

REVIEWER A

The authors present a follow up study in which certain methodological flaws observed in a previous study have been solved by means of better stimuli selection and an improved procedure. Participants were presented with sentences referring to motor actions that could either be simultaneously compatible or incompatible. The sentences were introduced with the adverbs “while” and “after” which made them sensible or not. According to the authors view, the observed relative difficulty for sentences with incompatible actions preceded by “while” can be accommodated by the indexical hypothesis but is in conflict with the temporal iconicity assumption. Hence, the study adds to evidence on the embodied nature of time-related semantics.

The study is interesting and seems appropriate for the journal, but the design and stimuli selection are quite complicated and I think some bits of the manuscript need clarification. Apart from this, the main problem of the study, in my view, is related to a possible overinterpretation of the results. The authors argue that longer reaction times in the experimental condition for while-paragraphs can be explained by embodied constraints. On the other hand, longer reaction times obtained in the control condition for after-paragraphs are (too briefly) related to a violation of the canonical order of events in the world. What I see here is that an effect of semantic incongruence is observed in both the experimental and control materials. This incongruence can be related to embodied incompatibilities in the case of experimental paragraphs, but not (at least not only) in the control paragraphs. So, in my opinion, the results point out an effect of semantic violations on sentence reading times, but it cannot be fully stated whether these effects are related to embodiment or are just reflecting cognitive cost due to “extra elaboration demanded by incongruent materials”. I think this possibility should be given more attention in the manuscript.

Other Comments:

In Spanish, two words are needed to express “after” (“después de”) and only one to express “while” (“mientras”). Also the tense of the verb that follows each adverb is different. I think these differences deserve some comment, although, given the results of the first-clause reading times, the adverbs themselves did not seem to have any impact on sentence processing.

I do not fully understand the last paragraph of the introduction, specially the second interpretation of the predictions of the iconicity assumption (“*after*-paragraphs entail a temporal shift that is not necessary in the *while*-paragraphs”) and the use of control sentences (“By contrast, for the embodiment approach, the reading times will be larger for the *after*-paragraphs than for the *while*-paragraphs”). Examples or more detailed explanations would be welcome.

In the sentence “a third clause describing the purpose of the action.”, the purpose of which of the actions (the first, the second or both...) it refers to should be specified.

Table 2 with the examples is difficult to follow. The “experimental paragraphs, third clause sensible” section could be split into two parts for “while” and “after” respectively. More generally, I think the text and the table would benefit from cross-references to each specific type of stimuli, something like: sensible while-paragraphs (see Table 2.C)...

Does the first example in Table 2 “Mientras limpiaba (Después de limpiar) la herida...” mean that this sentence appeared twice, once with each adverb? Or, did it appear with different adverbs for different participants? Other experimental items were not duplicated?

Given that 34 sensible and 6 non-sensible fillers were constructed there are more sensible paragraphs than non-sensible ones. Was there any reason for not doing 20 sensible and 20 non-sensible fillers?

The selection of probe words and the creation of the two counterbalance conditions should also be explained in more detail.

Length and frequency of nouns and verbs of the second clause were matched between experimental and control paragraphs, but experimental and control paragraphs are not compared with each other (“Experimental paragraphs and control paragraphs were analyzed separately”) so this doesn’t seem to be very important. On the contrary, we don’t know whether while and after clauses are fully matched (see comment above regarding whether the same sentences were used for “while” and “after” conditions or not).

Minor comments:

I cannot find the keywords

The "fan effect" should be briefly explained.

In the analyses of the first-clause reading times, I think the use of \leq is incorrect.

In the analyses of sensibility judgments I have the impression that degrees of freedom in “ $t_1(28) = 21.87; p \leq .0001$ ” should be 38.

An “r” is missing in “Mientas limpiabas” in the abstract.

In page 4, “expresses” should be “express”

Table 2 appears before Table 1.

Table 3 is not referenced in the text.

In the tables 3 and 4, a legend explaining the * should also be included.

Javier Rodríguez-Ferreiro
Universidad de Barcelona

REVIEWER B

La recomendación de este revisor es que el manuscrito sea ACEPTADO para su publicación en *Psicológica*, una vez subsanados unos puntos menores.

El trabajo está bien fundamentado desde el punto de vista teórico y metodológico. La discusión de los resultados se presenta convenientemente asentada en la literatura científica. Por otra parte, el planteamiento del problema, así como la interpretación de los datos, sigue un hilo argumental claro y comprensible para cualquier lector de *Psicológica*, más allá del campo de la psicolingüística. Este revisor (al no ser nativo de inglés) no entra a valorar el estilo de la redacción.

Como aspectos subsanables, se destacan los siguientes:

1. En la introducción (p. 7, abajo) se declara: “Specifically, we selected manual actions forming part of the ordinary human repertoire, because they have wider range of variety and complexity and involve larger neural networks than actions involving the legs or mouth”, esta afirmación necesita un apoyo referencial, probablemente de Pulvermüller (2005).

2. En el procedimiento, debería mencionarse que el ordenador registraba los tiempos de reacción.

3. Es presumible que, por la naturaleza de la tarea y las variables dependientes, los participantes fueran instruidos en la conveniencia de responder de manera acertada y rápida; de ser así, indíquese de forma explícita en el procedimiento (“speed and accuracy were emphasized...” o similar).

4. En la Tabla 3 no acaban de encajar los porcentajes de *Sensibility* en la condición de Control.

Julio González
Universitat Jaume I, Castellón