

Transcription-less analysis of five discourse tasks in Laurentian French persons with post-stroke aphasia: adaptation and reliability

by Amélie Brisebois | Véronique Desjardins | Eva Marois | Julie Bélanger | Simona Maria Brambati | Karine Marcotte | Université de Montréal | Université de Montréal | Université de Montréal | Centre de recherche du CIUSSS du Nord-de-l'île-de-Montréal | Université de Montréal | Université de Montréal

Abstract ID: 206

Event: SoA 2025 Copenhagen posters

Topic: Clinical and experimental work on aphasia and related disorders

Introduction and aims

Discourse analysis is essential in aphasia assessment and outcome measurement, yet transcription-based methods are time-consuming and resource-intensive. Transcription-less approaches such as Main Concept Analysis (MCA; Nicholas & Brookshire, 1993), thematic Units (TUs; Brisebois et al., 2020), and coherence/reference (Ulatowska et al., 2003) offer feasible alternatives. MCA integrates lexical and propositional content, while coherence/reference and TUs assess discourse structure and informativeness.

Despite promising use, their psychometric properties remain underexplored in Laurentian French. Guided by the LUNA framework (Dipper et al., 2021), which emphasizes multi-level discourse analysis, this study aimed to: (1) evaluate intra- and inter-rater and test-retest reliability of these transcription-less measures across five tasks and (2) assess construct validity. Secondary goals included the reliability of microstructural (i.e., linguistic) variables and linking measures across discourse levels.

Methods

As part of a larger ethics-approved study (CIUSSS NIM # 2020-1900), twenty-three French-speaking adults (age: $x = 58.2$, $SD = 12.8$ years old ; 13 females) with chronic post-stroke aphasia were assessed twice (mean interval = 11.7 days, $SD = 6.2$). Five monologic discourse tasks were administered: the WAB-R picture description (Kertesz, 2006), Cinderella storytelling using the administration method of AphasiaBank (MacWhinney et al., 2011) and (3) three narrative sequences from AphasiaBank (Broken Window, Refused Umbrella, and Cat Rescue).

Transcription-less measures included adapted main concepts (MCs), TUs (WAB-R only), and coherence/reference scoring. Microstructural variables (e.g., mean length of utterance (MLU), correct information units (CIUs)) were extracted from CLAN transcriptions. Raters scored 20% of samples twice (intra-rater) and cross-coded for inter-rater reliability. Reliability was evaluated using intraclass correlation coefficients (ICCs), Spearman's rho,

and Wilcoxon tests (not reported). Construct validity was assessed through correlations between macrostructural, propositional and linguistic variables.

Results

Intra- and inter-rater reliability was between good and excellent for most measures of each discourse tasks. It was poor to moderate for some main concepts with lower occurrences such as accurate incomplete.

For the transcription-less measures of the combined tasks, MCs yielded an ICC of .799 (95% CI [.582, .909]) and Spearman's ρ of .755 ($p < .001$), while the coherence/reference score demonstrated excellent reliability (ICC = .923, 95% CI [.828, .967], Spearman's $\rho = .911$, $p < .001$). The microstructural variables with the highest test-retest reliability were total words (ICC = .934, 95% CI [.852, .972], Spearman's $\rho = .835$, $p < .001$) and CIUs (ICC = .937, 95% CI [.857, .973], Spearman's $\rho = .811$, $p < .001$).

For the Cinderella retell task, test-retest reliability was excellent for MCs (ICC = .910, 95% CI [.787, .963], Spearman's $\rho = .903$, $p < .001$) and for absent MCs (ICC = .921, 95% CI [.812, .968], Spearman's $\rho = .708$, $p < .001$). The microstructural measures presenting the strongest test-retest reliability were: total words (ICC = .926, 95% CI [.824, .970], Spearman's $\rho = .880$, $p < .001$), words per minute (ICC = .908, 95% CI [.782, .962], Spearman's $\rho = .890$, $p < .001$), and CIUs (ICC = .919, 95% CI [.806, .967], Spearman's $\rho = .889$, $p < .001$).

For the WAB-R picture description, the coherence/reference scores yielded excellent reliability (ICC = .967, 95% CI [.925, .986], Spearman's $\rho = .887$, $p < .001$) whereas TUs showed good reliability (ICC = .791, 95% CI [.569, .906], Spearman's $\rho = .733$, $p < .001$). Among the microstructures variables, the highest reliability was found with MLU (ICC = .960, 95% CI [.909, .983], Spearman's $\rho = .825$, $p < .001$), total words (ICC = .953, 95% CI [.893, .980], Spearman's $\rho = .818$, $p < .001$) and CIUs (ICC = .941, 95% CI [.867, .975], Spearman's $\rho = .879$, $p < .001$).

Regarding construct validity, coherence/reference scores were strongly correlated with main concept scores across most tasks and time points. For example, at Time 2, the correlation between these variables reached Spearman's $\rho = .907$ ($p < .001$, p adjusted $< .001$).

Discussion

This study provides the first comprehensive psychometric evaluation of transcription-less macrostructural and propositional discourse measures in Laurentian French. MCs, TUs, and

coherence/reference demonstrated good to excellent intra- and inter-rater reliability across most tasks, supporting their feasibility for clinical and research use in chronic aphasia.

Among transcription-less variables, coherence/reference scores showed excellent reliability for the picture description of the WAB-R and the narrative sequences from AphasiaBank, and strong correlations with MCs. This relationship highlights how macrostructural discourse organization is closely tied to lexical and propositional informativeness, aligning with multilevel discourse models (e.g., Marini et al., 2011), and also support the construct validity of the measures under the LUNA framework (Dipper et al., 2021)

Microstructural variables such as CIUs and total words demonstrated excellent test-retest reliability, especially for the WAB-R and Cinderella tasks, consistent with previous findings (Stark et al., 2023). Although some variability was observed by task and aphasia severity, overall reliability levels support their use in longitudinal designs.

The use of combined narrative sequences (e.g., Broken Window, Refused Umbrella, Cat Rescue) and the Cinderella story retell yielded reliable scores for both coherence and main concept analysis. These tasks offer flexible yet psychometrically sound options for capturing discourse change over time. In contrast, while the WAB-R picture description limits the application of MCA, it provides efficient and highly reliable scores coherence/reference, Thematic Units and CIUs.

Taken together, the results validate transcription-less approaches for capturing key dimensions of discourse in aphasia. The development of culturally adapted discourse psychometric data further support their integration into clinical workflows as these three protocols are quickly administered. Overall, this work enhances the accessibility and psychometric rigor of discourse assessment tools for French-speaking PWA and enables more consistent tracking of recovery and treatment outcomes.

References

Brisebois, A., et al. (2020). The importance of thematic informativeness in narrative discourse recovery in acute post-stroke aphasia. *Aphasiology*, 34(4). <https://doi.org/10.1080/02687038.2019.1705661>

Dipper, L., et al. (2021). Creating a Theoretical Framework to Underpin Discourse Assessment and Intervention in Aphasia. *Brain Sciences*, 11(2), 183. <https://doi.org/10.3390/brainsci11020183>

Kertesz, A. (2006). *Western Aphasia Battery- Revised*. Pearson.

MacWhinney, B., Fromm, D., Forbes, M., & Holland, A. (2011). Aphasiabank: Methods for studying discourse. *Aphasiology*, 25(11), 1286-1307.

Marini, A., et al. (2011). Narrative language in traumatic brain injury. *Neuropsychologia*, 49(10), 2904-2910.

Nicholas, L. E., & Brookshire, R. H. (1993, August 8). *A System for Scoring Main Concepts in the Discourse of Non-brain-damaged and Aphasic Speakers*.

Stark, B. C., et al. (2023). Test-Retest Reliability of Microlinguistic Information Derived From Spoken Discourse in Persons With Chronic Aphasia. *Journal of Speech, Language, and Hearing Research*, 1-30.

Ulatowska, H., et al. (2003). Relationship between discourse and Western Aphasia Battery performance in African Americans with aphasia. *Aphasiology*, 17(5), 511-521.