Are object state changes represented during language comprehension? A non-replication and extension

Channing E. Hambric (c.hambric@bowdoin.edu)
Abhilasha A. Kumar (a.kumar@bowdoin.edu)
Department of Psychology, Bowdoin College
5600 College Station, Brunswick, ME 04011 USA

Kevin J. Holmes (kjholmes@reed.edu)
Department of Psychology, Reed College
3203 SE Woodstock Blvd, Portland, OR 97202 USA

Abstract

Previous work suggests that non-visual object properties like weight are automatically integrated into event models during language comprehension. Horchak and Garrido (2021) found that Portuguese speakers' response times were faster when the state of a presented object (e.g., a smashed tomato) matched the event implied by the preceding sentence (e.g., You drop a bowling ball on a tomato). In an exact replication in English (Experiment 1), we failed to replicate this weight-state match effect. In Experiment 2, we examined the potential role of sentence focus, manipulating whether the target item served as the subject or direct object in the sentence. Response times revealed a weight-state match effect, but only when the target object was the focus (i.e., subject) of the sentence. Overall, these findings suggest that the representation of object state changes during language comprehension may depend on the interaction of object properties and language-specific syntactic constraints.

Keywords: state change, mental representation, implicit object properties, situation model, language comprehension

Introduction

Imagine that you read a sentence about a child *dropping* their ice cream cone on the sidewalk. What kind of image comes to mind? It would likely involve the ice-cream-side down and the cone pointing up in the air. However, if the sentence had described the child *holding* an ice cream cone, a different image would likely come to mind, with the ice cream cone upright and intact. During language comprehension, readers and listeners use sentential context to form the appropriate mental representations of objects (Glenberg, 1997; Glenberg et al., 1999; Hoeben Mannaert et al., 2019). They also use preexisting conceptual knowledge to inform their interpretation of the described event. In the ice cream example, suppose the sentence were extended to *the child dropped their ice cream cone on the sidewalk on a hot summer day*. Based on sentential context alone, one's representation of the ice cream's state might not differ from before. However, given the knowledge that ice cream is prone to melting in the heat, the ice cream would be more accurately represented as both *upside down* and *melted*. People thus use a variety of cues to construct "situation models" (Zwaan & Radvansky, 1998) or "mental models" (Johnson-Laird, 1983) during language comprehension. Yet it remains unclear what types of information are automatically represented in these models, how this information is updated as events unfold, and what syntactic properties may moderate these factors.

The current body of evidence suggests that surface-level object characteristics are activated and integrated into mental models of events. The sentence-picture verification task is typically used in studies investigating these phenomena, where participants are presented with a sentence to read, followed by an image of an object. The task is to determine whether the pictured object was mentioned in the preceding sentence, and the response time to make this judgment is recorded. Early work by Stanfield and Zwaan (2001) found that after reading sentences like *The carpenter pounded the nail into the wall* and *The carpenter pounded the nail into the floor*, participants were faster to verify the object when it was shown in the orientation implied by the sentence (e.g., a horizontal vs. vertical nail). Kang et al. (2020) adapted this paradigm to manipulate the *state* of the object (rather than its orientation) by asking participants to read sentences that described either substantial change to an object's state (e.g., *The woman dropped the trash can*) or minimal to no change (e.g., *The woman placed the trash can*). Participants were faster to respond to the object in a modified state (i.e., a tipped-over trash can) when it matched, rather than mismatched, the preceding sentential context. Together, these findings suggest that readers construct representations of objects based on the syntactic, lexical, and deep structure of raw linguistic input.

There is also evidence that readers automatically incorporate *implicit* object properties into mental models during comprehension. In a recent series of experiments, Horchak and Garrido (2021) investigated whether one such property, object weight, is integrated into event representations and whether both modified and original object states remain active in working mental models (Bransford et al., 1972; Briner et al., 2014; Sachs, 1967). The latter question was inspired by Altmann and Ekves (2019), who posit that event representation entails encoding "ensembles of intersecting object histories," such that readers maintain access to the varying states of an object implied by the sentence narrative (Hindy et al., 2012; Solomon et al., 2015). In Horchak and Garrido's study, participants read sentences in which differently-weighted objects (e.g., a heavy *bowling ball* or a light *balloon*) were dropped onto another malleable object (e.g., *tomato*). Critically, the weight of the dropped objects was not explicitly expressed in the text, but could only be implicitly inferred based on pre-existing knowledge. After reading each sentence, participants were presented with a picture of either an intact (canonical) version of the malleable object or a smashed (non-canonical) version of the same object. Participants were faster to verify non-canonical objects after reading a heavy sentence than a light sentence, suggesting that they had mentally transformed the object into a smashed state upon integrating the dropped object's weight into their mental model. By contrast, no such effect was observed for canonical objects.

Horchak and Garrido (2021) offered two possible interpretations of their findings: either only the end-state of the object is encoded in the mental model (Zwaan & Pecher, 2012), or both the original state and its altered state actively compete (Altmann & Ekves, 2019; Hindy et al., 2012; Solomon et al., 2015). Horchak and Garrido (2021) suggest that the latter interpretation was supported by the finding that canonical objects were verified no faster after a light sentence than a heavy sentence, suggesting that both states remained accessible. The authors replicated the match effect for non-canonical objects, and the lack thereof for canonical objects, across several manipulations, including experiments that used different acting objects (e.g., *brick* vs. *sponge*), different tenses (*You drop* vs. *You dropped*), and different phrasal structures (*You dropped a bowling ball on a tomato* vs. *A brick fell on a tomato*).

This sentence-picture match effect is reminiscent of the widely studied action-sentence compatibility effect (ACE), in which response times are facilitated when the action implied by a sentence (e.g., pulling) matches the action used by participants to make their response. Evidence for the ACE has been taken as support for embodied theories of language comprehension (Glenberg & Kaschak, 2002; but see Morey et al., 2022; Papesh, 2015, for failures to replicate this effect). For example, Glenberg et al. (1999) posited that perceptual representations – not abstract semantic nodes or schemas – constitute the background knowledge used to generate event models. Although this view has been highly influential, we still know little about how *physical objects*, as opposed to motor actions, are represented in mental models of events. Moreover, although previous research has examined cross-language differences in the representation of object shape during comprehension (Sato et al., 2013), to date there has been no work investigating how language-specific factors may modulate the representation of *implicit* object properties. Therefore, it is critical to replicate existing sentence-picture match effects across populations, particularly those that use preexisting semantic knowledge to enact object state change.

To examine this issue, we conducted two experiments probing the Horchak and Garrido (2021) finding that object weight is automatically integrated into mental models of described state change events. Experiment 1 was an exact replication, with the exception that the original study was conducted in Portuguese and ours was conducted in English. To preview the results, we failed to replicate the sentence-picture match effect observed for non-canonical objects after reading heavy sentences. In Experiment 2, we investigated a potential explanation for this result: cross-language differences in syntactic structure may influence sentential focus and subsequent mental representations of object state changes. To test this possibility, we manipulated the sentence structure to place the malleable target object as either the focal subject or the direct object of the sentence. We found the same match effect originally observed by Horchak and Garrido (2021), but only when the target object was the focus of the sentence.

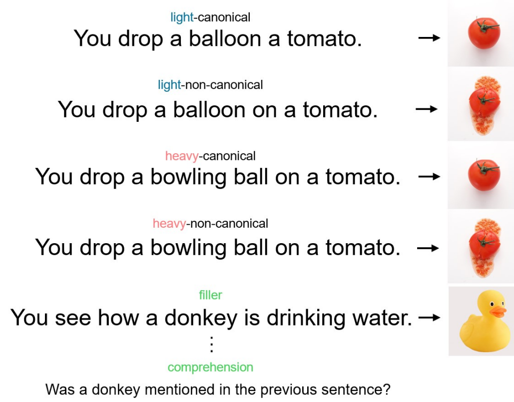

Figure 1: Sample stimuli from English replication of Horchak and Garrido (2021), Experiment 1.

Experiment 1: English Replication

Methods

We replicated the methods from Experiment 1 in Horchak and Garrido (2021). The replication was preregistered (anonymized link: `https://aspredicted.org/yr4sm.pdf`). Unless otherwise noted, our procedure was the same as in the original work.

Participants We relied on the power analysis by Horchak and Garrido (2021) to set the minimal target sample size of 77 participants. To account for exclusions, we recruited 99 participants from Prolific. Recruitment was restricted to native English speakers aged 18 and older. There were no exclusions based on comprehension performance, resulting in a final sample of 99 participants for the accuracy analysis. After excluding six participants for low task accuracy ($< 80\%$, as in Horchak & Garrido, 2021) the final sample size for the response times (RT) analysis was 93 participants.

Stimuli, Materials, and Design The experiment was programmed in *jsPsych* (De Leeuw, 2015) and conducted via *Cognition* (https://www.cognition.run/). We reused all stimuli from Horchak and Garrido's (2021) Experiment 1, with their English-translated sentence stimuli in place of the original Portuguese stimuli. The sentence–picture verification task contained 72 trials, with 24 experimental trials and 48 fillers. Of the experimental trials, 12 sentences were *heavy* (*You drop a bowling ball...*), and the other 12 were *light* (*You drop a balloon...*). Each sentence had a different *target object* that was easily squashable (e.g., *tomato*). The stimuli also included filler sentences that described seeing an event (e.g., *You see how a donkey is drinking water*) or dropping a different object (e.g., *You drop a cushion on a magazine*). All participants saw the same filler trials in randomized order.

In experimental trials, the target objects in the images were either intact (canonical) or smashed (non-canonical). Thus, the sentence–picture pairs fit into one of four *sentence type* × *picture state* conditions: heavy–canonical, heavy–non-canonical, light–canonical, or light–non-canonical (see Figure 1), with six trials in each condition. Four randomly assigned lists were used to counterbalance the condition in which each object was presented across participants.

Finally, 24 comprehension questions were randomly dispersed throughout the main task to ensure participants fully read and understood the sentences. Based on feedback from a pilot experiment, we modified comprehension trials to have a black background with blue text and the questions were slightly altered to make them easier to understand in English.

Procedure Participants completed the experiment online. The task began with six practice trials, followed by a 72-trial main task with randomly dispersed comprehension questions. In each trial of the sentence-picture verification task, participants were presented with a sentence followed by the picture of an object and asked to indicate if the object was mentioned in the preceding sentence. A fixation cross appeared for 1000 ms before the sentence stimulus. Participants pressed the space bar after fully understanding the sentence, prompting a fixation cross to appear for 500 ms followed by the target image. Upon seeing the image, participants responded with the 'S' key for yes and the 'N' key for no, which mapped onto the onsets of yes and no in the original Portuguese study. Response accuracy and response time from picture onset were recorded.

Results

Data Treatment We conducted all data treatment and analyses in R Studio version 4.2.0 (R Core Team, 2024). Before analyses, filler and comprehension trials were removed from the data. There were no exclusions due to poor comprehension performance. For RT analyses, we removed incorrect ('no') responses and excluded participants that responded correctly on less than 80% of experimental trials (N=6). Following Horchak and Garrido (2021), a log-10 transformation was used to normalize RT data and RTs that were more than three median absolute deviations away from its condition median were removed.

Accuracy Analysis Overall task accuracy was 95.24%. We fit a logit mixed-effects model using sentence type, picture type, and the interaction as fixed effects; random intercepts for participants and items; and by-participant random slopes for sentence type, picture type, and the interaction term. As shown in Figure 2 (left panel), there were no main effects,

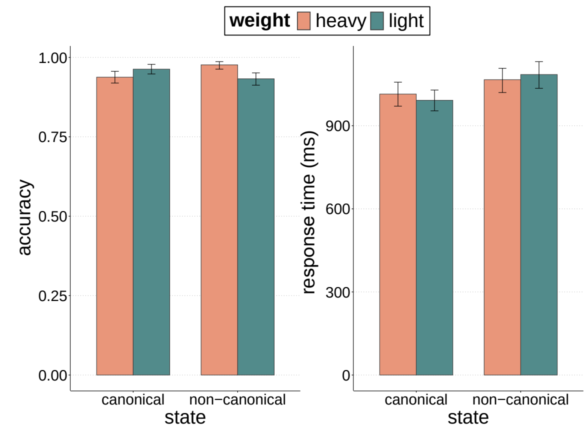

Figure 2: Mean accuracy and response times in Experiment 1. Unlike Experiment 1 in Horchak and Garrido (2021), the *weight* × *state* interaction was not significant in the response time data. Error bars represent bootstrapped 95% confidence intervals.

and the interaction between sentence and picture type was not significant (b = 0.26, z = 0.19 p = .849). This finding is consistent with Horchak and Garrido (2021), who did not observe any effects in accuracy measures.

Reaction Time Analysis We fit a linear mixed-effects model to the transformed RT data with the same fixed and random effects structure as the accuracy model. Again, as shown in Figure 2 (right panel), there were no main effects and critically, the interaction between sentence and picture type was not significant ($b < .01$, t = -0.38, p = .704). We thus failed to replicate the sentence-picture match effect found by Horchak and Garrido (2021).

Discussion

The present study sought to replicate Horchak and Garrido (2021), Experiment 1. Although we followed their exact procedure and used the same (translated) stimuli, we failed to replicate their finding that response times were facilitated for non-canonical objects after reading sentences that used implicit object properties to imply state change. This was unexpected given that Horchak and Garrido (2021) replicated the finding over a number of variations. Regardless, it appears that participants in our experiment either (1) did not activate the implicit weight-related properties of the dropped objects or (2) did not use these properties to form altered representations of the target objects. Why might readers not be attending to this information? We turned to examine the major factor that distinguished the two studies: language.

The original study was conducted in Portuguese. While we used the translated equivalents of the sentence stimuli provided by the authors, the syntactic structure varied across the two versions of the stimuli. Specifically, (European)

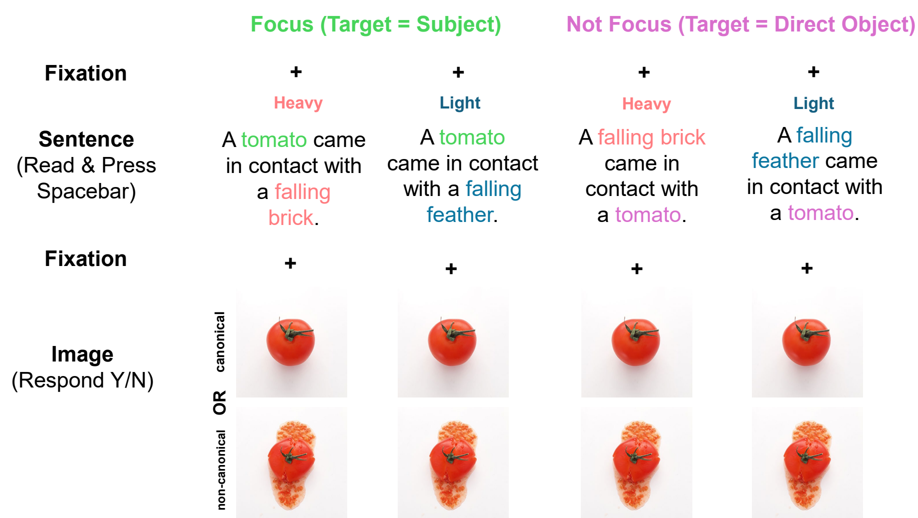

Figure 3: Target stimuli conditions, examples, and participant interactions in Experiment 2. Sentences manipulated the weight of the falling object (heavy vs. light) and the focus of the sentence (target either subject or direct object). The object in the image stimulus was either in its canonical (i.e., intact) or non-canonical (i.e., smashed) state.

Portuguese is considered a *null subject* language, whereas English requires a pronominal subject to be expressed in all finite clauses (Cognola et al., 2018; Soares et al., 2019). Consider the two sentences below:

Deixas *cair uma bola de bowling num tomate.*
You *drop a bowling ball on a tomato.*

In the Portuguese version, the sentence begins with the conjugated form of the verb *deixar* ("to let [fall]"), whereas the English translation begins with the explicit subject pronoun *you*. That is, in Portuguese it is not necessary to include the pronoun *you* (as it is implied by the conjugation), but this is grammatically obligatory in English.

Although these structures may not differ in their underlying meaning, their surface differences may have attentional consequences during reading. Specifically, we posit that presenting an explicit subject prior to the verb may influence what readers consider to be the focal noun phrase of the sentence, thereby modulating the relative salience of simulated objects during comprehension. Consistent with this possibility, linguists have proposed that speakers of languages where both null and overt subject structures are acceptable (which includes Brazilian Portuguese, Italian, and Spanish) use overt subjects to disambiguate antecedents and to shift attention to a specific referent (Frascarelli, 2007; Serratrice, 2008; Soares et al., 2019). Moreover, experimental and discourse analyses in such languages have indicated that readers tend to regard overt pre-verbal subjects as the focal topic of the sentence (Alonso-Ovalle et al., 2002; Santos, 2005). Thus, the *exclusion* of the pronominal subject in the original Portuguese experiments may have allocated more focus to the acting and target objects of the sentence, rendering them more prominent in readers' mental models. In our English replication, readers may have instead placed preferential focus on the beginning noun phrase *You dropped* rather than the subsequent modifying phrase *a bowling ball on a tomato.* To assess the possible role of sentential focus more directly, we conducted a follow-up experiment in which we manipulated whether the target object was presented as the pre-verbal subject of the sentence or as the post-verbal direct object.

Experiment 2: Sentence Focus Manipulation

The experiment was preregistered, and all data and analysis scripts can be found on the Open Science Framework (https://osf.io/8v6yz/). Unless otherwise noted, the methodology was the same as the previous experiment.

Methods

Participants We conducted a simulation-based power analysis using the *simr* package (Green & MacLeod, 2016). The simulations indicated that a sample size of approximately 150 participants was needed to achieve 80% power. To account for future exclusions, we recruited 239 participants from Prolific and Bowdoin College. Ten non-native English speakers were excluded from analysis. There were no exclusions for inadequate comprehension. The final sample for the accuracy analyses was 229 participants. After excluding 43 participants for low accuracy on experimental trials, the final sample size for RT analyses was 186 participants.

Stimuli, Materials, and Design The design largely mirrored the previous, with the exception that we added sentence focus as a factor. To account for the added factor and ensure adequate statistical power, we increased the number of target stimuli from 24 to 40. Additional target stimuli were taken

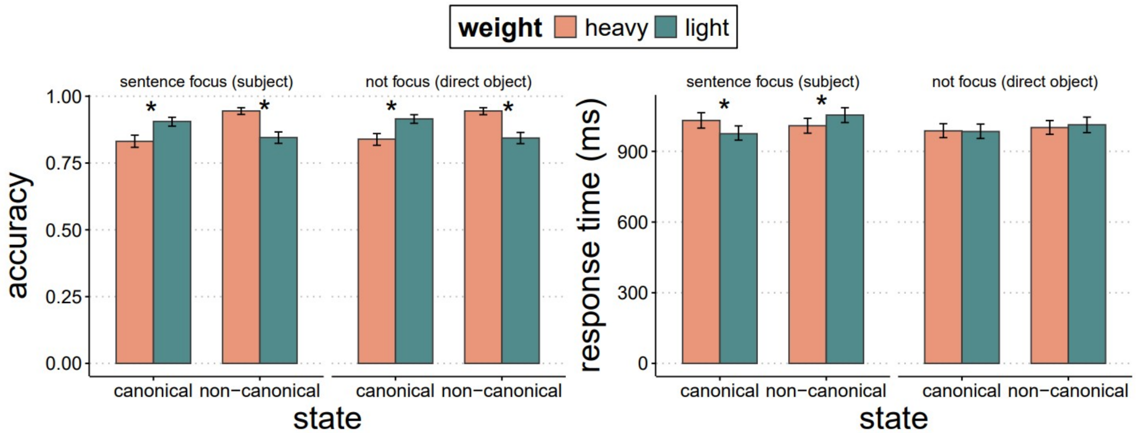

Figure 4: Mean accuracy and response times (RTs) in Experiment 2. Accuracy was higher for non-canonical objects following heavy sentences and for canonical objects following light sentences, regardless of sentence focus. RTs to non-canonical objects were faster following heavy sentences, but *only* when the target object was the focus (subject) of the sentence. Contrary to Horchak and Garrido (2021), RTs to canonical objects were faster after light sentences, but again, only when targets were the focus of the sentence. Error bars represent bootstrapped 95% confidence intervals and asterisks indicate a significant effect.

from other experiments (2-7) in Horchak and Garrido (2021), and were all transformable objects (e.g., *blueberry, teapot, bottle*). There were 88 total trials, with 40 experimental trials and 48 filler. Filler trials were the same as in Experiment 1. Twenty sentences were *heavy* sentences and the other 20 were *light* sentences. Within each weight condition, 10 sentences were presented with objects in their *canonical* state and 10 in their *non-canonical* state. Critically, of the trials within each *sentence* × *picture* condition, 5 sentences placed the target object as the focus (subject) of the sentence and 5 sentences placed it as the direct object at the end of the sentence. Eight randomly assigned lists were used to counterbalance the condition in which each object was presented across participants. Sample stimuli can be seen in Figure 3.

Three additional changes were made to the format of the sentences. First, we used *falling brick* and *falling feather* (instead of *bowling ball* and *balloon*) from Experiment 6 in Horchak and Garrido (2021), as they were more commonly used items with more intuitive weights. Second, we used the sentence structure (*A brick fell on a tomato*) from Experiment 6 to instantiate the sentence focus manipulation. Within this structure, we substituted a passive verb phrase that could be used with either focus condition (*A... came in contact with a...*). We note that this introduced a change in verb in addition to the change in sentence focus; however, it was necessary to use a non-directional verb in order to accommodate both sentence weight and focus manipulations. Filler trials were adapted to follow a similar format. Finally, to minimize encoding of target objects outside of the experimental trial context, we replaced any target items used as lures in filler trials with unrelated non-target objects.

Procedure The procedure was identical to Experiment 1 with the exception that participants used the 'Y' key for 'yes' responses and 'N' for 'no' responses during the task.

Results

Data Treatment Data were processed as in Experiment 1.

Accuracy Analysis Mean accuracy was 88.38%. We fit a logit mixed-effects model using sentence type, picture type, sentence focus, and their interactions as fixed effects; random intercepts for participants and items; and by-participant random slopes for sentence type, picture type, and focus. As shown in Figure 4 (left panel), there was a significant two-way interaction between sentence and picture type (b = -3.61, z = 11.15, $p < .001$), such that responses were more accurate when canonical objects were preceded by light sentences and when non-canonical objects were preceded by heavy sentences. The three-way interaction was not significant (b = -0.11, z = -0.29, p = .77). Thus, in contrast with both Horchak and Garrido (2021) and Experiment 1, we observed the standard sentence-picture match effect in accuracy measures.

Reaction Time Analysis We fit a linear mixed-effects model to the transformed RT data, with the same fixed and random effects structure described for the accuracy model. As shown in Figure 4 (right panel), there was a significant two-way interaction between sentence and picture type (b = .04, t = 4.17, $p < .001$). Critically, this was qualified by a significant three-way interaction. Follow-up contrasts indicated that, as in Horchak and Garrido (2021), non-canonical objects were verified faster following heavy than light sentences (b = 0.02, t = 2.50, p = .01), but *only* when the target object was the focus of the sentence. In contrast to Horchak and Garrido (2021), canonical objects were also verified faster following light compared to heavy sentences (b = 0.02, t = 3.39, $p <.001$), but again only when the target object was the focus of the sentence. This key interaction supports our interpretation of the results of Experiment 1. The English translations of the original Portuguese sentence stimuli may have directed readers' attention toward the pronominal subject (*You*) and away from the acting and target objects, reducing their likelihood of incorporating the object state change into their mental model of the event.

General Discussion

Accumulating evidence suggests that mental models are used to simulate described events during language comprehension (Glenberg, 1997; Glenberg et al., 1999; Horchak & Garrido, 2024; Zwaan & Pecher, 2012). This entails activating both surface-level and implicit properties of objects and tracking how interactions between objects may impact the state of a given object (Altmann & Ekves, 2019). Horchak and Garrido (2021) found that after reading sentences in which a heavy (versus light) object was dropped onto a malleable target object, participants were faster to verify that a smashed, non-canonical depiction of the target object had been mentioned in the preceding sentence. The present work sought to replicate this finding. Experiment 1 was an exact English replication of the first study of Horchak and Garrido (2021), but we failed to observe the same weight-state match effect. We posited that the differing syntactic structures of Portuguese (used in the original study) and English may have led readers to place focus on different noun phrases in the sentence. Specifically, the null subject form of Portuguese may have directed relatively more attention to the acting and target objects, thus accentuating their representation in the mental model. We examined this possibility in a follow-up study that manipulated whether the target object was the focus (subject) of the sentence or appeared in sentence-final position (direct object). Responses to non-canonical target objects were facilitated when preceded by a sentence implying state change (i.e., heavy sentences), but *only* when this object was the focus of the sentence.

This pattern of results suggests that language-specific syntax may modulate how entities are represented in mental models of events. Whereas the sentence structure used in the original Portuguese work did not include an explicit pre-verbal subject, the English translations required one (*you*). As a result, Horchak and Garrido's (2021) Portuguese readers may have focused more on the described objects and incorporated the implicit properties of these objects into their mental model of the described event, whereas English readers in our first experiment may have focused more on simulating the first noun phrase of the sentence (i.e., themselves dropping something). In Experiment 2, however, the sentence structure omitted the pre-verbal pronoun, thus directing readers' attention more to the state-change event itself. When the first noun phrase of the sentence highlighted the object that underwent the state change, participants shifted their mental model to simulate the impact of the acting object. Additional work could help substantiate our interpretation of the findings. For example, one could assess readers' conscious impressions of the relative prominence accorded by the sentence stimuli to the acting and target objects, the state-change event, and themselves. It would also be useful to replicate Experiment 2 in Portuguese, especially given the change in verb in Experiment 2.

Two other findings from Experiment 2 also deviated from the original Portuguese results. First, we observed a weight-state match effect for accuracy, regardless of sentential focus. One potential explanation is that the sentence structures, though similar, nevertheless differed in the specific verb used to generate expectancies of state change. Specifically, we modified the original verb phrase (*dropped*) to accommodate either the acting or target object as the subject of the sentence (*came in contact with*). Although this introduced another variation in sentence structure, the lack of a match effect when the target object was not the focus of the sentence suggests that the present findings cannot be fully attributed to the change in verb. Rather, *came in contact with* may generate weaker expectancies of state change than *dropped*, and this weaker effect may only manifest in reaction times when focus is explicitly placed on the target object. That said, the average accuracy was significantly lower in Experiment 2 (88.38% compared to 95.24% in Experiment 1), likely due to the increased number of experimental trials. It is possible that there was simply more room for accuracy effects to manifest. More work is needed to disentangle these possibilities. Second, we found that responses to canonical target objects were facilitated when preceded by a sentence implying no state change (i.e., light sentences; see Figure 4). Horchak and Garrido (2021) argued that such an effect would support the "constant scenario" theory of object representation, according to which readers form a single context-dependent representation of a given object (Zwaan & Pecher, 2012). They did not observe this effect for canonical objects, leading them to conclude that both original and modified object states actively compete in mental models. The present work casts doubt on this conclusion, and may require more careful investigation of which mental representations are activated during the comprehension of object state changes during reading.

Broadly, our findings suggest that the way individuals represent objects depends on language-specific syntax and which objects are attended to during comprehension. Moreover, the significant match effect observed in the focus condition deviates from multiple failures to replicate the ACE (Morey et al., 2022; Papesh, 2015). More work is needed to clarify if the way object state changes are simulated in mental models of events fundamentally differs from the simulations studied via the ACE or if syntax also plays a key role in the embodiment of described motor movements in language comprehension.

References

Alonso-Ovalle, L., Fernández-Solera, S., Frazier, L., & Clifton, C. (2002). Null vs. overt pronouns and the topic-focus articulation in spanish. *Italian Journal of Linguistics*, *14*, 151–170.

Altmann, G., & Ekves, Z. (2019). Events as intersecting object histories: A new theory of event representation. *Psychological Review*, *126*(6), 817.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology*, *3*(2), 193–209.

Briner, S. W., Virtue, S. M., & Schutzenhofer, M. C. (2014). Hemispheric processing of mental representations during text comprehension: Evidence for inhibition of inconsistent shape information. *Neuropsychologia*, *61*, 96–104.

Cognola, F., Casalicchio, J., et al. (2018). On the null-subject phenomenon. *Null subjects in Generative Grammar. A Synchronic and Diachronic Perspective*, 1–28.

De Leeuw, J. R. (2015). jspsych: A javascript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, *47*, 1–12.

Frascarelli, M. (2007). Subjects, topics and the interpretation of referential pro: An interface approach to the linking of (null) pronouns. *Natural Language & Linguistic Theory*, *25*, 691–734.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, *20*(1), 1–19.

Glenberg, A. M., Robertson, D. A., Jansen, J. L., & Johnson-Glenberg, M. C. (1999). Not propositions. *Cognitive Systems Research*, *1*(1), 19–33.

Green, P., & MacLeod, C. J. (2016). Simr: An r package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498.

Hindy, N. C., Altmann, G. T., Kalenik, E., & Thompson-Schill, S. L. (2012). The effect of object state-changes on event processing: do objects compete with themselves? *Journal of Neuroscience*, *32*(17), 5795–5803.

Hoeben Mannaert, L. N., Dijkstra, K., & Zwaan, R. A. (2019). How are mental simulations updated across sentences? *Memory Cognition*, *47*(6), 1201–1214.

Horchak, O. V., & Garrido, M. V. (2021). Dropping bowling balls on tomatoes: Representations of object state-changes during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(5), 838–857.

Horchak, O. V., & Garrido, M. V. (2024). Language comprehenders are sensitive to multiple states of semantically similar objects. *Journal of Memory and Language*, *135*, 104478.

Johnson-Laird, P. N. (1983). *Mental models. towards a cognitive science of language, inference and consciousness*. Cambridge, UK: Cambridge University Press.

Kang, X., Eerland, A., Joergensen, G. H., Zwaan, R. A., & Altmann, G. T. M. (2020). The influence of state change on object representations in language comprehension. *Memory Cognition*, *48*(3), 390–399.

Morey, R. D., Kaschak, M. P., Díez-Álamo, M., et al. (2022). A pre-registered, multi-lab non-replication of the action-sentence compatibility effect (ace). *Psychonomic Bulletin Review*, *29*(2), 613–626.

Papesh, M. H. (2015). Just out of reach: On the reliability of the action-sentence compatibility effect. *Journal of Experimental Psychology: General*, *144*(6), e116–e141.

R Core Team. (2024). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics*, *2*(9), 437–442.

Santos, A. L. (2005). Getting in focus: The role of the nsr in children's interpretation of sentences with focused preverbal material. In *Language acquisition and development: Proceedings of gala* (pp. 487–98).

Sato, M., Schafer, A. J., & Bergen, B. K. (2013). One word at a time: mental representations of object shape change incrementally during sentence processing. *Language and Cognition*, *5*(4), 345–373.

Serratrice, L. (2008). Null and overt subjects at the syntax-discourse interface: evidence from monolingual and bilingual acquisition. *LOT Occasional Series*, *8*, 181–200.

Soares, E. C., Miller, P. H., & Hemforth, B. (2019). The effect of verbal agreement marking on the use of null and overt subjects: a quantitative study of first person singular in brazilian portuguese. *Fórum Linguístico*, *16*(1), 3579–3600.

Solomon, S. H., Hindy, N. C., Altmann, G. T., & Thompson-Schill, S. L. (2015). Competition between mutually exclusive object states in event comprehension. *Journal of cognitive neuroscience*, *27*(12), 2324–2338.

Zwaan, R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension: Six replication attempts. *PloS one*, *7*(12), e51382.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162–185.