

Reply to Reviewer 2

2023-12-20

Dear Reviewer 2,

Thank you for your thoughtful feedback regarding this manuscript. This has been very helpful and I am grateful for the opportunity to improve the quality of this work. I have taken care to address each of your comments and suggestions below. Your comments have been copied to this document in plain text, and my replies are below in italics.

Introduction

I would exercise caution when using “L3 development” in the first line. Notice that the vast majority of work has focused on the initial stages and has not delved much into development per se.

I reworded this to “the initial stages of L3 development”.

I believe the debate has shifted slightly from holistic structural similarity to the order of acquisition. Considerable evidence has emerged indicating that the order of acquisition might not be a key factor. The central debate now revolves around holistic vs. structural similarity. That said, I can see how this replication study adds an interesting dimension to the discussion. It’s worth emphasizing that while there is substantial evidence against the order of acquisition, the current design enables us to draw more generalizable findings.

I added this idea in the section in that mentions mirror image groups.

Methods and Participants:

While I commend the authors for obtaining such a large sample size, I wonder whether there is enough control regarding the participants’ language backgrounds. The level of Spanish proficiency in the L1 English group appears lower than their proficiency in Portuguese, which poses challenges for the questions being asked.

Added this observation to the discussion section.

The low level of Spanish proficiency in the L1 English group complicates the interpretation of the results. Recent literature advocates for testing all languages to ensure that the underlying grammar for a specific domain is acquired.

I added a note on this matter to the participant description.

Results & Discussion:

In the pre-posed interpretation task, the L1 Spanish group exhibits a significantly lower proportion of correct responses than the L1 English group. It would be beneficial if the authors could delve deeper into this discrepancy. Why do we observe this difference? Why does the L1 English group outperform the other group? It may be worthwhile to consider linking this to CITH (Cabrelli and Iverson, 2023), especially given that participants in this study are not at the initial stages of learning L3 Brazilian Portuguese (BP). It’s not immediately apparent why the L1 English group would perform better than the other group.

I added a short paragraph addressing this point. I am not sure why this surprising result was found, and suggest that this could be a focus for future work. Additionally, I mention the CITH, but I believe it would predict the opposite of the result found here.

I believe the statement the authors make on line 21 needs further refinement.

I added information regarding the interpretation of the simultaneous presence an equivalence test and t-test, which reviewer 1 also suggested.

How does the LPM explain the results? In its current form, it seems capable of explaining everything, but if the authors wish to discuss the results in relation to the study, they should derive predictions from it. A couple of thoughts below: o a. Spanish and BP seem to function similarly in this domain, so we would always expect to see transfer from Spanish, specifically for this one domain. Why does one group perform better than the other? The LPM wouldn't predict differences here.

Thank you for this point, and I agree. I have changed this conclusion and I added some nuance here about the prediction of intermediate performance by the LPM, though this is outside the scope of the present design. My point is largely that the LPM and TPM both would indeed predict Spanish to influence Portuguese, but it appears that this influence is predicted to occur to distinct degrees, since the LPM predicts that the presence of a second background grammar (L3 processing) causes less successful performance than a single background grammar (L1 processing). So, I agree that this design does not allow for the teasing apart of the TPM and LPM and I have amended this in the manuscript.

o Once again, the final statement from the TPM is misleading. The TPM doesn't make predictions regarding mismatches between perception and production; in fact, it argues that these are two distinct processes likely to yield different outcomes.

I agree and have edited this also to remove the mismatch between perception and production.