ASHA American Speech-Language-Hearing Association

AJSLP

Research Article

# Discourse Task Type–Specific Linguistic Characteristics in Anomic Aphasia and Healthy Controls: Evidence From Mandarin–Chinese AphasiaBank

**Bao-Mei Deng,[a] Jing Gao,[b,c] Li-Si Liang,[a] Jia-Xin Zhao,[a] Feng Lin,[b] Ming-Yu Yin,[a] Hai-Qing Zheng,[a] and Xi-Quan Hu[a]**

[a] Department of Rehabilitation Medicine, The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou, Guangdong, China [b] School of Rehabilitation Medicine, Nanjing Medical University, Jiangsu, China [c] Sir Run Run Hospital, Nanjing Medical University, Jiangsu, China

ARTICLE INFO

Article History:
Received March 1, 2023
Revision received May 22, 2023
Accepted November 28, 2023

Editor-in-Chief: Katherine C. Hustad
Editor: Jacqueline Laures-Gore

https://doi.org/10.1044/2023_AJSLP-23-00078

ABSTRACT

**Purpose:** This study was designed to examine the hypothesis that discourse task types influence language performance in Mandarin Chinese–speaking people and to reveal the discourse task-specific linguistic properties of persons with anomic aphasia compared to neurotypical controls.
**Method:** Language samples from persons with aphasia ($n$ = 31) and age- and education-matched controls ($n$ = 31) across four discourse tasks (sequential-picture description, single-picture description, story narrative, and procedural discourse) were collected from Mandarin AphasiaBank. Task-specific distributions of parts of speech were analyzed using mosaic plots. The main effects of tasks in each group and the between-group differences within each task for several typical linguistic variables were evaluated, including the mean length of utterance, tokens, moving-average type-token ratio, words per minute, propositional density, noun–verb ratio, noun percentage, and verb percentage.
**Results:** The results revealed an impact of discourse tasks on most language variables in both groups. In the healthy controls, story narratives yielded the highest total words and lowest verb percentage. In the aphasia group, procedural discourse elicited the fewest total words and densest expressions, whereas their single-picture descriptions had the highest noun–verb ratio. For all tasks, the aphasia group performed worse than the control group in the mean length of utterance, tokens, moving-average type-token ratio, and words per minute. For noun–verb ratio, noun percentage, and verb percentage, only one task (i.e., single-picture description) showed significant between-group differences.
**Conclusion:** The selection of discourse tasks should be addressed in assessments and interventions for Mandarin Chinese–speaking individuals with aphasia to obtain more accurate and feasible outcomes.

Deficits in spoken discourse negatively affect the social interactions and quality of life of a person with aphasia (Wallace et al., 2017). In recent years, discourse has attracted substantial attention as an essential component of clinical practice and research (Bryant, Ferguson, & Spencer, 2016; Cruice et al., 2020; Dipper et al., 2021; Dipper & Pritchard, 2017; Kintz & Wright, 2017). Discourse generally refers to a unit of language beyond the level of a sentence and provides a flow of information more extensively than the sum of its parts (Armstrong, 2000; Bryant, Spencer, & Ferguson, 2016; Dipper & Pritchard, 2017). Discourse assessment offers a valid method for assessing the daily verbal skills of a person with aphasia (Bryant, Spencer, &

Correspondence to Xi-Quan Hu: huxiquan@mail.sysu.edu.cn. **Disclosure:** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

Ferguson, 2016). Traditional standardized aphasia batteries typically evaluate the linguistic domains in isolation and overlook language functions above the sentence level and complex cognitive-linguistic behaviors, failing to detect the subtle language deficits of some people with aphasia and identify overall communicative performance (Marini et al., 2011). Discourse assessment complements the evaluation of language and language-related cognitive processes at the discourse level, allowing for the identification of clinically significant residual impairments and the actual language performance under more natural conditions (Linnik et al., 2015).

However, the results of discourse production research are mixed, given the heterogeneity of discourse tasks (Stark, Dutta, Murray, Bryant, et al., 2021; Stark, Dutta, Murray, Fromm, et al., 2021). To collect spoken discourse samples, a variety of elicitation tasks have been used (Bryant, Ferguson, & Spencer, 2016), such as single-picture (Berube et al., 2022; Hameister & Nickels, 2018; Marcotte et al., 2022) or sequential-picture descriptions (Dalton & Richardson, 2015; Kong et al., 2016), story-retelling (Kong & Wong, 2018; Litovsky et al., 2022; Richardson et al., 2018, 2021), personal recounts (Law et al., 2018; Saling et al., 2017), and procedural discourse (Pritchard et al., 2015; Ryan et al., 2010; Stubbs et al., 2018). Nevertheless, recent evidence has revealed that different discourse tasks may impose distinct cognitive and linguistic demands, leading to variations in linguistic features such as lexical diversity (Fergadiotis & Wright, 2011; Fergadiotis et al., 2011), speech rate (Alyahya et al., 2020; Sahraoui & Nespoulous, 2012), word class (Stark & Fukuyama, 2020), and coherence (Alyahya et al., 2022; Olness, 2006; Wright et al., 2014). Procedural discourse, for instance, was found to be associated with the least lexical diversity in cognitively healthy adults, compared with recounts, event casts, and story narratives (Fergadiotis et al., 2011). Personal recounts were reported to have lower scores for global coherence in cognitively healthy individuals than procedural discourse, story narrative, and picture description (Wright et al., 2014). Alyahya et al. (2020) observed that picture-supported story narratives tended to elicit a higher quantity and lexical diversity of content words than composite picture descriptions and unsupported procedural discourse in both persons with aphasia and healthy adults. Stark (2019) analyzed a series of linguistic variables across four elicitation tasks, revealing that narrative discourse generated the greatest propositional density, whereas procedural discourse yielded the simplest syntax of the sentence. Stark and Fukuyama (2020) further demonstrated the importance of choosing the most sensitive discourse task to examine the specific aspects of linguistic abilities. These studies suggested that elicitation methods influence various linguistic properties and patterns produced during connected speech. Therefore, exploring the correlation between discourse tasks and language performance will aid in the standardization of discourse assessment administration, facilitating its large-scale implementation in medical settings, and providing guidance regarding discourse treatment in aphasia rehabilitation. Language performance refers to the actual use of language in concrete situations, including oral expression, understanding of oral language, reading, and writing. We focus on performance in oral expression, which can be calculated by the measures related to language productivity (e.g., total words, lexical diversity), information content (e.g., cohesion, semantic), and grammatical complexity (e.g., word class, syntactic; Bryant, Ferguson, & Spencer, 2016).

In addition to the heterogeneity of discourse tasks across studies, heterogeneity across participant characteristics is another factor possibly making comparisons between studies difficult. Although individuals with different aphasia types often exhibit considerable variation in language ability, many previous studies investigating discourse in aphasia have combined all speakers with aphasia into a single group (Alyahya et al., 2020; Stark, 2019). Studies have shown that aphasia type may serve as a moderator, contextualizing the effect of discourse tasks on language output (Dalton & Richardson, 2019; Fromm et al., 2016; Stark & Fukuyama, 2020). Hence, in the current study, we focused on individuals with anomic aphasia, the mildest type of aphasia, who were expected to produce discourse that could be more easily identified compared to other types of aphasia. This allowed us to acquire a relatively homogeneous group with a sufficient sample size to enable adequate and reliable analysis in relative isolation from confounding factors.

Although research has shown the impact of elicitation tasks on the breadth of linguistic features captured, it has mainly focused on English speakers (Alyahya et al., 2020; Sahraoui & Nespoulous, 2012; Stark, 2019; Stark & Fukuyama, 2020). A dearth of investigation remains for other languages, such as Mandarin. The profiling of linguistic features is implicitly linked to a specific language structure. Mandarin has linguistic characteristics that are drastically different from English, such as little inflectional morphology; no person or tense marking on verbs; and no number, gender, or case agreements between a noun and its modifier. Hence, the profile of impairment in one language cannot be simply applied to another (Bates et al., 1991). This study focused on the effects of different discourse task types on language production, specifically in Mandarin.

AphasiaBank was established to provide researchers and clinicians with an extensive, shared multimedia database of uniform discourse samples from individuals with and without aphasia (Forbes et al., 2012; Fromm et al., 2020; MacWhinney, 2019; MacWhinney & Fromm, 2016). Researchers and clinicians can use a standardized protocol to collect and analyze language samples. Currently, the AphasiaBank database contains a range of languages, such as English, German, Japanese, and Cantonese. Mandarin AphasiaBank has recently been developed (Jiang et al., 2023), providing a database for studies of connected speech in Mandarin-speaking persons with aphasia (https://aphasia.talkbank.org/access/Protocol.html). Using data from the Mandarin AphasiaBank, the current study had two primary aims: (a) to explore the effect of discourse task type on measures of oral expressive language in Mandarin-speaking persons with anomic aphasia and healthy adults and (b) to compare performance on measures of oral expressive language between anomic aphasia and healthy control groups within the same discourse type. We hypothesized that discourse task types influence language performance in both groups and that the aphasia group generally performed poorly compared to the control group on the majority of linguistic variables. These findings are expected to contribute significantly to discourse research on Mandarin-speaking persons with aphasia. The findings may facilitate a better understanding of the discourse task-specific microlinguistic characteristics of persons with anomic aphasia and guide clinicians in selecting appropriate discourse tasks when conducting discourse assessments and interventions.

## Method

### *Participants*

Language samples from persons with anomic aphasia and healthy individuals were extracted from the Mandarin AphasiaBank. The aphasia group consisted of participants with anomic aphasia who developed aphasia resulting from stroke, as verified through neuroimaging or a clear medical diagnosis. Anomic aphasia was classified using standard behavioral criteria according to the Chinese version of Western Aphasia Battery–Revised (Wang, 1997). Participants were native Mandarin speakers with normal or corrected-to-normal vision and hearing. Exclusion criteria were a history of dementia or comorbidities associated with serious cognitive consequences (such as brain injury). As we focused on language production instead of speech (i.e., motor) production in this study, we excluded persons with aphasia with moderate or severe dysarthria or apraxia of speech. Dysarthria was evaluated using the Frenchay Dysarthria Assessment (Enderby & Palmer, 2008) adapted for Chinese speakers, while apraxia of speech was assessed using the assessment standards for apraxia of speech from China Rehabilitation Research Center (Wong et al., 2023; You et al., 2019). Finally, the aphasia group included a total of 31 persons with anomic aphasia.

The control group consisted of 31 age- and education-matched healthy individuals who provided the discourse samples. None of them had any neurological conditions (e.g., stroke or head injury), history of neurodegenerative diseases (e.g., Alzheimer's disease or Parkinson's disease), hearing and/or visual impairments, or depression at the time of testing. All the participants were native Mandarin speakers.

The demographic information and statistical comparisons between the two groups are presented in Table 1. This study was approved by the ethical committee of The Third Affiliated Hospital of Sun Yat-sen University (approval number: [2019]02–585-01). Prior to participation in the Mandarin AphasiaBank protocol all participants signed an informed consent.

### *Discourse Elicitation*

The elicitation of language samples followed the Mandarin AphasiaBank protocol (https://aphasia.talkbank.org/protocol/languages/Mandarin/). Four different tasks were

**Table 1.** Demographic information of the aphasia and control groups.

| Group | Aphasia ($n$ = 31) | Control ($n$ = 31) | Comparisons |
|---|---|---|---|
| Age (years) | 42.26 (13.01) | 42.32 (12.92) | $t$ (60) = −0.020, $p$ = .984 |
| Education (years) | 15 (11–16) | 15 (12–16) | $U$ = 467.000, $p$ = .847 |
| Gender (male/female) | 22/9 | 17/14 | $\chi^2$ (1) = 1.728, $p$ = .189 |
| Time postonset (months) | 6.00 (2.87–9.00) | — | — |
| AQ | 86.90 (80.50–90.30) | — | — |
| Average duration (s) | 70.00 (45.50–108.50) | 55.00 (39.25–83.50) | $U$ = 5,987.000, $p$ = .003* |
| Average tokens | 61.00 (37.25–97.50) | 107.00 (74.25–171.00) | $U$ = 3,855.000, $p$ < .001** |

*Note.* Values shown are means (standard deviation) and median (interquartile range), unless stated otherwise. Statistical comparisons are two-tailed. Em dash indicates data not applicable. $t$ = independent-samples $t$ test; $U$ = Mann–Whitney $U$ test; $\chi^2$ = Pearson chi-square test; AQ = Aphasia Quotient from the Western Aphasia Battery–Revised.

*Significant effect at $p$ < .05. **Significant effect at $p$ < .001.

selected from the database in this study: (a) the sequential-picture description task of "Refused Umbrella," (b) the single-picture description task of "Cat Rescue," (c) the story narrative task of "The Tortoise and the Hare," and (d) the procedural discourse task of making ham and egg fried rice. Discourse task instructions and pictures are openly available in the Mandarin protocol section on the AphasiaBank website (https://aphasia.talkbank.org/protocol/languages/Mandarin/discourse_Ttsks.doc).

The sequential-picture description instructed participants to describe a story based on a series of line-drawing pictures presented in sequential order. This task required participants to create a story using the provided visual stimuli and deliver it logically and coherently. The "Refused Umbrella" stimuli used in this study portray the story of a student rejecting the umbrella that his mother offered him as he departed for school, only to be met by rain on his way to school and having to return home to take the umbrella.

The single-picture description involved the presentation of a scene in which various animate and inanimate objects, implying a timeline of events, were combined to create a central theme. The elements of an image interact with each other to create a dynamic and engaging visual story. Here, "Cat Rescue" refers to a young girl's struggle to save her cat from a tree; her father attempts to save it but is left dangling, so the fire department has to be called in to assist in the rescue.

Regarding the story narrative, participants were initially allowed to view sequential pictures as long as they wished. After the pictures were taken away, they were asked to produce a narrative in their own words. In this study, the tortoise and the hare was used. These stimuli depict the fable that a tortoise and a hare race to determine who is faster. The hare dozes off along the way, while the tortoise keeps moving forward despite its slow pace and eventually emerges victorious.

In the procedural discourse, the participants were asked to explain the procedures necessary to complete the activity. Photos and written texts of the components were provided when the participant was unable to respond. The task (ham and egg fried rice) employed in this study required participants to provide a step-by-step description of how to prepare ham and egg fried rice.

The examiner only provided nonverbal encouragement throughout the tests, without offering any prompts or queries. Language samples were recorded using a video recorder for further analysis.

## *Transcription, Coding, and Reliability*

Video-recorded language samples were transcribed orthographically and coded by experienced trained transcribers in Codes for the Human Analysis of Transcripts (CHAT) format (Macwhinney et al., 2011). CHAT is designed to operate closely with the computerized language analysis (CLAN) program, a tool designed to facilitate the automatic computation of indices in language interactions (MacWhinney, 2000). The detailed instructions of transcription and coding procedures can be available at the AphasiaBank website (http://aphasia.talkbank.org).

To obtain inter- and intrarater reliability of transcription and coding, eight of the language samples from each group (i.e., approximately 25.8% of the data for the control and aphasia group, respectively) were randomly selected and retranscribed and recoded by the first author and another experienced trained rater. These 16 transcripts were retranscribed and recoded by the first author to assess the intrarater reliability. The point-to-point agreement, calculated by the formula [total agreements/(total agreements + total disagreements) × 100%], was used to test reliability.

## *Linguistic Measures*

The EVAL, FREQ, and MORTABLE programs in CLAN were used to calculate linguistic measures. The EVAL command was applied to yield a composite discourse profile, including metrics such as mean length of utterance (MLU), words per minute (WPM), and propositional density. The FREQ program was used to calculate lexical diversity using the moving-average type-token ratio (MATTR). Unlike TTR, this index has the benefit of being invariant to the length of the text and computes TTRs for every consecutive window with a fixed-length segment. The MORTABLE program was used to generate a frequency and percentage chart of parts of speech (POS), such as adjectives, adverbs, nouns, and verbs. These programs produced an array of analytical results, which were automatically exported to an Excel spreadsheet. Detailed instructions on how to execute the commands in CLAN and the codes used for analysis are available at AphasiaBank (https://talkbank.org/manuals/Clin-CLAN.pdf).

To capture the multidimensional nature of the discourse, variables that touched on language quantity, quality, fluency, syntactic complexity, and informativeness were calculated. The following measures were employed to reflect the range of functional discourse skills of each speaker.

*MLU.* The MLU was defined as the average number of words produced per sentence, excluding morphemes in utterances with xxx, yyy, or www codes. This captures the structural and syntactic aspects of language production.

*Tokens.* Tokens refer to the total number of words produced by the participants, excluding repetitions and revisions. This was counted as a measure of production quantity and gross output.

*MATTR.* The MATTR was implemented in this study to gauge lexical diversity because it has recently been confirmed to be a reliable, valid, and unbiased measure of vocabulary richness in aphasia (Cunningham & Haley, 2020; Fergadiotis et al., 2013). Similar to TTR, it calculates the ratio of different words (types) to total words (tokens) in the sample but has the advantage of eliminating the effect of text length. We selected an analysis window of ten words in this study, as it was equal to the minimum number of words for a discourse task in the participant sample.

*WPM.* WPM was obtained by dividing the total number of words by the total duration of the sample. It was coded as a measure of the speech rate.

*Propositional density.* Propositional density was derived using the EVAL command and was approximated by the number of verbs, adjectives, adverbs, prepositions, and conjunctions divided by the total number of words. It is calculated as a measure of information content.

*Noun–verb ratio.* This was calculated by dividing the total number of nouns by the total number of verbs (excluding auxiliaries and modals). This index reflects syntactic complexity. A higher noun–verb ratio indicates a simpler syntactic complexity, whereas a lower noun–verb ratio denotes a more complicated syntactic complexity.

*Noun percentage.* This was computed by dividing the number of nouns by the total number of words, resulting in a proportion. This measure indicates whether greater or fewer nouns were produced. A higher noun percentage indicates a larger production of nouns, while a lower noun percentage implies a reduced retrieval of nouns.

*Verb percentage.* The percentage of verbs produced in each sample was calculated by dividing the total number of verbs by the total number of words. Similar to the noun percentage, this measure indicates whether greater or fewer verbs were produced. A higher verb percentage indicates a larger production of verbs, while a lower verb percentage implies a reduced retrieval of verbs.

## Data Analysis

Data analysis was performed using SPSS software version 27. The Shapiro–Wilk normality test was used to determine if a variable was normally distributed. Variables were normally distributed if the Shapiro–Wilk test $p$ values were greater than 0.05. Continuous variables with a normal distribution are presented as mean (standard deviation), and nonnormal variables are reported as median (interquartile range).

To determine the influence of the discourse task on linguistic components in both the control and aphasia groups, repeated-measures analyses of variance were conducted for data that followed a normal distribution, and Friedman tests were employed for nonnormally distributed variables. Post hoc tests were performed if there was a significant main effect of the discourse task. $P$ values were adjusted for multiple comparisons using the Bonferroni correction. R 4.2.0 (R Core Team, 2022) was used to construct mosaic plots to visualize the frequency distribution of POS.

To compare the linguistic characteristics of the aphasia and control groups within each discourse task, independent-samples $t$ tests or Mann–Whitney $U$ tests were performed, depending on whether the data were normally distributed. The Benjamini–Hochberg correction was applied to obtain a more stringent critical value to reduce the risk of type I error. The significance level was set at $p < 0.05$.

## Results

### Demographic Differences Between Groups

Demographic information of the aphasia and control groups is shown in Table 1. The two groups did not differ significantly in terms of age, $t$ (60) = −0.020, $p$ = .984; years of education ($U$ = 467.00, $p$ = .847); or sex ratio, $\chi^2$ (1) = 1.728, $p$ = .189. The discourse duration of the aphasia group was found to be significantly longer than that of the control group ($U$ = 5987.000, $p$ = .003). The tokens of the aphasia group were significantly lower compared to the control group ($U$ = 3855.000, $p$ < .001).

### Reliability on Transcription and Coding

For the control group, inter- and intrarater agreements for transcription were 95% and 97%, respectively; inter- and intrarater agreements for coding were 94% and 95%, respectively. For the aphasia group, inter- and intrarater agreements for transcription were 93% and 94%, respectively; inter- and intrarater agreements for coding were 91% and 93%, respectively.

### The Effect of the Discourse Task in the Control Group

The results indicated a significant effect of the discourse task for nearly all linguistic measures in the control group except for the noun–verb ratio (see Table 2). Post

**Table 2.** The within-subject effects of the discourse task in different groups.

| Linguistic variable | Aphasia (*n* = 31) | | Control (*n* = 31) | |
|---|---|---|---|---|
| | Main effect of discourse task | Significant post hoc tests | Main effect of discourse task | Significant post hoc tests |
| MLU | $^b\chi^2$ (3) = 10.084, | Sequential > Procedural, *p* = .026* | $^b\chi^2$ (3) = 13.026, | Single > Procedural, *p* = .035* |
| | *p* = .018* | | *p* = .005* | Story > Procedural, *p* = .007* |
| Tokens | $^b\chi^2$ (3) = 30.910, | Sequential > Procedural, *p* = .003* | $^b\chi^2$ (3) = 46.779, | Story > Procedural, *p* < .001** |
| | *p* < .001** | Single > Procedural, *p* < .001** | *p* < .001** | Story > Sequential, *p* < .001** |
| | | Story > Procedural, *p* < .001** | | Story > Single, *p* < .001** |
| MATTR | $^b\chi^2$ (3) = 6.600, | | $^b\chi^2$ (3) = 12.382, | Sequential > Procedural, *p* = .004* |
| | *p* = .086 | | *p* = .006* | |
| WPM | $^b\chi^2$ (3) = 11.893, | Story > Single, *p* = .008* | $^aF$ (7.000, 71.759) = 8.727, | Sequential > Single, *p* = .022* |
| | *p* = .008* | | *p* < .001** | Story > Single, *p* < .001** |
| | | | | Story > Procedural, *p* = .001* |
| Propositional density | $^b\chi^2$ (3) = 38.419, | Story > Single, *p* = .035* | $^aF$ (3.000,90) = 20.835, | Story > Sequential, *p* = .004* |
| | *p* < .001** | Procedural > Single, *p* < .001** | *p* < .001** | Procedural > Sequential, *p* < .001** |
| | | Procedural > Sequential, *p* < .001** | | Story > Single, *p* < .001** |
| | | Procedural > Story, *p* = .007* | | Procedural > Single, *p* < .001** |
| Noun–verb ratio | $^b\chi^2$ (3) = 16.476, | Single > Story, *p* = .003* | $^b\chi^2$ (3) = 3.480, | |
| | *p* < .001** | Single > Procedural, *p* = .003* | *p* = .323 | |
| | | Single > Sequential, *p* = .047* | | |
| Noun percentage | $^b\chi^2$ (3) = 8.592, | Single > Story, *p* = .026* | $^b\chi^2$ (3) = 11.477, | Sequential > Story, *p* = .010* |
| | *p* = .035* | | *p* = .009* | Sequential > Procedural, *p* = .047* |
| Verb percentage | $^b\chi^2$ (3) = 12.561, | Sequential > Single, *p* = .026* | $^b\chi^2$ (3) = 17.594, | Single > Story, *p* = .026* |
| | *p* = .006* | Procedural > Single, *p* = .007* | *p* < .001** | Procedural > Story, *p* = .003* |
| | | | | Sequential > Story, *p* = .001* |

*Note.* Repeated-measures analysis of variance (ANOVA) was performed for normally distributed values, and the Friedman test was used for nonnormally distributed values. Nonsphericity was corrected using Greenhouse–Geisser correction when necessary. Post hoc tests were performed if there was a significant main effect of discourse tasks. *P* values were corrected using Bonferroni correction for multiple comparisons. MLU = mean length of utterance; Sequential = sequential-picture description; Procedural = procedural discourse; Single = single-picture description; Story = story narrative; MATTR = moving-average type-token ratio; WPM = words per minute.

[a]Repeated-measures ANOVA. [b]Friedman test.

*Significant effect at *p* < .05; **Significant effect at *p* < .001.

hoc tests showed that the discourse task influenced spoken language across linguistic properties (see Table 2 for significant post hoc test results and Figure 1 for a general view of the comparisons). Regarding the quantity of language production (tokens), story narrative induced the greatest verbal output among the four discourse tasks. Regarding propositional density, story narrative and procedural discourse were denser than single- and sequential-picture descriptions. In addition, the sequential-picture description elicited a higher noun percentage than the story narrative (*p* = .010) and procedural discourse (*p* = .047). Moreover, the lowest percentage of verbs was evoked by story narratives among all the discourse tasks.

Discourse task-specific distributions of POS were presented using mosaic plots (see Figure 2). The area in a cell indicates the actual frequency of each POS. The larger the area of the embedded rectangle, the higher the output of the POS. The colors in the mosaic plot represent the outcomes of the comparison between the actual and anticipated frequencies using Pearson's residual test for each rectangle. As shown in Figure 2B, the story narrative elicited the highest total verbal output across all discourse tasks in the control group. Furthermore, each discourse task influenced specific POS distribution patterns. Some POS may have a bias in the corresponding discourse task. For instance, a high proportion of nouns was present in the sequential-picture description. The adjective and adverb proportions were overrepresented in the story narrative but underrepresented in the picture descriptions (both sequential and single pictures).

## The Effect of the Discourse Task in the Aphasia Group

The discourse task also significantly influenced language performance in the aphasia group, except for MATTR (see Table 2 for significant post hoc test results

**Figure 1.** Comparison of linguistic variables across four discourse tasks within each group (aphasia, control). The center lines within the boxplots represent the median; boxplot bounds represent the interquartile range; whiskers refer to minimum and maximum values. Asterisks indicate a significant difference (*$p < 0.05$, **$p < 0.001$). MLU = mean length of utterance; MATTR = moving-average type-token ratio; WPM = words per minute; Sequential = sequential-picture description; Single = single-picture description; Story = story narrative; Procedural = procedural discourse.

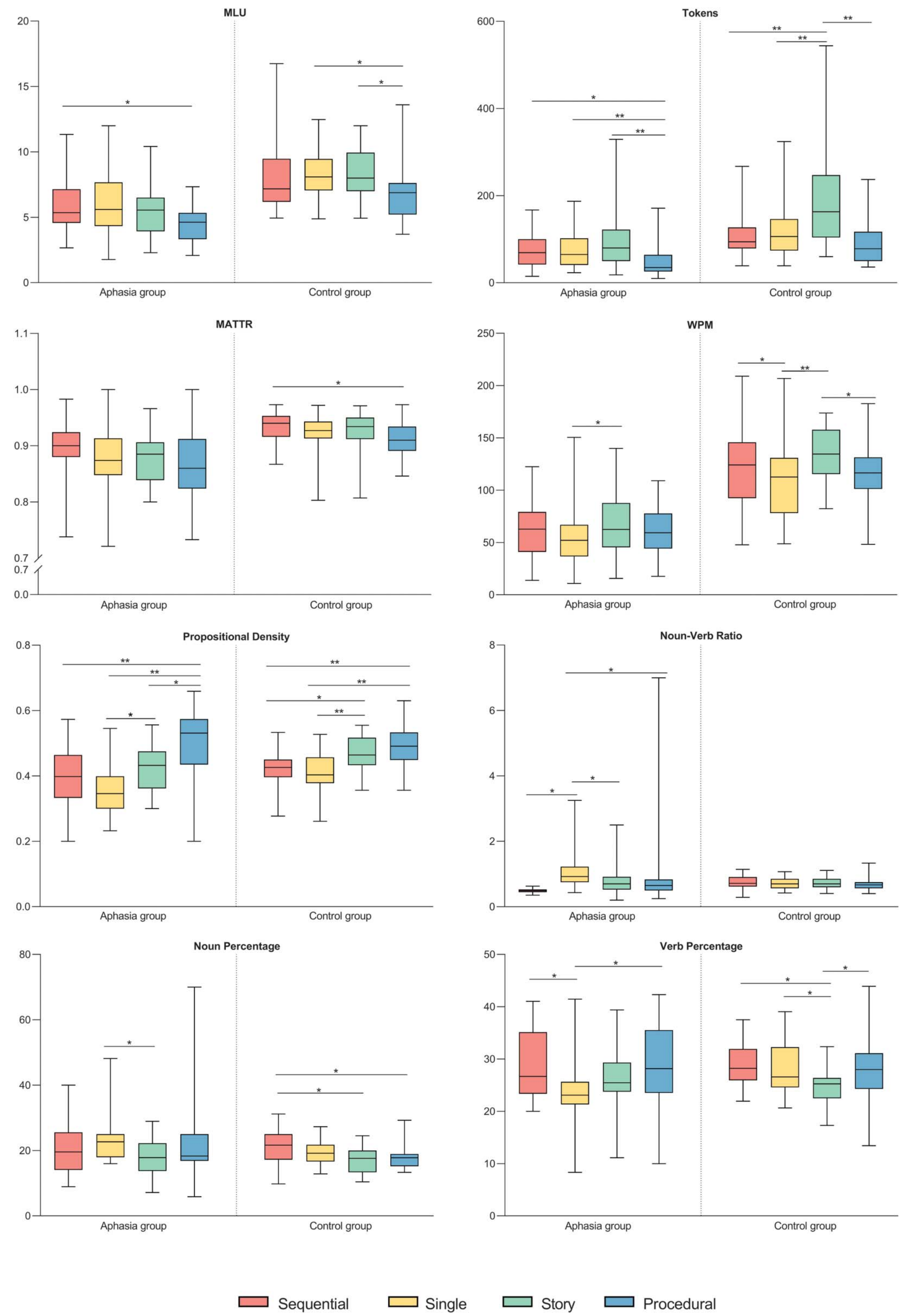
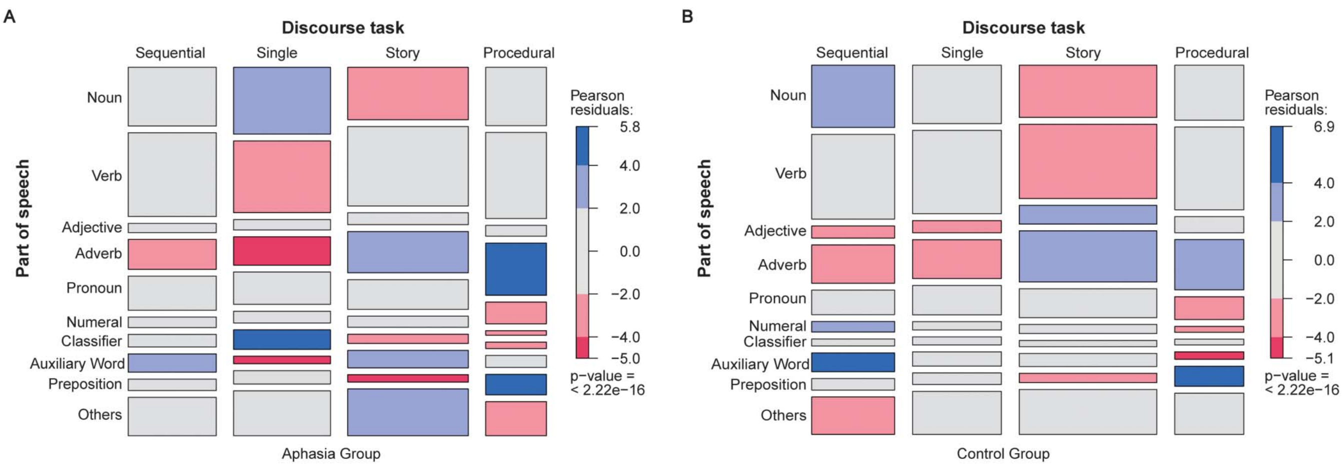


**Figure 2.** Mosaic displays for the overall frequency distribution of parts of speech in the (A) aphasia group and (B) control group across four discourse tasks (Sequential = sequential-picture description; Single = single-picture description; Story = story narrative; Procedural = procedural discourse). The size of a cell represents the frequency of the part of speech. The residual shading was based on the null model of frequency distribution being independent of discourse tasks. The size of residuals determines the lightness of a cell: very dark color for large residuals (> 4, corresponding to α = .0001), lighter color for medium-sized residuals (< 4 and > 2), gray for small residuals (< 2, corresponding to α = .05). Blue colors indicate significantly higher than expected, while red indicates significantly lower than expected. For example, the data for nouns under the single-picture description in the aphasia group have a blue color intensity greater than 2.0. This indicates that the proportion of nouns in the aphasia group is greater than expected for this discourse task.




and Figure 1 for a general view of the comparisons). Regarding the quantity of verbal output, procedural discourse elicited the fewest words (i.e., tokens) among the four discourse tasks in the aphasia group. Interestingly, however, it elicited the densest proposition for all the discourse tasks. Regarding the noun–verb ratio, the single-picture description had the highest value among the four discourse tasks as it tended to elicit a higher proportion of nouns and a lower proportion of verbs.

The distribution of POS for each discourse task in the aphasia group was shown in a mosaic plot (see Figure 2A). The single-picture description in the aphasia group elicited a high proportion of nouns and a lower proportion of verbs and adverbs. Similar to the control group, the story narrative elicited a higher proportion of adverbs and a significantly lower frequency of nouns.

### Between-Group Comparison of Linguistic Variables Within Each Discourse Task

Table 3 presented details on the performance on measures of oral expressive language in the aphasia and control groups, as well as comparison results. The results revealed significant differences between the two groups in most of the linguistic variables, except for propositional density during the sequential-picture description and procedural discourse, the noun–verb ratio, and noun and verb percentages during the sequential-picture description, story narrative, and procedural discourse. As expected, better performances were found for the control group than for the aphasia group in terms of syntactic structure (MLU), quantity of language production (tokens), lexical diversity (MATTR), and speech rate (WPM). Regarding informativeness, the aphasia group showed lower propositional density than the control group during the single-picture description ($t = -2.530$, $p = .022$) and story narrative ($t = -3.015$, $p = .006$). However, no significant differences were observed in sequential-picture description ($t = -1.692$, $p = .142$) or procedural discourse ($U = 410.500$, $p = .415$). The differences in the noun–verb ratio and noun and verb percentages between the two groups were not significant across discourse tasks, with the notable exception of the single-picture description. In the single-picture description, it was interesting to see that the aphasia group had a higher percentage of nouns than the control group ($U = 314.000$, $p = .029$), while the verb percentage noticeably decreased ($U = 248.000$, $p = .002$), resulting in a higher noun–verb ratio in the aphasia group than in the control group ($U = 229.000$, $p < .001$).

## Discussion

In this study, we leveraged the recently constructed language corpus, Mandarin AphasiaBank, to investigate the effect of elicitation methods on discourse performance and compare the discourse task-specific linguistic characteristics

**Table 3.** Between-group comparison of linguistic variables within each discourse task.

| Linguistic variable | Discourse tasks | Aphasia ($n$ = 31) | Control ($n$ = 31) | Difference |
|---|---|---|---|---|
| MLU | Sequential | 5.35 (4.57–7.14) | 7.18 (6.18–9.47) | $^{b}U$ = 223.00, $p$ < .001** |
| | Single | 5.93 (2.37) | 8.23 (1.74) | $^{a}t$ = −4.362, $p$ < .001** |
| | Story | 5.57 (1.98) | 8.36 (1.93) | $^{a}t$ = −5.620, $p$ < .001** |
| | Procedural | 4.62 (3.33–5.33) | 6.88 (5.21–7.62) | $^{b}U$ = 160.00, $p$ < .001** |
| Tokens | Sequential | 69.00 (42.00–100.00) | 94.00 (79.00–127.00) | $^{b}U$ = 252.00, $p$ = .002* |
| | Single | 65.00 (41.00–102.00) | 106.00 (74.00–146.00) | $^{b}U$ = 271.00, $p$ = .006* |
| | Story | 80.00 (50.00–122.00) | 163.00 (104.00–247.00) | $^{b}U$ = 201.00, $p$ < .001** |
| | Procedural | 35.00 (26.00–64.00) | 78.00 (50.00–117.00) | $^{b}U$ = 187.00, $p$ < .001** |
| MATTR | Sequential | 0.90 (0.88–0.92) | 0.94 (0.92–0.95) | $^{b}U$ = 229.00, $p$ < .001** |
| | Single | 0.87 (0.85–0.91) | 0.93 (0.91–0.94) | $^{b}U$ = 199.00, $p$ < .001** |
| | Story | 0.89 (0.84–0.91) | 0.93 (0.91–0.95) | $^{b}U$ = 145.50, $p$ < .001** |
| | Procedural | 0.86 (0.07) | 0.91 (0.03) | $^{a}t$ = −3.739, $p$ = .001* |
| WPM | Sequential | 61.65 (27.98) | 123.99 (36.58) | $^{a}t$ = −7.537, $p$ < .001** |
| | Single | 52.08 (36.69–66.86) | 112.62 (78.21–130.91) | $^{b}U$ = 94.00, $p$ < .001** |
| | Story | 65.76 (32.14) | 134.93 (24.88) | $^{a}t$ = −9.477, $p$ < .001** |
| | Procedural | 59.35 (24.42) | 118.69 (28.63) | $^{a}t$ = −8.780, $p$ < .001** |
| Propositional density | Sequential | 0.39 (0.09) | 0.42 (0.05) | $^{a}t$ = −1.692, $p$ = .142 |
| | Single | 0.36 (0.08) | 0.40 (0.07) | $^{a}t$ = −2.530, $p$ = .022* |
| | Story | 0.42 (0.07) | 0.47 (0.05) | $^{a}t$ = −3.015, $p$ = .006* |
| | Procedural | 0.53 (0.44–0.57) | 0.49 (0.45–0.53) | $^{b}U$ = 410.500, $p$ = .415 |
| Noun–verb ratio | Sequential | 0.77 (0.31) | 0.74 (0.20) | $^{a}t$ = 0.419, $p$ = .747 |
| | Single | 0.92 (0.75–1.23) | 0.70 (0.57–0.85) | $^{b}U$ = 229.000, $p$ < .001** |
| | Story | 0.70 (0.53–0.92) | 0.70 (0.60–0.85) | $^{b}U$ = 466.500, $p$ = .900 |
| | Procedural | 0.65 (0.50–0.83) | 0.67 (0.57–0.75) | $^{b}U$ = 474.500, $p$ = .963 |
| Noun percentage | Sequential | 20.81 (7.51) | 20.79 (5.14) | $^{a}t$ = 0.014, $p$ = .989 |
| | Single | 22.64 (17.95–25.00) | 19.17 (16.67–21.74) | $^{b}U$ = 314.000, $p$ = .029* |
| | Story | 18.53 (6.15) | 17.44 (3.77) | $^{a}t$ = 0.836, $p$ = .501 |
| | Procedural | 18.31 (16.87–25.00) | 17.81 (15.22–18.92) | $^{b}U$ = 404.000, $p$ = .375 |
| Verb percentage | Sequential | 26.67 (23.35–35.14) | 28.21 (25.93–31.92) | $^{b}U$ = 445.000, $p$ = .731 |
| | Single | 23.08 (21.31–25.64) | 26.56 (24.59–32.26) | $^{b}U$ = 248.000, $p$ = .002* |
| | Story | 26.35 (5.65) | 24.54 (3.36) | $^{a}t$ = 1.533, $p$ = .182 |
| | Procedural | 28.74 (7.09) | 27.95 (6.64) | $^{a}t$ = 0.457, $p$ = .742 |

*Note.* Values are presented as mean (standard deviation) and median (interquartile range), unless otherwise noted. Two groups were compared by independent-samples $t$ test for normally distributed data and Mann–Whitney $U$ test for nonnormally distributed data. $P$ values listed here are adjusted after Benjamini–Hochberg correction. MLU = mean length of utterance; Sequential = Sequential-picture description; Single = Single-picture description; Story = Story narrative; Procedural = Procedural discourse; MATTR = moving-average type-token ratio; WPM = words per minute.

[a]Independent-samples $t$ test. [b]Mann–Whitney $U$ test.

*Significant effect at $p < .05$. **Significant effect at $p < .001$.

of people with anomic aphasia and healthy adults to facilitate the understanding of discourse deficits specific to the Mandarin population. For our first study aim, as we hypothesized, the discourse task types did affect language performance in both groups. We found quantitative disparities in the volume and diversity of language production probed by distinct discourse elicitation approaches, which is consistent with previous studies of English speakers (Alyahya et al., 2020; Fergadiotis et al., 2011; Olness, 2006; Sahraoui & Nespoulous, 2012; Stark, 2019; Stark et al., 2022; Stark & Fukuyama, 2020). In the healthy controls, story narratives yielded the highest overall word count and the lowest proportion of verbs. In the anomic aphasia group, procedural discourse produced the fewest total words and densest expressions, whereas their single-picture descriptions had the highest noun–verb ratio. With respect to the second study aim, as predicted, the anomic aphasia group demonstrated worse performance than the healthy participants in terms of the MLU, tokens, MATTR, and WPM for all tasks. Unexpectedly, for noun percentage, verb percentage, and noun–verb ratio, the aphasia and control groups only differed in the single-picture description task.

The results for the control group showed that story narratives elicited the most overall word count and lowest verb percentage. The larger quantity of verbal output elicited by story narratives from healthy adults is consistent with previous research in English (Fergadiotis et al., 2011; Stark, 2019). It suggested that the story narrative is useful in prompting more language production for both English and Mandarin Chinese–speaking healthy populations. Interestingly, despite eliciting the maximum number of words, the story narrative contained a lower proportion of verbs in healthy participants. More in-depth observations can be made through the mosaic plot (see Figure 2), which provides a comprehensive overview of task-specific distributions of POS. It can be seen from the mosaic plot (see Figure 2B) that adverbs and adjectives made up a more significant proportion of the story narrative than the other discourse tasks in the control group, whereas nouns and verbs were less frequent. In this study, the story narrative involved participants telling a familiar fable, which may have resulted in speakers exhibiting more feelings and personal opinions in their narration. Adjectives and adverbs are essential components of a vivid narrative, providing details and context to characters and settings. Using adjectives and adverbs, one can convey character traits, emphasize a story's mood and tone, and create a sense of suspense or drama. This indicates that retelling familiar stories might lead to greater use of modifiers such as adverbs, which in turn reduces the proportion of verbs.

In the aphasia group, procedural discourse elicited the fewest tokens and densest expressions among the four discourse tasks. This indicates that procedural discourse induces less language output than other discourse tasks in persons with anomic aphasia. However, even with a limited amount of verbal output, procedural discourse produced by them provides considerable information (high-propositional density). With respect to the noun–verb ratio, it should be noted that, in the aphasia group, the single-picture task elicited the highest noun–verb ratio compared to other discourse tasks, indicating the least syntactic complexity. However, in the control group, no significant effect of discourse task types was found for the noun–verb ratio. On the one hand, these results were inconsistent with those of Stark (2019), who reported that procedural discourse exhibited the highest noun–verb ratio compared with picture descriptions (sequential and single pictures) and story narratives in both the aphasia and control groups. Instead, procedural discourse produced by persons with anomic aphasia in our study tended to elicit a larger proportion of verbs, resulting in a lower noun–verb ratio. It is likely that Mandarin emphasizes verbs more than English as a pro-drop language in which sentence subjects are often optional and frequently omitted (Li & Kiran, 2023; Tardif et al., 1997). Dropping the subject of a sentence is permitted in Mandarin but not in English. Procedural discourse requires more action words to explain the steps involved in an activity or a practice. As it is unnecessary to specify a particular character or subject in procedural discourse, null subjects in Mandarin may lead to a higher verb percentage, resulting in a lower noun–verb ratio. On the other hand, prior research has indicated that single-picture stimuli elicit more descriptive discourse than narratives in adults with and without aphasia, but this tendency seems more pronounced in persons with aphasia (Olness et al., 2002). This phenomenon can be attributed to narratives' greater linguistic demands for expression of temporality and reference; therefore, people with aphasia may find it simpler to construct descriptive discourse in response to single-picture stimuli (Olness, 2006). Sequential picture stimuli consist of a comic strip of consecutive pictures that are sequenced and essentially serve as a framework, providing detailed information about the development of the story or the course of the characters' activities. By contrast, single-picture stimuli do not provide such clear temporal cause-and-effect relationships. Describing a single picture may require more reasoning skills because speakers need to infer the logical linkages of the characters' events implied behind the single-scene image. Speakers may tend to ignore implied logical relationships between picture events, preferring to simply list objects or activities in pictures (Wright & Capilouto, 2009). As a result, single-picture stimuli are more likely to lead speakers to merely enumerate events and characters, resulting in a higher noun–verb ratio (Li & Kiran, 2023), which is associated with simpler syntax. In the context of clinical assessment, if the goal is to evaluate syntactic organization in people with mild aphasia (e.g., anomic aphasia) according to normal standards, it is plausible that employing more demanding discourse tasks may enhance sensitivity in detecting any residual difficulties they may have with syntactic structure. The results of this study indicated that there was no significant variation in syntactic complexity (as measured by noun–verb ratio) among neurotypical controls across discourse tasks. The participants with anomic aphasia had more difficulty producing narratives and organizing more complicated syntactic structure in response to single-picture prompts. Consequently, the utilization of single-picture descriptions, which pose a greater challenge for them, may be a more suitable strategy than other tasks to identify remaining syntax problems. When it comes to discourse treatment, if the objective is to enhance the syntactic complexity of people with aphasia, it is advised to select discourse tasks that gradually increase in difficulty. For example, it is easier for people with aphasia to construct a narrative and employ more independent clauses and complex sentences when describing sequential pictures than when describing a single picture. Clinicians who wish to target syntactic complexity may choose to elicit descriptions using sequential pictures. As an individual with aphasia makes progress, it may be appropriate to apply single-picture descriptions for the purpose of more advanced training. Superior skills in organizing sophisticated syntax are required for this task. By progressively raising the level of difficulty, this approach facilitates improved syntactic complexity, which then may be carried over to tasks where it is more challenging for individuals with aphasia to produce this level of complexity.

As anticipated, the aphasia group generally demonstrated lower values than the control group for most linguistic variables, including MLU, tokens, MATTR, and WPM. This indicates the capacity of these linguistic variables to distinguish between normal and impaired discourse in the four discourse tasks. Surprisingly, for noun percentage, verb percentage, and noun–verb ratio, single-picture description was the only discourse task that showed a noticeable difference between the aphasia and control groups. Research has indicated that cognitive deficits may contribute to language difficulties, especially in individuals with mild aphasia (Armstrong et al., 2013; Cavanaugh & Haley, 2020; DeDe & Salis, 2020; Frankel et al., 2007). Cognitive impairments, such as memory disturbance and executive dysfunction, may contribute to the simpler syntax (as indicated by a higher noun–verb ratio) observed in the single-picture description of individuals with anomic aphasia compared to healthy controls. As mentioned earlier, single-picture descriptions require speakers to have a higher cognitive ability to understand, judge, and logically reason about the antecedents and consequences of a picture. Unlike neurotypical adults, individuals with anomic aphasia who have cognitive deficits might have difficulty inferring the underlying story plots, remembering multiple pieces of information, constructing long sentences and processing complicated syntax (Andreetta et al., 2012; Salis & DeDe, 2022). As a result, they may exhibit simplified syntax when responding to single-picture stimuli. These results suggested that, when measured by noun–verb ratio, noun percentage or verb percentage, the single-picture description may have greater discriminatory potential for detecting discourse deficits and, thus, could be used to differentiate mild aphasia like anomic aphasia from neurotypical controls. Future research should include measures of cognitive abilities as a possible contributor to discourse difficulty in mild aphasia.

Although discourse in aphasia based on Indo-European languages has been extensively studied (Dipper et al., 2021; Stark, Dutta, Murray, Bryant, et al., 2021; Stark, Dutta, Murray, Fromm, et al., 2021), investigations in Mandarin have received less attention. This study focused on Mandarin Chinese speakers and supported the suggestion that discourse assessments adopted in clinical practice should select specific elicitation methods based on the aspects of the language components examined. The prompts and stimuli used to collect discourse samples may impose different linguistic and cognitive demands, resulting in discrepancies in the quantity and quality of the verbal output. The findings of this study provide further evidence for selecting the appropriate discourse task to evaluate specific linguistic elements and have critical implications for both discourse-based clinical assessments and interventions in Mandarin Chinese–speaking individuals with aphasia. It appears that a comprehensive assessment should encompass all types of discourse tasks, but that some will be favored over others depending on the goal of the session (Leaman & Edmonds, 2023). In clinical assessment, if we want to discern how people with mild aphasia differ from neurotypical controls in terms of syntactic complexity (as measured by noun–verb ratio), the single-picture description might be considered because it showed higher discriminating sensitivity than other tasks for detecting residual syntactic dysfunctions. When it comes to discourse treatment, if we aim to motivate a greater quantity of language output from persons with anomic aphasia, procedural discourse might not be an effective approach because it elicits the least overall number of words among tasks. Instead, narrating a familiar story might be more efficient to inspire them to produce more language. Moreover, the task-specific distributions of POS can provide valuable insights into determining which discourse task is more suitable for specialized training on a particular part of speech. For example, if our objective is to elicit more adverbs and improve the skill of adverb usage in people with anomic aphasia, employing story narratives or procedural discourse might be more effective than picture description. This is because, as shown in Figure 2 of the study, the percentage of adverbs in story narrative and procedural discourse is notably higher than other POS for both the aphasia and control groups. This suggests that adverbs are more frequently used in these tasks. Therefore, opting for the task of narrating a story or describing a procedure might be more effective in generating adverbs and providing intensive training in their usage. Additionally, if we desire to examine the effects after training for particular language or cognitive components, it would be more relevant to use discourse tasks that specifically require those particular language or cognitive attributes. To illustrate, when measuring the treatment outcome of individuals with mild aphasia subsequent to training on recognizing temporal–causal relationships, reasoning abilities or related cognitive skills, employing the single-picture description task might be more responsive. This is because this task tends to tax more cognitive demand on the skills of judging, identifying and inferring the implied meanings that are depicted in a pictorial scene (Capilouto et al., 2005).

## *Limitations and Future Directions*

Our study has several noteworthy limitations. First, we focused on collecting preliminary evidence solely from persons with anomic aphasia, hoping to provide a more accurate and clearer understanding of the relationship between language performance and elicitation strategies with less variability. Therefore, the conclusions of this study may not be applicable to other types of aphasia. Future work should investigate different types of aphasia to generalize the results to discourse deficits using a range of elicitation methods. Second, this study examined the microlinguistic measures of discourse. We did not study coherence, main concepts, or story structure, which would provide insights into the macrostructural aspects. Such macrolinguistic analyses and other discourse measurements are needed in future research to construct a more comprehensive framework of the range of discourse deficits that may be present in aphasia (Boyle, 2020; Linnik et al., 2015; Pritchard et al., 2017). Third, the sample size of the study was limited. Although we extracted and analyzed data from Mandarin AphasiaBank, the amount of data it contains is currently limited. As the Chinese Mandarin AphasiaBank database continues to grow, these findings should be updated to yield more precise and comprehensive conclusions. Finally, this study could be extended by including cognitive measurements to further explore the connection between language performance and cognitive demands in different discourse tasks.

## Conclusions

Discourse tasks influence the microstructural language components of Mandarin Chinese speakers with and without aphasia. Implementing assessments and interventions for discourse in individuals with aphasia may require a multifactorial approach or consideration of discourse task types to construct a more comprehensive understanding of connected speech in aphasia.

## Author Contributions

**Bao-Mei Deng:** Conceptualization (Equal), Data curation (Equal), Formal analysis (Lead), Investigation (Lead), Methodology (Lead), Project administration (Equal), Resources (Equal), Supervision (Lead), Validation (Equal), Visualization (Lead), Writing – original draft (Lead), Writing – review & editing (Lead). **Jing Gao:** Conceptualization (Equal), Data curation (Equal), Funding acquisition (Supporting), Project administration (Supporting), Writing – review & editing (Equal). **Li-Si Liang:** Investigation (Equal), Project administration (Supporting), Resources (Supporting), Writing – review & editing (Supporting). **Jia-Xin Zhao:** Data curation (Supporting), Investigation (Equal), Project administration (Supporting), Writing – review & editing (Supporting). **Feng Lin:** Conceptualization (Equal), Data curation (Equal), Project administration (Lead), Resources (Lead), Supervision (Lead), Validation (Lead). **Ming-Yu Yin:** Funding acquisition (Supporting), Project administration (Supporting), Validation (Equal), Writing – review & editing (Supporting). **Hai-Qing Zheng:** Project administration (Supporting), Validation (Equal), Writing – review & editing (Supporting). **Xi-Quan Hu:** Conceptualization (Supporting), Funding acquisition (Supporting), Project administration (Supporting), Supervision (Equal), Writing – review & editing (Supporting).

## Data Availability Statement

The data sets generated and/or analyzed during this study are not publicly available because of privacy or ethical restrictions; however, the corresponding author can share a limited amount of data upon reasonable request.

## Acknowledgments

This work was supported by National Key R&D Program of China (Grant 2022YFC3601200), awarded to Xi-Quan Hu; National Natural Science Foundation of China (Grant 82202805), awarded to Ming-Yu Yin; and Nanjing Medical University Science and Technology Development Fund (Grant NMUB20210350), awarded to Jing Gao. We thank Xiang-Wei Kong's contribution to data analysis. We are profoundly grateful to the developers and contributors of AphasiaBank for their dedication and commitment.

## References

**Alyahya, R. S. W., Halai, A. D., Conroy, P., & Lambon Ralph, M. A.** (2020). A unified model of post-stroke language deficits including discourse production and their neural correlates. *Brain, 143*(5), 1541–1554. https://doi.org/10.1093/brain/awaa074

**Alyahya, R. S. W., Lambon Ralph, M. A., Halai, A., & Hoffman, P.** (2022). The cognitive and neural underpinnings of discourse coherence in post-stroke aphasia. *Communications, 4*(3), Article fcac147. https://doi.org/10.1093/braincomms/fcac147

**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

**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)

**Bates, E., Wulfeck, B., & MacWhinney, B.** (1991). Cross-linguistic research in aphasia: An overview. *Brain and Language, 41*(2), 123–148. https://doi.org/10.1016/0093-934x(91)90149-u

**Berube, S. K., Goldberg, E., Sheppard, S. M., Durfee, A. Z., Ubellacker, D., Walker, A., Stein, C. M., & Hillis, A. E.** (2022). An analysis of right hemisphere stroke discourse in the modern cookie theft picture. *American Journal of Speech-Language Pathology, 31*(5S), 2301–2312. https://doi.org/10.1044/2022_AJSLP-21-00294

**Boyle, M.** (2020). Choosing discourse outcome measures to assess clinical change. *Seminars in Speech and Language, 41*(01), 001–009. https://doi.org/10.1055/s-0039-3401029

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

**Bryant, L., Spencer, E., & Ferguson, A.** (2016). Clinical use of linguistic discourse analysis for the assessment of language in aphasia. *Aphasiology, 31*(10), 1105–1126. https://doi.org/10.1080/02687038.2016.1239013

**Capilouto, G., Wright, H. H., & Wagovich, S. A.** (2005). CIU and main event analyses of the structured discourse of older and younger adults. *Journal of Communication Disorders, 38*(6), 431–444. https://doi.org/10.1016/j.jcomdis.2005.03.005

**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

**Cruice, M., Botting, N., Marshall, J., Boyle, M., Hersh, D., Pritchard, M., & Dipper, L.** (2020). UK speech and language therapists' views and reported practices of discourse analysis in aphasia rehabilitation. *International Journal of Language and Communication Disorders, 55*(3), 417–442. https://doi.org/10.1111/1460-6984.12528

**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 and Hearing Research, 63*(3), 710–721. https://doi.org/10.1044/2019_JSLHR-19-00226

**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

**Dalton, S. G. H., & Richardson, J. D.** (2019). A large-scale comparison of main concept production between persons with aphasia and persons without brain injury. *American Journal of Speech-Language Pathology, 28*(1S), 293–320. https://doi.org/10.1044/2018_AJSLP-17-0166

**DeDe, 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, 11*(2). https://doi.org/10.3390/brainsci11020183

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

**Enderby, P., & Palmer, R.** (2008). *Frenchay Dysarthria Assessment–Second Edition (FDA-2)*. Pro-Ed.

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

**Fergadiotis, G., Wright, H. H., & Capilouto, G. J.** (2011). Productive vocabulary across discourse types. *Aphasiology, 25*(10), 1261–1278. https://doi.org/10.1080/02687038.2011.606974

**Fergadiotis, G., Wright, H. H., & West, T. M.** (2013). Measuring lexical diversity in narrative discourse of people with aphasia. *American Journal of Speech-Language Pathology, 22*(2), S397–S408. https://doi.org/10.1044/1058-0360(2013/12-0083)

**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

**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., & MacWhinney, B.** (2020). Using AphasiaBank for discourse assessment. *Seminars in Speech and Language, 41*(01), 010–019. https://doi.org/10.1055/s-0039-3399499

**Fromm, D., Greenhouse, J., Hou, K., Russell, G. A., Cai, X., Forbes, M., Holland, A., & MacWhinney, B.** (2016). Automated proposition density analysis for discourse in aphasia. *Journal of Speech, Language, and Hearing Research, 59*(5), 1123–1132. https://doi.org/10.1044/2016_JSLHR-L-15-0401

**Hameister, I., & Nickels, L.** (2018). The cat in the tree – using picture descriptions to inform our understanding of conceptualisation in aphasia. *Language, Cognition and Neuroscience, 33*(10), 1296–1314. https://doi.org/10.1080/23273798.2018.1497801

**Jiang, Y. E., Liao, X. Y., & Liu, N.** (2023). Applying core lexicon analysis in patients with anomic aphasia: Based on Mandarin AphasiaBank. *International Journal of Language and Communication Disorders, 58*(6), 1875–1886. https://doi.org/10.1111/1460-6984.12864

**Kintz, S., & Wright, H. H.** (2017). Discourse measurement in aphasia research. *Aphasiology, 32*(4), 472–474. https://doi.org/10.1080/02687038.2017.1398807

**Kong, A. P.-H., Whiteside, J., & Bargmann, P.** (2016). The main concept analysis: Validation and sensitivity in differentiating discourse produced by unimpaired English speakers from individuals with aphasia and dementia of Alzheimer type. *Logopedics, Phoniatrics, Vocology, 41*(3), 129–141. https://doi.org/10.3109/14015439.2015.1041551

**Kong, A. P.-H., & Wong, C. W.** (2018). An integrative analysis of spontaneous storytelling discourse in aphasia: Relationship with listeners' rating and prediction of severity and fluency status of aphasia. *American Journal of Speech-Language Pathology, 27*(4), 1491–1505. https://doi.org/10.1044/2018_AJSLP-18-0015

**Law, S. P., Kong, A. P., & Lai, C.** (2018). An analysis of topics and vocabulary in Chinese oral narratives by normal speakers and speakers with fluent aphasia. *Clinical Linguistics and Phonetics, 32*(1), 88–99. https://doi.org/10.1080/02699206.2017.1334092

**Leaman, M. C., & Edmonds, L. A.** (2023). Analyzing language in the picnic scene picture and in conversation: The type of discourse sample we choose influences findings in people with aphasia. *American Journal of Speech-Language Pathology, 32*(4), 1413–1430. https://doi.org/10.1044/2023_AJSLP-22-00279

**Li, R., & Kiran, S.** (2023). Noun and verb impairment in single-word naming and discourse production in Mandarin-English bilingual adults with aphasia. *Aphasiology, 38*(2), 337–365. https://doi.org/10.1080/02687038.2023.2189994

**Linnik, A., Bastiaanse, R., & Höhle, B.** (2015). Discourse production in aphasia: A current review of theoretical and methodological challenges. *Aphasiology, 30*(7), 765–800. https://doi.org/10.1080/02687038.2015.1113489

**Litovsky, C. P., Finley, A. M., Zuckerman, B., Sayers, M., Schoenhard, J. A., Kenett, Y. N., & Reilly, J.** (2022). Semantic flow and its relation to controlled semantic retrieval deficits in the narrative production of people with aphasia. *Neuropsychologia, 170,* Article 108235. https://doi.org/10.1016/j.neuropsychologia.2022.108235

**MacWhinney, B.** (2000). *The CHILDES Project: Tools for analyzing talk*. (3rd ed.). Erlbaum.

**MacWhinney, B.** (2019). Understanding spoken language through TalkBank. *Behavior Research Methods, 51*(4), 1919–1927. https://doi.org/10.3758/s13428-018-1174-9

**MacWhinney, B., & Fromm, D.** (2016). AphasiaBank as BigData. *Seminars in Speech and Language, 37*(01), 10–22. https://doi.org/10.1055/s-0036-1571357

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

**Marcotte, K., Lachance, A., Brisebois, A., Mazzocca, P., Desilets-Barnabe, M., Desjardins, N., & Brambati, S. M.** (2022). Validation of videoconference administration of picture description from the Western Aphasia Battery-Revised in Neurotypical Canadian French Speakers. *American Journal of Speech-Language Pathology, 31*(6), 2825–2834. https://doi.org/10.1044/2022_AJSLP-22-00084

**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

**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

**Olness, G. S., Ulatowska, H. K., Wertz, R. T., Thompson, J. L., & Auther, L. L.** (2002). Discourse elicitation with pictorial stimuli in African Americans and Caucasians with and without aphasia. *Aphasiology, 16*(4–6), 623–633. https://doi.org/10.1080/02687030244000095

**Pritchard, M., Dipper, L., Morgan, G., & Cocks, N.** (2015). Language and iconic gesture use in procedural discourse by speakers with aphasia. *Aphasiology, 29*(7), 826–844. https://doi.org/10.1080/02687038.2014.993912

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

**R Core Team.** (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org

**Richardson, J. D., Dalton, S. G., Fromm, D., Forbes, M., Holland, A., & MacWhinney, B.** (2018). The relationship between confrontation naming and story gist production in aphasia. *American Journal of Speech-Language Pathology, 27*(1S), 406–422. https://doi.org/10.1044/2017_AJSLP-16-0211

**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), Article 110. https://doi.org/10.3390/brainsci11010110

**Ryan, J., Woodyatt, G., & Copeland, D.** (2010). Procedural discourse in intellectual disability and dual diagnosis. *Journal of Intellectual Disability Research, 54*(1), 70–80. https://doi.org/10.1111/j.1365-2788.2009.01229.x

**Sahraoui, H., & Nespoulous, J.-L.** (2012). Across-task variability in agrammatic performance. *Aphasiology, 26*(6), 785–810. https://doi.org/10.1080/02687038.2011.650625

**Saling, L. L., Laroo, N., & Saling, M. M.** (2017). Automatization of personal and impersonal discourse with narrative retelling as a function of age. *Aging, Neuropsychology and Cognition, 24*(6), 649–661. https://doi.org/10.1080/13825585.2016.1247780

**Salis, C., & DeDe, 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

**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., Bryant, L., Themistocleous, C., den Ouden, D.-B., & Roberts, A. C.** (2022). Best practice guidelines for reporting spoken discourse in aphasia and neurogenic communication disorders. *Aphasiology, 37*(5), 761–784. https://doi.org/10.1080/02687038.2022.2039372

**Stark, B. C., Dutta, M., Murray, L. L., Bryant, L., Fromm, D., MacWhinney, B., Ramage, A. E., Roberts, A., den Ouden, D. B., Brock, K., McKinney-Bock, K., Paek, E. U., Harmon, T. G., Yoon, S. O., Themistocleous, C., Yoo, H., Aveni, K., Gutierrez, K., & Sharma, S.** (2021). Standardizing assessment of spoken discourse in aphasia: A working group with deliverables. *American Journal of Speech-Language Pathology, 30*(1S), 491–502. https://doi.org/10.1044/2020_AJSLP-19-00093

**Stark, B. C., Dutta, M., Murray, L. L., Fromm, D., Bryant, L., Harmon, T. G., Ramage, A. E., & Roberts, A. C.** (2021). Spoken discourse assessment and analysis in aphasia: An international survey of current practices. *Journal of Speech Language and Hearing Research, 64*(11), 4366–4389. https://doi.org/10.1044/2021_JSLHR-20-00708

**Stark, B. C., & Fukuyama, J.** (2020). Leveraging big data to understand the interaction of task and language during monologic spoken discourse in speakers with and without aphasia. *Language, Cognition and Neuroscience, 36*(5), 562–585. https://doi.org/10.1080/23273798.2020.1862258

**Stubbs, E., Togher, L., Kenny, B., Fromm, D., Forbes, M., MacWhinney, B., McDonald, S., Tate, R., Turkstra, L., & Power, E.** (2018). Procedural discourse performance in adults with severe traumatic brain injury at 3 and 6 months post injury. *Brain Injury, 32*(2), 167–181. https://doi.org/10.1080/02699052.2017.1291989

**Tardif, T., Shatz, M., & Naigles, L.** (1997). Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language, 24*(3), 535–565. https://doi.org/10.1017/s030500099700319x

**Wallace, S. J., Worrall, L., Rose, T., Le Dorze, G., Cruice, M., Isaksen, J., Pak Hin Kong, A., Simmons-Mackie, N., Scarinci, N., & Gauvreau, C. A.** (2017). Which outcomes are most important to people with aphasia and their families? An international nominal group technique study framed within the ICF. *Disability and Rehabilitation, 39*(14), 1364–1379. https://doi.org/10.1080/09638288.2016.1194899

**Wang, Y. H.** (1997). The introduction of Western Aphasia Battery (WAB). *Chinese Journal of Rehabilitation Theory and Practice, 3*(2), 87–89.

**Wong, E. C. H., Wong, M. N., & Velleman, S. L.** (2023). Assessment and diagnostic standards of apraxia of speech in Chinese-speaking adults and children: A scoping review. *American Journal of Speech-Language Pathology, 32*(1), 316–340. https://doi.org/10.1044/2022_AJSLP-21-00355

**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., Koutsoftas, A. D., Capilouto, G. J., & Fergadiotis, G.** (2014). Global coherence in younger and older adults: Influence of cognitive processes and discourse type. *Aging, Neuropsychology and Cognition, 21*(2), 174–196. https://doi.org/10.1080/13825585.2013.794894

**You, L., Wang, Y., Chen, W., Zhang, S., Rao, J., Liu, L., & Shan, C.** (2019). The effectiveness of action observation therapy based on mirror neuron theory in Chinese patients with apraxia of speech after stroke. *European Neurology, 81*(5–6), 278–286. https://doi.org/10.1159/000503960