The effect of dependency length minimization on ordering acceptability

Motivation: Dependency Length Minimization (DLM; Temperley, 2007), which predicts a preference for shorter constituents to be closer to their syntactic heads, is an example of how syntactic variation is shaped by processing efficiency (Hawkins, 2014). Liu (2020) conducted a corpus study investigating this principle across 34 languages, using the double PP construction as a test case (e.g. *we **talked** [PP1 about PPs] [PP2 with the reviewers]*). The results showed that, although there is a crosslinguistic preference for DLM, this preference is weaker or does not exist in head-final contexts (e.g. preverbal orderings in Hindi) as compared to head-initial ones (e.g. postverbal orderings in English). In other words, the findings indicated that while shorter dependencies are possibly more preferred, DLM may affect acceptability differently in languages with distinct typological features.

Current study: We ask whether this pressure to shorten dependency lengths is reflected in acceptability judgments. This is an empirical question different from previous corpus and comprehension/production studies (Gibson, 1998) that has yet to be addressed across languages. We compare the double PP construction in English and Hindi, building on previous work in two ways: First, most evidence for DLM comes from corpora; it remains an open question if the preference for shorter dependencies holds in acceptability experiments. Second, most studies on constituent order have taken data from written texts or used written stimuli, yet ordering preferences can differ in spoken and written domains (Kramer, 2020, 2021), and the influence of DLM is weaker in spontaneous speech (Liu, 2019). This suggests that the processing constraints of DLM could operate differently with spoken language stimuli.

Methods: We conducted a preregistered acceptability judgment study using 20 sets of audio stimuli (Sedarous & Namboodiripad, 2020). Data collection from Hindi is ongoing; we report the English results here. Items had animate subjects and intransitive head verbs immediately followed by two adjacent PP dependents (Table 1). Each stimulus set included four conditions: the baseline was the SHORT-SHORT condition where both PPs had a length of 3; the lengths of the PPs were then varied in the other three conditions via attaching a relative clause (length 3) which modified the nominal head within the PP. 128 English-speaking participants heard 5 items from each condition (plus 60 fillers of varying acceptability) and rated them on a 1-7 Likert scale. The ratings were transformed into by-subject z-scores, then subjected to Bayesian mixed-effect analyses (predicting z-scored ratings as a function of condition with long-long as the reference level; random effects for items and participants).

Results & Discussion: We expected (1) LONG-LONG sentences would have the lowest mean ratings, and correspondingly the lowest coefficient value in the mixed-effect models; (2) the second lowest mean ratings and coefficient for LONG-SHORT sentences; (3) the potential rating differences between the SHORT-SHORT and SHORT-LONG conditions are less straightforward, since although the former has shorter dependencies, the latter abides by DLM. Results for English (Fig. 1 & 2) show LONG-LONG sentences do have the lowest mean ratings and coefficient value (mean z-score = -0.09; β = -0.09); SHORT-SHORT sentences are the most acceptable (mean z-score = 0.78; β = 0.88). However, we found no significant difference between the LONG-SHORT (mean z-score = 0.25; β = 0.35) and SHORT-LONG conditions (mean z-score = 0.22; β = 0.32), indicating that dependency length may not play a strong role in acceptability, at least with audio stimuli, as compared to how it predicts online processing behavior with written stimuli or patterns in corpora.

Table 1. Sample Stimuli in the English Experiment; relative clauses are underlined; the NPs within the PPs were all definite.

Condition	Sample Stimulus
short-short	The researcher looked [PP1 through the lens] [PP2 at the sky].
long-long	The researcher looked [PP1 through the lens that was adjusted] [PP2 at the sky that was darkening].
short-long	The researcher looked [PP1 through the lens] [PP2 at the sky that was darkening].
long-short	The researcher looked [PP1 through the lens that was adjusted] [PP2 at the sky].

Figure 1. Density plot of the z-scored acceptability ratings in the four conditions; the dashed line represents the mean z-scored rating.

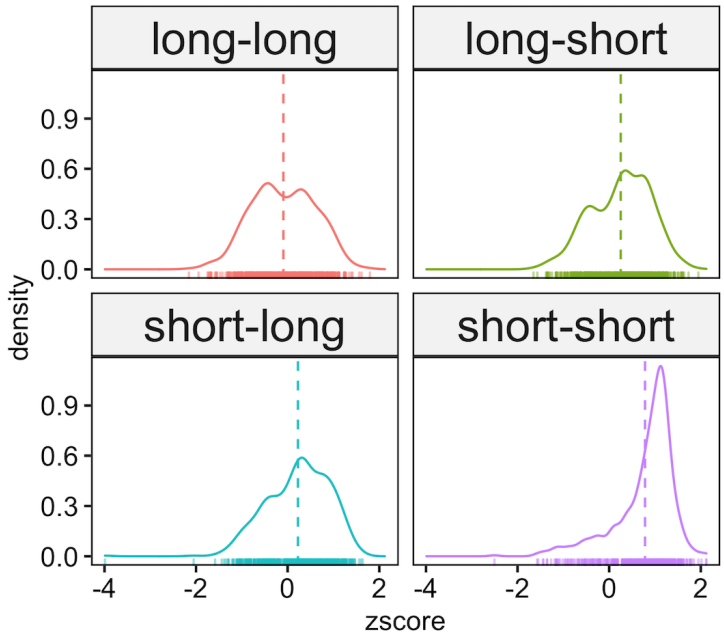

Figure 2. Coefficients of the four conditions predicting acceptability ratings; 95% confidence intervals were derived from the posterior of each parameter in the Bayesian mixed-effect model.