

Verb position and flexible constituent order processing: Comparing verb-final and verb-medial languages

Savithry Namboodiripad, Lorenzo Garcia-Amaya, Alex Kramer, Stephen Tobin, Yourdanis Sedarous, Nicholas Henriksen, Julie Boland & Andries Coetzee (University of Michigan) savithry@umich.edu

Based on findings that longer dependencies result in increased processing effort, dependency-length minimization (DLM) has been shown to be a significant factor in explaining crosslinguistic tendencies in constituent order linearization (Futrell et al. 2015). However, this work also finds cross-linguistic variation with regard to how “optimal” languages are (cf. Gildea & Temperley 2010). We focus on two proposed explanatory factors for this cross-linguistic variation: headedness (discussed by Futrell et al. 2015), and flexibility (Levshina 2019; Namboodiripad 2019), and suggest that processing differences may be one source of a reduced effect of DLM in some flexible languages.

Previous research on processing in flexible constituent order languages has shown that non-canonical constituent orders result in increased processing effort (e.g. Kaiser & Trueswell 2004); this effort can be measured via lowered acceptability ratings (Weskott & Fanselow 2011). Here, we use formal acceptability experiments to compare Spanish, an SVO language in which each of the six logical orders is grammatical and attested (see 1), to previous work on English, Malayalam, Korean, and Avar, the latter three of which are flexible verb-final languages (Namboodiripad 2019; Namboodiripad, Kim, & Kim 2018; Namboodiripad & Zaslansky 2019). Because Spanish is both flexible and canonically verb-medial, this allows us to test whether flexibility attenuates the preference for shorter dependencies (at least, in simple sentences) in non-verb-final languages.

117 participants in Salamanca, Spain rated transitive sentences with animate subjects and inanimate objects on a 1-7 Likert scale. Stimuli were distributed among lists pseudorandomly using a Latin Square. Each participant heard 5 tokens of each condition (recorded by a Spanish speaker from Salamanca), and 60 filler items which ranged in structure; participants heard sentences just once before being asked to rate (cf. speeded acceptability judgment experiments). We expected SVO would be rated highest, followed by verb-initial orders (as Spanish is PRO-drop), and then verb-final orders. We also expected that, because these sentences were presented without a discourse context, subject-object order would be preferred to object-subject order, though this preference may be slight (Kaan 2001).

The results (Fig.1; transformed into z-scores) show that SVO was indeed rated highest, followed VOS, VSO, SOV, OSV, and finally, OVS (means in Table 1). A linear mixed effects model was fitted (random effects = RATING & CONDITION; fixed effects = PARTICIPANT & ITEM, scaled to SVO). Model comparisons showed that CONDITION was a significant predictor of RATING ($\chi^2(5)=249$; $p>0.001$). Pairwise t-tests (Bonferroni corrected) showed no significant difference between the verb-final orders ($p=1.00$), and a marginal difference between verb-initial orders ($p=0.078$).

A comparison across languages (Fig. 2) reveals a striking result: When the verb is in non-canonical position (non-verb-final positions for Avar, Korean, and Malayalam, and non-verb-medial positions for Spanish) argument order is not a significant predictor of acceptability. This holds in every language except English, in which only SVO and OSV are grammatical. These results call into question explanations for cross-linguistic similarities based on headedness. Further, Spanish is more flexible than Korean, despite Korean having case-marking on all arguments; these results support recent corpus-based findings that case-marking does not correlate with flexibility, given a non-categorical approach to crosslinguistic variation (Levshina, submitted). Processing-based explanations for language-specific patterns, and implications for approaches which assume that syntactic dependencies directly correspond to processing effort (e.g. Miyamoto & Takahashi 2001), will be discussed further in the presentation.

- (1) 'the boy cleaned the bathroom'
- a. el chico limpia el baño
the boy cleaned the bathroom
SVO
 - b. el baño limpia el chico
the bathroom cleaned the boy
OVS
 - c. limpia el chico el baño
cleaned the boy the bathroom
VSO
 - d. cleaned limpia el baño el chico
the bathroom the boy
SOV
 - e. el chico el baño limpia
the boy the bathroom cleaned
OSV
 - f. el baño el chico limpia
the bathroom the boy cleaned
OSV

CONDITION	MEAN 1-7 RATING	MEAN Z-SCORE
SVO	6.51	1.16
OVS	3.16	-0.30
VSO	4.86	0.44
VOS	5.10	0.54
SOV	4.35	0.21
OSV	4.28	0.19

Table 1: Mean ratings by condition (ordered as in (1))

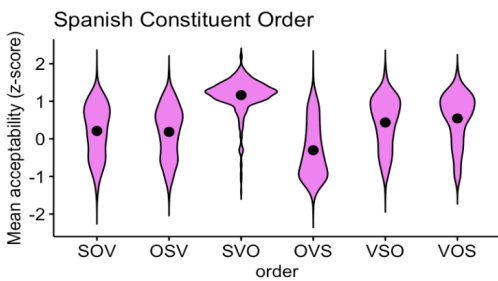


Figure 1: Violin plot of Spanish constituent order; dots represent mean z-scored responses

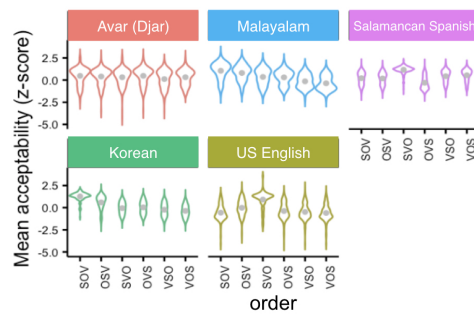


Figure 2: Crosslinguistic comparison, ordered by degree of flexibility; dots represent mean z-scored responses