

An Eye-tracking Study on Word Order Preference in Korean Comparative Constructions: An Integrative Examination of Language-Internal and Language-External Factors

Yunju Nam(yjnam05@hanyang.ac.kr)¹, Gaeun Lim¹

¹Department of German Language and Literature, Hanyang University

Background: Korean comparative constructions allow flexible word order: nominative-first (Subject-Comparative [Sub-Comp]) or comparative-first (Comparative-Subject [Comp-Sub]). However, word order processing and production vary with cognitive and linguistic factors: age affects the comprehension of Korean comparative constructions [1], and head-final languages show a Long-before-Short (LbS) production preference, whereby longer constituents tend to precede shorter ones [2]. Moreover, conceptually more accessible constituents tend to be produced earlier in a sentence [3]. The present study used eye-tracking to examine how multiple factors influence word order preferences in Korean comparative sentence production, assessing the relative contributions of visual processing—the visual-spatial arrangement of the two referents (left/right) and perceptual salience (figure size)—and language-internal constraints—NP length extension (adjectival/relative clause) and predicate markedness (e.g., unmarked: *keu-da* ‘big’; marked: *jak-da* ‘small’).

Method: Twenty-nine native Korean speakers participated in a picture-sentence production task with eye-tracking. Two labeled geometric shapes were first presented individually and then displayed simultaneously in different sizes while gaze data were recorded across three conditions: a basic comparative condition and two NP length-extended conditions (see Table 1). After arithmetic operations as a cognitive distractor task, participants freely produced a comparative sentence using a sentence frame with one of four predicates provided (*keu-da* ‘big’, *jak-da* ‘small’, *gil-da* ‘long’, *jjalb-da* ‘short’; see Figure 1). Fixation Time (FT), Entry Time (ET), and Fixation Time Proportion (FTP) for each AOI (Area of Interest) of each item were aligned with production data. The effects of gaze-based predictors (first-fixated item, longer-fixated item, last-fixated item, larger item, and left-located item) and linguistic predictors (LbS preference and predicate markedness) were evaluated using generalized linear mixed-effects models (GLMMs; Order ~ Predictors + (1|Participant) + (1|Item); $N=3,189$).

Results: A strong overall preference for Sub-Comp order was observed across all conditions (93.8%). Gaze-based predictors did not significantly influence word order choice (see Table 2), although the first-produced item reliably matched some of the gaze-based predictors, such as the last-fixated item predictor and the larger item predictor. In contrast, linguistic predictors showed significant effects: the LbS preference lowered Sub-Comp choice probability, and marked predicates (*jak-da*, *jjalb-da*) each reduced it relative to *keu-da* ‘big’. In particular, these markedness effects remained significant even after gaze variables were included in the model (*jak-da* ‘small’: $\beta = -1.47$; *jjalb-da* ‘short’: $\beta = -1.07$; both $ps < .001$), indicating that unmarked predicates (*keu-da* ‘big’, *gil-da* ‘long’) maintained a stronger Sub-Comp preference, whereas marked predicates (*jak-da* ‘small’, *jjalb-da* ‘short’) weakened it.

Discussion: The current findings suggest that word order preferences in Korean comparative constructions are based on internally represented linguistic preferences rather than on real-time visual input alone. Specifically, when the longer NP occupied the Comp position, the LbS preference and the Sub-Comp preference operated as competing constraints, leading to reduced Sub-Comp order choice (SbL: 95.7% vs. LbS: 91.6%). Importantly, the LbS effect remained significant even after gaze-based predictors were included ($ps < .05$), suggesting that the LbS preference reflects an independent structural preference rather than a gaze-driven effect. Predicate markedness further modulated word order choice: marked predicates reduced Sub-Comp choices by making the visually salient larger or longer item more likely to be assigned the role of comparison standard rather than subject. This suggests that word order in Korean comparative constructions is modulated primarily by constituent length and predicate-based grammatical role assignment, rather than by visual accessibility alone. These findings indicate that the Sub-Comp frame (*A-ga B-boda keu-da*) is strongly established as the default comparative construction in Korean.

Figure 1. Critical experimental examples from successive stimulus presentations

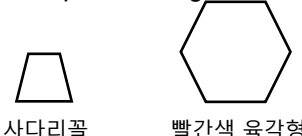
(a) Display of two shapes (e.g., Left: Trapezoid / Right: Red Hexagon)  사다리꼴 빨간색 육각형	(b) Distractor task 35+4	(c) Sentence production task (e.g., 크다: <i>keu-da</i> 'big') () () 크다
--	--	---

Table 1. Experimental conditions and example materials

Condition	NP1 (Left)	NP2 (Right)	Predicate cues
Basic	Triangle [larger]	Square	<i>Keu-da</i> ('big'); <i>Jak-da</i> ('small'); <i>Gill-da</i> ('long'); <i>Jjalb-da</i> ('short')
	Triangle	Square [larger]	
Adjective Extension (Color)	Blue triangle [larger]	Square	
	Blue triangle	Square [larger]	
	Square [larger]	Blue triangle	
	Square	Blue triangle [larger]	
Relative Clause Extension (Rel)	Triangle with a thick side [larger]	Square	
	Triangle with a thick side	Square [larger]	
	Square [larger]	Triangle with a thick side	
	Square	Triangle with a thick side [larger]	

Table 2. Statistical Analysis: GLMM of Predictors of Sub-Comp Order Choice

Predictor	β	<i>p</i> -value	OR
Gaze-based Factors (<i>N</i> =3,189)			
First_spoken_is_First-fixated item	-.161	.369	.852
First_spoken_is_Last-fixated item	.296	.089 (.)	1.344
First_spoken_is_Longer-fixated item	-.090	.611	.914
First_spoken_is_Larger item	-.16	.364	.848
First_spoken_is_Left-located item	-.003	.988	.997
Linguistic Factors			
LbS (<i>N</i> =2,545)	-.544	<.01**	.581
Extension type: Rel vs. Color	-.216	.277	.806
Predicate Markedness Model (<i>N</i> =3,189)			
Intercept (<i>keu-da</i> ; big)	6.394	<.001 ***	598.1
<i>gil-da</i> (long) vs. <i>keu-da</i> (big)	-.098	.730	.906
<i>jak-da</i> (small) vs. <i>keu-da</i> (big)	-1.162	<.001 ***	.313
<i>jjalb-da</i> (short) vs. <i>keu-da</i> (big)	-.694	.009**	.500

References

- [1] 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, 1335536.
- [2] Yunju Nam and Upyong Hong. (2013). Constituent length and word order preference in language production. *Korean Journal of Cognitive Science*, 24(1), 25-47.
- [3] Bock, J. K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, 21(1), 47-67.

Optional Supplemental Information

Korean comparative constructions allow flexible word order, with two noun phrases appearing before the predicate in either a nominative-first order (Subject-Comparative [Sub-Comp]) or a comparative-first order (Comparative-Subject [Comp-Sub]). The subject NP is marked by the nominative case marker *-i/-ga*, while the comparative NP is marked by the comparative marker *-boda* ('than'). For example, *both A-ga B-boda keu-da* (A-NOM B-than big-PRED; Sub-Comp) and *B-boda A-ga keu-da* (B-than A-NOM big-PRED; Comp-Sub) are grammatically allowed, both meaning 'A is bigger than B'.