ASHA American Speech-Language-Hearing Association

JSLHR

Research Note

Understanding How Older Adults Comprehend Simple Comparative Sentences in a Predicate-Final Language

Jimin Park,[a] Seunghun J. Lee,[b,c] and Jee Eun Sung[a,d,e]

[a] Department of Communication Disorders, Ewha Womans University, Seoul, South Korea [b] Department of Psychology and Linguistics, International Christian University, Tokyo, Japan [c] Department of African Languages, University of Venda, Thohoyandou, South Africa [d] Department of Brain and Cognitive Sciences, Ewha Womans University, Seoul, South Korea [e] Ewha Brain Institute, Ewha Womans University, Seoul, South Korea

ARTICLE INFO

Article History:
Received December 11, 2024
Revision received August 25, 2025
Accepted March 30, 2026

Editor-in-Chief: Julie A. Washington
Editor: William S. Evans

https://doi.org/10.1044/2026_JSLHR-24-00874

ABSTRACT

Purpose: This study examined the effects of aging on the comprehension of simple comparative sentences in a predicate-final language. Additionally, we explored individual variability among older adults by categorizing them based on chance-level performance.

Method: A total of 221 participants completed a sentence–picture matching task involving simple comparative constructions, which were manipulated by word order (comparee-first, standard-first). Response accuracy and response time (RT) were analyzed across age groups (young, old) and performance group within older adults (above-chance, chance).

Results: For age group comparisons, significant effects of group and word order emerged in RT. Young adults responded significantly faster than older adults in both word orders, with a larger group difference observed for comparee-first than for standard-first constructions. For both age and performance group comparisons, a significant interaction between word order and working memory (WM) scores emerged in RT. Participants with higher WM scores showed faster responses to comparee-first constructions, suggesting that cognitive resources play a role in processing syntactically complex comparative structures.

Conclusions: The current findings indicate that age-related differences in sentence processing can arise even in comparatively simple syntactic forms, revealing their sensitivity to cognitive aging. Importantly, the observed interplay between WM capacity and syntactic structure highlights the necessity of integrating both cognitive and linguistic dimensions in psycholinguistic models of sentence comprehension across the adult lifespan.

Previous studies on sentence processing have demonstrated that older adults experience difficulties in comprehending syntactically complex sentences, such as those with passive constructions (Liu, 2018; Mack et al., 2013; Sung et al., 2017) and center-embedding clauses (Caplan & Waters, 2005; Goral et al., 2011; Payne et al., 2014). These difficulties are typically attributed to both cognitive and linguistic factors—such as declines in working memory (WM) capacity and processing speed—as sentence comprehension relies heavily on cognitive resources (Just & Carpenter, 1992). In particular, syntactic complexity places increased demands on WM, making it more difficult to maintain structural dependencies or revise initial interpretations during incremental parsing (Caplan & Waters, 1999). Consequently, older adults tend to show higher error rates and slower processing (Salthouse, 1992, 1996). However, recent studies have shown that even syntactically simple sentences can reveal processing difficulties in older adults (Hwang et al., 2024; Obler et al., 1991; Sung et al., 2022).

Simple sentences refer to a sentence structure that consists of a single independent clause, containing a subject (noun phrase [NP]) and a predicate (verb phrase) to express

Correspondence to Jee Eun Sung: jeesung@ewha.ac.kr. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

a complete proposition. Simple sentences conform to a basic syntactic structure, with only the essential elements required to convey a complete thought, without any additional subordinate or dependent clauses. For instance, Sung et al. (2022) examined simple sentences with instrumental phrases (e.g., "Point to the book with the pencil") in three different language groups: English, Korean, and Japanese. They found that older adults in all three languages exhibited greater difficulty with instrumental sentences compared to their young counterparts, particularly under less plausible conditions. In a different type of simple sentences, Obler et al. (1991) found that English-speaking older adults have difficulty comprehending comparative constructions (e.g., "The horsehair rug was rougher than the baby blanket," p. 438) compared to active sentences (e.g., "The fierce wolf attacked the lost sheep in the woods," p. 438). More recently, Hwang et al. (2024) found that Korean-speaking older adults relatively perform worse on simple comparatives (e.g., "The dog is bigger than the cat," p. 3) compared to young adults, as demonstrated by both end-of-sentence comprehension and eye-tracking results. These results suggest that older adults tend to struggle with comprehending even syntactically simple sentences, underscoring the importance of further investigation into the effects of aging on simple sentence structures.

The current study examines the effects of age and word order on simple comparative constructions in a predicate-final language. Comparative constructions are included in various standardized language assessments for clinical populations, such as the Western Aphasia Battery–Revised (WAB-R; Kertesz, 2007) and the Test for Reception of Grammar (Bishop, 2003). Although syntactically simple, comparatives often pose challenges due to the semantic processing required to evaluate two distinct objects on a predicative scale, utilizing specific morphological markers (e.g., "–er," "than") or comparative phrases (e.g., "compared to"). According to Dixon (2008), comparative constructions typically comprise four components: (a) two comparable NPs, referred to as the comparee and the standard of comparison (comparee-NP and standard-NP, respectively); (b) a parameter of comparison, which are adjectives in most languages; (c) an index of comparison, involving comparative morphological cues (e.g., "more," "–er"); and (d) a marker of the standard-NP (e.g., "than," "compared to"). For example, in the sentence "The dog is bigger than the cat," the comparee-NP is "The dog," the standard-NP is "the cat," the parameter of comparison is "big," the index of comparison is "–er" within "bigger," and the marker of the standard-NP is "than."

The linguistic characteristics of the Korean language offer a unique perspective in existing comparative studies, which have predominantly focused on subject–verb–object (SVO) languages such as English (Bishop & Bourne, 1985; Obler et al., 1991; Rasuki, 2020). Korean, as a predicate-final language, permits relatively flexible word order—SOV or object–subject–verb (OSV). This flexibility is enabled by the case marking system in Korean, where case markers (CMs) are explicitly attached to the NPs, assigning thematic roles to each NP (Sohn, 2001). In Korean comparative constructions, a comparative-CM ("–poda") and a nominative-CM ("–ka") are attached to each NP, with an adjective in the final position functioning as a predicate (AdjP). The NP with comparative-CM is identified as the standard-NP, and the NP with a nominative-CM is identified as the comparee-NP. Similar to Hwang et al. (2024), this study manipulates word order in comparative constructions: standard-NP coming first (standard-first) versus comparee-NP coming first (comparee-first). For instance, in the standard-first construction, the sentence would be structured as "The dog-$_{\text{COMP}}$ the cat-$_{\text{NOM}}$ big-$_{\text{AdjP}}$ (Compared to the dog, the cat is big)," whereas the comparee-first is represented as "The cat-$_{\text{NOM}}$ the dog-$_{\text{COMP}}$ big-$_{\text{AdjP}}$ (The cat is big compared to the dog)."

Although standard-first and comparee-first sentences in Korean convey the same meaning because of their relatively free word order, they exhibit different syntactic and semantic traits. Syntactically, in the comparee-first structure, a nominative-NP (The cat-$_{\text{NOM}}$) is initially introduced, followed by the standard-NP marked with the comparative-CM. This structure is considered syntactically unmarked because the nominative-CM comes first. Conversely, the standard-first structure is syntactically marked, as the NP with the comparative-CM precedes the nominative-NP. Existing research demonstrated that Korean speakers significantly outperformed in processing unmarked structures (e.g., nominative-NP-first; SOV) compared to their marked counterparts (e.g., accusative-NP-first; OSV; Sung, 2015; Sung et al., 2017). This performance difference suggests that, despite the flexibility in word order, Koreans prefer nominative-NP-first structures, which are considered syntactically unmarked. This preference is presumed to reflect comparative constructions, favoring comparee-first structures with nominative-NP-first arrangements.

However, semantically, this may not align with the isomorphic mapping hypothesis (O'Grady & Lee, 2001, 2005), which proposes that sentence processing is easier when the word order of linguistic elements mirrors the cognitive sequence of events being described—a principle known as isomorphism. In comparative constructions, a standard-first structure (e.g., "The cat is big compared to the dog") follows this principle: The standard-NP (cat) comes first and serves as the reference point for evaluating the comparee-NP (dog), reflecting the natural direction of comparison. In Korean, however, speakers tend to favor a nominative-NP-first (comparee-first) structure, which aligns with syntactic preferences but violates isomorphism, since the mental comparison requires holding the first-mentioned item in the memory until the reference point is revealed. This creates a semantic–syntactic mismatch: The isomorphic (standard-first) structure aids conceptual processing but is less syntactically natural, while the preference (comparee-first) structure is syntactically aligned but less semantically transparent. This study examines how such mismatches in simple comparative constructions affect sentence processing, particularly in relation to age-related differences in sentence comprehension.

This hypothesis posits that sentence processing is more straightforward when the word order aligns with the temporal sequence of events, a concept referred to as isomorphism. In comparative constructions, placing the standard-first with a comparative-CM aligns with isomorphic principles, as the standard-NP serves as the reference point for comparison with the comparee-NP, reflecting the temporal sequence of events. However, this conflicts with syntactic preferences, where Korean speakers tend to favor a nominative-NP-first (comparee-first) structure. Thus, while the standard-first structure is isomorphic, its syntactic marking creates a mismatch between the semantic and syntactic features. In contrast, a comparee-first structure with nominative-CM adheres to the preferred word order but lacks isomorphism, as the sequence does not correspond to the temporal progression of the comparison. This study examines the impact of semantic–syntactic mismatches in simple comparative constructions on age-related changes.

This study investigates age-related differences (young vs. older) in the comprehension of comparative constructions using a sentence–picture matching paradigm, focusing on word order manipulations (comparee-first vs. standard-first). Additionally, to explore individual variability among older adults, participants were divided into two subgroups based on their performance levels (above-chance vs. chance). We hypothesize that individuals with stronger cognitive resources—particularly younger adults—will rely more on morphosyntactic structure (i.e., the nominative-NP-first template) to process Korean comparative sentences, making comparee-first constructions easier to comprehend despite the syntactic–semantic mismatch. In contrast, individuals with weaker cognitive resources—more commonly observed among some older adults—will rely more heavily on semantic isomorphism, making standard (isomorphic) constructions easier to process than comparee-first constructions. Accordingly, we predict that accuracy on comparee-first sentences will be positively associated with cognitive resources.

Method

Participants

The study included 221 Korean individuals—115 young adults and 106 older adults. All participants completed the Health Screening Questionnaire (Christensen et al., 1991) and a Korean version of the Mini-Mental State Examination (K-MMSE; Kang, 2006). They reported no neurological or psychiatric diseases or vision/hearing problems. Older adults also undertook the Seoul Verbal Learning Test (SVLT) from the Seoul Neuropsychological Screening Battery–Second Edition (Kang et al., 2012), scoring within normal age- and education-adjusted ranges (≥ 16 percentile) on both the K-MMSE and SVLT scores (see Table 1). All participants underwent hearing screening using pure-tone audiometry (PTA) at 0.5, 1, 2, 4, and 8 kHz. Only those with hearing thresholds below 25 dB HL in both ears were included. While individual audiometric data were not retained, this screening ensured that participants met the inclusion criteria.

Older adults were further subdivided based on their performance on the comparative sentence–picture matching

Table 1. Demographic and cognitive screening information of participants.

(A) Age groups			
	Young (*n* = 115)	**Older (*n* = 106)**	***t*-test results**
Age (*SD*)	24.61 (5.21)	67.51 (8.06)	$t = -47.36$, $p < .001$***
Education (*SD*)	13.94 (1.47)	12.83 (4.50)	$t = 2.50$, $p < .05$*
K-MMSE (*SD*)	29.55 (0.79)	28.10 (1.96)	$t = 7.29$, $p < .001$***
WM (*SD*)	22.49 (4.59)	16.13 (5.74)	$t = 9.13$, $p < .001$***
(B) Performance groups			
	Above-chance (*n* = 95)	**Chance (*n* = 11)**	***t*-test results**
Age (*SD*)	66.52 (7.61)	76.09 (6.88)	$t = -4.32$, $p < .001$***
Education (*SD*)	13.60 (3.73)	6.18 (5.25)	$t = 4.55$, $p < .001$***
K-MMSE (*SD*)	28.38 (1.59)	25.73 (3.10)	$t = 4.64$, $p < .001$***
WM (*SD*)	16.57 (5.24)	11.82 (7.11)	$t = 2.15$, $p = .03$*

Note. WM scores were obtained by summing the scores from the Digit-Forward and -Backward Test from the Seoul Neuropsychological Screening Battery–Second Edition (Kang et al., 2012). K-MMSE = Korean version of the Mini-Mental State Examination (Kang, 2006); WM = working memory.

*$p < .05$. ***$p < .001$.

task. Classification was determined using a binomial distribution, with chance-level accuracy defined as 11–20 correct responses out of 32 trials. This range corresponds to the 95% confidence interval around chance performance (50%) for a binomial distribution with 32 trials. Participants whose scores fell within this range for at least one condition were categorized as performing at chance level, whereas those who exceeded this range in both conditions were classified as performing above chance. Based on this criterion, 95 older adults were assigned to the above-chance-level range group and 11 to the chance-level range group. Young adults were not subdivided, as none performed at chance-level range. The institutional review board of Ewha Womans University approved this study (No. 2022-0112), and written informed consent was obtained from all participants.

Stimuli

The stimuli consist of 32 sentences with three phrases: standard-NP, comparee-NP, and an AdjP (see Table 2). The CM of the first NP determines whether the sentence is standard-first (NP1-comparative CM) or comparee-first (NP1-nominative CM), with 16 sentences in each condition. All sentences were semantically reversible, with NPs counterbalanced as standard-NP and comparee-NP. Stimuli included four semantic adjective categories: size, weight, length, and thickness, each featuring two opposing adjectives (e.g., “big” and “small”).

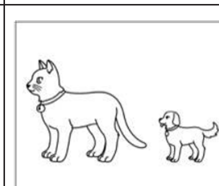
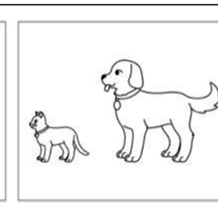
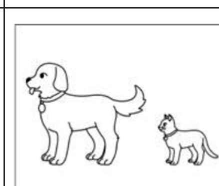
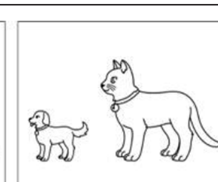
Experimental Procedure

The experiment was conducted in a quiet room, with all participants tested individually. Before beginning the comparative sentence–picture matching task, participants completed the screening procedures, which included cognitive assessments (e.g., K-MMSE, WM) and a PTA screening to confirm thresholds below 25 dB HL in both ears. Only participants who met all inclusion criteria proceeded to the main task. In each trial, participants listened to a target sentence while viewing the target and foil pictures on a monitor and indicated their choice by pressing “1” for left or “3” for right on the numeric keypad. Four practice trials followed the instructions to ensure task comprehension. The item presentation order was fixed across participants and presented in a single block without breaks. The entire experiment, including instructions and practice, lasted approximately 5–10 min. Auditory stimuli were presented at a minimum of 65 dB SPL, with volume adjustments made during practice trials if requested.

Statistical Analysis

For accuracy, a generalized linear mixed-effects model with a binomial distribution and a logit link function was fitted using the glmer function from the lme4 and lmerTest packages in R (R Core Team, 2022). Fixed effects were sum-coded as follows: age group (young = +1, older = −1) and word order (comparee-first = +1, standard-first = −1). For the older adult subgroup analysis, fixed effects were sum-coded as older performance group (above-chance = +1, chance = −1) and word order (comparee-first = +1, standard-first = −1). WM *z* scores were included as a continuous predictor, with a WM × Word Order interaction term added to test whether WM moderated the word order effect. Random intercepts were specified for both subjects and

Table 2. Examples of sentence stimuli and pictures for each condition.

Word order	Predicate		Sentence stimuli	Picture example	
Standard-first (syntactically unmarked, isomorphic)	Size	Small	Compared to the cat, the dog is small. Koyangi-ka kangaci-pota cakta. The cat-$_{COMP}$ the dog-$_{NOM}$ small-$_{AdjP(Size)}$.		
Comparee-first (syntactically marked, non-isomorphic)		Big	The cat, compared to the dog, is big. Koyangi-ka kangaci-pota khuta. (The cat-$_{NOM}$ the dog-$_{COMP}$ big-$_{AdjP(Size)}$.)		

Note. COMP = comparative case marker; NOM = nominative case marker; AdjP(Size) = adjective as a predicate of “size” category.

items. Additionally, by-subject random slopes for word order were included to account for individual variability in sensitivity to word order, in line with recommendations for a maximal random effects structure (Barr et al., 2013). For the age group analysis, years of education were included as a covariate to control for significant group differences.

For response time (RT) analyses, only correct responses were included. RT measurement began after sentence offset (i.e., after the audio file ended). Outliers were removed using Tukey's fences, excluding values beyond 1.5 times the interquartile range (IQR) below the first quartile (Q1 − 1.5 × IQR) or above the third quartile (Q3 + 1.5 × IQR; Tukey, 1977). As a result, 578 of the 6,848 trials (8.4%) were excluded for the age group comparison (young vs. older), and 446 of the 3,213 trials (13.9%) were excluded for the older group comparison (above-chance vs. chance-level). To normalize the distribution and reduce positive skew in RTs, raw RTs were log-transformed (logRT) prior to analysis.

LogRT was used as the dependent variable in all linear mixed-effects models involving RT. Two linear mixed-effects models were fitted separately, as age group and performance group could not be entered simultaneously. The first model included age group, word order, WM z scores, education, and the interactions of Age Group × Word Order and WM z Scores × Word Order as fixed effects. The second model included performance group, word order, WM z scores, education, and the interactions of Performance Group × Word Order and WM z Scores × Word Order as fixed effects. Both models included random intercepts for subjects and items. Models that additionally included by-participant random slopes for word order produced a singular fit in both cases, indicating that the random slope variance could not be reliably estimated from the data. This issue persisted when analyses were repeated with ±3 SD participant-level outlier trimming, exclusion of participants contributing fewer than 50% of usable trials, and inclusion of both correct and incorrect responses with accuracy as a fixed effect, and it was not resolved by alternative optimizers. Random intercepts alone were therefore retained in the final models. Residual diagnostics were examined both visually (e.g., Q–Q plots, residuals vs. fitted values) and quantitatively to ensure that assumptions of normality and homoscedasticity were met, following recommendations by Meteyard and Davies (2020).

Results

Effects of Age Group, Word Order, and WM on Accuracy

A generalized linear mixed-effects model was fitted to predict accuracy from age group (young vs. older), word order (comparee-first vs. standard-first), and WM z scores, including random intercepts for subjects and items and by-subject random slopes for word order. Results from a simpler model without the WM term are provided in Appendix A, and the full model is reported in Appendix B.

The model revealed a significant main effect of age group (estimate = .60, SE = .10, z = 5.96, p < .001), with young adults performing significantly better than older adults (young: M = .97, SD = .18; older: M = .88, SD = .33). The main effect of word order was not statistically significant (estimate = −.02, SE = .15, z = −0.11, p = .91), indicating no reliable difference in performance between standard-first (M = .92, SD = .27) and comparee-first (M = .93, SD = .26). The interaction between age group and word order also did not reach significance (estimate = .08, SE = .08, z = 1.09, p = .27), suggesting that the effect of word order on accuracy did not differ reliably across age groups (see Table 3A and Figure 1A).

WM z scores significantly predicted accuracy (estimate = .33, SE = .10, z = 3.28, p = .01), with higher WM associated with greater accuracy overall. However, the interaction between WM and word order was not significant (estimate = .02, SE = .07, z = 0.23, p = .81), indicating that WM did not modulate sensitivity to word order. Education was included as a covariate and showed a small but significant positive effect on performance (estimate = .08, SE = .02, z = 3.57, p < .001).

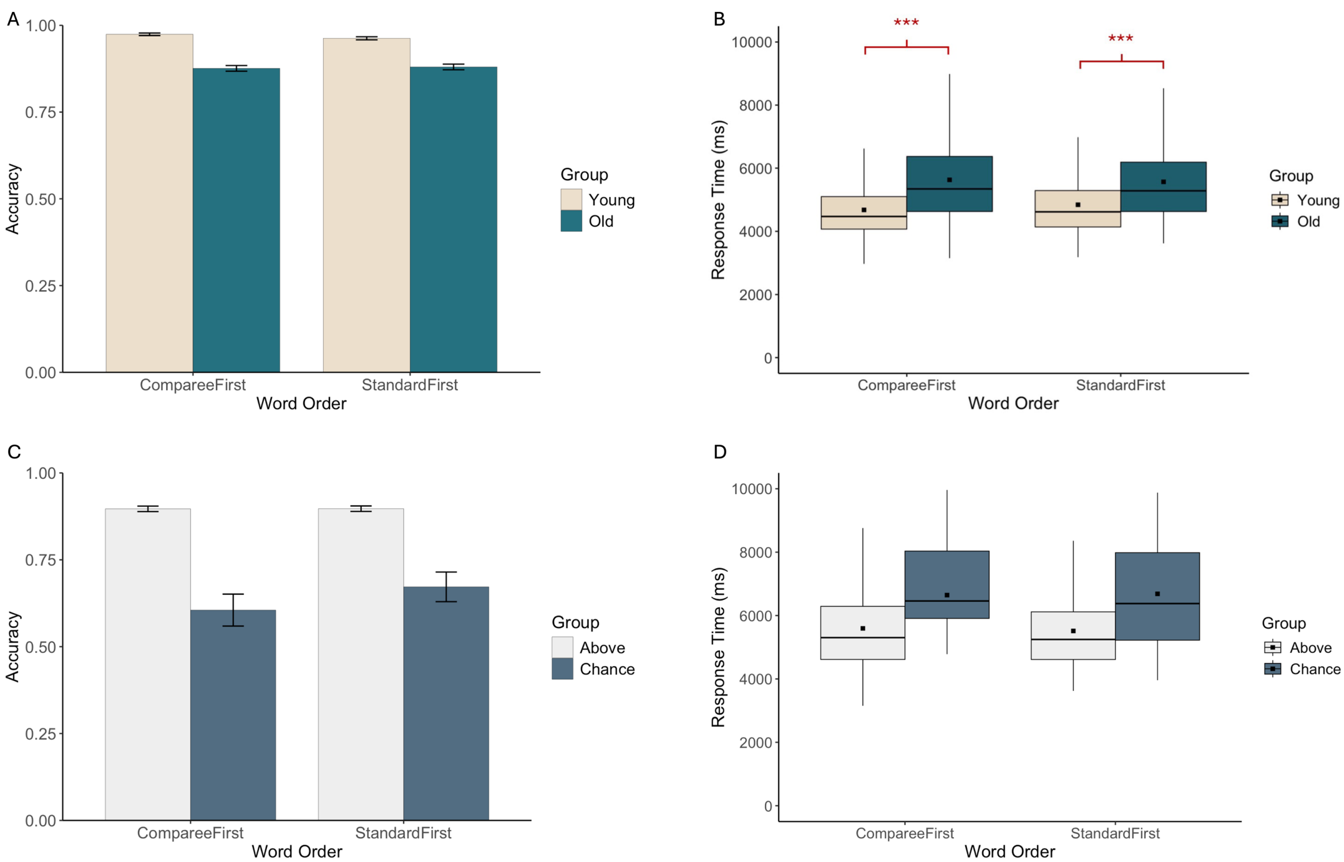
Effects of Age Group, Word Order, and WM on RT

A linear mixed-effects model was fitted to predict logRT from age group, word order, and WM z scores, including random intercepts for subjects and items (see Appendix C). The analysis revealed a significant main effect of age group (estimate = −.07, SE = .01, t = −6.24, p < .001), with older adults showing significantly longer RTs than young adults (older: M = 5,757.73, SD = 1,571.07; young: M = 4,761.33, SD = 923.58). The main effect of word order was not significant (estimate = −.01, SE = .01, t = −0.47, p = .65), indicating no overall RT

Table 3. Accuracy of sentences by word order between groups.

(A) Age groups: young vs. older		
	Young	**Older**
Standard-first (SD)	.96 (.19)	.88 (.33)
Comparee-first (SD)	.97 (.16)	.88 (.33)
(B) Performance groups within older adults: above-chance vs. chance		
	Above-chance	Chance
Standard-first (SD)	.92 (.27)	.63 (.48)
Comparee-first (SD)	.90 (.30)	.71 (.46)

Figure 1. Accuracy and response time across age and performance groups by word order. (A) Mean item-level accuracy (with standard error bars) for young and old adults across word order conditions (comparee-first vs. standard-first). (B) Box-and-whisker plots of response time (RT, in milliseconds) for young and old adults across word orders. (C) Mean item-level accuracy (with standard error bars) for performance-based subgroups within the old adult group (above-chance vs. chance) across word orders. (D) Box-and-whisker plots of RTs for above-chance-level and chance-level range performers across word order conditions. ***p < .001.



difference between standard-first and comparee-first trials (standard-first: M = 5,218.32, SD = 1,279.18; comparee-first: M = 5,179.26, SD = 1,405.56).

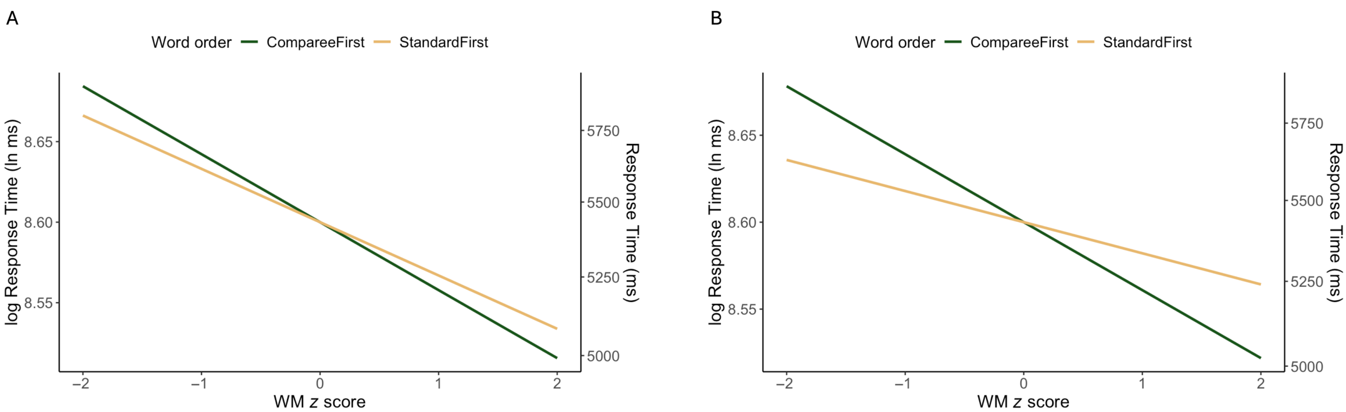
However, the interaction between age group and word order was significant (estimate = −.01, SE = .00, t = −3.94, p < .001). Post hoc comparisons confirmed that young adults responded faster than older adults in both word order conditions, with the gap being larger for the comparee-first condition (estimate = −.16, SE = .02, z = −6.86, p < .001) than for the standard-first condition (estimate = −.13, SE = .02, z = −5.42, p < .001; see Table 4A and Figure 1B).

There was also a significant main effect of WM z score (estimate = −.04, SE = .01, t = −3.03, p < .01), indicating that individuals with higher WM capacity responded faster overall. Notably, the interaction between WM z score and word order was also significant (estimate = −.00, SE = .00, t = −1.98, p < .05), suggesting that WM influenced RTs more strongly in comparee-first than standard-first trials. Trend analysis confirmed that the negative WM effect on RT was significantly steeper in the comparee-first condition (estimate = −.04, SE = .01) compared to standard-first (estimate = −.03, SE = .01, p < .05; see Figure 2A).

Table 4. Response time of sentences by word order between groups.

(A) Age groups: young vs. older		
	Young	**Older**
Standard-first (*SD*)	4,843.86 (940.53)	5,692.64 (1,478.81)
Comparee-first (*SD*)	4,679.77 (899.37)	5,822.82 (1,656.61)
(B) Performance groups within older adults: above-chance vs. chance		
	Above-chance	Chance
Standard-first (*SD*)	5,582.73 (1,379.95)	7,207.28 (2,494.42)
Comparee-first (*SD*)	5,668.66 (1,486.89)	7,570.79 (2,737.35)

Figure 2. Interaction between working memory and word order on response time by age and performance group within old adults. Log-transformed response times (in milliseconds) as a function of working memory (WM) *z* scores and word order condition (comparee-first vs. standard-first) among older adults. Panel A shows the interaction by age group (young vs. old adults). Panel B shows the interaction by performance group within the old adult sample (above-chance vs. chance performers). In both panels, response times decreased with higher WM scores, with a steeper decline in the comparee-first condition (green line). These results indicate that individuals with higher WM capacity benefit more from the comparee-first word order, processing it faster than the standard-first structure.



Effects of Performance Group, Word Order, and WM on Accuracy

Accuracy was modeled as a function of performance group (above-chance vs. chance), word order (comparee-first vs. standard-first), and WM *z* scores, with random intercepts for subjects and items and by-subject random slopes for word order. Results from the simpler model without the WM term are presented in Appendix A, and results from the full model are presented in Appendix B.

A significant main effect of performance group was found (estimate = .69, *SE* = .16, *z* = 4.32, *p* < .001), with the above-chance group showing higher accuracy (*M* = .90 *SD* = .30) than the chance group (*M* = .64, *SD* = .48). The main effect of word order was not significant (estimate = −.23, *SE* = .19, *z* = −1.17, *p* = .24), indicating comparable performance between comparee-first (*M* = .88, *SD* = .33) and standard-first (*M* = .88, *SD* = .33) conditions. WM *z* score was significant (estimate = .35, *SE* = .11, *z* = 3.27, *p* < .01), indicating that individuals with higher WM scores performed better overall in comparative constructions. Neither the interaction between performance group and word order (estimate = .05, *SE* = .10, *z* = 0.55, *p* = .58) nor the interaction between word order and WM *z* score (estimate = −.06, *SE* = .07, *z* = −0.77, *p* = .44) reached statistical significance (see Table 3B and Figure 1C).

Effects of Performance Group, Word Order, and WM on RT

A linear mixed-effects model examined logRT as a function of performance group (above-chance vs. chance), word order (comparee-first vs. standard-first), and WM *z* scores, with random intercepts for subjects and items (see Appendix C). There was a significant main effect of performance group (estimate = −.12, *SE* = .03, *t* = −3.84, *p* < .01), with the chance group responding more slowly (*M* = 8,121.54, *SD* = 2,904.34) than the above-chance group (*M* = 5,675.64, *SD* = 1,485.47).

The main effect of word order was not significant (estimate = −.00, *SE* = .02, *t* = −0.29, *p* = .77), as RT were similar for comparee-first (*M* = 5,863.15, *SD* = 1,754.49) and standard-first (*M* = 5,739.57, *SD* = 1,596.02) conditions. The interaction between group and word order was also not significant (estimate = .01, *SE* = .01, *t* = 0.71, *p* = .48), indicating that word order did not differentially affect RT across performance groups (see Table 4B and Figure 1D).

The main effect of WM *z* score was not significant (estimate = −.03, *SE* = .02, *t* = −1.60, *p* = .11). However, the interaction between word order and WM *z* score was significant (estimate = −.01, *SE* = .00, *t* = −2.70, *p* < .01), suggesting that the effect of WM on RT differed by word order. Trend analysis showed that the WM effect was significantly greater in the comparee-first condition (estimate = −.04, *SE* = .02) compared to standard-first (estimate = −.02, *SE* = .02), demonstrating that those with higher WM processed comparee-first faster than standard-first constructions (*p* = .01; see Figure 2B).

Discussion

The present study investigated whether even structurally simple comparative sentences can reveal age-related effects, alongside the influence of WM capacity, on sentence processing using a sentence–picture matching paradigm in a predicate-final language. Older adults exhibited significantly lower accuracy and longer RT for comparative sentences compared to young adults. This pattern is consistent with prior research documenting age effects on relatively simple sentence types, such as comparatives (Hwang et al., 2024; Obler et al., 1991), implausible instrumental sentences (Sung et al., 2022), and pseudoverbs in pronoun–verb constructions (Poulisse et al., 2019).

We also found significant effects of WM on both accuracy and RT, such that individuals with lower WM capacity exhibited reduced accuracy and longer RTs—paralleling the aging effects observed earlier. Because WM capacity declines with age and significant group differences were evident in WM scores, the WM effects appear to align closely with the aging effects—consistent with previous studies reporting age-related reductions in WM capacity (Caplan & Waters, 2005; Sung et al., 2017; Waters & Caplan, 2001). Due to these features, WM effects should be interpreted with caution, as they may simply echo aging effects. Consistent with this interpretation, we also observed significant interactions both between word order and age group and between word order and WM.

Young adults were faster on comparee-first constructions compared to standard-first constructions, whereas older adults showed the reverse pattern. A similar trend emerged in the interaction between word order and WM, with individuals with higher WM capacity processing comparee-first constructions more quickly, while those with lower WM capacity performed better on standard-first structures. Taken together, these results indicate that age effects and WM effects operate in a parallel manner, consistent with the decline of WM capacity in aging. The significant interaction between word order and age group effects indicates that, in Korean, young adults' preference for comparee-first constructions aligns with the language's syntactically favored nominative-NP-first order, as shown in production studies (Sung et al., 2019). This suggests that young adults rely more on CM-driven syntactic templates, whereas older adults tend to rely on isomorphic mapping strategies that follow the semantic sequence of the comparison (O'Grady & Lee, 2001, 2005). Considering the significant interaction between word order and WM capacity, aging-related declines in WM may contribute to a greater reliance on semantic sequencing rather than syntactic markers. These findings are partially consistent with Hwang et al. (2024), who employed the same comparative sentence–picture paradigm with online eye-tracking measures. In that study, a significant interaction between age group (young vs. older) and word order was observed in eye-tracking data, as reflected in target advantage scores. Specifically, at the predicate position, young adults showed higher target fixations for comparee-first sentences, whereas older adults showed higher fixations for standard-first sentences. Notably, Hwang et al. (2024) did not find significant interactions between word order and age group for either RT or accuracy, possibly due to their limited sample size (n = 26 per group). By contrast, with a larger sample in the present study, we were able to detect these patterns as significant in offline RT measures, although not in accuracy. Interestingly, word order effects as a function of age group or WM emerged only in RT, suggesting that RT—by reflecting processing speed—may be more sensitive than accuracy in capturing subtle influences of word order, aging, and WM. These findings align with theoretical accounts proposing that processing speed declines with age and serves as a critical index of WM capacity (Salthouse, 1992, 1996).

To further examine individual differences among older adults, we compared above-chance and chance subgroups. Only 11 participants (approximately 10% of the total sample of 106) fell into the chance-level range, suggesting that the current experiment was not overly challenging for the participants overall. No significant interactions between group and word order were observed for either accuracy or RT within the older adult group. However, a significant interaction emerged between word order and WM capacity for RT, indicating that WM effects were greater in the comparee-first than in the standard-first condition. Specifically, older adults with higher WM capacity processed comparee-first constructions more quickly than standard-first constructions, whereas those with lower WM capacity showed the reverse pattern. This mirrors the findings from the age group analyses (young vs. older; see Appendices D and E for statistical details on the young vs. above-chance vs. chance comparison). As illustrated in Figure 2B, WM effects within the older group analyses appear aggregated and strengthened, replicating the same pattern observed in the age group comparisons (Figure 2A). These within-group results reconfirm the hypothesis that the syntactic template–based comparee-first construction (nominative-NP-first) imposes greater processing demands due to its lack of isomorphism, which requires rearranging thematic roles to align with the temporal progression of comparison. Such demands may explain why individuals with limited cognitive resources struggle more with these constructions despite their structural simplicity.

In conclusion, this study demonstrates that even simple comparative constructions can serve as sensitive indicators of age-related cognitive decline in sentence processing. The observed interaction between WM capacity and syntactic structure reinforces the importance of considering both cognitive and linguistic factors when evaluating language comprehension across the lifespan. This finding aligns with prior work showing that demographic and cognitive variables influence comprehension of syntactically complex forms such as center-embedded clauses (Caplan & Waters, 2005; Goral et al., 2011) and passives (Sung et al., 2017). Our results extend this evidence by showing that such influences are not only limited to complex syntax but also shape the processing of syntactically simple structures—underscoring the nuanced interplay between cognitive resources and linguistic form in aging.

Author Contributions

Jimin Park: Formal analysis, Data curation, Writing – original draft. **Seunghun J. Lee:** Conceptualization, Methodology, Formal analysis, Writing – review & editing. **Jee Eun Sung:** Conceptualization, Methodology, Supervision, Writing – original draft, Writing – review & editing.

Data Availability Statement

The data sets generated and/or analyzed during the current study are not publicly available due to restrictions imposed by our institutional review board. However, detailed model specifications (R formulae) for all analyses are provided in the Appendices. Further information can be requested from the corresponding author, subject to institutional and ethical approval.

Acknowledgments

This work was supported by National Research Foundation of Korea grants funded by the Ministry of Science and Information and Communication Technology (RS-2022-NR070151, RS-2024-00461617) and by the Ewha Global Excellence Program awarded to Jee Eun Sung.

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bishop, D. (2003). *Test for Reception of Grammar–Version 2 (TROG-2)*. Pearson. https://www.pearsonclinical.ca/en-ca/Store/Professional-Assessments/Speech-%26-Language/Test-for-Reception-of-Grammar%E2%80%94Version-2/p/P100008258

Bishop, D., & Bourne, E. (1985). Do young children understand comparatives? *British Journal of Developmental Psychology, 3*(2), 123–132. https://doi.org/10.1111/j.2044-835X.1985.tb00964.x

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22*(1), 77–94. https://doi.org/10.1017/s0140525x99001788

Caplan, D., & Waters, G. S. (2005). The relationship between age, processing speed, working memory capacity, and language comprehension. *Memory, 13*(3–4), 403–413. https://doi.org/10.1080/09658210344000459

Christensen, K. J., Multhaup, K. S., Nordstrom, S., & Voss, K. (1991). A cognitive battery for dementia: Development and measurement characteristics. *Psychological Assessment, 3*(2), 168–174. https://doi.org/10.1037//1040-3590.3.2.168

Dixon, R. M. W. (2008). Comparative constructions: A cross-linguistic typology. *Studies in Language, 32*(4), 787–817. https://doi.org/10.1075/sl.32.4.02dix

Goral, M., Clark-Cotton, M., Spiro, A., III, Obler, L. K., Verkuilen, J., & Albert, M. L. (2011). The contribution of set switching and working memory to sentence processing in older adults. *Experimental Aging Research, 37*(5), 516–538. https://doi.org/10.1080/0361073X.2011.619858

Hwang, J., Lee, S. J., & Sung, J. E. (2024). Eye-movement reveals word order effects on comparative sentences in older adults using a verb-final language. *Frontiers in Psychology, 15,* Article 1335536. https://doi.org/10.3389/fpsyg.2024.1335536

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122–149. https://doi.org/10.1037/0033-295X.99.1.122

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://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artiId=ART001153430

Kang, Y., Jang, S., & Na, D. (2012). *Seoul Neuropsychological Screening Battery–Second Edition (SNSB-II)*. Human Brain Research & Consulting.

Kertesz, A. (2007). *Western Aphasia Battery–Revised*. The Psychological Corporation. https://doi.org/10.1037/t15168-000

Liu, X. (2018). Effects of working memory load and age on the comprehension of passive sentences. *International Journal of Psychological Studies, 10*(3), 13–20. https://doi.org/10.5539/ijps.v10n3p13

Mack, J. E., Meltzer-Asscher, A., Barbieri, E., & Thompson, C. K. (2013). Neural correlates of processing passive sentences. *Brain Sciences, 3*(3), 1198–1214. https://doi.org/10.3390/brainsci3031198

Meteyard, L., & Davies, R. A. I. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language, 112,* Article 104092. https://doi.org/10.1016/j.jml.2020.104092

Obler, L. K., Fein, D., Nicholas, M., & Albert, M. L. (1991). Auditory comprehension and aging: Decline in syntactic processing. *Applied Psycholinguistics, 12*(4), 433–452. https://doi.org/10.1017/S0142716400005865

O'Grady, W., & Lee, M. (2001). The isomorphic mapping hypothesis: Evidence from Korean. *Brain and Cognition, 46*(1–2), 226–230. https://doi.org/10.1016/S0278-2626(01)80072-5

O'Grady, W., & Lee, M. (2005). A mapping theory of agrammatic comprehension deficits. *Brain and Language, 92*(1), 91–100. https://doi.org/10.1016/j.bandl.2004.05.009

Payne, B. R., Grison, S., Gao, X., Christianson, K., Morrow, D. G., & Stine-Morrow, E. A. (2014). Aging and individual differences in binding during sentence understanding: Evidence from temporary and global syntactic attachment ambiguities. *Cognition, 130*(2), 157–173. https://doi.org/10.1016/j.cognition.2013.10.005

Poulisse, C., Wheeldon, L., & Segaert, K. (2019). Evidence against preserved syntactic comprehension in healthy aging. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(12), 2290–2308. https://doi.org/10.1037/xlm0000707

Rasuki, M. (2020). *The effectiveness of pre-modified input and interactional feedback during task-based language teaching on the uptake and automaticity of English comparative structures*

[Doctoral dissertation, Curtin University]. http://hdl.handle.net/20.500.11937/84218
R Core Team. (2022). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.r-project.org/
Salthouse, T. A. (1992). Influence of processing speed on adult age differences in working memory. *Acta Psychologica, 79*(2), 155–170. https://doi.org/10.1016/0001-6918(92)90030-H
Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*(3), 403–428. https://doi.org/10.1037/0033-295x.103.3.403
Sohn, H.-M. (2001). *The Korean language*. Cambridge University Press. https://www.cambridge.org/gb/universitypress/subjects/languages-linguistics/asian-language-and-linguistics/korean-language
Sung, J. E. (2015). Effects of syntactic structure on sentence comprehension ability as a function of the canonicity of word-order and Its relation to working memory capacity in Korean-speaking elderly adults. *Communication Sciences & Disorders, 20*(1), 24–33. https://doi.org/10.12963/csd.15229
Sung, J. E., DeDe, G., & Lee, S. J. (2022, October 23–25). *Cross-linguistic validations of semantic and syntactic interface for aging-related changes in sentence processing* [Poster presentation]. 60th Annual Meeting of the Academy of Aphasia, Philadelphia, PA, USA.
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., Yoo, J. K., Lee, S. E., & Eom, B. (2017). Effects of age, working memory, and word order on passive-sentence comprehension: Evidence from a verb-final language. *International Psychogeriatrics, 29*(6), 939–948. https://doi.org/10.1017/S1041610217000047
Tukey, J. W. (1977). *Exploratory data analysis*. Reading/Addison-Wesley.
Waters, G. S., & Caplan, D. (2001). Age, working memory, and on-line syntactic processing in sentence comprehension. *Psychology and Aging, 16*(1), 128–144. https://doi.org/10.1037/0882-7974.16.1.128

Appendix A

Mixed-Effects Logistic Regression With Random Slopes for Word Order on Accuracy

(A) Age groups: young vs. older adults						
	Estimate	***SE***	***z***	***p***	***OR***	**95% CI**
(Intercept)	1.81	.34	5.27	< .001***	6.09	[3.11, 11.91]
Education	.11	.02	4.94	< .001***	1.12	[1.07, 1.17]
Age group	.77	.09	8.53	< .001***	2.15	[1.81, 2.57]
Word order	–.02	.15	–0.11	.915	0.98	[0.73, 1.32]
Age Group × Word Order	.09	.07	1.40	.162	1.10	[0.96, 1.25]
(B) Performance groups within older adults: above-chance vs. chance						
(Intercept)	1.02	.31	3.33	< .001***	2.78	[1.52, 5.07]
Education	.07	.02	3.02	.002**	1.07	[1.02, 1.12]
Performance group	.70	.17	4.17	< .001***	2.01	[1.45, 2.79]
Word order	–.17	.18	–0.96	.338	0.84	[0.59, 1.20]
Performance Group × Word Order	.04	.10	0.39	.695	1.04	[0.86, 1.26]

Note. R model equation for (A): accuracy ~ age group × word order + education + (1 + word order | subject) + (1 | item). R model equation for (B): accuracy ~ performance group × word order + education + (1 + word order | subject) + (1 | item). Education is entered as a covariate factor. *SE* = standard error; *OR* = odds ratio; CI = confidence interval.

$^{**}p < .01$. $^{***}p < .001$.

Appendix B

Mixed-Effects Logistic Model Examining the Impact of Group, Working Memory, and Word Order on Accuracy

(A) Age groups: young vs. older adults						
	Estimate	***SE***	***z***	***p***	***OR***	**95% CI**
(Intercept)	2.19	.35	6.20	< .001***	8.98	[4.49, 17.97]
Education	.08	.02	3.57	< .001***	1.09	[1.04, 1.14]
Age group	.60	.10	5.96	< .001***	1.82	[1.49, 2.22]
Word order	–.02	.15	–0.11	.911	0.98	[0.73, 1.32]
WM *z* score	.33	.10	3.28	< .01**	1.39	[1.14, 1.70]
Age group × word order	.08	.08	1.09	.274	1.09	[0.94, 1.26]
Word Order × WM *z* Score	.02	.07	0.23	.815	1.02	[0.89, 1.16]
(B) Performance groups within older adults: above-chance vs. chance						
(Intercept)	1.62	.35	4.63	< .001***	5.04	[2.54, 1.00]
Education	.04	.02	1.63	.104	1.04	[1.00, 1.09]
Performance group	.69	.16	4.32	< .001***	1.99	[1.46, 2.73]
Word order	–.23	.19	–1.17	.241	0.80	[0.55, 1.16]
WM *z* score	.35	.11	3.27	< .01**	1.42	[1.15, 1.75]
Performance Group × Word Order	.05	.10	0.55	.583	1.06	[0.87, 1.28]
Word Order × WM *z* Score	–.06	.07	–0.77	.441	0.95	[0.82, 1.09]

Note. R model equation for (A): accuracy ~ age group × word order + WM *z* score × word order + education + (1 + word order | subject) + (1 | item). R model equation for (B): accuracy ~ performance group × word order + WM *z* score × word order + education + (1 + word order | subject) + (1 | item). Education is entered as a covariate factor. *SE* = standard error; *OR* = odds ratio; CI = confidence interval; WM = working memory.

$^{**}p < .01$. $^{***}p < .001$.

Appendix C

Mixed-Effects Logistic Model Examining the Impact of Group, Working Memory, and Word Order on Response Time

(A) Age groups: young vs. older adults							
	Estimate	***SE***	***df***	***t***	***p***	**β**	**95% CI**
(Intercept)	8.76	.05	2.44	188.56	< .001***	.08	[–.05, .21]
Education	–.01	.00	220.50	–4.59	< .001***	–.18	[–.26, –.10]
Age group	–.07	.01	208.00	–6.24	< .001***	–.32	[–.42, –.22]
Word order	–.01	.01	30.03	–0.47	.645	–.03	[–.12, .07]
WM *z* score	–.04	.01	210.20	–3.03	.003**	–.15	[–.25, –.05]
Age Group × Word Order	–.01	.00	601.00	–3.94	< .001***	–.04	[–.06, –.02]
Word Order × WM *z* Score	–.00	.00	601.80	–1.98	.048*	–.02	[–.04, .00]
(B) Performance groups within older adults: above-chance vs. chance							
(Intercept)	8.90	.06	112.10	150.90	< .001***	.46	[.21,. 71]
Education	–.01	.00	102.60	–2.86	< .01**	–.19	[–.32, –.06]
Performance group	–.12	.03	111.80	–3.84	< .001***	–.45	[–.68, –.22]
Word order	–.00	.02	51.73	–0.29	.774	.00	[–.12, .12]
WM *z* score	–.03	.02	100.20	–1.60	.113	–.10	[–.22, .02]
Performance Group × Word Order	.01	.01	2643.00	0.71	.476	.02	[–.04, .08]
Word Order × WM *z* Score	–.01	.00	2631.00	–2.70	.007**	–.04	[–.06, –.01]

Note. R model equation for (A): response time ~ age group × word order + WM *z* score × word order + education + (1 | subject) + (1 | item). R model equation for (B): response time ~ performance group × word order + WM *z* score × word order + education + (1 | subject) + (1 | item). *SE* = standard error; *df* = degrees of freedom; β = standardized coefficient; CI = confidence interval; WM = working memory.

**p* < .05. ***p* < .01. ****p* < .001.

Appendix D

Mixed-Effects Logistic Model Examining the Impact of Group, Working Memory, and Word Order on Accuracy (Three-Group Comparison)

Age groups: young vs. above-chance vs. chance						
	Estimate	***SE***	***z***	***p***	***OR***	**95% CI**
(Intercept)	2.27	.41	5.50	< .001***	9.66	[4.30, 21.68]
Education	.03	.02	1.30	.192	1.03	[0.98, 1.08]
Chance	–1.53	.35	–4.39	< .001***	0.22	[0.11, 0.43]
Young	1.23	.22	5.54	< .001***	3.43	[2.22, 5.31]
Standard-first	.23	.32	0.73	.466	1.26	[0.67, 2.37]
WM *z* score	.33	.10	3.18	.001**	1.39	[1.13, 1.70]
Chance ×Standard-First	.17	.41	0.41	.678	1.18	[0.53, 2.64]
Young × Standard-First	–.29	.30	–0.98	.327	0.75	[0.41, 1.34]
Standard-First × WM *z* Score	–.01	.13	–0.08	.935	0.99	[0.76, 1.28]

Note. R model equation: accuracy ~ group × word order + WM *z* score × word order + education + (1 + word order | subject) + (1 | item). Reference level: group = above-chance, word order = comparee-first. Education is entered as a covariate factor. *SE* = standard error; OR = odd ratios; CI = confidence interval; WM = working memory.

$^{**}p < .01$. $^{***}p < .001$.

Appendix E

Mixed-Effects Logistic Model Examining the Impact of Group, Working Memory, and Word Order on Response Time (Three-Group Comparison)

Age groups: young vs. above-chance vs. chance							
	Estimate	***SE***	***df***	***t***	***p***	***β***	**95% CI**
(Intercept)	8.79	.06	244.10	153.86	< .001***	.26	[.09, .44]
Education	–.01	.00	216.20	–3.36	< .001***	–.14	[–.22, –.06]
Chance	.30	.06	249.50	5.19	< .001***	1.08	[.67, 1.49]
Young	–.16	.02	232.40	–6.24	< .001***	–.56	[–.74, –.39]
Standard-first	–.01	.03	33.24	–0.29	.774	–.02	[–.20, .15]
WM *z* score	–.05	.01	232.30	–3.67	< .001***	–.17	[–.26, –.08]
Chance × Standard-First	.01	.03	6,595.00	0.41	.685	.04	[–.14, .22]
Young × Standard-First	.03	.01	6,590.00	3.13	.002**	.12	[.04, .19]
Standard-First × WM *z* Score	.01	.01	6,590.00	2.32	.021*	.04	[.01, .08]

Note. *R* model equation: response time ~ group × word order + WM *z* score × word order + education + (1 | subject) + (1 | item). Reference level: group = above-chance, word order = comparee-first. Education is entered as a covariate factor. *SE* = standard error; *df* = degrees of freedom; β = standardized coefficient; CI = confidence interval; WM = working memory.

*$p < .05$. **$p < .01$. ***$p < .001$.