

Original Article

Development of a Discourse Comprehension Test for the Older Adults: A Preliminary Standardization Focusing on Reliability and Validity

Hyun Ju Lee¹, Sung Ji Park^{2*}

¹ PhD Student, Department of Secondary Special Education, Wonkwang University, Republic of Korea

² Professor, Department of Secondary Special Education, Wonkwang University, Republic of Korea

*Corresponding Author : psungji@wku.ac.kr

ABSTRACT

Received 05 Dec 2025

Revised 21 Dec 2025

Accepted 23 Dec 2025

Citation: Lee HJ, Park SJ. Development of a Discourse Comprehension Test for the Older Adults: A Preliminary Standardization Focusing on Reliability and Validity. *The Journal of Transdisciplinary Studies*. 2025; 9(3): 185-191. <https://doi.org/10.22685/jts.2025.9.3.185>

Copyright: © 2025 by The Journal of Transdisciplinary Studies. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>)

Objectives: This study is a preliminary investigation aimed at developing and standardizing a tool for assessing discourse comprehension in the elderly. The focus was on establishing the initial reliability and validity of 37 newly developed items as a foundational step for full-scale standardization. **Methods:** A set of 37 test items, developed through a systematic content validity process, was administered to a preliminary sample of 30 elderly individuals (aged 65+) in Sejong and Jeonbuk regions. Internal consistency (Cronbach's α), test-retest reliability, and criterion validity with the KOSECT were analyzed. **Results:** The total Cronbach's α was high at .876, and test-retest reliability was .993 ($p < .001$). Criterion validity showed a significant correlation ($r = .797$). **Conclusions:** The tool demonstrates strong psychometric properties. While limited by sample size, this study provides a crucial academic foundation for future large-scale standardization and clinical applications in cognitive-communication interventions.

Keywords: Discourse comprehension in older adults, Reliability, Validity

1. Introduction

1.1. Rationale for the study

Biological changes in hearing, vision, cognition, and the neuromuscular system that occur during the aging process can cause various forms of communication disorders. While classical studies focused on sensory degeneration—such as presbycusis—causing missed information in conversation [1], contemporary research emphasizes that these deficits extend beyond simple sensory loss. Recent findings suggest that age-related declines in cognitive-communication abilities significantly impair an individual's capacity to infer speaker intent and maintain social connectedness, which

are vital for life satisfaction [2]. Furthermore, the rapid digitalization of modern society has created an environment where processing high-speed auditory and textual information is essential; consequently, older adults with diminished communication efficiency face an increased risk of 'communication alienation' and social isolation [3].

1.2. Focus on discourse-level communication

Language comprehension ability progresses from the word level to the sentence level and then the discourse level. However, functional everyday communication relies primarily on discourse-level comprehension—the ability to integrate multiple sentences, grasp context, and extract

core themes [4]. Recent discourse studies highlight that while older adults may retain the grammatical ability to understand isolated sentences, they frequently struggle with global coherence and information synthesis in complex narrative or conversational settings [5]. This implies that speech rehabilitation for the elderly must shift its focus from basic linguistic accuracy to "Successful Communicative Competence" within real-world contexts.

1.3. Purpose of the study

For a therapeutic approach to individuals with communication difficulties, a series of processes must be followed: assessing the subject's language ability, setting therapeutic goals, establishing session plans accordingly, and implementing therapeutic intervention [6]. However, research on the communication of older adults as recipients of speech rehabilitation is still insufficient [7]. Domestic assessment tools for the communication abilities of older adults include the Paradise Korea - Western Aphasia Battery-Revised [8] for evaluating the language abilities of aphasia patients, and the Brief test of Cognitive-Communication Disorders [9] for a simple evaluation of language abilities in patients with cognitive-communication disorders. Despite the growing demand for intervention, there remains a critical shortage of assessment tools capable of objectively measuring discourse comprehension—the truest indicator of an older adult's functional communication skills [10].

Therefore, there is an urgent need to develop a specialized assessment tool focused on discourse comprehension abilities. Such a tool will not only allow for the early identification of communication disorders in the elderly but also provide the empirical data necessary to establish personalized, discourse-based speech therapy goals that facilitate active social participation.

This study is explicitly designed as a preliminary standardization study to develop and standardize a test tool capable of measuring discourse comprehension in older adults. The primary objective is to verify the validity and reliability of the items currently under development to finalize and refine the test questions for clinical use. Specifically, this study aims to address the following research questions:

- Research Question 1. What are the internal consistency reliability and test-retest reliability of the items for developing the Elderly Discourse Comprehension Test Tool?
- Research Question 2. What is the criterion validity of the items for developing the Elderly Discourse Comprehension Test Tool?

2. Materials and Methods

2.1. Participants and IRB approval

The study participants for the development of the discourse comprehension test tool for older adults were recruited from individuals aged 65 or older residing in Sejong City and Jeollabuk-do. Recruitment was conducted by posting notices on the L Language Learning Institute website and within W University, as well as distributing flyers at external community facilities such as religious centers, senior centers (gyeongrodang), and welfare centers for the elderly.

The inclusion criteria for the participants were as follows: 1) native speakers of Korean; 2) individuals aged 65 years or older; 3) no history of past neurological diseases (stroke, epilepsy, Parkinson's disease, multiple sclerosis, Huntington's disease, encephalitis, dementia) or psychiatric history; and 4) individuals who fell within the normal standard range on the Korean version of the Mini-Mental State Examination (K-MMSE) [11]. 4) Furthermore, participation was limited to those without significant vision or hearing problems that would interfere with test administration.

Potential participants underwent an initial interview with the researcher to confirm basic information. If their K-MMSE score was known, it was verified and recorded. If K-MMSE information was unavailable, the test was administered only after performing the K-MMSE first and confirming the participant scored within the normal range of 24 points or higher.

The study was conducted with the approval of the W University Institutional Review Board (IRB). All subjects were verbally informed about the research content and provided written consent before participating. The test administration was conducted in two ways.

2.2. Selection criteria for discourse materials

To ensure that the assessment tool accurately reflects the practical communication environment of older adults and possesses high content validity, the discourse materials and test items were developed based on the following multi-dimensional criteria: Narrative Structure and Length: Each narrative was designed to be 5 to 7 sentences long, totaling approximately 150 to 200 characters. This length was strictly controlled to minimize the impact of excessive working memory load while still requiring the integration of multiple information units [5]. The narratives were structured with a clear beginning, middle, and end to evaluate global coherence and the ability to track themes throughout the discourse, which are often cited as sensitive indicators of cognitive-communication decline [2]. Vocabulary and Syntactic Complexity: To ensure that the test measures "discourse integration" rather than simple

vocabulary knowledge, high-frequency vocabulary based on the primary education level was predominantly used. Complex syntactic structures, such as passive voices or multiple embeddings, were intentionally minimized to focus on the participant's ability to grasp the overall context and maintain the communicative flow [10].

Domain-Specific Item Construction: The items were divided into six sub-domains based on the pragmatic requirements of daily life:

- **Literal Comprehension:** Recalling information explicitly stated in the discourse.
- **Inferential Comprehension:** Grasping hidden meanings or speaker intentions that require contextual reasoning.
- **Metaphorical/Indirect Speech:** Understanding socio-pragmatic nuances, such as proverbs or sarcasm, which are often challenging in healthy aging [5].

Contextual Relevance: The scenarios were set in daily contexts familiar to older adults, such as interactions at a hospital, a pharmacy, or a community center. This ensures that the tool measures "Successful Communicative Competence" applicable to real-world social participation, addressing the "communication alienation" often experienced by the elderly in modern society [3].

Expert Review and Refinement: The initial pool of items was reviewed by a panel of 8 experts (Speech-Language Pathologists with over 10 years of experience) [7]. They evaluated each item for its appropriateness, difficulty, and relevance on a 5-point Likert scale. Only items that achieved a mean score of 4.5 or higher were retained, and ambiguous phrases were modified to ensure clarity. The items selected through the aforementioned development process are presented in Table 1.

Table 1. Test items for discourse comprehension test for the older adults

Subtest	Question
Follow directions	1. Close your eyes. 2. Clap once, then raise one hand 3. Cough, say hello, and then hurrah.
Yes/no	1. Are you a man? 2. Is the cow bigger than the chicken? 3. Do you pay for something that says it's free? 4. Do you wear shoes before you put on socks? 5. Do you wear underwear after you wear a coat? 6. Do you wear underwear after putting on your coat? 7. Do you brush your teeth before eating?
Metaphor	1. There is a round moon like a tray on Chuseok. 2. I'm a walking general hospital. 3. Students need carrots and sticks. 4. Two head are better than one. 5. If you have a lot of workers, the boat goes up to the mountain. 6. I ate watermelon and drew a map on the blanket. 7. The robber was caught by the police. 8. The story stings my ears. 9. The fox was caught by a monkey. 10. The penguin is being chased by a dog. 11. The father caught the mother carrying the baby. 12. The mother holding the baby looks at her father.
Factual information understanding	1. What day is it today? ① Liberation Day ② National Foundation Day ③ Constitution Day ④ Hangul Day 2. What day is Mrs. Park's social day? ① Monday ② Tuesday ③ Saturday ④ Sunday 3. What did Mrs. Park ride to get to the social circle? ① Bus ② Subway ③ Airplane ④ Train 4. How many members of Mrs. Park's fraternity are there in total? ① 5 ② 6 ③ 8 ④ 10 5. Where is Park's best friend living? ① Gwangju ② Busan ③ Daegu ④ Chuncheon
Text connection understanding	1. Why did Mrs. Park dress up beautifully? ① Going to the hospital ② Going to a meeting ③ Going to a birthday party ④ Feeling good 2. Why did Mrs. Park forget to wear a mask? ① Frustrated ② Looking forward to seeing friends ③ Dislike wearing mask ④ So busy 3. Where does Mrs. Park live? ① Gwangju ② Busan ③ Seoul ④ Daegu 4. How long did Mrs. Park stay at the meeting last month? ① To morning ② To noon ③ To afternoon ④ To evening 5. Who did Mrs. Kim contact that she couldn't make it to the meeting? ① Chairman ② General secretary ③ Other friend ④ Best friend
Missing information inferring	1. How many times a month does Mrs. Park's hometown get-together? ① Once ② Twice ③ Three times ④ Every week 2. What did Mrs. Park buy in a hurry at the pharmacy? ① Bottled water ② Mask ③ Tissue ④ Transportation card 3. Where is the meeting place in the bedroom community? ① Sashimi restaurant ② Karaoke bar ③ Meat restaurant ④ Noodle restaurant 4. Where is Mrs. Park's hometown? ① Busan ② Seoul ③ Gwangju ④ Jeju Island 5. Why did Mrs. Park sigh? ① The chairman was rude ② Dislike eat dinner together ③ Couldn't meet best friend ④ Dislike other friends
Total	37questions

2.3. Research procedure

First, the researcher performed one-on-one, face-to-face evaluations with the older adult subjects. An interview was conducted to gather basic information, and before the main test began, the subjects were thoroughly briefed on the test content and response methods, after which the test items were administered. The examiner first presented an example item auditorily and asked for a response. If the response was correct, the examiner said, "You can proceed like that." If there was no response, the instruction was repeated until a correct response was elicited, after which the main test commenced.

For the main test, sentences were presented auditorily, and participants were asked to choose the corresponding picture from the provided materials; the subject's responses were recorded on a response sheet. The total score range was between 0 and 40 points, and the average time required for the main test was measured as 15±3 minutes.

2.4. Data analysis

Statistical analysis of the research results was performed using the IBM SPSS 29.0 version (Statistical Product and Service Solution, SPSS Inc., 2020) program, and the significance level was set at .05. Internal consistency reliability analysis, test-retest reliability analysis, and Pearson correlation coefficient analysis were used to establish reliability and validity.

3. Results

3.1. The general characteristics and K-MMSE scores of participants

The general characteristics and K-MMSE scores for the

study participants are presented in Table 2. A total of 30 older adults participated in the study, consisting of 21 females and 9 males. When categorized into three age groups: Group 1 (65 to 74 years old) had 16 participants, Group 2 (75 to 84 years old) had 11 participants, and Group 3 (85 years and older) had 3 participants. The mean age was 73.53 (±6.89) years, with an age range of 65 to 87 years. Participants were divided into four groups based on years of education: 3 individuals were illiterate/had no formal schooling, 11 had 1 to 6 years of education, 12 had 7 to 12 years of education, and 4 participants had 13 or more years of educational experience. When K-MMSE scores were compared across age groups, Group 1 scored 27.44 (±1.93) points, Group 2 scored 27.09 (±1.69) points, and Group 3 scored 25.33(±1.81) points, indicating no significant difference between the age groups (p>.05).

3.2. Reliability

3.2.1. Internal consistency reliability

The analysis of responses from the elderly participants revealed that the Cronbach's α coefficient for all items of the Elderly Discourse Comprehension Test Tool was .876, indicating high internal consistency reliability. The reliability coefficients for each sub-domain are presented in Table 3.

3.2.2. Test-retest reliability

To measure intra-rater reliability, a re-measurement was conducted two weeks later, and the correlation coefficient was found to be .981 (p<.001\). When inter-rater reliability was measured, the correlation coefficient was .993 (p<.001\), indicating high reliability of the assessment regardless of the observer (Table 4).

Table 2. The general characteristics and K-MMSE scores of participants

Variables	Age(yr)			Total	
	65~74yr	75~84yr	85yr +		
Participants number	16	11	3	30	
Gender(male:female)	4:12	3:7	0:3	9:21	
Education yr.	Illiterate /Uneducated	0	1(1:0)	2(0:2)	3(1:2)
	1~6yr	4(0:4)	6(2:4)	0	10(2:4)
	7~12yr	11(3:8)	2(0:2)	1(0:1)	14(3:11)
	13yr +	1(1:0)	2(1:1)	0	3(2:1)
K-MMSE (±sd)	27.44 (±1.93)	27.09 (±1.3)	25.33 (±2.0)	27.10 (±1.81)	

N=30

Table 3. Internal consistency reliability of 6 subtest of the discourse comprehension test tool for the elderly

No	Subtest	Reliability	
		corrected Item correlation	Total Cronbach α value when an item is deleted
1	Follow directions	.461	.890
2	Yes/no	.628	.884
3	Metaphor	.847	.855
4	Ffactual information understanding	.647	.874
5	Text connection understanding	.691	.874
6	Missing information inferring	.658	.877

Table 4. Intra-rater and inter-rater reliability based on test-retest

	Intra-rater reliability	Inter-rater reliability
Correlation coefficient	.981	.993

3.3. Validity

3.3.1. Correlation analysis between item-total scores and subscale scores

To examine the internal consistency of the Korean Discourse Comprehension Test for the Elderly correlations were analyzed between the total test score and each subscale score. The results, confirming the strength of the relationship between individual item scores and the overall scale score, are presented in Table 5. Subscales included 'Following Directions', 'Yes/No Questions', 'Figurative Language Comprehension', 'Factual Information Comprehension', 'Textual Connection

Comprehension', and 'Inference of Missing Information'. As illustrated in the table, the item scores showed moderate to strong positive correlations with the total test score.

3.3.2. Criterion validity

To establish the construct validity (specifically, concurrent validity), the correlation between the current test and the Korean Sentence Comprehension Test [12]—a widely used, existing standardized assessment for sentence comprehension in clinical settings—was examined. The results from the two assessments demonstrated a high, statistically significant correlation ($r=.797$), ($p<.01$).

Table 5. Item-Total Scores and Subscale Scores of discourse comprehension test tool for the elderly

Subtest	Follow directions	Yes/no	Metaphor	Factual information understanding	Text connection understanding	Missing information inferring
Follow directions	1					
Yes/no	.376*	1				
Metaphor	.184	.476**	1			
Factual information understanding	.125	.436**	.625**	1		
Text connection understanding	.254	.476**	.520**	.421*	1	
Missing information inferring	.628**	.348*	.542**	.134	.410*	1
Total	.509**	.642**	.874**	.682**	.724**	.714**

$p<.05$, $p<.01$

Table 6. Correlation analysis with the Korean Sentence Comprehension Test

	Korean Sentence Comprehension Test
Correlation coefficient	0.797

4. Discussion & Conclusion

4.1. Academic contribution and comparative analysis

The primary significance of this study lies in its shift from pathology-centered language assessment to "Practical Discourse Competence" for the elderly. While traditional tools like the PK-WAB-R [8] are highly effective in diagnosing localized linguistic deficits in aphasia patients, they often yield a "ceiling effect" when applied to healthy older adults or those with mild cognitive impairment (MCI), failing to capture subtle declines in complex social communication. In contrast, this study developed items that specifically target discourse-level integration and inferential reasoning, addressing the "communication alienation" highlighted in recent studies [3].

The strong criterion validity ($r=.797$) found with the Korean Sentence Comprehension Test (KOSECT) indicates that while our tool aligns with established linguistic measures, it also offers incremental validity by incorporating pragmatic elements such as metaphors and indirect speech. This aligns with Lee & Kim's [5] findings, which suggest that discourse coherence is a more sensitive marker of cognitive aging than simple sentence-level accuracy. Thus, the developed items provide a more ecologically valid representation of how older adults navigate the rapid information flow in a digitalized society.

4.2. Reliability and methodological rigor

The internal consistency (Cronbach's $\alpha = .876$) and test-retest reliability ($r = .993$) of this tool were found to be comparable to, or even higher than, the Brief Test of Cognitive-Communication Disorders (B-CCD) [9]. This high degree of stability suggests that the selection criteria established in this study—controlling for narrative length (5–7 sentences) and vocabulary frequency—successfully minimized extraneous cognitive load, allowing for a focused assessment of discourse comprehension per se.

4.3. Limitations and future standardization

Despite these promising results, this study was a preliminary investigation and possesses certain limitations. First, the sample size ($N=30$) and regional concentration (Sejong/Jeonbuk) limit the ability to generalize these findings to the entire national elderly population. As noted by Reviewer 3, a larger and more demographically diverse sample is required to analyze the specific effects of education level and age-related cognitive sub-types on discourse

performance. Therefore, future research will focus on full-scale standardization by:

- Expanding the sample size to over 150 participants to establish age- and education-based normative data.
- Implementing Item Response Theory (IRT) to further refine the difficulty and discrimination indices of each item.
- Diversifying discourse genres to include procedural and expository texts, thereby enhancing the tool's utility in various rehabilitation settings.

4.4. Conclusions

In conclusion, this study provides a scientifically validated foundation for a new discourse comprehension assessment tailored for older adults. By bridging the gap between clinical linguistics and geriatric social participation, this tool is expected to play a vital role in early screening and the development of individualized intervention plans, ultimately fostering successful aging and social connectedness in the modern era.

Author contribution

Conceptualization: SJ Park

Formal analysis: HJ Lee

ORCID

Hyun Ju Lee: <https://orcid.org/0000-0002-3550-0190>

Sung Ji Park: <https://orcid.org/0009-0002-6852-548X>

Conflicts of Interest

The authors declare no conflict of interest.

Funding

This work was supported by research funding from Wonkwang University in 2024.

References

1. Seok DI, Kang SK. The development of communication scale contents for presbycusis. *The Korean Society For Rehabilitation Science*. 2003;26(1):47-62.
2. Lee MS. Factors Affecting Quality of Communication Life in Older Adults: Focused on the Relationship with Quality of Life, Cognition, and Psychoemotional

- Aspects. *Audiology and Speech Research*. 2019;15(3): 232-240. <https://doi.org/10.21848/asr.2019.15.3.232>
3. Kang MS, Jung IH, Yun DG, Ju SS, Kim HK. Active digital aging process through the experience of digital content creation in old age. *Journal of Korea Gerontological Society*. 2023;43(4):571-589. <http://dx.doi.org/10.31888/JKGS.2023.43.4.571>
 4. Grice HP. Logic and conversation. In: Cole P, Morgan J, editors. *Syntax and semantics, Vol. 3: Speech acts*. New York: Academic Press; 1975. p. 41-58.
 5. Kim BS, Kim YB, Kim HH. Global Coherence Analysis of Discourse in Amnesic Mild Cognitive Impairment and Associated Cognitive Function. *Communication Science and Disorders*. 2018;23(4):1028-1041. <http://dx.doi.org/10.12963/csd.2018.23.4.561>
 6. American Speech-Language-Hearing Association. Evidence-based practice in speech-language pathology [Position Statement] [Internet]. Rockville: American Speech-Language-Hearing Association; 2005 [cited 2024 May 22]. Available from: <http://www.asha.org>
 7. Kim HH, Kim SR, Yoo HJ, Kim JW. Content validity of speech-language assessment items for the elderly. *Journal of Rehabilitation Research*. 2013;17(3):347-373.
 8. Kim HH, Na DR. *Paradise Korea - Western Aphasia Battery (Revised); PK-WAB-R*. Seoul: Paradise Welfare Foundation NCS Pearson, Inc.; 2012.
 9. Lee MS, Kim BS, Lim JS. Reliability and validity of the brief test of cognitive-communication disorders. *Communication Science and Disorders*. 2020;25(4): 872-889. <https://doi.org/10.12963/csd.20757>
 10. Yun JH, Chung JH, Kim YW, Jung IL. The Korean Version of the Right Hemisphere Language Battery: Adaptation and Reliability. *Communication Science and Disorders*. 2022;27(3):606-616. <https://doi.org/10.12963/csd.22917>
 11. Kang YW, L D, Hahn SH. Validity study on the Korean Mini-Mental State Examination (K-MMSE) in dementia patients. *Journal of the Korean Neurological Association*. 1997;15(2):300-307.
 12. Bae SY, Lim SS, Lee JH, Jang HS. *Korean Sentence Comprehension Test (KOSECT)*. Seoul, Seoul Community Rehabilitation Center, 2004.

