

Research Article

Core Lexicon Analysis of a Storytelling Task in Korean-Speaking Individuals With Aphasia

Jimin Park,^a  Sujin Choi,^a  So Eun Kim,^a  Hana Kim,^b  and Jee Eun Sung^{a,c,d} 

^aDepartment of Communication Disorders, Ewha Womans University, Seoul, South Korea ^bDepartment of Communication Sciences & Disorders, University of South Florida, Tampa ^cDivision of Brain & Cognitive Sciences, Ewha Womans University, Seoul, South Korea ^dEwha Brain Institute, Ewha Womans University, Seoul, South Korea

ARTICLE INFO**Article History:**

Received September 17, 2025

Revision received March 10, 2026

Accepted July 7, 2026

Editor-in-Chief: Catherine A. Off

Editor: Brielle C. Stark

https://doi.org/10.1044/2026_AJSLP-25-00421

ABSTRACT

Purpose: This study aimed to validate newly developed, age-stratified Korean core lexicon (CL) checklists and evaluate their clinical utility in assessing discourse production in persons with aphasia (PWA) using a culturally familiar storytelling task.

Method: Thirty-six PWA and 36 age- and education-matched cognitively healthy adults (CHAs) participated in a Korean picture-based storytelling task. Percentage of core nouns and verbs were calculated using age-specific checklists and compared across group (CHAs vs. PWA) and lexical category (noun vs. verb). We also examined whether core noun and verb percentages were associated with aphasia severity and which word class significantly predicted severity.

Results: PWA exhibited significantly lower percentage of CL than CHAs, with a larger noun–verb difference observed in PWA than in CHAs. Pearson correlation analyses showed that both core nouns and verbs were significantly and positively associated with aphasia severity. Hierarchical regression analyses revealed that core verb percentage was a stronger predictor of aphasia severity than core noun percentage, accounting for significant variance beyond demographic factors, whereas core noun percentage did not explain additional variance when core verb percentage was already included in the model.

Conclusions: This study presents the first empirical validation of Korean CL checklists and supports their clinical utility in identifying discourse impairments in PWA. The findings emphasize the central role of verbs in discourse assessment for verb-salient languages and underscore the broader need for culturally and linguistically tailored assessment tools in aphasia evaluation.

Discourse refers to language that extends beyond single sentences, encompassing cohesive and coherent stretches of communication (Dipper & Pritchard, 2017; Ulatowska et al., 2004). Clinically, discourse is important because it offers a way to understand aspects of communication ability that are not often captured by standardized test batteries (Cavanaugh & Haley, 2020; Dutta & Mohapatra, 2024; Fromm et al., 2017; Wright, 2011). From the perspective of

persons with aphasia (PWA), discourse represents a primary target for intervention, highlighting the need for language to support daily communication (Wallace et al., 2017; Worrall et al., 2011).

Traditionally, discourse samples have been analyzed using linguistic outcome measures such as mean length of utterance, lexical diversity, and correct information units (CIUs). While these measures yield detailed profiles, they require multiple steps of transcription and coding that demand substantial time and effort (Duncan & Murray, 2012; Verna et al., 2009). Survey studies confirm that these demands create significant barriers in clinical practice (Bryant et al., 2017; Cruice et al., 2020; Stark et al., 2021). Bryant et al. (2017), for example, reported that nearly half of clinicians in five English-speaking countries

Correspondence to Jee Eun Sung: jeesung@ewha.ac.kr and Hana Kim: hanakim1@usf.edu. **Publisher Note:** This article is part of the Special Issue: Select Papers From the 54th Clinical Aphasiology Conference. **Disclosure:** The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.

rarely or only occasionally analyze discourse samples. Likewise, Cruice et al. (2020) found that only 30% of clinicians in the United Kingdom regularly conduct discourse analysis, with uptake largely restricted to those with greater experience in working with PWA. Instead, approximately 70% of clinicians rely on discourse tasks embedded within standardized assessments (e.g., Western Aphasia Battery–Revised [WAB-R]; Kertesz, 2007), which provide general and often subjective ratings of content and fluency, rather than systematic scoring or detailed discourse analyses (Fromm et al., 2017; Maddy et al., 2015). As a result, examiners who wish to evaluate discourse in more detail must undertake additional transcription and analysis, further limiting feasibility in routine clinical practice. Taken together, these findings highlight the gap between the recognized value of discourse assessment and its limited clinical implementation, underscoring the need for clinician-friendly methods that allow efficient, real-time evaluation.

One approach that directly addresses this gap is core lexicon (CL) analysis, a clinician-friendly method for evaluating discourse performance (Dalton & Richardson, 2015; H. Kim et al., 2019, 2023; H. Kim & Wright, 2020a; MacWhinney et al., 2010). CL quantifies semantically important words produced in a given discourse task using a checklist of words predetermined from cognitively healthy adults (CHAs; MacWhinney et al., 2010). By providing a norm-referenced standard, CL enables efficient evaluation of PWA's discourse production without the extensive transcription and coding required by traditional methods. Over the past decade, CL checklists have been developed for a range of elicitation contexts: picture description (e.g., Cookie Theft, Dalton et al., 2024; Cat in Tree, Kintz & Kim, 2023, and Dalton et al., 2020), procedural discourse (e.g., How to Make a Peanut Butter and Jelly Sandwich, Fromm et al., 2013; Dalton et al., 2020), and narrative tasks (e.g., Cinderella; Dalton et al., 2020; MacWhinney et al., 2010). These task-specific checklists have since been validated across several languages, such as English and Mandarin, and evidence consistently supports their reliability and sensitivity (Chen & Chang, 2024; Dalton & Richardson, 2015; Jiang et al., 2023; H. Kim & Kintz, 2024; H. Kim et al., 2019, 2021; Kintz et al., 2024).

Validation studies demonstrate that CL aligns with traditional discourse measures. For example, H. Kim and Wright (2020b) examined 11 PWA performing two narrative tasks (*Good Dog*, *Carl*, Day, 1985; *Picnic*, McCully, 1984) and found that CL scores were significantly correlated with both macrolinguistic (e.g., coherence) and microlinguistic (e.g., lexical diversity) measures. Similarly, Chen and Chang (2024) compared Mandarin-speaking CHAs and PWA across seven discourse tasks, including picture description, procedural discourse, and storytelling, and reported significant correlations between CL, lexical diversity, and CIU

across all tasks. Beyond convergent validity, CL has also proven sensitive in distinguishing discourse performance within and across PWA subtypes. H. Kim et al. (2019, 2021) developed age-stratified CL checklists across five word classes (nouns, verbs, adverbs, adjectives, and function words) from 470 CHAs and applied them to 11 PWA, showing that verb and function word scores significantly correlated with Aphasia Quotient (AQ) from the WAB-R (Kertesz, 2007). Dalton and Richardson (2015) further analyzed *Broken Window* narratives (Menn et al., 1998) from the AphasiaBank protocol (MacWhinney et al., 2011), creating CL checklists from 92 CHAs across four age groups and comparing them with 235 PWA; CL scores not only differentiated CHAs from PWA but also distinguished fluent and Broca's aphasia subtypes. Extending this evidence, Stark et al. (2025) demonstrated that CL can even detect latent aphasia—individuals performing within the normal range on the WAB-R but exhibiting subtle discourse deficits—with 74% classification accuracy based on a *Cinderella* storytelling task, underscoring the measure's sensitivity to subtle group differences. Collectively, these findings highlight CL as a valid and sensitive index of lexical and discourse-level performance in PWA.

Although CL measures have been increasingly adopted in discourse research for languages such as English (e.g., Dalton & Richardson, 2015; H. Kim et al., 2019) and Mandarin (e.g., Chen & Chang, 2024; Jiang et al., 2023), they have yet to be developed for Korean. Most Korean studies continue to rely on conventional discourse metrics, such as lexical diversity, mean length of utterance, and information units (H. Choi, 2022; Jin et al., 2016; H. Kim, Sung, & Jeong, 2022; H. S. Kim & Sung, 2022; H. Lee et al., 2024; S. J. Lee et al., 2009; J. H. Park & Choi, 2024; Sung et al., 2016). However, these approaches are particularly labor-intensive and time-consuming in Korean due to both the typological complexity of the language and the limited availability of reliable automated analysis tools. Korean permits subject and object omission (pro-drop), allowing predicates to function as syntactically complete sentences (H. Lee, 2015; Sohn, 2001); its case-marking system enables flexible word order (Sohn, 2001; Sung et al., 2019); and serial verb constructions express nuanced meanings without explicit conjunctions (Sohn, 2001). While these linguistic features are informative, they complicate morphosyntactic parsing and have limited the development of computational tools. For example, Computerized Language Analysis (MacWhinney, 2000)—a widely used tool for automated speech and morphosyntactic analysis—does not support Korean and currently provides morphological parsing capabilities for only a limited set of languages, such as English, Spanish, and Chinese. Korean-specific morphological analyzers such as *KoNLPy* (E. L. Park & Cho, 2014) have been developed, but they often produce inaccurate or

inconsistent results due to issues with morpheme segmentation and part-of-speech disambiguation. For example, the term *Heungbuka* can ambiguously be parsed as either the title of a traditional song itself (proper noun) or as the noun “Heungbu” with the nominative case marker “-ka,” leading to parsing errors. Given these linguistic and technological constraints, the development of CL measures tailored to Korean represents a promising and practical alternative, offering a standardized and clinically meaningful framework for discourse analysis that reduces the dependence on error-prone automated tools.

In addition to language-specific features, cultural familiarity critically influences discourse production across languages (Degani & Onysko, 2024; Mazumdar et al., 2023; Yoo et al., 2026). For instance, Yoo et al. (2026) compared two AphasiaBank picture tasks (*Cat Rescue* and *Refused Umbrella*) between Korean and English speakers (CHAs and PWA) using main concept analysis, which measures essential elements within a concept (e.g., subject, verb; Dalton & Richardson, 2019; Nicholas & Brookshire, 1995; Richardson & Dalton, 2016). To account for Korean-specific pro-drop features, the authors modified the original scoring method by awarding full credit for omitted subjects when they were referable from the preceding utterance. Even with this adjustment, Korean CHAs and PWA still scored significantly lower on *Cat Rescue* than their English-speaking counterparts, but not on *Refused Umbrella*. Furthermore, stepwise discriminant analyses between two languages revealed that only *Cat Rescue* yielded significant discriminants for both controls and PWA, whereas none were identified in *Refused Umbrella*. This pattern likely reflects cultural familiarity: Rescuing a cat from a tree is not a common scenario in Korea, making it more difficult for Korean speakers to identify the central theme of the picture, whereas a mother telling her son to take an umbrella on a rainy day represents a culturally universal event. These findings underscore the importance of cultural sensitivity in discourse assessment and highlight the need to incorporate culturally relevant tasks into the development of CL checklists to ensure validity and clinical applicability.

Considering Korean linguistic and cultural characteristics (Sung et al., 2024), we have developed the first Korean CL checklists using narratives from 180 CHAs, divided equally into three age groups—young ($n = 60$), middle-aged ($n = 60$), and older adults ($n = 60$; J. H. Park, Choi, et al., 2025). Discourse was elicited using the folktale *Heungbu and Nolbu* (*Bad Older Brother and Good Younger Brother*), selected for its high familiarity across generations (S. Choi et al., 2022) and its frequent use in Korean clinical research (S. N. Kim & Shin, 2024; H. Kim, Sung, & Jeong, 2022; J. H. Park & Choi, 2024). The checklists were constructed separately for nouns and verbs, as these open-class categories are highly sensitive to generational variation and

together account for approximately 80% of the Korean lexicon (G. Lee et al., 2018). Within each age group, nouns and verbs produced by at least 40% of participants were identified as core items. Comparisons of core item lists revealed that percent agreement across age groups was consistently higher for nouns (76%–93%) than for verbs (50%–72%). As a preliminary step toward developing clinically acceptable discourse measures for the Korean population and prior to applying them in clinical contexts, we explore Least Absolute Shrinkage and Selection Operator logistic regression (Tibshirani, 1996; Tibshirani et al., 2018) to determine whether an age-stratified checklist would be appropriate for Korean speakers. J. H. Park, Choi, et al., 2025 found the noun “sister-in-law” and the verb “split” as distinguishing the young group from both middle-aged and older adults, and the noun “younger brother” as distinguishing middle-aged adults from older adults. These findings align with prior work (H. Kim et al., 2019; H. Kim, Schoemann, & Wright, 2022) emphasizing the need for age-specific CL checklists by word class rather than a single unified list.

Building on these age-stratified Korean CL lists, the primary aim of this study was to determine whether CL word class distribution (nouns vs. verbs) differentiates PWA from CHAs. This serves as a test of construct validity, assessing whether a specific word class (nouns vs. verbs) is more sensitive to group differences, in line with theoretical expectations about verb vulnerability in PWA. A secondary aim was to examine whether CL usage predicts aphasia severity, as measured by AQ. This addresses criterion-related validity, which evaluates the extent to which CL measures are significantly associated with an established clinical indicator of language impairment.

Method

Participant

A total of 72 individuals participated in the study, including 36 PWA and 36 age- and education-matched CHAs. All participants were native Korean speakers and provided written informed consent prior to participation. The study was approved by the institutional review board of Ewha Womans University (No. 2022–0140).

PWA met the following inclusion criteria: (a) diagnosis of aphasia secondary to a left-hemisphere cortical or subcortical stroke; (b) the aphasia classification was based on the Paradise-Korean version of the Western Aphasia Battery–Revised (H. Kim & Na, 2012); (c) no prior history of brain injury or other neurological conditions before stroke onset; (d) no difficulties with vision, hearing, or color perception on a health screening questionnaire

(Christensen et al., 1991); and (e) production of at least one meaningful word relevant to each picture scene, consistent with the minimum utterance criterion outlined in Korean clinical guidelines (H. Kim et al., 1998). This criterion ensured that discourse samples reflected a basic level of lexical–semantic engagement while permitting inclusion across a broad range of aphasia severity.

Detailed lesion location data were not available for all participants due to community-based recruitment. All individuals, however, had a confirmed clinical diagnosis of aphasia secondary to a left-hemisphere cortical or subcortical stroke, as verified through medical records and/or documentation from referring clinicians. Three participants met criteria for latent aphasia, with AQ scores of 94.3, 95.0, and 96.4 (all with 16 years of education). These individuals were included in the sample. Detailed clinical characteristics are provided in the Appendix.

All CHAs scored within the normal range (i.e., above the 16th percentile, adjusted for age and education) on the Korean version of the Mini-Mental State Examination (Kang, 2006) and the Seoul Verbal Learning Test–Immediate Recall (Kang et al., 2012). They reported no history of neurological or psychiatric disorders and no difficulties with vision, hearing, or color perception on a health screening questionnaire (Christensen et al., 1991). Group demographic characteristics are presented in Table 1.

Task

Participants completed a storytelling task based on a culturally familiar picture stimulus depicting the traditional Korean folktale Heungbu and Nolbu (*Bad Older Brother and Good Younger Brother*). This story is widely recognized across age groups in Korea (S. Choi et al., 2022) and is frequently employed in both research and clinical contexts

(H. S. Kim & Sung, 2022; S. N. Kim & Shin, 2024; J. H. Park & Choi, 2024; see Figure 1). The picture stimuli used in this study are the same as those described in S. Choi et al. (2022); readers are referred to that publication for a visual representation of the stimuli.

The stimulus consisted of eight black-and-white illustrations arranged in a fixed sequence, visually depicting key events in the story. Heungbu (good younger brother) is expelled from home by his greedy brother Nolbu (bad older brother); driven by hunger, he visits Nolbu's house and is beaten with a rice paddle. After being turned away, Heungbu tends to a swallow with a broken leg. In return, the swallow brings him a gourd seed, which grows into a magical gourd filled with gold, silver, and treasures, making Heungbu's family wealthy. Nolbu, attempting to replicate this outcome, deliberately injures a swallow and obtains a gourd seed, but the resulting gourd releases goblins who destroy his house. The story concludes with Nolbu in poverty, repenting, and reconciling with Heungbu.

Elicitation Procedure

The task was administered using the following instructions: “I will now show you a series of pictures. Please tell the story of Heungbu and Nolbu while looking at the pictures.” The eight picture scenes were presented sequentially on a monitor. Participants were encouraged to produce a natural and coherent version of the story while viewing the pictures, with sufficient time provided for extended and uninterrupted speech production.

Transcription and Reliability

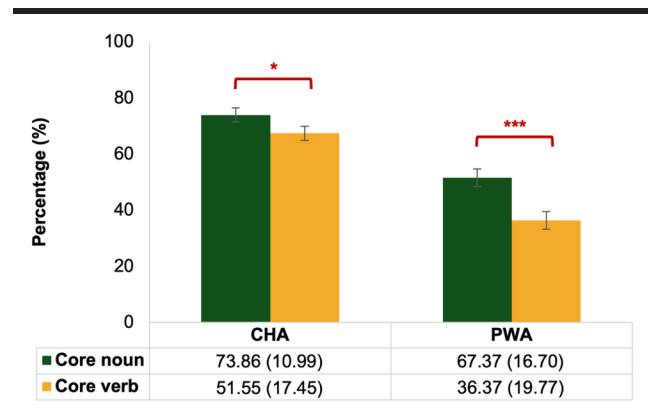
All storytelling responses were audio-recorded and subsequently transcribed verbatim by three certified Korean

Table 1. Demographic characteristics of each group.

Characteristics	PWA (<i>n</i> = 36)	CHAs (<i>n</i> = 36)	<i>F</i>	<i>p</i>
Gender (male:female)	20:16	20:16	—	—
Age (yrs) [range]	51.67 (17.02) [24–82]	51.61 (16.16) [26–77]	0.000	.989
Young	31.42 (5.18) [24–39]	32.25 (4.81) [24–38]	0.167	.687
Middle-aged	53.33 (5.33) [44–59]	52.83 (4.32) [43–59]	0.064	.803
Old	70.25 (6.24) [62–82]	69.75 (4.78) [64–77]	0.051	.824
Education (yrs) [range]	13.78 (2.73) [6–18]	13.72 (1.56) [11–16]	0.011	.916
K-MMSE (/30) [range]	—	28.78 (1.25) [25–30]		
SVLT-IM (/36) [range]	—	21.92 (4.04) [14–31]		
AQ (/100) [range]	76.44 (13.04) [16–97]	—	—	—

Note. Values are presented as mean (*SD*). Em dashes indicate data not applicable. PWA = persons with aphasia; CHAs = cognitively healthy adults; yrs = years; K-MMSE = Korean version of the Mini-Mental State Examination (Kang, 2006); AQ = Aphasia Quotient from the Paradise-Korean version of the Western Aphasia Battery–Revised (H. Kim & Na, 2012); SVLT-IR = Seoul Verbal Learning Test–Immediate Recall (Kang et al., 2012).

Figure 1. Percent agreement for nouns and verbs in cognitively healthy adults (CHAs) and persons with aphasia (PWA). The CHA group (green bars) produced a significantly higher proportion of core items than the PWA group (yellow bars) across both word classes. Notably, the group difference was greater for core verbs than for core nouns, as indicated by a significant interaction effect ($p < .05$). Error bars represent standard errors. Mean (*SD*) values for each condition are shown below the bars.



speech-language pathologists trained in discourse analysis. To establish transcription reliability, 10% of the samples were randomly selected for independent retranscription to calculate inter- and intrarater reliability across the entire transcription. Intrareliability was 96%, and interreliability was 97%. CL scoring was conducted manually by three trained raters using the finalized transcripts. A separate random 10% of samples were selected to evaluate scoring reliability. Both intra- and interrater agreement for CL scoring reached 98%.

CL Scoring Procedures

This study employed age-specific CL checklists developed by J. H. Park, Choi, et al. (2025), based on discourse samples elicited using the same Heungbu and Nolbu storytelling task. Following the procedure outlined in J. H. Park, Choi, et al. (2025), we applied a 40% cutoff to identify core lexical items—meaning that a word had to be produced by at least 40% of participants to be included in the checklist. Although previous studies typically used a 50% threshold (e.g., Dalton & Richardson, 2015) or an arbitrarily determined number of lexical items (e.g., Chen & Chang, 2024; H. Kim et al., 2019), applying this to Korean speakers resulted in quite limited verb sets (e.g., only six to seven core verbs per age group). The 40% threshold provided a more balanced and representative CL (e.g., nine to 12 core verbs per group), better reflecting the distribution of commonly produced lexical items in Korean discourse.

Percentage scores for CL items were calculated using these checklists (see Table 2). For each participant, the

number of core nouns (or verbs) produced was divided by the total number of core nouns (or verbs) on the checklist and multiplied by 100. This was computed separately for nouns and verbs to provide a standardized measure of lexical informativeness. Scoring was based on the written transcripts.

Statistical Analysis

The statistical analyses were designed to address the study's two primary aims: (1) to examine whether CL word class distributions—specifically, core nouns versus core verbs—differ between PWA and CHAs, thereby evaluating construct validity, and (2) to determine whether CL performance predicts aphasia severity, providing evidence for criterion-related validity. Analyses for Aim 1 were conducted using IBM SPSS Statistics (Version 29.0.2.0; IBM Corp., 2022), whereas analyses for Aim 2 were conducted using RStudio (Version 2022.07.0 Build 548; RStudio, PBC).

To address the first aim, a two-way mixed analysis of variance (ANOVA) was conducted with participant group (PWA vs. CHAs) as the between-subjects factor and word class (core nouns vs. core verbs) as the within-subject factor. The dependent variable was the percentage of CL items produced. Prior to analysis, assumptions for parametric tests were evaluated. Normality of core noun and verb percentages was assessed separately for each group using Shapiro–Wilk tests, along with visual inspection of Q–Q plots. Homogeneity of variance was examined using Levene's tests. Although minor deviations from normality were observed in the PWA group, Q–Q plots and box plots indicated no severe violations or influential outliers. Given the balanced sample sizes and the robustness of ANOVA to modest normality violations, parametric tests were retained.

To address the second aim, Pearson's correlation coefficients were calculated within the PWA group to assess the relationship between the percentages of core nouns and verbs and AQs. To further evaluate the relative contributions of each word class, multiple regression analyses were conducted. Four regression models were estimated: a demographic-only baseline model, a demographics plus core verb model, a demographics plus core noun model, and a combined model including both core verb and core noun percentages. Age and education were included in all models as demographic covariates. Model fit was evaluated using R^2 , adjusted R^2 , Akaike information criterion (AIC), and Bayesian information criterion (BIC). Multicollinearity diagnostics, including variance inflation factor (VIF) and tolerance values, were additionally examined to assess potential collinearity among predictors.

Table 2. Core noun and verb checklists for each age group.

A. Young						
Rank	Core nouns	N	%	Core verbs	N	%
1	Nolbu (bad older brother)	60	100.00	Fix	58	96.67
2	Heungbu (good younger brother)	60	100.00	Break	52	86.67
3	Gourd	57	95.00	Expel	50	83.33
4	Treasure	56	93.33	Go out	49	81.67
5	Leg	55	91.67	Hit	35	58.33
6	Goblin	53	88.33	Bring back	33	55.00
7	Wife	53	88.33	Injure	32	53.33
8	Scoop	50	83.33	Open	29	48.33
9	Family	46	76.67	Go (come)	27	45.00
10	Swallow	45	75.00			
11	Gourd seed	44	73.33			
12	House	43	71.67			
13	Rice	35	58.33			
14	Cheek	34	56.67			
B. Middle-aged						
1	Nolbu (bad older brother)	60	100.00	Break	53	88.33
2	Heungbu (good younger brother)	60	100.00	Expel	52	86.67
3	Leg	59	98.33	Fix	51	85.00
4	Treasure	58	96.67	Go out	41	68.33
5	Goblin	57	95.00	Bring back	46	76.67
6	Gourd	56	93.33	Go (come)	41	68.33
7	Scoop	54	90.00	Hit	35	58.33
8	House	52	86.67	Open	34	56.67
9	Swallow	51	85.00	Listen	29	48.33
10	Gourd seed	48	80.00	Injure	27	45.00
11	Family	45	75.00	Seek out	25	41.67
12	Wife	43	71.67	Look	25	41.67
13	Rice	38	63.33			
14	Cheek	34	56.67			
15	Older brother	30	50.00			
C. Older						
1	Gourd	60	100.00	Expel	51	85.00
2	Treasure	60	100.00	Fix	51	85.00
3	Nolbu (bad older brother)	59	98.33	Break	50	83.33
4	Heungbu (good younger brother)	59	98.33	Go out	48	80.00
5	Swallow	55	91.67	Go (come)	45	75.00
6	Leg	55	91.67	Bring back	38	63.33
7	Scoop	55	91.67	Hit	35	58.33
8	Gourd seed	43	71.67	Split	33	55.00
9	House	43	71.67	Grow	27	45.00
10	Goblin	41	68.33	Plant	25	41.67
11	Wife	39	65.00			
12	Rice	38	63.33			
13	Cheek	37	61.67			
14	Family	33	55.00			
15	Older brother	28	46.67			
16	Sister-in-law	27	45.00			
17	Younger brother	24	40.00			

Note. N = number of participants (max = 60).

Result

Group Differences in Core Noun and Verb Percentages

To evaluate construct validity, a two-way mixed ANOVA was conducted on CL percentages, with group (PWA vs. CHAs) as the between-subjects factor and word class (core nouns vs. core verbs) as the within-subject factor. There was a main effect of group, $F(1, 70) = 61.217, p < .001, \eta^2 = .467$, and a main effect of word class, $F(1, 70) = 32.562, p < .001, \eta^2 = .317$. These main effects were further qualified by a significant Group $\times$ Word Class interaction, $F(1, 70) = 5.225, p = .025, \eta^2 = .069$, indicating that the difference between core noun and core verb percentages was larger in PWA than in CHAs. Post hoc analyses revealed that core nouns were produced at significantly higher percentages than core verbs in both groups. The effect was larger in PWA ($p < .001$, Cohen's $d = 0.956$, 95% CI [0.566, 1.347]) than in CHAs ($p = .023$, Cohen's $d = 0.397$, 95% CI [0.055, 0.734]; see Figure 1).

CL Production and Aphasia Severity in PWA

To assess criterion-related validity, we examined the relationship between CL production and aphasia severity in PWA. Pearson's correlation revealed significant positive correlations for both core nouns ($r = .407, p = .014$, 95% CI [.090, .649]) and core verbs ($r = .454, p = .005$, 95% CI [.147, .681]), indicating that higher CL percentages was associated with less severe aphasia (see Figure 2).

A series of hierarchical multiple regression analyses were conducted to examine whether core verb and noun production predicted AQs beyond demographic variables. Four regression models were estimated: a demographic-only baseline model (Model 0), a demographics plus core verb model (Model 1), a demographics plus core noun model (Model 2), and a combined model including both core verb and core noun percentages (Model 3). Adjusted R^2 , AIC, and BIC values were used for model comparison (see Table 3).

Model 0 showed limited explanatory power, adjusted $R^2 = .053, F(2, 33) = 1.986, p = .131$. In Model 1, the inclusion of core verb percentage significantly improved model fit, adjusted $R^2 = .318, F(3, 32) = 6.432, p = .001$. Core verb percentage was a significant positive predictor of AQs ($B = .350, \beta = .531, p < .001$, 95% CI [.158, .543]). Age was also a significant negative predictor ($B = -.293, \beta = -.382, p = .016$, 95% CI [-.527, -.059]), whereas education was not significant ($p = .575$).

In Model 2, the inclusion of core noun percentage also yielded a significant model, adjusted $R^2 = .270$,

$F(3, 32) = 5.316, p = .002$. Core noun percentage was a significant positive predictor of AQs ($B = .360, \beta = .480, p = .002$, 95% CI [.140, .590]). Age was a significant negative predictor ($B = -.320, \beta = -.420, p = .011$, 95% CI [-.565, -.075]), whereas education was not significant ($p = .888$).

In Model 3, both core verb and core noun percentages were entered simultaneously, resulting in the highest adjusted R^2 (.333), $F(4, 31) = 5.374, p = .001$. Core verb percentage remained a marginal predictor of AQs ($B = .246, \beta = .372, p = .053$, 95% CI [-.004, .495]), whereas core noun percentage was not significant ($B = .182, \beta = .244, p = .195$, 95% CI [-.099, .463]). Multicollinearity diagnostics indicated no evidence of problematic collinearity among predictors (all VIFs < 2.0 ; all tolerances $> .50$).

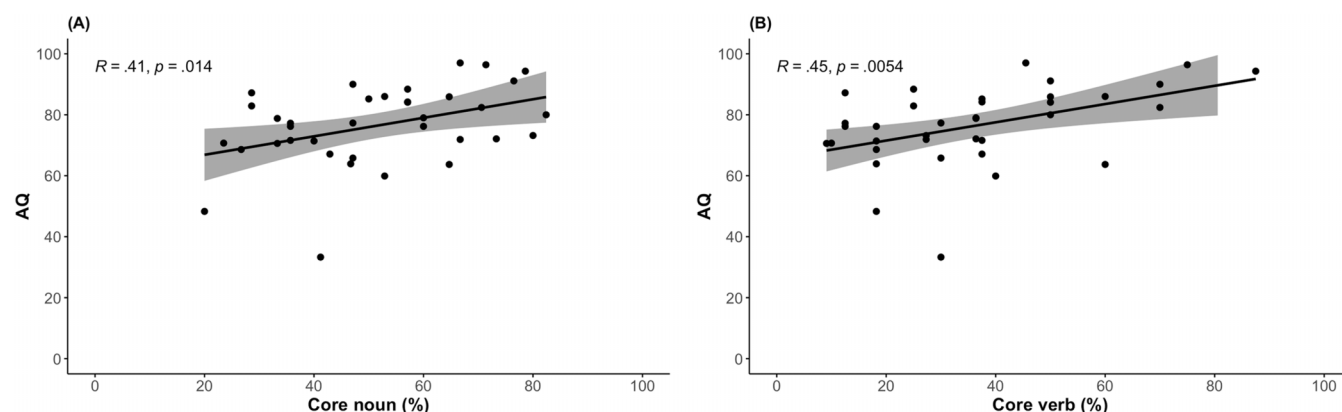
Overall, model comparisons using AIC and BIC indicated that Model 1 provided the best balance between model fit and parsimony (AIC = 279.06, BIC = 286.98). Although Model 3 yielded the highest adjusted R^2 , the inclusion of core noun percentage did not meaningfully improve predictive utility beyond core verb percentage alone, suggesting that core verb production is the stronger and more parsimonious predictor of AQs.

Discussion

The present study provides the first validation of age-stratified Korean CL checklists using a culturally familiar storytelling task with PWA, offering evidence for their clinical utility in assessing discourse production within a verb-salient language. First, a significant main effect of group indicated that PWA produced significantly lower percentage of CL items compared to CHAs. This finding supports the construct validity of the Korean CL lists and is consistent with prior research demonstrating reduced CL production in PWA (Chen & Chang, 2024; Dalton & Richardson, 2015; Jiang et al., 2023; H. Kim et al., 2019). Second, a significant main effect of word class revealed that verbs were produced less frequently than nouns across participants. This overall reduction in verb production is consistent with prior findings suggesting that verbs impose greater processing demands, including argument structure complexity and morphosyntactic integration (Berndt et al., 2002; Cruice et al., 2014).

Critically, the significant Group $\times$ Word Class interaction indicated that the difference between noun and verb production was greater in PWA than in CHAs, with both groups producing fewer core verbs than nouns but this pattern being more pronounced in PWA. These findings suggest that aphasia disproportionately exacerbates difficulty in verb production, amplifying general verb-related processing vulnerability rather than producing a qualitatively distinct lexical

Figure 2. Scatter plots depicting the relationship between core lexicon production and aphasia severity (Aphasia Quotient [AQ]) among individuals with aphasia. Panel A shows the correlation between core noun percentage and AQ ($r = .41, p = .014$). Panel B shows the correlation between core verb percentage and AQ ($r = .45, p = .005$). Shaded regions represent 95% confidence intervals around the regression line.



impairment. This pattern is broadly consistent with previous CL studies in English (Dalton & Richardson, 2015; H. Kim et al., 2019) and Mandarin (Chen & Chang, 2024; Jiang et al., 2023). For example, H. Kim et al. (2019) demonstrated that core verb production was more strongly associated with aphasia severity than other word classes, and Jiang et al. (2023) reported that Mandarin-speaking PWA performed significantly worse than CHAs on both core nouns and verbs, with verbs produced less frequently than nouns across multiple discourse tasks (picture description, procedural discourse, and storytelling). Additional studies further support these patterns. Dalton and Richardson (2015) reported qualitative difficulties in verb retrieval, showing that individuals with aphasia often rely on nonspecific “light” verbs rather than target lexical items, suggesting increased vulnerability in verb production. Similarly, Chen and Chang (2024) documented reduced CL production across word classes in PWA, indicating a more general lexical retrieval deficit. By directly examining the interaction between group and word class, the present study provides clearer evidence that PWA experience greater difficulty producing core verbs than core nouns, relative to CHAs, highlighting the specific vulnerability of verb retrieval in PWA.

To further examine the relationship between CL production and aphasia severity, we analyzed the association between core noun and verb production and AQs. Both core nouns and verbs were positively correlated with AQ, indicating that greater production of core lexical items was associated with milder aphasia severity. These findings suggest that CL production at the discourse level reflects overall language impairment in PWA. However, regression analyses controlling for demographic variables revealed a more specific pattern: Among the word-class variables examined in this study, core verb percentage

demonstrated a stronger and more consistent association with AQs than core noun percentage. This pattern is broadly consistent with prior work (e.g., H. Kim et al., 2019), which has also examined the role of core verb production in relation to aphasia severity in English-speaking PWA. Notably, their study also reported that core verbs were produced significantly more frequently in fluent than in nonfluent aphasia, further suggesting that verb production is particularly informative for predicting aphasia severity and differentiating aphasia subtypes (H. Park, Kim, & Kintz, 2026). By contrast, Jiang et al. (2023) found no significant correlation between core nouns or verbs and AQs in a Mandarin-speaking cohort. This discrepancy may be attributable to differences in aphasia subtype composition, severity, and sample size: Jiang et al. examined only anomic PWA ($n = 11$) with relatively mild impairment (average AQs = 82.5; range not reported), whereas H. Kim et al. (2019) included a mixed group of fluent and nonfluent PWA ($n = 11$; average AQs = 68.8; range: 54.4–90.7). The present study extended this line of inquiry with a larger sample ($n = 36$) with a broader severity range (average AQs = 76.38; range: 16–96.4) of Korean speakers, providing a broader basis for examining AQ-related variability.

The prominence of verbs as predictors of aphasia severity is particularly noteworthy given the typological characteristics of Korean. Korean is a verb-salient, pro-drop language in which verbs can function as complete utterances and carry substantial syntactic and semantic load (H. Lee, 2015; Sohn, 2001). Given this structural centrality, one might expect verb production to be relatively preserved or contextually supported. However, the present findings indicate that verbs remain especially vulnerable in aphasia, even within a linguistic system that foregrounds verbal morphology and argument structure. This vulnerability is all the more striking when considered against empirical

Table 3. Multiple regression models predicting aphasia severity.

Model 0: Demographics only						
Predictor	<i>B</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>	95% CI
Intercept	89.32	15.58	—	5.73	< .001	[57.62, 121.02]
Age	−0.25	0.14	−.33	−1.87	.071	[−0.53, 0.02]
Education	0.01	0.84	.00	0.01	.994	[−1.70, 1.71]
Model fit						
R^2	.107					
Adjusted R^2	.053					
<i>F</i>	1.99					
AIC	289.96					
BIC	296.29					
Model 1: Demographics + verb						
Intercept	73.21	13.92	—	5.26	< .001	[44.86, 101.57]
Age	−0.29	0.12	−.38	−2.55	.016	[−0.53, −0.06]
Education	0.41	0.72	.09	0.57	.575	[−1.06, 1.88]
Verb	0.35	0.09	.53	3.71	.001	[0.16, 0.54]
Model fit						
R^2	0.376					
Adjusted R^2	0.318					
<i>F</i>	6.43					
AIC	279.06					
BIC	286.98					
Model 2: Demographics + noun						
Intercept	76.04	14.27	—	5.33	< .001	[46.98, 105.10]
Age	−0.32	0.12	−.42	−2.70	.011	[−0.57, −0.08]
Education	−0.11	0.74	−.02	−0.14	.888	[−1.61, 1.40]
Noun	0.36	0.11	.48	3.29	.002	[0.14, 0.59]
Model fit						
R^2	0.333					
Adjusted R^2	0.270					
<i>F</i>	5.32					
AIC	281.49					
BIC	289.41					
Model 3: Demographics + verb + noun						
Intercept	71.32	13.84	—	5.16	< .001	[43.11, 99.54]
Age	−0.32	0.12	−.41	−2.76	.010	[−0.55, −0.08]
Education	0.23	0.73	.05	0.32	.751	[−1.25, 1.71]
Verb	0.25	0.12	.37	2.01	.053	[−0.00, 0.50]
Noun	0.18	0.14	.24	1.32	.195	[−0.10, 0.46]
Model fit						
R^2	0.409					
Adjusted R^2	0.333					
<i>F</i>	5.37					
AIC	279.08					
BIC	288.58					

Note. *B* = unstandardized coefficient; *SE* = standard error; β = standardized coefficient; CI = confidence interval; AIC = Akaike information criterion; BIC = Bayesian information criterion.

evidence of the prominence of verbs in Korean discourse production. Sung et al. (2016), for example, reported that Korean PWA produced more verbs per utterance than English PWA, reflecting the structural prominence of verbs

in Korean. In addition, J. Park et al. (2024) demonstrated age-related differences in single- and serial-verb production using the same storytelling task employed in the present study, linking verb production to semantic memory capacity.

Taken together, these findings underscore the functional importance of verbs in Korean discourse. Against this backdrop, the disproportionate vulnerability of verbs observed in the present study suggests that verb production constitutes a particularly sensitive locus of breakdown in Korean aphasia, underscoring the value of word class–based CL analysis in typologically verb-prominent languages.

The findings of this study carry several important clinical implications, particularly for discourse assessment in Korean PWA. First, the newly developed Korean CL checklists offer a clinically feasible and sensitive tool for detecting linguistic impairments in PWA using discourse tasks. Unlike conventional Korean discourse assessments, which rely heavily on metrics such as CIUs, the CL checklists enable clinicians to capture lexical-semantic access within ecologically valid, connected speech, offering richer and more functionally relevant insights into a speaker's communicative abilities. Second, the results provide empirical support for the construct validity of the CL framework in Korean, a verb-salient language. Our analysis showed that core verbs were produced significantly less frequently than core nouns across both groups and that this gap was greater in PWA. Importantly, core verb production alone significantly predicted AQs, emphasizing the central role of verbs in Korean discourse structure. These findings suggest that clinicians working with Korean-speaking individuals should pay particular attention to verb retrieval in discourse-based assessments and intervention. Third, while the approach has been explored in other languages (e.g., English, Mandarin), this study extends its application to Korean and demonstrates the need to adapt assessment tools to the typological and grammatical characteristics of the target language. Assessment frameworks developed in noun-prominent languages such as English may not fully capture clinically meaningful deficits in verb-salient languages such as Korean. Clinicians should therefore exercise caution when applying English-based discourse tools in non-English contexts, especially in languages where verbs serve a central organizing role in sentence construction and information packaging. Finally, because the CL checklists were developed using a culturally familiar narrative, the current approach demonstrates the value of using discourse stimuli that are contextually relevant to the speaker population. Clinicians can draw on such materials to elicit more naturalistic and representative language samples in both assessment and treatment, particularly when targeting discourse-level lexical access.

Despite these contributions, several limitations should be acknowledged. First, the sample included mixed aphasia subtypes without sufficient numbers to examine subtype-specific performance. Future studies should recruit larger and more clinically diverse samples with well-characterized aphasia profiles (e.g., fluent vs. nonfluent) to explore whether CL measures differentially capture deficits across subtypes, as

suggested by prior research (e.g., Richardson et al., 2018). Second, the current CL checklists are limited to nouns and verbs, capturing only a subset of word classes. Expanding the checklists to include additional categories such as adjectives, adverbs, and function words would allow for a more comprehensive profile of discourse (e.g., H. Kim et al., 2019, 2021). Third, as this is the first application of CL framework in Korean, further research is needed to replicate and extend these findings across tasks, contexts, and clinical populations.

Future directions should prioritize robust validation of the Korean CL through systematic psychometric evaluation. Specifically, examining concurrent validity between the CL and established linguistic measures, as well as assessing test–retest reliability in Korean-speaking individuals with aphasia, will strengthen the empirical foundation of this assessment tool. Building on this validation work, the Korean CL could then be systematically extended to other clinical populations—mirroring the trajectory of CL application across diverse English-speaking clinical groups, including individuals with mild cognitive impairment and dementia (e.g., Kintz et al., 2024; Stark et al., 2025), while simultaneously investigating treatment sensitivity to discourse-level lexical changes. Furthermore, adapting the CL for use in spontaneous and conversational discourse contexts will enhance ecological validity and clinical utility. One possible approach is to develop context-general CL lists that are not tied to a single discourse task or topic but are instead derived from lexical items that occur consistently across a range of discourse contexts. For example, high-frequency light verbs such as *do*, *take*, *have*, *make*, and *go*, for example, may serve as potential cross-context lexical targets because they support many everyday expressions and are likely to appear across diverse communicative situations. Selected function words could also be considered as part of an expanded framework for evaluating discourse performance across contexts. This approach may allow CL analysis to be applied more flexibly to spontaneous or conversational discourse while maintaining the practical advantages of checklist-based scoring. Importantly, future expansion of the CL framework should remain grounded in language-specific development using typologically appropriate and culturally situated tasks, rather than relying solely on translated or generalized English-language protocols. This approach may help ensure that subsequent validation efforts remain contextually meaningful and theoretically coherent across linguistic and cultural contexts.

Conclusions

The present study provides the first empirical validation of age-stratified Korean CL checklists in a verb-salient language, applying them to PWA through a culturally familiar storytelling task. The findings demonstrate that developing discourse tools aligned with language's

linguistic structure and cultural context enhances their clinical utility and interpretability. In particular, regression analyses controlling for demographic variables showed that core verb production demonstrated a stronger and more consistent association with aphasia severity than core noun production, underscoring the importance of language-specific assessment targets in languages where verbs serve as central elements of utterances. Overall, these findings underscore the importance of tailoring assessment frameworks to the linguistic and cultural characteristics of the target population and pave the way for more inclusive, globally relevant approaches to aphasia assessment and treatment.

Author Contributions

Jimin Park: Formal analysis, Writing – original draft. **Sujin Choi:** Data curation, Writing – original draft. **So Eun Kim:** Data curation, Writing – original draft. **Hana Kim:** Conceptualization, Methodology, Supervision, Writing – review & editing. **Jee Eun Sung:** Conceptualization, Methodology, Supervision, Writing – review & editing.

Data Availability Statement

Further information can be requested from the corresponding author, subject to institutional and ethical approval.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grants funded by the Korean government Ministry of Science and ICT (MSIT) in 2026 (RS-2022-NR070151, RS-2024-00461617, NRF-2026S1A5A2A03006789; RS-2024-00461617, Project Title: Global Convergent Research on Healthcare Solutions to Overcome Hearing Loss and Dementia, 50%; RS-2022-NR070151. Project Title: Development of language-hearing program and integrative neural network studies for early assessment and treatment of hearing loss-induced dementia, 25%; NRF-2026S1A5A2A03006789, Project Title: Cross-linguistic and multicultural aspects of language processing in aging and neurogenic communication disorders: Evidence from Korean, Chinese, English, and Spanish, 25%; awarded to Jee Eun Sung). This work was supported by Research Initiative funds by the University of South Florida (awarded to Hana Kim).

References

Berndt, R. S., Burton, M. W., Haendiges, A. N., & Mitchum, C. C. (2002). Production of nouns and verbs in aphasia:

- Effects of elicitation context. *Aphasiology*, 16(1–2), 83–106. <https://doi.org/10.1080/02687040143000212>
- Bryant, L., Spencer, E., & Ferguson, A. (2017). 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>
- 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
- Chen, J., & Chang, H. (2024). Core lexicon analysis of spoken discourse production by Mandarin Chinese speakers with aphasia. *Aphasiology*, 38(11), 1766–1785. <https://doi.org/10.1080/02687038.2024.2325169>
- Choi, H. (2022). Correlation between ratios of word-finding behavior and CIU in patients with mild cognitive impairment and dementia of Alzheimer's type. *Communication Sciences & Disorders*, 27(3), 617–628. <https://doi.org/10.12963/csd.22910>
- Choi, S., Lee, H., Jo, E., Lim, Y., Choi, Y., Han, J., & Sung, J. E. (2022). Preliminary procedures to develop storytelling-based assessment for aging and neurogenic disorders. *Communication Sciences & Disorders*, 27(3), 589–605. <https://doi.org/10.12963/csd.22907>
- Christensen, K. J., Multhaup, K. S., Nordstrom, S., & Voss, K. (1991). A cognitive battery for dementia: Development and measurement characteristics. *Psychological Assessment*, 3(2), 1681–174. <https://doi.org/10.1037/1040-3590.3.2.168>
- 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 & Communication Disorders*, 55(3), 417–442. <https://doi.org/10.1111/1460-6984.12528>
- 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>
- Dalton, S. G., Al Harbi, M., Berube, S., & Hubbard, H. I. (2024). Development of main concept and core lexicon checklists for the original and modern Cookie Theft stimuli. *Aphasiology*, 38(12), 1975–1999. <https://doi.org/10.1080/02687038.2024.2340794>
- Dalton, S. G. H., Kim, H., Richardson, J. D., & Wright, H. H. (2020). A compendium of core lexicon checklists. *Seminars in Speech and Language*, 41(1), 45–60. <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
- 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
- Day, A. (1985). *Good dog, Carl*. Simon & Schuster. <https://www.simonandschuster.com/books/Good-Dog-Carl/Alexandra-Day/9780689817717>
- Degani, M., & Onysko, A. (2024). Cultural variation in New Zealand English stories about place. *World Englishes*, 43(3), 399–413. <https://doi.org/10.1111/weng.12659>
- Dipper, L., & Pritchard, M. (2017). Discourse: Assessment and therapy. In F. D. M. Fernandes (Ed.), *Advances in speech-*

- language pathology. IntechOpen. <https://doi.org/10.5772/intechopen.69894>
- Duncan, E. A. S., & Murray, J. (2012). The barriers and facilitators to routine outcome measurement by allied health professionals in practice: A systematic review. *BMC Health Services Research*, 12(1), Article 96. <https://doi.org/10.1186/1472-6963-12-96>
- Dutta, M., & Mohapatra, B. (2024). Expanding the scope: Multimodal dimensions in aphasia discourse analysis—Preliminary findings. *Frontiers in Human Neuroscience*, 18, Article 1419311. <https://doi.org/10.3389/fnhum.2024.1419311>
- 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. A., Forbes, M., Holland, A., & MacWhinney, B. (2013, May 28–June 2). *PWAs and PBJs: Language for describing a simple procedure* [Paper presentation]. Clinical Aphasiology Conference, Tucson, AZ, USA. <http://aphasiology.pitt.edu/id/eprint/2491>
- IBM Corp. (2022). *IBM SPSS Statistics* (Version 29.0.2.0) [Computer software]. <https://www.ibm.com/products/spss-statistics>
- Jiang, Y. E., Liao, X. Y., & Liu, N. (2023). Applying core lexicon analysis in patients with anomia aphasia: Based on Mandarin AphasiaBank. *International Journal of Language & Communication Disorders*, 58(6), 1875–1886. <https://doi.org/10.1111/1460-6984.12864>
- Jin, C., Choi, H., & Lee, J.-Y. (2016). Usefulness of spontaneous speech analysis scales in patients with mild cognitive impairment and dementia of Alzheimer's type. *Communication Sciences & Disorders*, 21(2), 284–294. <https://doi.org/10.12963/csd.16310>
- Kang, Y. (2006). A normative study of the Korean-Mini Mental State Examination (K-MMSE) in the elderly. *Korean Journal of Psychology*, 25(2), 1–12. https://uci.kci.go.kr/resolution/result.do?res_cd=G704-001037.2006.25.2.002
- Kang, Y., Jang, S., & Na, D. (2012). *Seoul Neuropsychological Screening Battery—Second Edition (SNSB-II)*. Human Brain Research & Consulting Co.
- Kertesz, A. (2007). *Western Aphasia Battery—Revised (WAB-R)*. Pearson. <https://www.pearsonassessments.com/en-us/Store/Professional-Assessments/Speech-%26-Language/Western-Aphasia-Battery-Revised/p/100000194>
- Kim, H., Berube, S., & Hillis, A. E. (2023). Core lexicon in aphasia: A longitudinal study. *Aphasiology*, 37(10), 1679–1691. <https://doi.org/10.1080/02687038.2022.2121598>
- Kim, H., & Kintz, S. (2024). Discourse production across the adult lifespan: Microlinguistic processes. *Topics in Cognitive Science*. Advance online publication. <https://doi.org/10.1111/tops.12729>
- Kim, H., Kintz, S., & Wright, H. H. (2021). Development of a measure of function word use in narrative discourse: Core lexicon analysis in aphasia. *International Journal of Language & Communication Disorders*, 56(1), 6–19. <https://doi.org/10.1111/1460-6984.12567>
- Kim, H., Kintz, S., Zelnosky, K., & Wright, H. H. (2019). Measuring word retrieval in narrative discourse: Core lexicon in aphasia. *International Journal of Language & Communication Disorders*, 54(1), 62–78. <https://doi.org/10.1111/1460-6984.12432>
- Kim, H., Kwon, M., Na, D. L., Choi, S., Lee, K., & Chung, C. S. (1998). Decision making in fluency measures of aphasic spontaneous speech. *Communication Sciences & Disorders*, 3, 5–19.
- Kim, H., & Na, D. L. (2012). *Paradise Korean Version of the Western Aphasia Battery Revised (PK-WAB-R)*. Paradise Welfare Foundation.
- Kim, H., Schoemann, A. M., & Wright, H. H. (2022). Quality of measurement in core lexicon measures. *Journal of Speech, Language, and Hearing Research*, 65(8), 2919–2930. https://doi.org/10.1044/2022_JSLHR-20-00722
- Kim, H., Sung, J. E., & Jeong, J. H. (2022). Non-transcription analysis of connected speech in mild cognitive impairment using an information unit scoring system. *Journal of Neurolinguistics*, 61, Article 101035. <https://doi.org/10.1016/j.jneuroling.2021.101035>
- Kim, H., & Wright, H. H. (2020a). A tutorial on core lexicon: Development, use, and application. *Seminars in Speech and Language*, 41(1), 20–31. <https://doi.org/10.1055/s-0039-3400973>
- Kim, H., & Wright, H. H. (2020b). Concurrent validity and reliability of the core lexicon measure as a measure of word retrieval ability in aphasia narratives. *American Journal of Speech-Language Pathology*, 29(1), 101–110. https://doi.org/10.1044/2019_AJSLP-19-0063
- Kim, H. S., & Sung, J. E. (2022). Production abilities of predicative adjectives in Korean-speaking individuals with aphasia. *Journal of Speech-Language & Hearing Disorders*, 31(4), 9–25. <https://doi.org/10.15724/jslhd.2022.31.4.009>
- Kim, S. N., & Shin, S. (2024). Storytelling ability of patients with dementia of the Alzheimer's type as a function of auditory cues. *Journal of Speech-Language & Hearing Disorders*, 33(4), 11–20. <https://doi.org/10.15724/jslhd.2024.33.4.011>
- Kintz, S., & Kim, H. (2023). Discourse processing in older adults: Considering discourse elicitation tasks. In C. A. Coelho, L. R. Cherney, & B. B. Shadden (Eds.), *Discourse analysis in adults with and without communication disorders: A resource for clinicians and researchers* (pp. 33–49). Plural. <https://www.pluralpublishing.com/publications/discourse-analysis-in-adults-with-and-without-communication-disorders-a-resource-for-clinicians-and-researchers>
- Kintz, S., Kim, H., & Wright, H. H. (2024). A preliminary investigation on core lexicon analysis in dementia of the Alzheimer's type. *International Journal of Language & Communication Disorders*, 59(4), 1336–1350. <https://doi.org/10.1111/1460-6984.12999>
- Lee, G., Park, H., & Lee, G. (2018). A study on selection of basic vocabulary for Hangeul learning. *Journal of CheongRam Korean Language Education*, 65, 213–235. <https://doi.org/10.26589/jockle.65.201803.213>
- Lee, H. (2015). Information structure, topic predictability and gradients in Korean case ellipsis: A probabilistic account. *Linguistic Research*, 32(3), 749–771. <https://doi.org/10.17250/khisli.32.3.201512.009>
- Lee, H., Choi, Y., & Sung, J. E. (2024). Age-related changes in connected speech production: Evidence from eye-tracking in the culturally adapted picture description task. *Frontiers in Psychology*, 15, Article 1334788. <https://doi.org/10.3389/fpsyg.2024.1334788>
- Lee, S. J., Kim, H., Seo, S., & Kim, M. (2009). Patterns of word class production between picture description and narrative tasks in aphasia. *Communication Sciences & Disorders*, 14(4), 470–483. <https://www.e-csd.org/journal/view.php?number=412>
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk, Volume II: The database*. Psychology Press. <https://www.taylorfrancis.com/books/mono/10.4324/9781315805641/childes-project-brian-macwhinney>
- 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>
- MacWhinney, B., Fromm, D., Holland, A., Forbes, M., & Wright, H. (2010). Automated analysis of the Cinderella story. *Aphasiology*, 24(6–8), 856–868. <https://doi.org/10.1080/02687030903452632>

- Maddy, K., Howell, D. M., & Capilouto, G. J. (2015). Current practices regarding discourse analysis and treatment following non-aphasic brain injury. *Journal of Interactional Research in Communication Disorders*, 6(2), 211–236. <https://doi.org/10.1558/jircd.v6i2.25519>
- Mazumdar, B., Donovan, N. J., & Duncan, E. S. (2023). Identifying an appropriate picture stimulus for a Bangla picture description task. *Journal of Speech, Language, and Hearing Research*, 66(4), 1334–1350. https://doi.org/10.1044/2022_jslhr-22-00152
- McCully, E. A. (1984). *Picnic*. HarperCollins. <https://www.harpercollins.com/products/picnic-emily-arnold-mccully>
- Menn, L., Reilly, K. F., Hayashi, M., Kamio, A., Fujita, I., & Sasanuma, S. (1998). The interaction of preserved pragmatics and impaired syntax in Japanese and English aphasic speech. *Brain and Language*, 61(2), 183–225. <https://doi.org/10.1006/brln.1997.1838>
- Nicholas, L. E., & Brookshire, R. H. (1995). Presence, completeness, and accuracy of main concepts in the connected speech of non-brain-damaged adults and adults with aphasia. *Journal of Speech and Hearing Research*, 38(1), 145–156. <https://doi.org/10.1044/jshr.3801.145>
- Park, E. L., & Cho, S. (2014). KoNLPy: Korean natural language processing in Python. In *Proceedings of the 26th Annual Conference on Human & Cognitive Language Technology* (pp. 133–136).
- Park, H., Kim, H., & Kintz, S. (2026). Heavy and light verb production in people with aphasia: Evidence from frequency rate and core lexicon approaches. *Aphasiology*, 40(9), 1659–1670. <https://doi.org/10.1080/02687038.2025.2519328>
- Park, J. H., & Choi, H. (2024). Characteristics of word-finding behaviors in healthy elderly adults according to the type of spontaneous speech task. *Communication Sciences & Disorders*, 29(2), 323–338. <https://doi.org/10.12963/csd.240019>
- Park, J., Choi, S., Kim, S. E., Kim, H., & Sung, J. E. (2024, October 18–20). *Comparative analysis of predicate types on young and older adults in a predicate-final language* [Paper presentation]. 62nd Academy of Aphasia 2024, Nara, Japan.
- Park, J., Choi, S., Kim, S. E., Kim, H., & Sung, J. E. (2025). Development of core lexicon lists for cognitively healthy adults across age groups in a Korean storytelling task. *Communication Sciences & Disorders*, 30(3), 579–592. <https://doi.org/10.12963/csd.250116>
- 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. H., 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
- Sohn, H.-M. (2001). *The Korean language*. Cambridge University Press.
- 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, Article 1500735. <https://doi.org/10.3389/fnhum.2024.1500735>
- 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
- Sung, J. E., DeDe, G., & Lee, S. E. (2016). Cross-linguistic differences in a picture-description task between Korean- and English-speaking individuals with aphasia. *American Journal of Speech-Language Pathology*, 25(4S), S813–S822. https://doi.org/10.1044/2016_AJSLP-15-0140
- Sung, J. E., Lee, S. J., & Eom, B. (2019). Aging-related differences in the resolution of ambiguity from case marker deletions in a verb-final language. *Communication Sciences & Disorders*, 24(3), 695–706. <https://doi.org/10.12963/csd.19642>
- Sung, J. E., Scimeca, M., Li, R., & Kiran, S. (2024). Cross-linguistic and multicultural considerations in evaluating bilingual adults with aphasia. *American Journal of Speech-Language Pathology*, 33(6), 2716–2731. https://doi.org/10.1044/2024_ajslp-23-00496
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1), 267–288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- Tibshirani, R. J., Rinaldo, A., Tibshirani, R., & Wasserman, L. (2018). *Uniform asymptotic inference and the bootstrap after model selection*. <https://doi.org/10.1214/17-AOS1584>
- Ulatowska, H. K., Olness, G. S., & Williams, L. J. (2004). Coherence of narratives in aphasia. *Brain and Language*, 91(1), 42–43. <https://doi.org/10.1016/j.bandl.2004.06.024>
- Verna, A., Davidson, B., & Rose, T. (2009). Speech-language pathology services for people with aphasia: A survey of current practice in Australia. *International Journal of Speech-Language Pathology*, 11(3), 191–205. <https://doi.org/10.1080/17549500902726059>
- Wallace, S. J., Worrall, L., Rose, T., Le Dorze, G., Cruice, M., Isaksen, J., Kong, A. P. H., 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>
- Worrall, L., Sherratt, S., Rogers, P., Howe, T., Hersch, D., Ferguson, A., & Davidson, B. (2011). What people with aphasia want: Their goals according to the ICF. *Aphasiology*, 25(3), 309–322. <https://doi.org/10.1080/02687038.2010.508530>
- Wright, H. H. (2011). Discourse in aphasia: An introduction to current research and future directions. *Aphasiology*, 25(11), 1283–1285. <https://doi.org/10.1080/02687038.2011.613452>
- Yoo, Y. R., Park, J., & Sung, J. E. (2026). Cross-linguistic and cross-cultural effects on picture description: A main concept analysis in Korean and English speakers with and without aphasia. *American Journal of Speech-Language Pathology*, 35(5S), 2544–2557. https://doi.org/10.1044/2026_AJSLP-25-00422

Appendix

Clinical Information of Persons With Aphasia

ID	Age	Sex	WAB-R AQ	Aphasia subtype
A1	38	M	84.2	Anomic
A2	74	M	33.3	Broca
A3	31	F	88.4	Anomic
A4	69	M	80	Anomic
A5	72	M	59.9	Transcortical sensory
A6	59	F	71.4	Anomic
A7	63	M	63.7	Wernicke
A8	38	M	85.2	Anomic
A9	79	M	85.9	Anomic
A10	57	F	71.9	Conduction
A11	59	M	48.3	Wernicke
A12	56	F	76.2	Anomic
A13	62	F	86	Anomic
A14	39	M	84.1	Anomic
A15	65	M	77.3	Transcortical motor
A16	74	F	65.8	Transcortical motor
A17	27	F	96.4	Latent
A18	25	F	82.9	Anomic
A19	65	M	90	Anomic
A20	59	M	78.8	Anomic
A21	35	M	67.1	Conduction
A22	82	F	82.4	Anomic
A23	45	F	63.9	Wernicke
A24	54	M	72.1	Anomic
A25	49	M	95	Latent
A26	56	M	68.6	Conduction
A27	44	F	79	Conduction
A28	50	M	73.2	Anomic
A29	52	F	70.6	Wernicke
A30	30	F	87.2	Anomic
A31	68	M	70.7	Transcortical sensory
A32	70	F	91.1	Anomic
A33	28	M	77.24	Anomic
A34	29	M	71.6	Conduction
A35	33	F	94.3	Latent
A36	24	F	76.2	Anomic

Note. Due to community-based recruitment, detailed lesion location data were not available for all participants. However, all individuals included in this study had a confirmed clinical diagnosis of aphasia resulting from a left-hemisphere cortical or subcortical stroke, as verified through medical records and/or documentation from referring clinicians. WAB-R = Western Aphasia Battery–Revised; AQ = Aphasia Quotient; A = aphasia; M = male; F = female.