Spanish L2 Learners' performance attending to gender grammar cues

Jorge L. Vargas Mutizabal[1]

[1] Rutgers - The State University of New Jersey

Author Note

Correspondence concerning this article should be addressed to Jorge L. Vargas Mutizabal, . E-mail: jorge.vargas@rutgers.edu

Spanish L2 Learners' performance attending to gender grammar cues

Data analysis

Reaction times

In preparation for data analysis, incorrect responses were removed, and the data distribution of reaction times was visually inspected. The long tail of the distribution suggested the presence of outliers, and the Interquartile Range (IQR) method was used to remove them. The difference between the first and third quartiles was computed to obtain the range of the middle 50% of the distribution. Values that fell below 1.5 x IQR of the Q1 or above 1.5 x IQR of the Q3 were removed. The original data set had 1,343 observations, and 44 observations were removed, keeping 1,299 data points. A series of nested linear-mixed model comparisons was used to assess the effect of session and conditions (fixed effects) on reaction times, and the random effects were defined as the interaction between the session and condition, item number, and participant. The null model was fitted without fixed effects, while model 1 included session. Then, Model 2 added the session and condition, while Model 3 tested the interaction between them.

Accuracy

Similar to reaction times, different models were used to assess how accuracy changed with session and conditions (fixed effect), and the random effects were also the interaction of session and condition, item, and participant. Although logistic regression was used for the binary outcome (correct or incorrect response), the family used was binomial with a logit link function. The null model included accuracy, and the second model included session, while the third model added session and condition. Lastly, the interaction between session and condition was tested in the fourth model.

Results

Reaction times

Reaction times for the pretest and posttest decreased across all four conditions: Cond. 1 (2,386 ms to 2,014 ms), for Cond 2 (2,580 ms to 2,143 ms), for Cond 3 (2,471 ms to 2,171 ms), and for Cond 4 (2,439 ms to 2,280 ms). Figure 1.**??** Illustrate the mean reaction times for the two sessions. However, the nested model comparisons (model 1) showed that session did not have a main effect ($X^2(2) = 3.87$, $p = 0.14$), suggesting that participants' reaction times did not significantly improve in the posttest. Then, adding conditions and session (model 2) showed a main effect ($X^2(3) = 19.66$, $p < .001$), suggesting that reaction times in Conditions 2, 3, and 4 significantly varied from Condition 1. Lastly, the interaction between session and condition (model 3) did not show a main effect ($X^2(6) = 6.99$, $p = .322$), indicating that participants did not significantly improve their reaction times across conditions from the pre to the posttest.

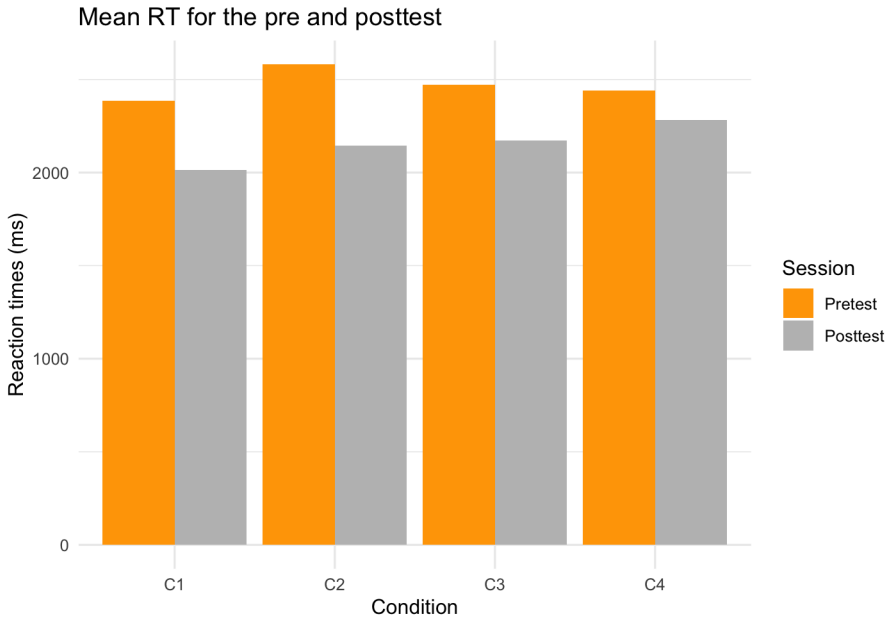

Figure 1

Accuracy

The mean accuracy also improved across the four conditions between the pre- and posttest. For Cond. 1 (96.5 % to 97.6 %), Cond 2 (94.8 % to 98.2 %), Cond 3 (93.7 % to 97.0 %), and in Cond 4 (72.7 % to 91.5 %). Figure 2. Shows the accuracy trends between the two sessions. However, the nested model comparisons did not report a main effect for session ($X^2(2) = 4.99$, $p = 0.083$), suggesting that participants did not significantly improve across the four conditions. Later, adding session and condition, did not show a main effect for condition, indicating that accuracy did not significantly vary with phrase condition ($X^2(3) = 5.49$, $p = 0.139$). Lastly, the interaction between session and condition did not report a main effect ($X^2(6) = 2.82$, $p = 0.832$), which implies that participants did not significantly improve from the pretest to the posttest.

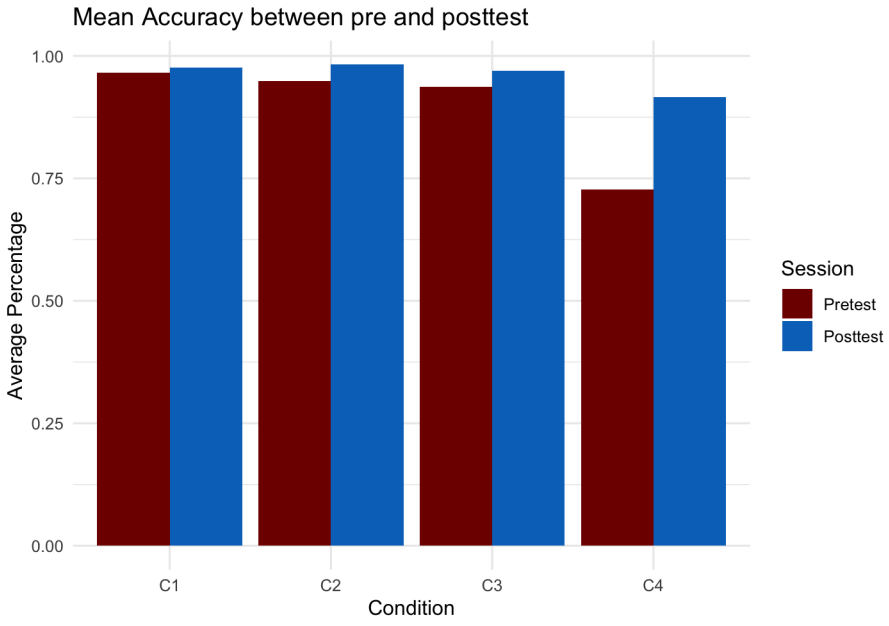

Figure 2