (2003). Longitudinal assessment of coherence in an adult with fluent aphasia: A follow-up study. *Aphasiology*, *17*(2), 173–182. https://doi.org/10.1080/729255216

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Cruice, M., Pritchard, M., & Dipper, L. (2014). Verb use in aphasic and non-aphasic personal discourse: What is normal? *Journal of Neurolinguistics*, *28*, 31–47. https://doi.org/10.1016/j.jneuroling.2013.12.002

Cruice, M., Worrall, L., & Hickson, L. (2006). Perspectives of quality of life by people with aphasia and their family: Suggestions for successful living. *Topics in Stroke Rehabilitation*, *13*(1), 14–24. https://doi.org/10.1310/4JW5-7VG8-G6X3-1QVJ

Cunningham, K. T., & Haley, K. L. (2020). Measuring lexical diversity for discourse analysis in aphasia: Moving-average type-token ratio and word information measure. *Journal of Speech, Language, & Hearing Research*, *63*(3), 710–721. https://doi.org/10.1044/2019_JSLHR-19-00226

Dalton, S. G. H., Kim, H., Richardson, J. D., & Wright, H. H. (2019). A compendium of core lexicon checklists. *Seminars in Speech and Language*, *41*(1), 045–060. https://doi.org/10.1055/s-0039-3400972

Dalton, S. G., & Richardson, J. D. (2015). Core-lexicon and main-concept production during picture-sequence description in adults without brain damage and adults with aphasia. *American Journal of Speech-Language Pathology*, *24*(4), S923–S938. https://doi.org/10.1044/2015_AJSLP-14-0161

De De, G., & Hoover, E. (2021). Measuring change at the discourse-level following conversation treatment: Examples from mild and severe aphasia. *Topics in Language Disorders*, *41*(1), 5–26. https://doi.org/10.1097/TLD.0000000000000243

De De, G., & Salis, C. (2020). Temporal and episodic analyses of the story of Cinderella in latent aphasia. *American Journal of Speech-Language Pathology*, *29*(1S), 449–462. https://doi.org/10.1044/2019_AJSLP-CAC48-18-0210

Dipper, L., Marshall, J., Boyle, M., Hersh, D., Botting, N., & Cruice, M. (2021). Creating a theoretical framework to underpin discourse assessment and intervention in aphasia. *Brain Sciences 2021*, *11* (2), 183. https://doi.org/10.3390/BRAINSCI11020183

Dipper, L., Pritchard, M., Walkden, E., & Cruice, M. (2018). How do speakers with and without aphasia use syntax and semantics across two discourse genres? *Aphasiology*, *32*(6), 720–738. https://doi.org/10.1080/02687038.2018.1447642

Dutta, M., Murray, L. L., Park, H., Burklow, E., Bose, A., Kim, H., Greenslade, K., Ramage, A. E., Combs, C., Balasubramanian, A., & Casilio, M. (2025). Let's chat about spoken discourse: A tutorial to support

use of spoken discourse analysis when providing aphasia clinical services. *American Journal of Speech-Language Pathology*, *34*(4), 1931–1966. https://doi.org/10.1044/2025_AJSLP-24-00534

Dutta, M., Murray, L. L., & Stark, B. C. (2024). The relationship between executive functioning and narrative language abilities in aphasia. *American Journal of Speech-Language Pathology*, *33*(5), 2500–2523. https://doi.org/10.1044/2024_AJSLP-23-00314

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

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

Fox, S., Armstrong, E., & Boles, L. (2009). Conversational treatment in mild aphasia: A case study. *Aphasiology*, *23*(7–8), 951–964. https://doi.org/10.1080/02687030802669526

Frankel, T., Penn, C., & Ormond-Brown, D. (2007). Executive dysfunction as an explanatory basis for conversation symptoms of aphasia: A pilot study. *Aphasiology*, *21*(6–8), 814–828. https://doi.org/10.1080/02687030701192448

Fromm, D., Forbes, M., Holland, A., Dalton, S. G., Richardson, J., & MacWhinney, B. (2017). Discourse characteristics in aphasia beyond the Western Aphasia Battery cutoff. *American Journal of Speech-Language Pathology*, *26*(3), 762–768. https://doi.org/10.1044/2016_AJSLP-16-0071

Fromm, D., MacWhinney, B., Forbes, M., & Holland, A. (2011). *Is the Cinderella task biased for age or gender*. Academy of Aphasia.

Glosser, G., & Deser, T. (1992). A comparison of changes in macrolinguistic and microlinguistic aspects of discourse production in normal aging. *Journal of Gerontology*, *47*(4), 266–272. https://doi.org/10.1093/geronj/47.4.P266

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. English Language Series, Longman.

Hazamy, A. A., & Obermeyer, J. (2020). Evaluating informative content and global coherence in fluent and non-fluent aphasia. *International Journal of Language & Communication Disorders*, *55*(1), 110–120. https://doi.org/10.1111/1460-6984.12507

Hickin, J., Mehta, B., & Dipper, L. (2015). To the sentence and beyond: A single case therapy report for mild aphasia. *Aphasiology*, *29*(9), 1038–1061. https://doi.org/10.1080/02687038.2015.1010474

Higgins, J. J. (2004). *An introduction to modern nonparametric statistics*. Brooks/Cole.

Hilari, K., & Northcott, S. (2017). "Struggling to stay connected": Comparing the social relationships of healthy older people and people with stroke and aphasia. *Aphasiology*, *31*(6), 674–687. https://doi.org/10.1080/02687038.2016.1218436

Holland, A. L. (1982). Observing functional communication of aphasic adults. *The Journal of Speech and Hearing Disorders*, *47*(1), 50–56. https://doi.org/10.1044/jshd.4701.50

Kertesz, A. (2007). Western aphasia battery-revised. PsychCorp.

Kevická, V., Kaščáková, L., Dančík, D., Brandoburová, P., & Marková, J. (2025). Adaptation of main concept analysis for Slovak SLPs: A pilot study. *Aphasiology*, *40*(4), 825–842. https://doi.org/10.1080/02687038.2025.2473035

Kintz, S. (2024). Discourse characteristics in aphasia. In A.-PAK.-HIN. Kong (Ed.), *Spoken discourse impairments in the neurogenic populations: A state-of-the art, contemporary approach* (pp. 23–36). Springer. https://doi.org/10.1007/978-3-031-45190-4_2

Kong, A. P.-H., Linnik, A., Law, S.-P., & Shum, W. W.-M. (2018). Measuring discourse coherence in anomic aphasia using rhetorical structure theory. *International Journal of Speech-Language Pathology*, *20*(4), 406–421. https://doi.org/10.1080/17549507.2017.1293158

Laks, A. B., DeMarco, A. T., & Turkeltaub, P. E. (2025). Optimizing detection of very mild aphasia with letter fluency tests. *American Journal of Speech-Language Pathology*, *34*(3), 1193–1202. https://doi.org/10.1044/2025_AJSLP-24-00471

Leaman, M. C., & Archer, B. (2023). Choosing discourse types that align with person-centered goals in aphasia rehabilitation: A clinical tutorial. *Perspectives of the ASHA Special Interest Groups*, *8*(2), 254–273. https://doi.org/10.1044/2023_PERSP-22-00160

Leaman, M. C., & Edmonds, L. A. (2021). Measuring global coherence in people with aphasia during unstructured conversation. *American Journal of Speech-Language Pathology*, *30*(1S), 359–375. https://doi.org/10.1044/2020_AJSLP-19-00104

Linnik, A., Bastiaanse, R., Stede, M., & Khudyakova, M. (2022). Linguistic mechanisms of coherence in aphasic and non-aphasic discourse. *Aphasiology*, *36*(2), 123–146. https://doi.org/10.1080/02687038.2020.1852527

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

Marini, A., Andreetta, S., Del Tin, S., & Carlomagno, S. (2011). A multi-level approach to the analysis of narrative language in aphasia. *Aphasiology*, *25*(11), 1372–1392. https://doi.org/10.1080/02687038.2011.584690

Marshall, R. C. (1993). Problem-focused group treatment for clients with mild aphasia. *American Journal of Speech-Language Pathology*, *2*(2), 31–37. https://doi.org/10.1044/1058-0360.0202.31

Martzoukou, M., Nousia, A., & Nasios, G. (2023). Undetected language deficits in left or right hemisphere post-stroke patients. *Applied Neuropsychology: Adult*, *32*(3), 606–614. https://doi.org/10.1080/23279095.2023.2195111

Mayer, J., & Murray, L. (2003). Functional measures of naming in aphasia: Word retrieval in confrontation naming versus connected speech. *Aphasiology*, *17*(5), 481–497. https://doi.org/10.1080/02687030344000148

Morris, J., Franklin, S., & Menger, F. (2011). Returning to work with aphasia: A case study. *Aphasiology*, *25*(8), 890–907. https://doi.org/10.1080/02687038.2010.549568

Murray, L. L., & Clark, H. M. (2015). Neurogenic disorders of language and cognition: Evidence-based clinical practice. Pro-Ed.

Olness, G. S. (2006). Genre, verb, and coherence in picture-elicited discourse of adults with aphasia. *Aphasiology*, *20*(2–4), 175–187. https://doi.org/10.1080/02687030500472710

Olsen, T. S., Bruhn, P., & Öberg, R. G. E. (1986). Cortical hypoperfusion as a possible cause of 'subcortical aphasia'. *Brain*, *109*(3), 393–410. https://doi.org/10.1093/brain/109.3.393

Pallickal, M., & Hema, N. (2020). Discourse in Wernicke's aphasia. *Aphasiology*, *34*(9), 1138–1163. https://doi.org/10.1080/02687038.2020.1739616

Park, H., Obermeyer, J., Kornisch, M., Hall, J., & Ontario, C. (2023). Semantic aspects of verb production in various discourse tasks in people with nonfluent aphasia. *American Journal of Speech-Language Pathology*, *32*(5S), 2418–2429. https://doi.org/10.1044/2023_ajslp-22-00293

Prins, R., & Bastiaanse, R. (2004). Analyzing the spontaneous speech of aphasic speakers. *Aphasiology*, *18*(12), 1075–1091. https://doi.org/10.1080/02687030444000534

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/10.1080/02687038.2015.1057891

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

Salis, C., & De De, G. (2022). Sentence production in a discourse context in latent aphasia: A real-time study. *American Journal of Speech-Language Pathology*, *31*(3), 1284–1296. https://doi.org/10.1044/2022_AJSLP-21-00232

Silkes, J. P., Zimmerman, R. M., Greenspan, W., Reinert, L., Kendall, D., & Martin, N. (2021). Identifying verbal short-term memory and working memory impairments in individuals with latent aphasia. *American Journal of Speech-Language Pathology*, *30*(1S), 391–406. https://doi.org/10.1044/2020_AJSLP-19-00105

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

Stark, B. C., Alexander, J. M., Hittson, A., Doub, A., Igleheart, M., Streander, T., & Jewell, E. (2023). Test-retest reliability of microlinguistic information derived from spoken discourse in persons with chronic aphasia. *Journal of Speech, Language, & Hearing Research*, *66*(7), 2316–2345. https://doi.org/10.1044/2023_JSLHR-22-00266

Stark, B. C., & Dalton, S. G. (2024). A scoping review of transcription-less practices for analysis of aphasic discourse and implications for future research. *International Journal of Language & Communication Disorders*, *59*(5), 1734–1762. https://doi.org/10.1111/1460-6984.13028

Stark, B. C., Dalton, S. G., & Lanzi, A. M. (2025). Access to context-specific lexical-semantic information during discourse tasks differentiates speakers with latent aphasia, mild cognitive impairment, and cognitively healthy adults. *Frontiers in Human Neuroscience*, 18. https://doi.org/10.3389/fnhum.2024.1500735

Stark, B. C., Dutta, M., Murray, L. L., Fromm, D., Bryant, L., Harmon, T. G., & Roberts, A. C. (2021). Spoken discourse assessment and analysis in aphasia: An international survey of current practices. *Journal of Speech, Language, & Hearing Research*, *64*(11), 4366–4389. ht tps://d oi-org.proxy.lib.pdx.edu/1 0.1044/2021_JSLHR-20-00708

Ulatowska, H. K., Olness, G. S., & Williams, L. J. (2004). Coherence of narratives in aphasia. *Brain and Language*, *91*(1), 42–43.

Wallace, G. L. (2010). Profile of life participation after stroke and aphasia. *Topics in Stroke Rehabilitation*, *17*(6), 432–450. https://doi.org/10.1310/tsr1706-432

Whitworth, A., Leitão, S., Cartwright, J., Webster, J., Hankey, G. J., Zach, J., Howard, D., & Wolz, V. (2015). Narnia: A new twist to an old tale. A pilot RCT to evaluate a multilevel approach to improving discourse in aphasia. *Aphasiology*, *29*(11), 1345–1382. https://doi.org/10.1080/02687038.2015.1081143

Wilson, S. M., Eriksson, D. K., Schneck, S. M., & Lucanie, J. M. (2018). A quick aphasia battery for efficient, reliable, and multidimensional assessment of language function. *PLOS ONE*, *13*(2), e0192773. https://doi.org/10.1371/journal.pone.0192773

Wright, H. H., & Capilouto, G. J. (2009). Manipulating task instructions to change narrative discourse performance. *Aphasiology*, *23*(10), 1295–1308. https://doi.org/10.1080/02687030902826844

Wright, H. H., & Capilouto, G. J. (2012). Considering a multi-level approach to understanding maintenance of global coherence in adults with aphasia. *Aphasiology*, *26*(5), 656–672. https://doi.org/10.1080/02687038.2012.676855

Wright, H. H., & Capilouto, G. J. (2016). Considering a multi-level approach to understanding maintenance of global coherence in adults with aphasia. In C. Code & D. Freed (Eds.), *The science of aphasia rehabilitation* (pp. 42–58). Routledge.

Wright, H. H., Capilouto, G. J., & Koutsoftas, A. (2013). Evaluating measures of global coherence ability in stories in adults. *International Journal of Language & Communication Disorders*, *48*(3), 249–256. https://doi.org/10.1111/1460-6984.12000

Wright, H. H., Koutsoftas, A., Fergadiotis, G., & Capilouto, G. (2010). Coherence in stories told by adults with aphasia. *Procedia - Social and Behavioral Sciences*, *6*, 111–112. https://doi.org/10.1016/j.sbspro.2010.08.056