Proposal: a study L3 production at first exposure

# General overview

The present proposal advocates for the use of highly proficiency late bilinguals (AOA > 12) in research on L3 phonology. At present, many studies examine L3 production in L3 learners who display a range of both L2 and L3 proficiencies. By examining bilinguals, rather than true L3 learners, evidence may be gathered and inform how L3 phonological acquisition begins. Taken together, the results of the present proposal and the results from the previous body of literature on L3 phonology can provide a roadmap of L3 phonological acquisition.

This proposal is a proposed replication of a pilot study which has already been carried out in Spanish-English bilinguals. Following the results of this pilot study, I propose a replication with Polish L1 English L2 speakers who will learn L3 Spanish. The proposal procedure and data analysis would remain the same, with the French set of words being changed for a Polish set of words.

# Pilot Study

A pilot study was carried out in Spanish-English bilinguals in which they were taught single /p/ initial words in French, a language that they did not know. 8 participants who speak L1 Spanish and L2 English completed a French shadowing task training phases followed by three elicited production tasks in Spanish, French, and English of /p/ initial words. Voice Onset Time (VOT) was then measured for each production in each participant and analyzed using a generalized mixed effects model (GLMM). The results provide evidence that these participants produced p-initial French words as more aspirated than their Spanish production. A post-hoc test of equivalence then suggested that the participants’ production patterns in their L2 (English) and their L3 (French) were statistically equivalent.

# Methods

## Participants

The participants were Spanish native speakers who learned English as a second language. 8 participants in the L1 Spanish group and completed the pilot study. These participants were recruited from the online platform prolific.

## Materials

The experiment consisted of a shadowing task for training and three elicited production tasks (experiment found here: <https://www.labvanced.com/player.html?id=18621>). The tasks involved producing single /p/ initial words in each of the languages. The word list used in the present study was a subset adopted from Llama and Cordoso (2011).

## Procedure

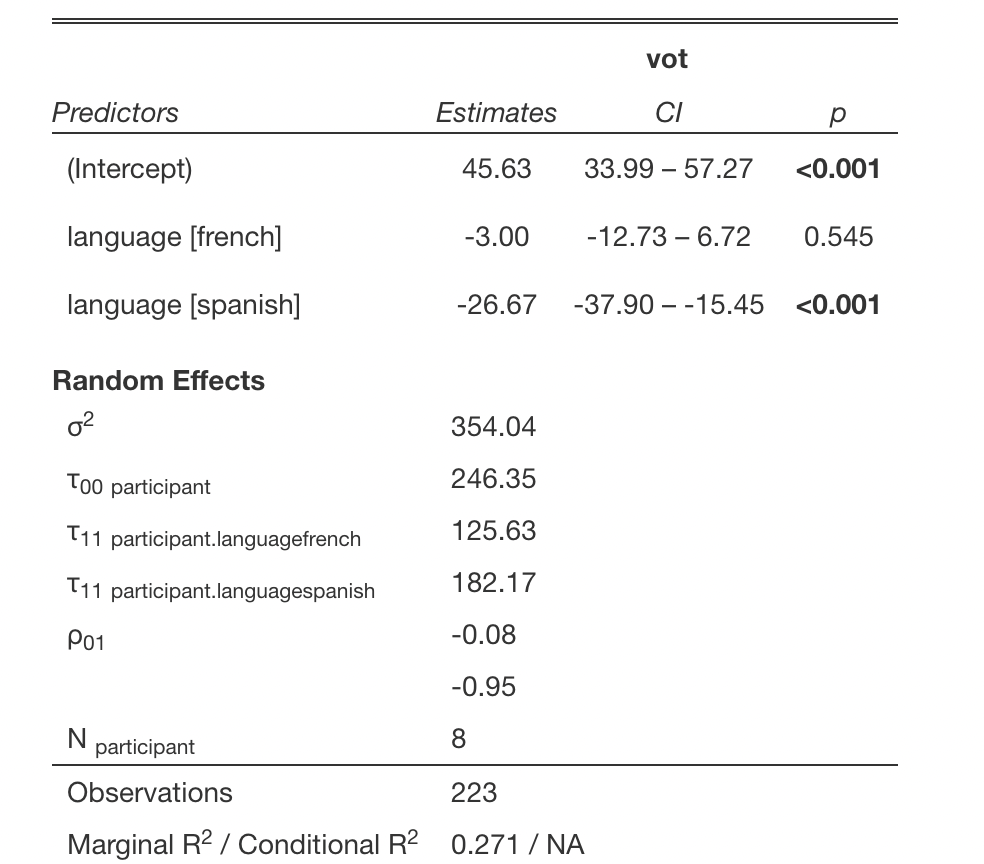
First, the participants completed a French shadowing task. During this task, one word was displayed on the screen at a time. The participants could then play an audio clip of each word while viewing the word on the screen. The participants had a separate button on the same screen which they then were instructed to use after listening to the target word. Next, the participants recorded words in Spanish and English in consecutive sessions. Words were presented in isolation and one at a time. Participants recorded themselves saying words out loud and submitted each word immediately after recording. Finally, participants saw the same French words as the shadowing tasks, and recorded themselves pronouncing each word. These recording were used for analysis. Each participant produced 10 tokens in English, 8 tokens in Spanish, and 10 tokens in French, resulting in a total of 79 tokens in English (1 token was removed due to poor audio quality), 80 tokens in French, and 64 tokens in Spanish.

## Analysis Plan

A generalized linear mixed effects model (GLMM) was used to analyze the data. VOT (in MS) was modeled as a function of language (Spanish, English, French). In order to account for the nested nature of the data, the random effects structure included random intercepts per participant and random slopes per language. Experiment-wise alpha was set to .05.

# Predicted and pilot results

The mean VOT for French productions was 42.62 (SD = 26.89), which was rather close to the mean English production (45.77 SD = 27.16). The mean VOT for Spanish was 18.95 (SD = 9.63). The results of the GLMM for the pilot data can be seen below.



The model output fixed effect parameters suggest that relative to the baseline of English VOT, the difference between French and English productions is 3ms. On the other hand, the difference between Spanish and French is 26ms.

## Post-hoc test of equivalence

Based on the output of the GLMM, it was of interest whether evidence could be provided that the participants were producing French equivalently to their L2 English. The equivalence test carried out suggests this is the case, t(156.92) = -2.417, p = 0.00839, given equivalence bounds of -13.513 and 13.513 (on a raw scale) and an alpha of 0.05. The raw equivalence bounds were created by setting the effect size to -.5 and .5 for the upper and lower bounds respectively.