Dear Editor,

We would like to start by thanking you for considering our paper entitled “Lexico-syntactic interactions in the resolution of relative clause ambiguities in a second language (L2): The role of cognate status and L2 proficiency” (ID 58889) for publication in the “Psicologia” Journal. We also want to thank you and the reviewers for the careful work done and for all the comments and suggestions, which we found to be highly valuable and useful to improve the manuscript.

As you can see in our response letter, we addressed all the issues raised by the two Reviewers, making all the necessary changes in this new version of the manuscript. We addressed each issue sequentially, and included both comments and responses for Reviewer #1 and Reviewer #2 separately. Changes to the manuscript are signaled by specifying the page number where the change has been made. Finally, edits and additions to the manuscript are highlighted in yellow.

Thank you for reviewing our work.

We look forward to hearing from you.

Best regards,

Ana Paula Soares

Department of Basic Psychology, School of Psychology
University of Minho, Campus de Gualtar
4710-057 BRAGA - PORTUGAL
Phone: +351 253 604 236 / Fax: +351 253 604221
E-mail: asoares@psi.uminho.pt
REVIEWER 1:

First, we would like to thank Reviewer 1 for the time spent doing this thorough and careful review. In our opinion, all suggestions were extremely relevant for the improvement of the manuscript. Thank you once again. All changes are highlighted in yellow in the body of the text. The page number in which they have been inserted is also reported to facilitate their location throughout the manuscript.

1) General comments and suggestions

My main point of concern revolves around the cognate effect expected (and found) in Experiment 2 (sentence completion in L2-English for native L1-Portuguese intermediate and advanced learners of English). First, it should be stressed that none of the three predictions made by the authors (see page 11 of manuscript) is fulfilled, namely, (1) a general interference effect of L1 preferences toward high attachment, caused by cognates, in ambiguous L2 RC resolution; (2) a more pronounced interference effect toward high attachment in L2 when the cognate is located at N1 position; and (3) a stronger interference effect from L1 in L2 RC interpretation for intermediate learners. In fact, exactly the opposite pattern was found to that laid out in predictions 1 and 2, and no difference was obtained between the two groups of learners (prediction 3).

Given these unexpected results, the authors try to account for them by using a two-step reasoning: (1) high attachment is more costly than low attachment generally; and (2) the presence of cognate words enhances the cognitive load participants endure due to the activation of the L1, thus ‘facilitating’ the supposedly easier low-attachment strategy in processing L2 (English) sentences, which happens to be the preferred resolution of the ambiguity for native speakers of this language. So we end up with the paradoxical result that the L2 learners in this study behave very much like native speakers. There are several problems with this explanation of the results, aside of their being inconsistent with the predictions. First, the arguments set forth by the authors in the General Discussion section sound inevitably ad-hoc, after the fact explanations. The difficulties the authors run into when trying to provide a convincing interpretation of the results are especially acute when they try to explain why the NC-C condition yields a greater percentage of high-attachment responses (though note that the figure is close to chance performance in attachment choices) (see end of page 27). Apparently, in an effort to make their explanations more plausible, the authors focus on two pairs of comparisons that seem to be better suited to their interpretation of the results: C-C vs NC-NC, and C-NC vs NC-C. In both cases, there is a complexity effect associated to cognates, with more low attachment (L2) preferences in sentences with cognates, in the former case, and in sentences with cognates at N1 position, in the latter case. However this overlooks the cumbersome fact that, contrary to what one should expect, the NC-C condition yields more low-attachment responses than the NC-NC condition, which should generate the
greatest amount of low-attachment choices according to the authors’ line of reasoning. However, they remain silent on this issue.

Thank you for pointing this out. Indeed, the results obtained in Experiment 2 did not support our hypotheses. However, it is important to note that even though they were exactly in the opposite direction of what was expected (note that the hypotheses formulated were tentative since, to the best of our knowledge, no previous studies were conducted on this topic), they were, in our point of view, very interesting. They clearly suggested that the lexical and syntactic levels of processing in L2 sentence comprehension seem to interact in a bilingual reading system, which is not only highly interactive within each level of processing (lexical and syntactic) as previous research has demonstrated, but, importantly, across levels of processing (see pp. 25-26 for the new information included on this issue, and also pp. 28-29). Therefore, we consider that, although the direction of the results was not as expected, they did support a more general hypothesis regarding the existence of cross-language activation between the lexical and syntactic levels of representation in the bilingual mind, a research topic that has been poorly explored in the bilingual and second language acquisition literature, as stated throughout the paper. Since previous research on this topic is scarce and, in addition, we use an offline task that is not sensitive to the temporal course of sentence processing, different explanations were presented, namely the cognitive load hypothesis, arguing that, since cognates could generate stronger cross-syntactic competition for RC attachment than noncognates, this could have overloaded the processor and stimulated the use of a local (LA) over a non-local (HA) strategy. We acknowledge that this hypothesis is only tentative and [requires] future studies using online techniques should be conducted in order to test if it can really account for the result (this is now clearly stated on p. 31, see also p. 32).

In the Discussion section, we have focused on C-C vs. NC-NC and on C-NC vs. NC-C comparisons because they were the critical comparisons at stake attending to the hypotheses. It is absolutely true, as you state, that more low attachment preferences were observed in sentences with cognates, in the former case, and in sentences with cognates at N1 position, in the latter case. However, contrary to what is said, the condition NC-C did not yield more low-attachment responses than the NC-NC condition, but the reverse: the NC-C condition produced 47.2% of high attachment responses and the NC-NC condition produced 38.4% HA responses (see Results section, p. 23). Although the comparison between these two conditions was not directly discussed, the truth is that NC-NC and NC-C were the conditions with more HA responses, compared both to the C-C and the C-NC conditions, which yielded 30.8% and 34.9% of HA responses, respectively.

Nevertheless, we acknowledge the relevance of discussing other cognate comparisons besides the critical ones mentioned above in order to fully analyze the extent to which the explanations advanced can account for the results. Therefore, in this version of the manuscript we have added new information concerning the NC-C/NC-NC and the C-C/C-NC comparisons (see p. 25-26). We have also discussed in more depth the limitations of each explanation advanced to fully account for the results (see pp. 31-32). Finally, it is also important to note that we have corrected some inaccurate
There are two other points of concern regarding the analysis and interpretation of the results of Experiment 2. One is the fact that significant differences among cognate status conditions were only found in the subject analysis, not in the item analysis. The authors do not try to give an account of this anomaly. I think the least they should do is find out the source of this null result by inspecting the behavior of the items used in their study across the experimental conditions.

Yes, you are right. The cognate effect in the F2 analysis failed to reach statistical significance. As requested, in this version of the manuscript we have added new information to account for this F2 null effect (see pp. 24-25). As you will see, we consider that this “anomaly” was due to the fact that in the item analysis the cognate factor was entered as a between-items factor, which has contributed to enhance the variability in the responses to the items within the same condition, thus making it hard for any cognate effect to emerge. Nevertheless, it is worth noting that in an attempt to analyze whether the inclusion of potential outlier items might underlie the null cognate effect observed, we conducted a second F2 analysis excluding items 2 SD below or above the mean HA responses in each condition, following a common procedure in the Psycholinguistics literature. This has led us to exclude one item from the C-C condition. In order to keep the number of items constant across conditions, we have also excluded one item from the other cognate conditions (the one that most closely approached the value of an outlier). Subsequently, we conducted the F2 ANOVA again considering the same design. A cognate null effect was still observed. It is important to note that based on this procedure we have also conducted a new F1 analysis excluding the same four items (one in each condition) from the participant data that were not considered in the second F2 analysis. Again, a strong cognate effect emerged, mimicking the F1 results reported in the manuscript. Therefore, we truly believe that the null F2 cognate effect emerges from both the limited number of items considered per condition (note that the strict control imposed to the materials has lead us to exclude several other sentence fragments that were initially constructed), and the high variability that exists in the responses to the items from the same condition, when the means in analysis were averaged over the participants and not over the items. As we also mentioned, the only way to increase the power to detect a cognate effect in the item analysis is to increase the number of sentence fragments per cognate condition (see p. 25), which should be considered in future studies (see p. 34).

The other issue I want to emphasize is the unexpected lack of differences in the responses of the two groups of participants. In this regard, the authors offer an explanation that strikes me as contradictory, since they first suggest that participants (even at intermediate levels of proficiency) seem to have attained a high level of proficiency in L2, while they subsequently propose that both groups
of participants follow a lexically-driven mode of processing, characteristic of non-native, low-proficiency speakers (see page 29 of the manuscript).

Thank you for pointing this out. We acknowledge that the discussion of this point is unclear and might even sound contradictory. Therefore, in this version of the manuscript we opted to rephrase the text (see pp. 33-34). We have also added new information on p. 23, where for the first time we present information about it in order to make it clear from the beginning.

They go on to suggest that in order to settle this and other questions, “additional studies [should be conducted] using other types of tasks, participants, and measures that are sensitive to the time course of lexical and syntactic processing”. I would like to capitalize on this remark, since I think that some additional source of evidence drawn from an online experiment should in fact be added to this paper, in order to clarify the questions that are left open. This leads to my last remark of this section of my review on general comments. Though it might be correct to say that offline tasks have yielded more reliable differences than online tasks as far as attachment choices is concerned (see page 10), this kind of task is not as reliable as online methods to make statements about differences in processing load, as the authors do in order to account for their results. Given that the results of their second experiment do not square with their predictions, and that, as I have tried to explain above, the kind of explanations they offer have some noticeable flaws, I would suggest that they run a new experiment with an online task (a self-paced reading experiment would do) and the same materials, to see whether the pattern of results they have found are consistent with the offline judgments. In my view, only on this condition could the processing claims they do in their discussion of the results be upheld.

We agree. In fact, we are currently conducting an eye-tracking study with the same materials to test the processing claims presented in the Discussion. Only an online technique such as this can allow for a better understanding of the processes and mechanisms involved in L2 RC processing, and, ultimately, shed some light on the hypotheses advanced in the paper to explain the stronger L1 RC interference effect for noncognates than for cognates. However, we agree with the Editor and Reviewer#2 that the presentation of a new online experiment is not strictly necessary (note that the inclusion of the eye-tracking experiment would increase the length of the paper in such a way that its publication would be problematic). As mentioned in the manuscript, the sentence completion task used in the experiments reported here was chosen as a first step to study the lexico-syntactic interactions on L2 RC processing because the differences between EP and English RC attachment preferences were more reliable in offline than online tasks (see p. 11, see also 31). But, evidently, the interesting results found here strongly recommend that future studies using online techniques are conducted in order to shed light on the nature of the lexico-syntactic interactions
established during L2 sentence processing, a research topic that has been poorly explored in the bilingual and L2 acquisition literature (see p. 34).

2) Specific comments and suggestions (by sections)

- **Introduction:**
  
  - The expression “the non-selectivity in the access to cognate words” (p. 4) should be explained, since it is only familiar to readers acquainted with the bilingual processing literature.

    Thank you for pointing this out. In order to make this clearer (particularly to readers not familiar with the bilingual literature) we have rephrased the sentence (please see p. 4).

    - It is taken for granted (e.g. on page 6) that low attachment is the strategy used by English native speakers, but this is just a tendency, not an absolute choice, much the same as the high attachment preference is typical, but not overwhelmingly preferred, by speakers of EP, among many other languages. This probabilistic nature of attachment preferences should be mentioned in the Introduction.

    You are absolutely right. To reinforce this probabilistic view of RC attachment preferences, new information has been added, and some words were replaced when characterizing the RC preferences observed in English and EP (please see pp. 5-6).

    - When discussing the studies conducted with Dutch-English bilinguals (pp. 8-9), the authors should explain the Dutch example involving pronouns ‘die’ and ‘dat’. Also, they should mention the differences in preferences between these two languages at the beginning of the discussion, before presenting the example. And finally, from the example they cite, it is not clear why the locus of the effect is syntactic in nature. This should also be explained.

      Following your suggestion, new information has been added to the manuscript in order to better explain the use of the pronouns ‘die’ and ‘dat’ from the Desmet and Declercq (2006) study. Moreover, as requested, new information has also been added to justify why the locus of the RC priming effect is considered syntactic in nature. Finally, the differences between Dutch and English in RC attachment preferences have also been included when presenting Desmet and Declercq’s (2006) study, as suggested (please see p. 9).

    - When presenting the Schoonbaert et al study (p. 9), it is not clear whether the authors of this study manipulated the cognate status of the translation equivalents in their sentences. Please clarify.

      In the Schoonbaert et al. study, the authors did not manipulate the cognate status of the translation equivalents used, although both cognate and noncognate words
(verbs) were used in the sentences. The main manipulation was between verbs that either shared or not semantics between languages (translation and non-translation equivalents) as in the example provided in the manuscript (give-geven vs. give-verkopen). Nevertheless, new information has been added to make this clearer for the reader (please see p. 10).

- **When presenting the materials and the experimental conditions of their experiments** (p. 10), the authors should state that the ambiguous RCs in their study had two potential attachment sites, and that the combinations based on the cognate status of words refer to the two host NPs in order.

  As suggested, this information has now been included at the end of the Introduction section (please see p. 11-12).

- **General method**

  - In the description of the materials (p. 12), it should be mentioned that the cognate status variable was only relevant in Experiment 2.

    Thank you for pointing this out. As suggested, this information has now been included (see p. 14 and also p. 17).

  - The remark that “number seems to disrupt the ‘natural’ processes involved in RC attachment” (p. 12) is completely unjustified and unnecessary for the argument authors put forth.

    We agree. This information has been deleted in the reviewed version of the manuscript (see p. 14).

  - Authors should clarify what they mean by the expression “judges with specialized knowledge in English” (found on p. 13, and again on p. 17).

    Information provided (see p. 15 and 19).

  - The section on materials (pp. 12-15) should include an explicit description of the design of the experiments.

    Following your suggestion the design of the experiments is now presented at the end of the materials section (see p. 17).

- **Experiment 1**

  - Participants: authors should reveal the name of the universities where the study was carried out.

    This information has been included (please see p. 18).
Results and discussion: in the tests of comparison with chance (p. 18), the expressions “EP native speakers choose significantly more HA than LA … while English native speakers choose significantly more LA than HA” are mistaken, since these two attachment options were not directly compared in these analyses.

Thank you for pointing this out, you are absolutely right. The mistake has been corrected (see p. 20). We have also reorganized the presentation of results reporting the data from the one-sample t-tests first, since it is more easily understandable for the reader (it is also more consistent with the results reported in Experiment 2).

- **Experiment 2**

Results and discussion: the sentence “both L2 groups revealed a clear tendency to resolve the RC ambiguities by adopting a general LA strategy” (p. 20) is an overstatement, since this is only true in three of the four experimental conditions.

This was effectively an overstatement. Since this information is redundant considering the information obtained from the one-sample t-tests reported, we chose to delete it from the manuscript (see p. 22).

Results and discussion: the last sentence of p 22, “cognates seem to facilitate (i.e., to stimulate a more native-like way of resolving L2 RC syntactic ambiguities) rather than hamper L2 RC resolutions” is a bit tendentious, since it assumes that cognates bear the burden of syntactic (attachment) decisions, which seems to me unwarranted.

We agree. This information has been deleted in this reviewed version of the manuscript (see p. 25).

- **General discussion**

Pp. 23-24: The summary and interpretation of results is focused only on Experiment 2. Later in this section Experiment 1 is presented as a control experiment, which has not been mentioned previously. If Experiment 1 was indeed intended as a control experiment, this should be mentioned from the beginning, and perhaps the order of experiments ought to be reversed, with the control experiment in second place.

It is true. In the discussion we did focus more on interpreting the data from Experiment 2 than Experiment 1 because the results from Experiment 2 were the ones that allowed us to discuss the hypotheses formulated concerning the lexicosyntactic interactions during L2 sentence processing. Experiment 1 was indeed assumed as a control in the manuscript as it was only conducted to ensure that the typical RC preferences observed in EP (HA) and in English (LA) were in fact observed in our materials, which allowed us to assume HA sentence completions as a
mark of L1 syntax interference. We believe this is now clearly stated on pp 12-13, 17, and 20 (see also p. 27). However, in spite of being a control study, we consider that, for this particular paper it would be more appropriate to present it first and not second, as you suggested. Indeed, we think that before presenting the results obtained with the L2 learner groups, it is imperative to show that the RC preferences typically observed in L1 (HA) and L2 (LA) were in fact observed in our materials. Only this would validate using the HA strategy as a mark of L1 syntactic interference during L2 sentence processing, as assumed in Experiment 2. For these reasons we chose to maintain the previous order in the presentation of the Experiments in the manuscript.

- **At the end of page 23**, it is said that the C-C condition should yield more HA decisions “relative to all other conditions”. This statement is perhaps too radical, since the C-NC condition should also yield in principle a high amount of HA responses. Besides this, no direct comparison is ever made between the C-C and the C-NC conditions.

  You are absolutely right. That statement has now been rephrased (see p. 27). Moreover, we have also added information concerning C-C and C-NC comparisons (pp. 32-33).

- **On page 28**, the “cognate facilitation effect” is again mentioned (see second comment on Experiment 2 above).

  Thanks. Because we have changed the way the LS model is introduced in the discussion section (see p. 32), that statement was deleted.

- **Also on page 28**, the authors claim that the results of Experiment 2 could also be explained by appealing to Hartsuiker’s LS model. However, they follow a strange line of reasoning when they allude to a “cognitive boost” in L2 syntactic processing in this connection. I think this expression is inappropriate: if the results of Experiment 2 are better explained in terms of cognitive load, it makes no sense to appeal to a cognitive boost. On anyone’s understanding, cognates could only boost L1 (syntactic) representations in L2 sentence processing. I think this argument should be better grounded.

  We agree. We have rephrased this part of the text in order to improve our line of reasoning (see p. 32).
REVIEWER 2:

We would like to start by thanking Reviewer #2 for his attentive and careful review. Your suggestions were invaluable for the improvement of this manuscript, and we found all your questions to be extremely relevant. All changes introduced to respond to your questions are highlighted in yellow throughout the text. The corresponding page numbers are also reported to facilitate location in the manuscript.

- **Literature review.** This is an excellent literature review and mostly comprehensive. The only additional note that I would ask the authors to include is that recent studies have also demonstrated the effect that L2 parsing can have on the L1 (e.g., Dussias & Sagarra, 2007), modulated by such factors like immersion and proficiency—thus on the other end of the Shallow Structures Hypothesis is a view in which bilingual sentence processing is highly interactive and permeable between the L2 and the L1.

  Thank you for the comment. Yes, we are aware of this interesting work by Dussias and Sagarra, which was already referenced in our manuscript. We did not highlight this idea because in this paper we are mainly interested in the influence of L1 parsing over L2 syntactic processing, and not the other way around. Nevertheless, following your suggestion we include a brief note that helps highlight the interactive and highly dynamic view of syntax in bilinguals. Thank you for the suggestion (please see p. 8).

- **Cognate status.** At least in the example presented on p. 12, teacher/professor are considered non-cognates, but of course, professor is a cognate between EP and English. Were there other instances like this? It would be helpful to have materials included in Appendix. How might this impact the results?

  Thank you for pointing this out. Indeed, the Portuguese word ‘professor’ could be translated into English both as ‘professor’ and ‘teacher’. However, the English ‘professor’ is not the most immediate translation equivalent for Portuguese, because the English word ‘professor’ has a more specific meaning. While the Portuguese sense may refer to a teacher from any grade, the English word refers to university professors only, and thus, strictly speaking, there isn’t a direct correspondence between the two. Nevertheless, it is important to highlight that L2 learners only performed the task in English (i.e., they only saw the word ‘teacher’), so there was hardly a cognate interference effect in this case. Of note, there aren’t any other cases like this one in the experimental materials (see Appendix A now included as suggested, see p. 14 and p. 48-50). Therefore, for the reasons mentioned above, we consider that this case did not affect the results. Yet, in order to avoid that potential question from an attentive reader, we opted for changing the example provided in the manuscript (see p. 13).

- **Materials statistics.** This section on pp. 14-15 is hard to read. I’d suggest that instead of underlining and using the full names for the conditions (i.e., NC-C and NC-NC), that the authors rewrite this section and the stats results broken by N1 and N2. For example: “Thus, the pairwise comparisons revealed that for N1, orthographic similarity did not differ between cognates or noncognates (ps = 1) but did differ between cognate status (ps< .001). The same pattern of results
were found for N2, i.e., no orthographic difference between cognates or noncognates…

We agree that the statistics of the materials section is hard to follow. Thus, we have re-written it following your suggestion. Thank you (please see p. 16).

• Fillers. Please include additional description of fillers. How did these fragments differ from the critical trials?

Following your suggestion, information regarding the fillers is now included in the manuscript (please see pp. 16-17). As you will see, the fillers have different grammatical structures unrelated with the target structure and were intentionally built to be simple and unambiguous in order to make the task easier for the participants.

• Experiment 2. Please also include the Table 2 results as a bar plot (like Figure 1).

Following your suggestion the information in Table 2 is now presented as a bar plot as in Experiment 1 (see please see p. 22 and p. 47).

• Interpretation. The unexpected finding that Cognates actually result in a stronger L2-like processing pattern is mainly interpreted as perhaps due to greater cognitive demands. I am skeptical of this account as it would require assuming that LA is a (universal) default position for RC attachment; otherwise a greater cognitive resources account should arguably lead to a default, L1-like parsing strategy, i.e. HA. I believe that there is an alternative and more compelling account for the results, which is briefly mentioned on p. 28. It could be that cognates, as lexical items that induce greater activation between L1 and L2 lexicons, may also induce greater activation of both syntactic parsing systems, thus making it easier to select an L2-like parsing strategy. This easier to override L1 strategy should occur in the condition in which L1 competition would be greatest, i.e. in the NC-C condition because this is the condition in which the L2 strategy is most easily acquired, if indeed cognates serve as a facilitatory cue for L2 parsing. In contrast, under a cognitive demands/resources account, C-NC should have been the condition that resulted in the greatest LA strategy because it would arguably be the source of greatest competition and hence, most demanding. Regardless, either account remains plausible at this stage.

Thank you for the suggestion. In this version of the paper we tried to elaborate a little more on what was previously presented in p. 28 concerning Hartsuiker’s proposal and which we believe to be in line with your proposal. Nevertheless, it should be pointed out that, although the NC-C condition seems in fact to produce greater competition between L1 (HA) and L2 (LA) parsing strategies, as you stated (note that it was precisely in this condition that the differences between HA and LA responses did not reach statistical significance), it is also true that in this condition the number of HA responses (indicative of L1 syntax interference) was greater. Therefore, the NC-C condition seems to represent more difficulties for subjects to override the L1 strategy and not the reverse as you suggested. Moreover, we would like to stress the fact that, although the C-NC condition produced the second highest score of LA responses (the highest was the C-C condition), the results indicate that
the differences between the C-C and the C-NC conditions did not reach statistical significance, which is not consistent with your proposal. The lexical-syntactic interplay that these results demonstrate is complex and requires further investigation using other techniques, particularly online techniques, as you next suggest.

- **A future follow-up (not necessary for this study) would have been to test the same participants in their L1 (EP).** If the participants had maintained a fully L1-like strategy (HA), then possibly the cognitive resources account would be supported. However, the participants may have similarly demonstrated an L2 on L1 effect demonstrating LA for certain C conditions even in their L1 (e.g., Dussias & Sagarra, 2007). I’d like to see the discussion expand on this alternative account and mention potential follow-ups to adjudicate between the accounts.

  We agree. In fact, we are currently conducting a new study where we explore precisely this question. The discussion is now enriched with potential follow-ups from the present study (see pp. 31-34).

- **In terms of null proficiency effect, but modulated by cognate status, the authors interpret this finding as partial support for SSH (Clahsen & Felser).** I think that this also suggests an alternative: that acquisition and parsing is not an all-or-nothing variable but rather is more gradient. Rather than being a lexical driven process as C&F argue, it could be that L2-activation of syntactic preferences is lexically driven, i.e. a usage-based account. This is not what C&F argue in their framework.

  We agree that parsing is not an all-or-nothing variable and should be conceptualized in a more gradient manner. However, we must confess that we were slightly confused with your suggestion, since the argument presented to justify the usage-based account is the same used to justify the SSH account.