**COMMENTS FOR THE AUTHOR:  
  
Reviewer #2: This manuscript investigates whether English learners of Spanish use prosodic cues to anticipate Spanish inflectional suffixes as well as monolingual Spanish speakers and whether the results correlate with learners' L2 proficiency and working memory. This is a well-designed study and a well-written paper and provides with detailed statistical data. However, I hope the author(s) clarify the following points in a potential revised version:**

**(1)     Since the carrier sentences were the same for each pair of the target words, could it be possible that leaners, esp. the beginners, simply focused on the target syllable (even ignored the sentence meaning) and only relied on the auditory cues to identify the target words instead of really processing the lexical information in the eye-tracking experiment?**

The carrier sentences were identical for all the trials in the gating task (*the person says*  + target word), but the they were unique to each trial in the eye-tracking task. The eye-tracking task followed a standard within-subjects design, such that the same subject was never exposed to both conditions of the same carrier sentence. Thus, if subject 1 listened to sentence 01 in condition 1 and sentence 02 in condition 2, subject 2 listened to sentence 01 in condition 2 and sentence 02 in condition 1. This was in the original manuscript: “Participants were randomly assigned to one of two versions. Each version contained one of the two conditions of a given word pair (e.g., if *firma* ‘(s)he signs’ appeared in version 1, then *firmó* ‘he signed’ appeared in version 2).” We think these sentences make it clear that a given subject was never exposed to 2 conditions of the same sentence, but we will be happy to add more information if needed.

They ask if it is possible that the learners only relied on auditory cues without processing lexical information. I think we should mention/remind the reviewer that the participants had to click the mouse after processing the sentence and that we only used correct responses.

**(2)     I don't quite understand why the correlation of the results in the WM task and the gating task was examined since the auditory carrier phrase was always the same and the subjects needed only to focus on such short target words and their competitor in the gating task.**

We investigated WM effects in anticipatory abilities (see RQ3), not in lexical processing. This could be obvious at a word level in a simple task. In fact, we found a marginal main effect for WM in accuracy. We added this sentence to our RQ3 hypotheses: “It is noteworthy that our hypotheses relate to WM and anticipatory abilities, not to WM and overall lexical processing.” We will be happy to further justify the inclusion of WM if requested.

**(3)     Better to provide an appendix with the complete set of the target words and their carrier sentences in the stimuli to make the manuscript reader-friendly.**

An appendix with target words was added. The carrier sentences were excluded because they were the same for the gating task and they would make the manuscript too long (the manuscript is already over 11,400 words).

**(4)     Some minor points:  
a.      Page 9, 1st paragraph, "Spanish, on the other hand, is considered a syllable-timed language, i.e., each syllable has roughly the same duration and vowel quality, regardless of stress placement." Better to say something like "the duration and vowel quality do not vary if the stress placement shifts in a word".**

Change implemented.

**b.      Page 21, 2nd paragraph, "Figure 3" must be "Figure 2", and so as to other figure numbers in the followed text.**

Joseph, do this.

**c.      Page 24, 2nd paragraph, "mean: 0.65" must be "mean = 0.65".**

Changed.

**d.      Page 27, 1st paragraph, better to provide references for the last sentence.**

Nuria adds this.

**e.      Page 30, the last 5th line, "probably to to" with an extra "to".**

Changed.