UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

What does action do to object? The case of metaphoric action.

Permalink

https://escholarship.org/uc/item/5rq492bs

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 47(0)

Authors

Cora, Cagatay
Desai, Rutvik H.
Lai, Vicky Tzuyin

Publication Date

2025

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

What does action do to object? The case of metaphoric action.

Cagatay Cora (cagataycora@arizona.edu)
Department of Psychology and Cognitive Science Program, University of Arizona,
1503 E University Blvd, Tucson, AZ 85721, USA

Rutvik Desai (RUTVIK@sc.edu)
Department of Psychology, University of South Carolina,
1512 Pendelton Street, Columbia, SC 29208, USA

Vicky Tzuyin Lai (tzuyinlai@arizona.edu)
Department of Psychology and Cognitive Science Program, University of Arizona,
1503 E University Blvd, Tucson, AZ 85721, USA

Abstract

We examined whether embodiment effects at a particular word influence subsequent words. We recorded EEG while participants read action sentences that were literal-concrete (LC), literal-abstract (LA), and metaphorical (MET). Prior work showed that at the verbs, both LC and MET elicited more negative N400s than LA, reflecting sensorimotor simulations. We found that at the object nouns, LC elicited a more negative negativity in the 500-700 ms time window than LA and MET, which may reflect combined influences from embodiment spilled over from the verbs and imagery linked to noun concreteness and imageability. These findings suggested that literal action embodiment can yield extended motor engagement extending into subsequent words, but metaphorical action verbs cannot. Moreover, lexical properties at the noun played a role.

Keywords: Embodiment; Metaphor; N400; N700; ERP

Introduction

Metaphorical language leverages semantic flexibility to construct meanings that extend interpretations (Semino et al., 2008; Rutter et al., 2012; Benedek et al., 2014). For instance, in *"*The secret swept across the campus*"*, the inanimate subject noun "secret" paired with an action verb evokes the notion of something spreading quickly, thereby transforming "sweep" into a figurative domain (Gentner & France, 1988). This illustrates the capacity of metaphor to impart rich, nuanced meanings (Cameron & Gibbs, 2008).

Some *embodied cognition* researchers argued that metaphors engage sensorimotor simulations akin to those triggered by literal action descriptions (Barsalou et al., 2003; Gibbs, 2006; Qu et al., 2013; Yin et al., 2013). Specifically, processing a sentence like "The church bent the rules" may involve mentally simulating the act of bending, with this spatial concrete concept supporting an abstract sense of "altering". Such metaphorically embodied simulation is supported by findings showing faster comprehension when participants perform or imagine congruent actions before reading metaphorical expressions (Wilson & Gibbs, 2007), as well as from imaging and electrophysiological studies revealing activation in sensorimotor-related brain regions (Bardolph & Coulson, 2014; Boulenger et al., 2012; Desai et al. 2011; Lai et al., 2019; Lauro et al., 2013; Li et al., 2022). However, other researchers suggested that verbal metaphor comprehension primarily engages abstract semantic systems rather than concrete sensorimotor features (Chatterjee, 2010; Forgács, 2022; Raposo et al., 2009), particularly if the figurative meaning is familiar and conventionalized.

Event-related potential (ERP) studies have probed the time course of metaphor comprehension and the interplay between concrete and abstract meanings. The ERP N400 component is a negative going waveform between 200 ms and 600 ms post stimulus onset and parietally distributed, is widely regarded as reflecting semantic processing (Kutas & Hillyard, 1980). In the same time window, there is also a frontal negativity that is modulated by concreteness (Holcomb et al., 1999). In addition, a late negativity at around 700 ms and frontally distributed has been linked to imagery-driven semantic elaboration (Barber et al., 2013; Gullick et al., 2013; West & Holcomb, 2000). Lai et al. (2019) used these ERP components to examine embodiment in metaphoric and literal sentences with action verbs (Table 1). Metaphoric (MET) verbs patterned with literal-concrete (LC) verbs, both elicited negativities in the 200-500 ms time window, more negative than the literal-abstract (LA) verbs, supporting an embodied view of metaphor at the verbs. The MET-LA difference was broadly distributed on the scalp, which differed from the LC-LA difference, which was more frontal, and also differed from the MET-LC difference which was parietal. This suggests that the topography could help differentiate literal embodiment vs. metaphoric embodiment.

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society* ©2025 the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

In contrast, Li et al. (2022) found that verbs in subject-verb metaphors (e.g., "The company grasped the opportunity") elicited a more positive P600/LPC than literal-concrete, suggesting a reanalysis to resolve semantic conflict and activate abstract meaning. But verbs in the verb-object metaphors (e.g., "The boss grasped the opportunity") elicited N400s similar to the literal-concrete condition, in comparison to the literal-abstract, indicating literal sensorimotor simulations. At the object position ("opportunity"), Li et al. (2022) reported that subject-verb metaphors elicited a more positive P600/LPC than the literal-abstract, which according to the authors reflects the cognitive effort required for integrating metaphorical meaning into sentence context. Directly relevant to the current study, Li et al. also reported N400 effects at the object noun position, showing greater negativity for literal-concrete objects compared to both literal-abstract and metaphorical ones, regardless of structure (subject-verb or verb-object).

Table 1: Example stimuli.

Condition	Example
Literal-abstract	*The church altered the **rules**.*
Literal-concrete	*The bodyguard bent the **rod.***
Metaphor	*The church bent the **rules.***

The current study asks whether the object nouns following the literal and metaphoric verbs also show embodied simulations. If embodiment effect at the verbs affects the following words, we predict that the object nouns in the literal-concrete and metaphorical conditions will elicit more negative N400 than the object nouns in the literal-abstract condition. Given that the N700 has been associated with mental imagery (Gullick et al., 2013), we also expect a more negative late negativity for the literal-concrete and metaphor than for the literal-abstract conditions. Finally, if metaphorical context affects object nouns in terms of more integration or reanalysis (Li et al., 2022), not embodiment, then object nouns in metaphoric condition will elicit a more positive LPC/P600 than the literal-abstract. If the action embodiment effect at the verb is short-lived and transient, we would observe no effect at the object noun.

Methods and Materials

Participants

Data are from Lai et al. (2019). Out of the original 34 participants, six were excluded due to low trial counts (< 50%) post-artifact rejection and EEG system errors. The remaining 26 participants (14 women) had a mean age of 19.25 years (range 18–22). Informed consent was obtained prior to participation, in line with the local Institutional Review Board.

Materials

The stimuli were 88 triplets of MET, LA, LC sentences (Table 1). Each sentence was in an NP-Verb-NP (NP means noun phrase) structure. The MET sentences featured agents with action verbs that couldn't be physically performed by the agents. In LA, the agents were paired with abstract verbs. In LC, agents performed concrete actions. Note that all sentences were semantically coherent and syntactically well-formed, with no anomalous items included.

Different from Lai et al. (2019), which focused on verb processing, the present study examined ERPs at the object nouns. To ensure comparability across conditions, we normed the target nouns on lexical frequency and length (SUBTLEXus; Brysbaert & New, 2009), cloze probability (using both human ratings and GPT-3.5-turbo estimates), concreteness (Brysbaert, Warriner, & Kuperman, 2013), and imageability (Wilson, 1988). No significant differences were found across conditions in word frequency, length, or predictability. Human cloze responses were obtained from 60 native English speakers across Latin-square rotated lists, and log probabilities were also estimated using 10 iterations per item with GPT-3.5-turbo. However, significant differences were found in concreteness and imageability: LC object nouns were rated significantly higher than those in both the LA and MET conditions (ps < .001), which did not differ from each other (see Table 2).

Table 2: Descriptive Statistics

Measure	LA (Mean ± SD)	LC (Mean ± SD)	MET (Mean ± SD)	p-value
Log Frequency	3.23 ± 0.51	3.02 ± 0.48	3.22 ± 0.49	= .07
Length	6.34 ± 1.20	5.84 ± 1.10	6.31 ± 1.30	=.33
Cloze (Human)	4.87 ± 0.82	4.88 ± 0.79	4.71 ± 0.85	= .53
Cloze (LLM)	-.00002	-.00001	-.00001	>.05
Concreteness	3.00 ± 0.80	4.60 ± 0.70	3.00 ± 0.90	< .001
Imageability	449.27 ± 83.93	567.42 ± 55.85	449.27 ± 83.93	<.001 (LC> MET)

Procedure

Participants completed the consent form and were set up for scalp EEG. The experiment took place in a quiet, dimly lit room, with participants seated 70–80 cm from a computer screen.They were instructed to attentively read each sentence for comprehension, as occasional yes/no comprehension questions would follow. Sentences were presented word-by-word in white Courier New font on a black background, with word durations dependent on length (mean = 851 ms). Comprehension questions appeared randomly after 25% of the sentences (i.e., catch trials), and participants responded using a button box. A "blink, or continue?" screen appeared between trials, allowing participants to take breaks as needed. The session lasted approximately 1.5 hours, with 2-minute breaks every 12–13 minutes for electrode impedance checks. .

Electrophysiological Acquisition and Processing

Continuous EEG was recorded using a 128-channel Hydrocel GSN SensorNet (Tucker et al., 1994) with the EGI system and NetStation Recording software (EGI Inc., Eugene, OR). The EEG data were amplified, digitized at a 250 Hz sampling rate, and referenced to the vertex during recording. Impedances were below 40 kΩ throughout.

Data processing was conducted using Brain Vision Analyzer 2.0. First, raw EEG data in NetStation format were converted to Brain Vision format using EEGLAB software (Delorme & Makeig, 2004). The EEG was then re-referenced offline to the average of the left and right mastoids. Next, the signals were band-pass filtered between 0.1 and 30 Hz (12 dB/octave slope), and a 60 Hz notch filter was applied to remove line noises. Ocular artifacts were detected via semiautomatic ICA procedure (Infomax algorithm; Makeig et al., 1997). A Value Trigger Algorithm was applied to the vertical EOG (VEOG) channel to detect blinks. ICA components contributing significantly to vertical (VEOG) and horizontal (HEOG) activity, determined by the square of their correlations with these channels, were identified and removed from the data. Segments were excluded if they contained signals exceeding ±100 μV within the segment (epoch). Data were then segmented from 200 ms before the onset of the object nouns to 1,000 ms after. Baseline correction was applied from -200 to 0 ms preceding the target word onset. The remaining trials were averaged for each condition for each participant, producing grand average waveforms.

Statistical Analyses

The selections of time windows and electrode locations were a priori based on prior studies (Lai et al., 2019; Barber et al., 2013; Gullick et al., 2013), as well as visual inspection of the current data. Mean amplitudes were exported from two time windows (N400: 300–500 ms; N700: 500–700 ms) and from four scalp regions (left anterior, right anterior, left posterior, right posterior), covering lateral frontal and posterior sites where N400 and N700 effects are typically observed in the concreteness and imagery literature. (Left anterior: E12, E13, E19, E20, E24, E28, E29; right anterior: E4, E5, E111, E112, E117, E118, E124, left posterior: E52, E53, E59, E60, E61, E66, E67; and right posterior: E77, E78, E84, E85, E86, E91, E92). Exported mean amplitudes were entered in separate Repeated-Measures ANOVAs of 3 Condition (LA, LC, MET) x 4 Location (LA, RA, LP, RP). All p values were adjusted for multiple comparisons using FDR.

Results

The comprehension question accuracy was 87.4% (SD = 6.7%), indicating that participants were doing the task.

N400 (300-500 ms)

There was no significant effect of Condition ($F(2,50)$ = 1.401, p = .256).

N700 (500–700 ms)

There was a main effect of Condition ($F(2, 50)=8.219$, $p<$ .001). Because condition did not interact with location, we combined all 4 quadrants. Pairwise comparisons showed that LC was more negative than LA ($\Delta=0.9038$, p = .0020, d = 0.59) and LC was more negative than MET (Δ =-0.8406, p = .002, d = 0.50). MET did not differ from LA (Δ = 0.0631, d = 0.04) (see Figure 1).

To examine whether the N700 amplitude followed a progressive increase from LA < MET < LC, we conducted a linear trend test. The contrast analysis revealed a significant linear trend (β = -0.4519, p = 0.046), indicating that N700 amplitude followed a stepwise increase from LA, MET, to LC. However, the effect size was small (R^2 = 0.051), indicating it is not the sole factor modulating N700.

Discussion

In this study, we investigated whether the literal and metaphoric embodiment effects observed at the verbs in Lai et al. (2019) would extend to subsequent object nouns. Our key findings were that first, at object nouns, LC elicited a more negative N700 than LA and MET. According to our hypothesis, this suggests that a literal action verb that engages motor related processes at the verb affects processing downstream at the object noun. However it is possible that it is co-determined by concreteness and imageability properties of the object noun. Second, MET did not differ from LA, but showed a linear trend to be different from LA, also in N700. Thus, a metaphoric action verb engages motor related processes at the verb and verb only.

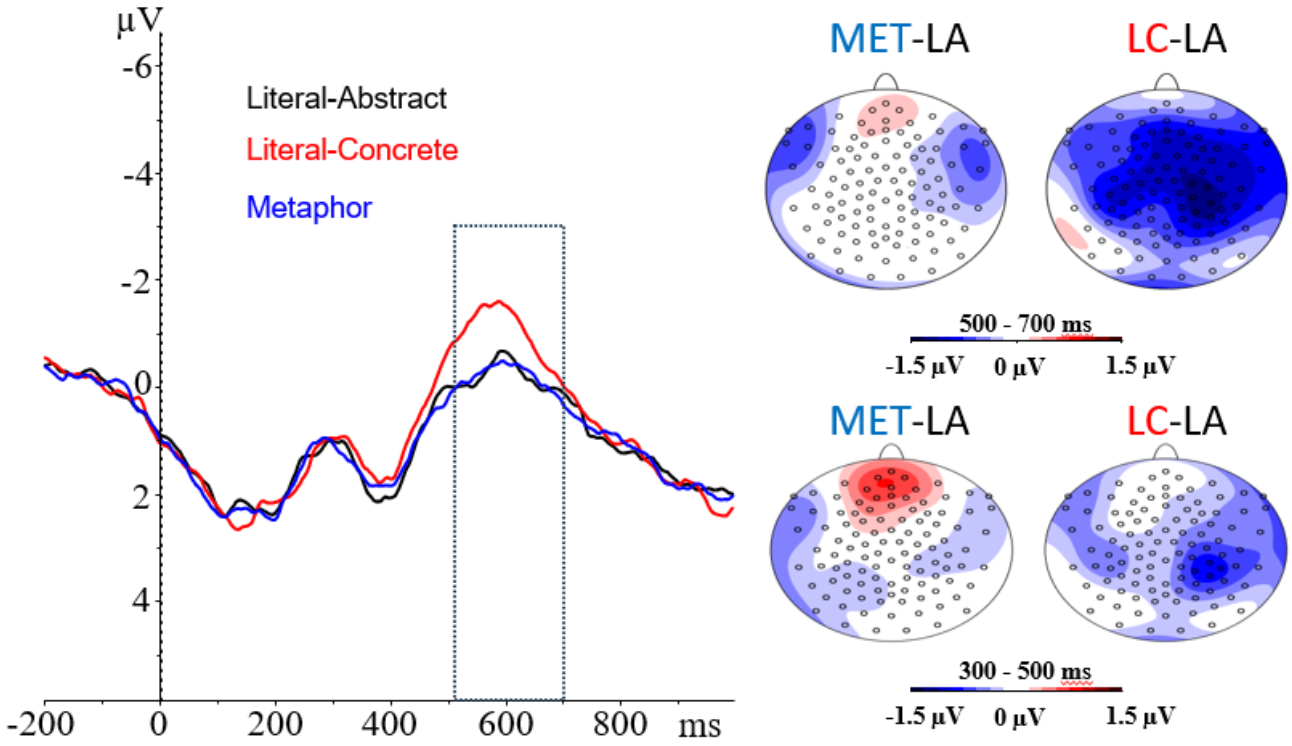


Figure 1: Grand-averaged ERPs for the metaphor (MET, blue), literal-abstract (LA, black), and literal-concrete (LC, red) conditions. Topographical maps on the right show condition differences (MET–LA and LC–LA) in each time window.

N700

In terms of literal embodiment, At the object nouns, LC elicited a more negative N700 than LA. As we suggested above, this was likely due to literal embodiment effect being carried over from the action verbs to the object nouns. However, note that in Lai et al. (2019), the literal embodiment effect at the verbs was revealed in the N400, not the N700. There are at least three possible interpretations. First, the literal embodiment effect trailed off as words went by. By the time the object noun appeared, the embodied activation was weakened. This was reflected via an ERP negativity delay, delayed from ~400 ms to ~700 ms. The second interpretation is that bottom-up, word-level semantic activation led to the N700. As reported in methods, the object nouns in the LC condition were rated as being more concrete and imageable than the object nouns in the LA condition. Since the N700 has been linked to imagery in the word processing literature (Gullick et al. 2013), it could be that these low-level word properties overwrote high-level contextually-driven embodiment and gave rise to the N700 effect. Third, both abovementioned possibilities played their roles. That is, there is a balance between a weakened embodied context and a developing influence from the object noun word level concreteness and imageability. We entertained all three interpretations, as the current data cannot tease them apart.

In terms of metaphoric embodiment, While pairwise comparisons showed that MET did not differ from LA, the linear trend analysis indicated a graded increase from LA, MET, to LC. This suggests that metaphoric action did affect the following object nouns to some extent. Note that this could not be accounted for by word level concreteness/imageability, because the object nouns in MET and LA were the same abstract nouns.

Summarizing, our key findings suggest that while concrete or action language can trigger sensorimotor simulations, the degree to which those simulations transfer further to subsequent words may depend on whether the context is metaphoric or literal and whether subsequent words are concrete and imageable.

Connections to Prior ERP Literature

Our results differ from Li et al. (2022), who reported more positive P600/LPC effects for metaphors at the object position of both subject-verb metaphors and verb-object metaphors. In contrast, our findings did not reveal any P600/LPC activity anywhere. One potential reason is that

such P600 is related syntactic expectedness: Their readers read both subject-verb vs. verb-object metaphors, whereas our readers only read subject-verb metaphors.

Our results align with single-word research on the concreteness N700 effect (Holcomb et al., 1999; Kanske & Kotz, 2007; West & Holcomb, 2000). The more negative N700 for literal-concrete nouns is consistent with prior evidence that concrete words elicit sustained frontal negativity compared to abstract words—even when controlling for frequency, familiarity, and context availability (Barber et al., 2013; West & Holcomb, 2000;). While initially attributed to imagery-based mechanisms (West & Holcomb, 2000), Barber et al. (2010) found a similar extended negativity when comparing nouns and verbs varying in sensory-motor richness. Our finding is consistent with the view that the N700 reflects multimodal activation for concrete words (Barber et al., 2013). In sentence-level ERP studies (Lai et al., 2019), also reported that concreteness N400 effects were modulated by context: More frontal for literal context and more broadly distributed for metaphoric/figurative context. These dissociable patterns suggest that figurative context may engage distinct neural generators for concreteness effects—an issue our study did not examine directly but one that could guide future topographic analyses of object noun processing."

Mechanism: Attention and Resource Allocation in Language Processing

One possible explanation for the transient nature of metaphor-driven sensorimotor simulations is the strategic allocation of attentional resources during sentence comprehension. According to capacity-limited models of language processing (Just & Carpenter, 1992), readers or listeners have a finite pool of attentional resources that can be deployed to interpret demanding linguistic elements. When a figurative verb first appears (e.g., "bend" used metaphorically), it may trigger a burst of sensorimotor simulation as the comprehender works to reconcile the literal action with a metaphoric context (Gibbs, 2006; Lai et al., 2019). Once the figurative meaning is resolved, attentional resources may be reallocated to the next word, especially if that word (e.g., the subsequent noun) necessitates new conceptual or lexical integration. This handoff of resources can reduce the likelihood that sensorimotor activation carries over, as the comprehender effectively "drops" the embodied representation to focus on more immediate semantic demands (Rayner & Clifton, 2009). While metaphor comprehension likely imposes an attentional bottleneck, our linear trend analysis suggests that some sensorimotor engagement persists beyond the verb—albeit in a weakened form. Rather than a binary switch from "embodied" to "non-embodied" processing, metaphor processing appears to gradually transition between abstract and concrete meaning, contributing to a graded sensorimotor effect.

Such a view agrees with theories of attentional control in language which propose that readers flexibly adjust their depth of processing depending on context and task goals (Huettig & Altmann, 2005; Van Petten & Luka, 2012). Metaphor comprehension may therefore act as an attentional bottleneck: once the initial interpretive challenge posed by the figurative expression is addressed, the system switches to a more routine, lexical-semantic mode for subsequent words. This perspective clarifies why we might observe strong, short-lived embodiment effects at the verb, followed by minimal effects downstream. It also highlights how top-down attentional control interacts with bottom-up lexical and semantic cues to shape the time course of sensorimotor grounding- a dynamic interplay that future studies could examine by manipulating task demands, metaphor complexity, or the presence of distractors to see how attention is redistributed in real-time.

Implications for Computational Modeling of Language

Recent advancements in natural language models, such as BERT (Devlin et al., 2019) and GPT (Brown et al., 2020), have yielded impressive results across a variety of language tasks by leveraging large-scale textual corpora. However, these models generally focus on the statistical relationships among words (i.e., distributional semantics) and have limited capacity to incorporate sensorimotor information. The findings from our study suggest that the sensorimotor underpinnings of language may be transient or context-bound. This poses new challenges for models aiming to capture the full richness of semantic processing. If certain types of embodied activation fade quickly, then simply encoding static, multimodal features (e.g., images or sound) may be insufficient. Instead, an ideal computational approach might incorporate a time-sensitive or "decaying" representation of sensorimotor activation, reflecting how real-time comprehension can shift as each new word arrives. Such dynamic embeddings could offer a more faithful simulation of human language processing and might better predict nuanced ERP patterns like the N700 effect observed in our data.

An emerging body of research supports this move toward incorporating grounded or sensorimotor constraints in computational frameworks. For instance, Bruni, Tran, and Baroni (2014) introduced multimodal distributional models that integrate visual features to improve word representations, while Andrews, Vigliocco, and Vinson (2009) combined experiential and distributional information to model conceptual knowledge. Building on these directions, future work could refine transformer architectures by introducing a temporal dimension to sensorimotor embeddings, one that models how embodied

traces wax and wane as context unfolds. Such models could provide a test bed for hypotheses about transient embodiment, thereby offering not only improved language understanding performance but also deeper insight into the oscillations and interplay of sensorimotor and symbolic representations in the human brain.

LLM-Based Measures and Predictability

In our norming, the human cloze scores and the GPT-based conditional probabilities were similar, showing low predictability for the critical nouns across conditions, as expected. Nevertheless, more refined distributional metrics derived from LLMs often outperform standard human cloze norms in predicting N400 amplitude (Michaelov et al., 2022, 2024). These findings suggest that advanced LLM-based surprisal measures, which capture subtle contextual constraints on a token-by-token basis, may sometimes reveal nuances missed by simpler probability or cloze measures. Future studies might benefit from employing more sophisticated LLM-based surprisal metrics to detect subtle semantic differences in contexts.

Considerations and Future Directions

There are several considerations. First, we matched low-level lexical variables such as word frequency, length, and cloze probabilities, to minimize confounds and ensure that observed ERP differences could be attributed to higher-level context of embodiment. However, word concreteness and imageability could not be matched, possibly due to the inherent nature of metaphor that uses concrete terms for abstract meaning metaphorically. Future studies should either match imageability and concreteness or take them into account to pin down the embodiment context effect.

Our study focused primarily on conventional (familiar) metaphors, yet the degree of novelty in metaphorical expressions can drastically alter neural responses (Cardillo et al., 2012). Novel or creative metaphors often demand additional conceptual integration, potentially evoking a different time course of embodiment effects that extend into subsequent words. If a metaphor is especially novel, sensorimotor simulations might persist longer, appearing not just at the verb but also at the noun. Future experiments could systematically manipulate metaphor novelty to see whether more innovative metaphors lead to distinct ERP outcomes at the object position, possibly amplifying or prolonging sensorimotor activation.

The metaphoric condition involved agents performing physically impossible actions. There might be increased cognitive effort required to process events that are unexpected, unusual, or violate real-world knowledge. This effect could be distinct from "metaphoricity," because it is not necessarily about the figurative nature of the expression but rather about how unexpected or implausible the action itself is. Consequently, these effects might also have overshadowed downstream to the object noun processing, modulating ERPs at the object nouns. Future studies could address this "processing difficulty issue" by matching or systematically manipulating the salience of metaphors (e.g., conventional vs. novel metaphors) to explore how varying degrees of "metaphorical complexity" influence metaphor processing at both the verb and the object noun positions. Here, "metaphorical complexity" is used as an umbrella term referring to the degree of cognitive effort required to process a metaphor, which can be influenced by factors such as novelty, conventionality, semantic distance, and conceptual mapping difficulty, among others.

Our task was passive reading with occasional comprehension checks. While this helps approximate natural reading, it provides limited insight into participants' metalinguistic or interpretive strategies. Including a more explicit imagery or context-related task (e.g., asking participants to rate the imageability of the final word after each sentence) might amplify or clarify the literal embodiment effects at the object nouns.

Finally, future research could employ time-frequency analyses or source localization to better characterize the neural oscillations and cortical areas involved in how embodiment processes affect the following words during sentence processing.

Conclusion

Our findings suggest that while literal action descriptions can evoke more sustained sensorimotor engagement that potentially carries over to subsequent words, metaphorical uses of the same verbs appear to trigger more transient and context-dependent embodiment effects. This underscores the dynamic interplay between sensorimotor richness, lexical-semantic properties, and attentional resources in shaping how language is processed in real time. By highlighting the limited "downstream" impact of metaphor-triggered embodiment, our results contribute to a hybrid view of language comprehension, where bodily simulations can be briefly activated for figurative expressions but remain more robust and persistent when describing literal actions.

Acknowledgments

We thank Dr. Steven Bethard for his guidance on tokenization issues and algorithmic approaches for obtaining stable log probabilities from ChatGPT, and Dr. John Allen for his assistance with data analysis and interpretation.

References

Andrews, M., Vigliocco, G., & Vinson, D. (7 2009). Integrating experiential and distributional data to learn semantic representations. *Psychological Review*, *116*, 463–498. doi:10.1037/a0016261

Barber, H. A., Otten, L. J., Kousta, S.-T., & Vigliocco, G. (4 2013). Concreteness in word processing: ERP and behavioral effects in a lexical decision task. *Brain and Language*, *125* doi:10.1016/j.bandl.2013.01.005

Bardolph, M., & Coulson, S. (2014). How vertical hand movements impact brain activity elicited by literally and metaphorically related words: An ERP study of embodied metaphor. *Frontiers in Human Neuroscience*, *8*. doi:10.3389/fnhum.2014.01031

Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2 2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, *7*. doi:10.1016/S1364-6613(02)00029-3

Benedek, M., Beaty, R., Jauk, E., Koschutnig, K., Fink, A., Silvia, P. J., … Neubauer, A. C. (4 2014). Creating metaphors: The neural basis of figurative language production. *NeuroImage*, *90*, doi:10.1016/j.neuroimage.2013.12.046

Boulenger, V., Shtyrov, Y., & Pulvermüller, F. (2 2012). When do you grasp the idea? MEG evidence for instantaneous idiom understanding. *NeuroImage*, *59*. doi:10.1016/j.neuroimage.2011.11.011

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., … Amodei, D. (5 2020). *Language Models are Few-Shot Learners*.

Bruni, E., Tran, N. K., & Baroni, M. (1 2014). Multimodal Distributional Semantics. *Journal of Artificial Intelligence Research*, *49*, 1–47. doi:10.1613/jair.4135

Brysbaert, M., & New, B. (11 2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*, 977–990. doi:10.3758/BRM.41.4.977

Brysbaert, M., Warriner, A. B., & Kuperman, V. (9 2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*, 904–911. doi:10.3758/s13428-013-0403-5

Cameron, L. (9 2008). *Metaphor and talk*. doi:10.1017/CBO9780511816802.013

Cardillo, E. R., Watson, C. E., Schmidt, G. L., Kranjec, A., & Chatterjee, A. (2 2012). From novel to familiar: Tuning the brain for metaphors. *NeuroImage*, *59*, 3212–3221. doi:10.1016/j.neuroimage.2011.11.079

Chatterjee, A. (3 2010). Disembodying cognition. *Language and Cognition*, *2*, 79–116. doi:10.1515/langcog.2010.004

Delorme, A., & Makeig, S. (3 2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21. doi:10.1016/j.jneumeth.2003.10.009

Desai, R. H., Binder, J. R., Conant, L. L., Mano, Q. R., & Seidenberg, M. S. (9 2011). The Neural Career of Sensory-motor Metaphors. *Journal of Cognitive Neuroscience*, *23*, 2376–2386. doi:10.1162/jocn.2010.21596

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (10 2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*.

Forgács, B. (2022). *The Pragmatic Functions of Metaphorical Language* (pp. 41–57). pp. 41–57. doi:10.1007/978-3-030-66175-5_4

Gentner, D., & France, I. M. (1988). *The Verb Mutability Effect: Studies of the Combinatorial Semantics of Nouns and Verbs*. doi:10.1016/B978-0-08-051013-2.50018-5

Gibbs, R. W. (6 2006). Metaphor Interpretation as Embodied Simulation. *Mind Language*, *21*, 434–458. doi:10.1111/j.1468-0017.2006.00285.x

Gullick, M. M., Mitra, P., & Coch, D. (5 2013). Imagining the truth and the moon: An electrophysiological study of abstract and concrete word processing. *Psychophysiology*, *50*, 431–440. doi:10.1111/psyp.12033

Holcomb, P. J., Kounios, J., Anderson, J. E., & West, W. C. (1999). Dual-coding, context-availability, and concreteness effects in sentence comprehension: An electrophysiological investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 721–742. doi:10.1037/0278-7393.25.3.721

Huettig, F., & Altmann, G. T. M. (5 2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition*, *96*, B23–B32. doi:10.1016/j.cognition.2004.10.003

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122–149. doi:10.1037/0033-295X.99.1.122

Kanske, P., & Kotz, S. A. (5 2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Research*, *1148*, 138–148. doi:10.1016/j.brainres.2007.02.044

Kutas, M., & Hillyard, S. A. (1 1980). Reading Senseless Sentences: Brain Potentials Reflect Semantic Incongruity. *Science*, *207*, 203–205. doi:10.1126/science.7350657

Lai, V. T., Howerton, O., & Desai, R. H. (7 2019). Concrete processing of action metaphors: Evidence from ERP. *Brain Research*, *1714*, 202–209. doi:10.1016/j.brainres.2019.03.005

Li, Y., Lu, X., Wang, Y., Wang, H., & Wang, Y. (7 2022). Is the Processing of Chinese Verbal Metaphors Simulated or Abstracted? Evidence From an ERP Study. *Frontiers in Psychology*, *13*. doi:10.3389/fpsyg.2022.877997

Makeig, S., Jung, T.-P., Bell, A. J., Ghahremani, D., & Sejnowski, T. J. (9 1997). Blind separation of auditory event-related brain responses into independent components. *Proceedings of the National Academy of Sciences*, *94*, 10979–10984. doi:10.1073/pnas.94.20.10979

Michaelov, J. A., Bardolph, M. D., Van Petten, C. K., Bergen, B. K., & Coulson, S. (4 2024). Strong Prediction: Language Model Surprisal Explains Multiple N400 Effects. *Neurobiology of Language*, *5*, 107–135. doi:10.1162/nol_a_00105

Michaelov, J. A., Coulson, S., & Bergen, B. K. (9 2023). So Cloze Yet So Far: N400 Amplitude Is Better Predicted by Distributional Information Than Human Predictability Judgements. *IEEE Transactions on Cognitive and Developmental Systems*, *15*, 1033–1042. doi:10.1109/TCDS.2022.3176783

Oldfield, R. C. (3 1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113. doi:10.1016/0028-3932(71)90067-4

Qu, F.-B., Yin, R., Zhong, Y., & Ye, H.-S. (5 2013). Motor Perception in Language Comprehension: Perspective from Embodied Cognition. *Advances in Psychological Science*, *20*, 834–842. doi:10.3724/SP.J.1042.2012.00834

Raposo, A., Moss, H. E., Stamatakis, E. A., & Tyler, L. K. (1 2009). Modulation of motor and premotor cortices by actions, action words and action sentences. *Neuropsychologia*, *47*, 388–396. doi:10.1016/j.neuropsychologia.2008.09.017

Rayner, K., & Clifton, C. (1 2009). Language processing in reading and speech perception is fast and incremental: Implications for event-related potential research. *Biological Psychology*, *80*, 4–9. doi:10.1016/j.biopsycho.2008.05.002

Rutter, B., Kröger, S., Hill, H., Windmann, S., Hermann, C., & Abraham, A. (12 2012). Can clouds dance? Part 2: An ERP investigation of passive conceptual expansion. *Brain and Cognition*, *80*, 301–310. doi:10.1016/j.bandc.2012.08.003

Semino, E., & Steen, G. (9 2008). *Metaphor in Literature*. doi:10.1017/CBO9780511816802.015

Tucker, D. M., Liotti, M., Potts, G. F., Russell, G. S., & Posner, M. I. (1 1994). Spatiotemporal analysis of brain electrical fields. *Human Brain Mapping*, *1*, 134–152. doi:10.1002/hbm.460010206

West, W. C., & Holcomb, P. J. (11 2000). Imaginal, Semantic, and Surface-Level Processing of Concrete and Abstract Words: An Electrophysiological Investigation. *Journal of Cognitive Neuroscience*, *12*, 1024–1037. doi:10.1162/08989290051137558

Wilson, M. (1 1988). MRC psycholinguistic database: Machine-usable dictionary, version 2.00. *Behavior Research Methods, Instruments, & Computers*, *20*, 6–10. doi:10.3758/BF03202594

Wilson, N. L., & Gibbs, R. W. (7 2007). Real and Imagined Body Movement Primes Metaphor Comprehension. *Cognitive Science*, *31*, 721–731. doi:10.1080/15326900701399962