Sensorimotor Grounding in Novel Word Acquisition

John Michael Hollander

Follow this and additional works at: https://digitalcommons.memphis.edu/etd

Recommended Citation
https://digitalcommons.memphis.edu/etd/3589

This Dissertation is brought to you for free and open access by University of Memphis Digital Commons. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of University of Memphis Digital Commons. For more information, please contact khggerty@memphis.edu.
SENSORIMOTOR GROUNDING IN NOVEL WORD ACQUISITION

By

John M. Hollander

A Dissertation

Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

Major: Experimental Psychology

The University of Memphis

August 2024
Copyright © 2024 John Michael Hollander

All rights reserved
Dedication

This work is dedicated to my family. To my parents, Kellie and Mark, whose love and support has shown me how to be compassionate, resilient, and to ask highly of myself. To my sister and brother, Ellen and Robert, who remind me to seek joy, to be creative, and to never stop growing. To my wife, Stephanie, who inspires me in all things. I am grateful.
Acknowledgements

I would like to express heartfelt appreciation and gratitude to my advisor, Dr. Andrew Olney. Both inside academia and out, his patience, dedication, wisdom, and support have been (and will continue to be) invaluable to me. He is a super cool guy, seriously. I am grateful for my committee members, Drs. Roger Kreuz, Shelbi Kuhlmann, and Joah Williams, who played essential roles in various capacities during my time as a graduate student. I would like to give special thanks to Dr. John Sabatini, whose guidance, encouragement, and belief in my abilities helped me learn, create, and produce much more than I thought I could. I would also like to thank the faculty and support staff of the Department of Psychology and the Institute for Intelligent Systems, who never hesitated to lend me help and support.

I am also immensely grateful for advisors and mentors from previous institutions along my academic journey. Special thanks to Dr. Kim Epting, who recognized my own interests and ambitions before I did, fostered my curiosities into lifelong career passions, and afforded me valuable experience and advice for “not if, but when” I go to grad school. Likewise, special thanks to Dr. Alissa Dark-Freudeman, whose guidance and support taught me how to be both analytical and creative in science, and Dr. Graciela Espinosa-Hernandez, whose warmth and encouragement have inspired me to broaden my horizons as an academic.
Abstract

Embodied models of language comprehension assume that words become associated with sensorimotor experiences during word learning. Novel word learning paradigms may provide insight into embodied effects, but studies in this domain have yet to account for how concepts and information known in first language (L1) might influence the sensorimotor encoding of new words and concepts. Further, research on embodiment and bilingualism suggests that second language semantic representations involve sensorimotor encoding, but because participants in these studies are already proficiently bilingual, the manner and degree to which L1 information influences newly learned word representations remain unclear. The experiments in this dissertation address a gap in this literature by examining embodiment in novel word learning while gradually introducing elements of L1 interference. Experiment 1 replicated and extended previous findings by Öttl et al. (2017), which explored the relationship between language and sensorimotor traces in novel word learning, in a fully online environment where participants learned novel words as names for novel (invented) animals. Experiment 2 introduced new, realistic visual features (wings and legs) to test if L1-associated vertical spatial cues would influence embodied processing involving new word representations. Experiment 3 further involved L1 representations by using images of recognizable animals to test the dominance of L1 concepts over newly formed word representations. The results of Experiments 1 and 2 suggest that embodied semantic representations can be formed nearly immediately during word acquisition, while Experiment 3 suggests that existing L1 semantic networks outperform such newly acquired representations. These findings provide insights into the rapid formation and activation of sensorimotor traces in language learning and highlight the need for further research into the interplay between L1 and L2 representations in embodied language processing.
Table of Contents

Chapter | Page
---|---
Dedication | iii
Acknowledgements | iv
Abstract | v
Table of Contents | vi
List of Tables | vii
List of Figures | viii

1. Introduction | 1
  Measuring embodiment in language | 3
  Embodiment in new word learning | 7
  The present studies | 10

2. Experiment 1 | 12
  Method | 12
    Participants | 12
    Materials | 13
    Procedure | 14
    Design | 17
  Results | 17
  Discussion | 24

3. Experiment 2 | 27
  Method | 28
    Participants | 28
    Materials | 29
    Procedure | 30
    Design | 30
  Results | 31
  Discussion | 36

4. Experiment 3 | 42
  Method | 43
    Participants | 43
    Materials | 43
    Procedure | 44
    Design | 44
  Results | 44
  Discussion | 50

5. General discussion | 54

References | 62
List of Tables

Table 1. Stimulus materials, adapted from Öttl et al. (2017).

Table 2. Mixed effects model predicting reaction times.

Table 3. Logistic mixed effects model predicting error rate.

Table 4. Modified object images.

Table 5. Mixed effects model predicting reaction times in Experiment 2.

Table 6. Mixed effects model predicting error rates in Experiment 2.

Table 7. Modified object images for Experiment 3.

Table 8. Mixed effects model predicting reaction times in Experiment 3.

Table 9. Mixed effects model predicting error rates in Experiment 3.
List of Figures

Figure 1. Example of a vertical Stroop-like test. 7

Figure 2. Examples of learning and test phase from Öttl et. al, (2017). 9

Figure 3. Examples of the learning phase and test phase. 17

Figure 4. Free response accuracy over successive learning cycles for present and model study. 18

Figure 5. Reaction time means grouped by location and response direction. 20

Figure 6. Error rates grouped by location and response direction. 21

Figure 7. Free response accuracy over successive learning cycles through Experiment 2. 31

Figure 8. Reaction time means grouped by location, response direction, and feature. 33

Figure 9. Error rate means grouped by location, response direction, and feature. 34

Figure 10. Free response accuracy over successive learning cycles for all experiments. 45

Figure 11. Reaction time means grouped by location, response direction, and animal type. 47

Figure 12. Error rates grouped by location, response direction, and animal type. 48

Figure 13. Reaction time results in Experiments 1, 2, and 3 collapsed to learning position and response direction. 57
Introduction

Word learning is often described with an illustrative scenario in which infants acquire lexical knowledge through the co-occurrence of a spoken word with its referent in the environment (Suarez-Rivera et al., 2022; Yu & Smith, 2012). This process involves the intricate coordination of cognitive operations that integrate information for referent disambiguation in rich, multisensory settings (Schroer & Yu, 2023). Word and concepts are developed over repeated pairings of lexical information and successful, physical interactions with a corresponding object. The sensorimotor information involved during these referent interactions is critical to developing early semantic knowledge (see Wellsby & Pexman, 2014 for review). Complex physical interactions serve not only to disambiguate referents in a complex scene, but they also help form the basis for features and affordances which inform the structure and organization of semantic memory. The role of sensorimotor information in language processing forms the central issue in “embodied” accounts of language (see Louwerse, 2011; Zwaan & Madden, 2005 for examples) which propose specific mechanisms and constraints on how and when perceptual information and language influence one another. Embodied applications in language acquisition research demonstrate that rich, multisensory experiences directly support word learning because embodied interactions provide more opportunities to learn from associative statistical regularities and encode them in semantic memory (Goldstein et al., 2010; Schroer, 2023; Seidl et al., 2023).

However, the role of embodiment in most word learning contexts is not always clear. Most language acquisition occurs outside of the simplified context of one-to-one concrete object naming; it often occurs during conversational speech, without the presence of the referents of each utterance. Abstract concepts without external, tangible referents also provide challenges to
embodied language theories (Löhr, 2019). Many models of word learning acknowledge that the acquisition of lexical semantic concepts involves both online, real-time referent interactions and associative learning through the relationships between words themselves (see McMurray et al., 2012; Regier, 2005 for examples). Although words can be learned from language alone, embodied relations are still encoded in their semantic representations. Studies in this domain suggest that new words learned without direct referential experience can become “grounded” in sensorimotor information indirectly through associative activation of already grounded words that are related to the novel concept (Davis & Yee, 2021; Günther et al., 2018), sometimes through physical metaphor for abstract concepts (Pecher et al., 2011). In addition to this spreading activation mechanism, research also shows that the social, affective, and other interoceptive contexts in which words are learned are influentially embedded in their representations (Snefjella & Kuperman, 2016). Importantly, for both real-time and indirect grounding, the influence of sensorimotor information on the representations of newly learned words appears to be immediate (Günther et al., 2020; Öttl et al., 2017).

However, the manner and degree to which embodied information affects language processing are highly task and context dependent, which poses a challenge to comprehensive theories of human language and cognition (Ostarek & Huettig, 2019). This challenge further extends into applications of this domain. Research into how embodied language processing might be leveraged to facilitate word learning has resulted in two major findings: first, that rich, multisensory environments increase vocabulary acquisition and retention during early development, and second, that richer environments increase engagement and motivation in educational settings (see Fugate et al., 2019; Jusslin et al., 2022 for reviews). However, gaps have been identified in this field of research which demand investigation. First, the majority of
embodied language learning research has been focused on first language (L1) learning in early childhood, while more advanced populations, including adults and second language (L2) learners, have been neglected. Because learning in childhood often involves learning words and corresponding concepts in tandem, it does not clearly extend to adults who are learning new labels for existing concepts, or who have a larger lexicon through which existing concepts are grounded. Second, the specific mechanisms and contexts of encoding and activation during embodied learning tasks have not been examined with enough granularity to draw firm conclusions about exactly how and to what degree specific types of sensorimotor information facilitate learning.

The overall goal of this project is to replicate and extend research in embodied language learning to examine some specific contexts and mechanisms that facilitate word learning in adults. To establish background and justify the experiments, a brief review of selected experimental paradigms in embodied language research is followed by an examination of embodiment in L2 and novel word representation. After synthesizing the literature in these domains, experiments are described that seek to contribute to both basic and applied research in embodied language learning research.

**Measuring embodiment in language**

Empirical evidence suggests that comprehending language involves the neurocognitive simulation of the meaning of an utterance. To illustrate, in one seminal study, participants read simple sentences, such as *the carpenter hammered the nail into the wall*. After reading each sentence, the participants were shown one of two pictures that either matched or mismatched the orientation of an object in the sentence. In the case of this example, they would be shown either a horizontally oriented nail (the matching condition), or a vertically oriented nail (the mismatching
condition). When viewing the picture, the participants were asked to determine whether that object was mentioned in the sentence they just read. Participants were significantly faster to respond to pictures in the matching condition than the mismatching condition. (Stanfield & Zwaan, 2001). In addition to orientation, subsequent studies have also found this to be true of the featural components of objects. For example, after reading *the ranger saw the eagle in the sky*, participants were faster to respond to a picture of an eagle with outstretched wings than to one with folded wings (Zwaan et al., 2002). The most common interpretation of these effects that listeners are mentally simulating the meaning of each sentence, which preactivates the featural representation of, in this example, a flying eagle more than a perched one, even though its wings were never explicitly mentioned. Semantic preactivation can be measured neurocognitively as well. For example, some word recognition studies involving neuroimaging techniques have shown that when participants read the word *kick*, some areas of the motor cortex involved in movement of the feet and legs activate higher than baseline, despite no leg movement taking place (Pulvermüller et al., 2005; Tettamanti et al., 2005). Other studies demonstrate consistent patterns of activation in perceptual brain regions when recognizing concrete objects, (Kiefer et al., 2008; Van Dam et al., 2012) and motor regions when processing abstract concepts (Dreyer & Pulvermüller, 2018; Moseley et al., 2012).

Experimental paradigms involving vertical spatial associations have been an especially prolific branch of embodied language processing research. These paradigms often involve the presentation of words at the upper and/or lower edges of a computer screen while participants are asked to make judgments about the words. For example, Zwaan and Yaxley (2003) found that participants rated the relatedness of word pairs faster when the words were displayed in their canonical physical positions on a computer screen (e.g., *sky* at the top and *ground* at the bottom).
compared to when they appeared in unexpected positions (e.g., *sky* at the bottom and *ground* at the top). Evidence supporting the concept of embodied cognition also arises from studies involving single words displayed in various positions on a computer screen. For instance, Šetić and Domijan (2007) presented words from a list of canonically high or low concepts (e.g., *bird*, *carpet*, etc.), and asked participants to judge whether words represented animate or inanimate objects. Participants made these judgments faster when the concepts matched their expected locations (e.g., *bird* at the top of the screen) compared to mismatches. Pecher et al., (2010) found similar results when participants were asked to judge whether animal words displayed were associated with the ocean or the sky. Alternative explanations of these findings have been explored and must be acknowledged; research suggests that task dependency and language statistics may account for some embodied language effects (Dudschig & Kaup, 2017; Louwerse & Jeuniaux, 2010), though other work has shown that neurocognitive simulation is a robust explanation of many findings in psycholinguistic research (Gozli et al., 2016; Körner et al., 2023). Researchers seeking to scrutinize the role of embodiment in language processing have identified a key conceptual condition to determine the strength of embodied theory: even shallow linguistic tasks should trigger sensorimotor activations that affect behavior (Vogt et al., 2019). In other words, embodied processing should be automatic because it is intrinsic to the cognitive system rather than a mode of processing that is engaged on demand. To create a sufficiently shallow task, researchers have adapted a Stroop-like paradigm which examines vertical spatial associations without requiring access to semantic information.

To illustrate, Lachmair et al., (2011) presented participants with nouns typically found in either lower or upper vertical space (e.g., *root* or *roof*). Each noun was shown centrally on the screen in one of four font colors. Participants responded using a specially constructed response
box which required upward or downward arm movements to respond. This device was equipped with four keys: two middle keys and a lower and upper response key. Two colors are associated with the upper key, and the two other colors are associated with the lower key (see Figure 1). Participants began a trial by depressing both middle keys simultaneously, one with each hand. For a downward (or upward) arm movement, they released the respective middle key, pressed either the lower (or upper) response key, and then returned to the middle key. The hand not involved in responding remained on its middle key. The time interval from when the noun appeared until one of the middle keys was released served as the dependent variable. This design yielded a congruency effect; when the typical location of the object matched the correct response direction, font color judgments were significantly faster. Critically, this task does not necessitate access to any semantic information. While other embodied language processing tasks required participants to make judgments about some property of the words’ referents, this Stroop-like task does not. Because this task can be completed without reference to any physical features or properties of the concepts, it may be considered linguistically shallow. Since congruence effects are still observed even in this shallow task, this study supports the theory that simulation is an automatic, core aspect of language comprehension. This task has since been adapted to explore grounding and spatial associations in many psychological domains, including numerical and social cognition (Ahlberg et al., 2018; Dudschig et al., 2015; Schütt et al., 2022; Thornton et al., 2013; Zhai et al., 2018).
Figure 1. Example of a vertical Stroop-like test.

Recently, researchers have used paradigms like these to investigate remaining questions about the mechanisms and constraints of embodiment in language processing. One key focus of this domain is whether, how, and when grounding occurs when learning new words, especially with respect to bilingualism and language education.

**Embodiment in new word learning**

The majority of embodied language learning research is comprised of studies examining adults (especially college students) in their first language. As a result, many questions remain about the nature and role of embodiment over the course of language learning. For example, are newly learned words grounded in sensorimotor experiences immediately, or does it take years of co-occurrence between words and experiences to create this effect? How does previous experience with related concepts integrate with the grounding of novel words? To answer questions like these, the field of embodied language processing research has recently turned to designs involving artificial language or bilingual participants.
To investigate whether sensorimotor grounding can occur immediately or must take years of co-occurrence, Öttl et al., (2017) designed an experiment in which participants learned novel words for new objects (see Figure 2). Specifically, these researchers fabricated 8 stuffed animals that did not directly correspond to any real animal. These objects were placed either high up on a wall (over 2 meters from the floor), or close to the ground (less than half a meter). During a learning phase, participants acquired the names of the objects through a pointing and naming procedure with corrective feedback. Afterwards, during a test phase, participants completed a Stroop-like spatial association task, similar to those described previously. In this task, participants were presented with the names of the objects they learned and classified the color of the printed font using upward or downward arm movements on a constructed response box. This study revealed a congruence effect; when the vertical position of the learned objects matched the response direction, font color judgments were faster. This study provides evidence in support of embodied accounts of language by demonstrating that spatial associations between words and sensorimotor experience can be established almost immediately in the word learning process, without the need for years of co-occurrence.
This behavioral research aligns with neurological studies that show rapid association between motor cortex activation and novel words (Bechtold et al., 2018; Branscheidt et al., 2017). Although the study by Öttl et al., (2017) examines key aspects of embodiment during word learning in some conditions, it does not necessarily reflect naturalistic language learning, especially L2 learning. When adults learn new words, especially in a second language, they very rarely learn a new word for an entirely new concept altogether. Rather, they often learn new labels for concepts they have already acquired in their first language, which are already grounded in previous sensorimotor experience. Concepts that might be considered “new” are often closely related to other, familiar concepts. In these cases, the embodied activation of features or properties of familiar concepts may interact with the sensorimotor encoding of novel words. How L1 embodied information influences the novel word learning process is a complex topic that is critical to investigate in order to develop more comprehensive theories of embodiment in language and may inform practical approaches to new word learning in educational settings.
On the complementary side of this problem space, research on embodiment in bilingual individuals suggests that L2 representations also elicit embodied language effects. For example, during a vertical Stroop-like procedure, German-English bilinguals were faster to identify font colors when the response movements matched word referent locations (e.g., vogel or bird requiring upward responses; this is also true of emotion nouns and prepositions) (Ahlberg et al., 2018; Dudschig et al., 2014; Kühne & Gianelli, 2019). However, it is unclear whether new words in L2 form new, separate sensorimotor encodings, or rely on “translational” encodings from previously grounded L1 symbols. While debates on the topic highlight critical research on multiple perspectives (see Monaco et al., 2019 for review), most study designs involve L2 bilinguals with at least moderate degrees of proficiency. This approach makes it difficult to address two outstanding challenges in this field. First, if L2 speakers are already relatively proficient, the mechanistic influence of L1 groundings on new word learning are not being observed in real time, but rather after years of experience with L1, L2, and their interaction. Second, since L2 most often involves alternative labels for familiar concepts, it can be difficult to find instances in which L1 and L2 representations are in competition (i.e., contain different or opposing sensorimotor groundings). Observing direct competition between groundings L1 and L2 could provide key evidence for theories of embodiment, language learning, and bilingualism.

**The present studies**

To address these challenges, a series of experiments were designed to investigate the role of embodiment in new word learning with a focus on immediate sensorimotor groundings and L1 interference. Since embodied language studies can be highly task and context dependent (Dudschig & Kaup, 2017; Khatin-Zadeh et al., 2021; Wilson & Golonka, 2013), Experiment 1 involves a computerized replication of the previously described experiment conducted by Öttl et
al., (2017). Experiments 2 and 3 adapt this paradigm to introduce elements of L1 activation to observe how such information interacts with embodiment in new word learning. Along these themes, this research addresses the following research questions:

1) How robust are the effects of embodiment during new word learning across new (e.g., computerized) contexts?

2) Do features of L1 concepts which imply vertical spatial associations interfere with or enhance newly built L2 embodied representations?

3) Does learning new labels for previously known concepts with vertical spatial associations in L1 interfere with or enhance newly built L2 embodied representations?
Experiment 1

The materials and procedure are intended to replicate a computerized version of the experiment originally conducted by Öttl et al. (2017). In the original experiment, participants learned novel words from an experimenter who pointed and verbally named objects which were placed either high up on a wall (over 2 meters from the floor), or close to the ground (less than half a meter). Afterwards, participants completed a vertical association Stroop-like task on a constructed response box. The researchers found a congruency effect between object location and response movement indicating embodied processing.

Embodied language processing effects are often highly specific to the context and task (Dudschig & Kaup, 2017; Khatin-Zadeh et al., 2021; Wilson & Golonka, 2013). Since this study involved motor movements in a large space and an emphasis on multimodal language input during the learning and testing phases, the replicability of these results must be examined. The present experiment was conducted in an attempt to replicate these findings in an alternate setting. To rigorously test the generalizability of this study, this experiment was conducted using a computerized version of the materials and a procedure with online data collection (i.e., involving smaller movements and less emphasis on multimodal input).

Method

Participants

Thirty-five participants ($M_{age} = 33.8$, $SD_{age} = 10.4$, 33% female) were recruited from the Prolific recruitment platform for financial compensation of $5 USD to participate in this study. All participants were native English speakers. No participants exhibited outliers of average reaction time or error rates during any part of the study, so no exclusion criteria were imposed. Participants completed the task online using their own computers.
Materials

Eight artificial 2-syllable words were used in this study, identical to those used by Öttl et al. (2017). These words were associated with images of the eight stuffed animal objects used in that study, using the same pairings (see Table 1). Object names were always presented in black, 40-point Arial font. Images were roughly 250x250 pixels in size. The objects and words were designed to avoid correspondence with any lexical concept so that they are not associated with any spatial location. Each object was assigned to either a high or low spatial location on the screen during the task. Four of the objects always appeared in a spatially upper location, while the other four objects always appeared in a spatially lower location. Upper location was defined as 5% of the screen below the top edge of the screen; lower location was defined as 5% of the screen above the bottom edge of the screen. All participants acquired the same four objects in the upper zone and the same four objects in the lower zone on the screen. All procedures were programmed using the jsPsych experimental library (De Leeuw, 2015).
Table 1

Stimulus materials, adapted from Öttl et al. (2017)

<table>
<thead>
<tr>
<th>Lower spatial location</th>
<th>Upper spatial location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word</td>
<td>Image</td>
</tr>
<tr>
<td>MORA</td>
<td>![MORA Image]</td>
</tr>
<tr>
<td>JOSTER</td>
<td>![JOSTER Image]</td>
</tr>
<tr>
<td>KAFOR</td>
<td>![KAFOR Image]</td>
</tr>
<tr>
<td>RESSEL</td>
<td>![RESSL Image]</td>
</tr>
</tbody>
</table>

**Procedure**

The experiment consisted of a learning phase followed by a test phase. The learning phase consisted of four cycles, described as follows.

**Learning Phase.** The learning phase began with an introductory round in which all eight images were displayed within their respective upper and lower locations, spaced horizontally evenly. One at a time, the phrase “this is a(n) x”, appeared in close proximity to each object, where x is the name of that object (see Figure 3). After this introductory round, participants completed a learning cycle. In the learning cycle, all eight images were displayed on the screen as they were in the introductory phase. The phrase “click on the x” appeared centrally on the screen, where x is one of the eight object names, chosen randomly without replacement. Participants received
performance feedback. In the case of an incorrect response, the word “incorrect” appeared along with the phrase “this is a x” next to the correct object. In the case of a correct response, the word “correct” appeared centrally on the screen. This procedure was completed three times in each learning cycle. After the third repetition, participants were instructed to freely name the objects by typing their names in text boxes that appeared next to each object. Participants also received feedback, including correct object names, after the free response round. After the free response round, the horizontal positions of the objects were randomized, though each object remained in its respective high or low zone. Participants completed four of these learning cycles. To mask the purpose of the study, participants completed a distractor task consisting of free responses to five simple arithmetic operations (adding, subtracting, or multiplying two single-digit positive integers) between the third and fourth learning cycle. The distractor task also served to separate the learning phase from the test phase as well as to extend word learning effects over irrelevant content.

**Test phase.** In the test phase, object names were used in an adapted Stroop-like paradigm. One at a time, words were presented centrally on the screen in one of four colors in 40-point Arial font in either red, green, blue, or orange text. Participants responded to the color of the word by clicking and dragging the word with their cursor. Participants were instructed that two colors (e.g., blue and orange) required moving the word to the upper zone of the screen; the other two colors (e.g., red or green) required moving the word to the lower zone (see Figure 3B). The colors that corresponded to upward or downward movements were randomized between participants. This procedure was programmed using a mouse-tracking plugin for jsPsych, which was designed as a web-based analogue for vertical Stroop tasks, and has been used in successful replications of similar tasks (Schütt et al., 2022) without the need of physical constructed
response boxes, such as the one originally used by Öttl et al. (2017). Response times were measured as the time elapsed between the presentation of the stimuli and the moment that participants initially clicked the word. This latency measurement reflects those used in both face-to-face and web-based studies (Lachmair et al., 2011; Schütt et al., 2022).

Participants first completed sixteen practice trials in which participants acquired the color-response mapping by responding to nonword letter strings (e.g., XXXXX) that appeared in one of the four colors. Performance feedback was provided. Following the practice trials, the object names were used as stimuli. The test phase consisted of eight consecutive sampling blocks; in each block, every word appeared once in each of the four colors. In other words, the eight words in all four colors created thirty-two unique trials, which were sampled without replacement, and this sampling cycle occurred eight times, for a total of 256 test trials. After the Stroop-like task, participants completed another round of freely naming the objects in order to assess the robustness of their learning.
A. Learning phase

![Learning phase image]

B. Test phase

![Test phase image]

Figure 3. Examples of the learning phase (A) and test phase (B).

**Design**

This experiment used a 2x2 within-subjects design. The factors include object location during learning (2 levels: up and down) and response direction (2 levels: upward and downward).

**Results**

In the learning phase, a paired samples t-test indicates that participants correctly identified more objects during the fourth free response round than the first round, $t(34) = 8.83, p < .001$. This result indicates that participants were generally able to learn the objects’ names.
before the test phase. The learning data over successive free response rounds generally reflects the observations of Öttl et al. (2017) (see Figure 4).

![Figure 4. Free response accuracy over successive learning cycles for present and model study. Error bars represent standard error of the mean.](image)

In the test phase, reaction times were analyzed using a linear mixed-effects model. Incorrect trials (2.6% of the data) were excluded from this analysis. Reaction times below 200ms and greater than 2.5 standard deviations more than each participant’s average were also excluded from this analysis (2.7% of the remaining data). Raw reaction times were positively skewed (skewness = 4.2) so reaction times were log transformed (skewness = .27), and then entered into the model as the dependent variable. Object location during learning (hereafter “location”) and the direction of response during the Stroop-like task (hereafter “response”) and their interaction were entered as fixed effects. The random effect structure was determined by entering the maximal structure (including subject intercepts) and systematically removing slope terms until convergence was achieved. If models with equally complex structures converged, AIC was used to select between them. This model selection procedure was designed to comply with the
recommendations of Barr et al. (2013) to use the maximal random effects structure supported by the data, as well as guidelines from Meteyard and Davies (2020) for best practices in reporting. This analysis revealed no significant main effects of response or location, but a significant interaction between those factors ($B = -.05$, $p = .009$) Figure 5 graphically represents these group means, using raw RT rather than log transformed RT for interpretability; Table 2 contains the model output. Planned, pairwise contrasts (Tukey-HSD) indicated no significant differences between estimated marginal means of log-transformed reaction time, though comparisons are in the expected directions (i.e., response directions that match the learning locations are faster than when they mismatch). A logistic mixed-effects model was constructed to analyze error rate. The model construction and selection processes were identical to that of the reaction time model. In this case, the model with only random intercepts for subjects did not converge, so random effects were removed. This analysis yielded no significant interaction, nor any significant effects of response or location. Table 3 and Figure 6 represent this analysis.
Figure 5. Reaction time means grouped by location and response direction. Figure represents group means and standard errors rather than marginal means of linear model.
Figure 6. Error rates grouped by location and response direction. Figure represents group means and standard errors rather than marginal means of linear model.
Table 2

Mixed effects model predicting reaction times

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Estimates</th>
<th>std. Error</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>6.69</td>
<td>0.05</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Response (0 = down)</td>
<td>0.02</td>
<td>0.01</td>
<td>0.078</td>
</tr>
<tr>
<td>Location (0 = down)</td>
<td>0.02</td>
<td>0.01</td>
<td>0.184</td>
</tr>
<tr>
<td>Response × Location</td>
<td>-0.05</td>
<td>0.02</td>
<td>0.009</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random Effects</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2$</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau_{00}$ subject</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau_{00}$ word</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N_{subject}$</td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N_{word}$</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>4248</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marginal R$^2$ / Conditional R$^2$</td>
<td>0.002 / 0.501</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Model equation: logRT ~ response * location + (1|subject)
$\tau$ = estimated variance of random effects
Table 3

Logistic regression model predicting error rate

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Odds Ratios</th>
<th>CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>0.03</td>
<td>0.02 – 0.05</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Response (0 = down)</td>
<td>0.78</td>
<td>0.46 – 1.29</td>
<td>0.332</td>
</tr>
<tr>
<td>Location (0 = down)</td>
<td>0.69</td>
<td>0.40 – 1.16</td>
<td>0.161</td>
</tr>
<tr>
<td>Response × Location</td>
<td>1.40</td>
<td>0.66 – 3.00</td>
<td>0.388</td>
</tr>
</tbody>
</table>

Observations 4360  
R² 0.001

*Note. Model equation: error ~ response * location, family = binomial*
Discussion

The primary purpose of this experiment was to replicate the findings of Öttl et al. (2017), which investigated the link between language and sensorimotor traces during novel word learning, in a completely online, computerized setting. Our results indicate a generally successful replication of the original study, thereby supporting the initial findings and extending their applicability to new contexts. Specifically, the learning rate observed in the free-response tasks in our study closely reflects that of the model study. Further, the significance of the interaction in the analysis of the Stroop-like task (and the lack of significance of main effects) matches the analyses reported in the model study. Regarding error rates, the present study found no significant interactions or main effects, while the model study found a significant main effect of response direction.

A significant portion of embodied language processing research involves physical tasks conducted in large spaces or requires participants to use large motor movements. These studies often necessitate a high degree of physical interaction. In the model study (Öttl et al., 2017), the learning phase occurred in a room where researchers and participants spoke and physically pointed at objects as high as two meters off the floor, and the test phase involved whole arm movements on a constructed response box. In contrast, this replication study involved a learning phase where objects were only several inches apart on a computer screen, and test phase responses were made using only hand and/or finger movements. This replication study demonstrates that the effects observed in these larger, more physically involved settings can also be replicated in smaller, computerized settings, and supports the robustness of the sensorimotor traces in word learning and comprehension. This finding broadens the applicability of embodied
language processing theories and encourages further endeavors to investigate how sensorimotor experiences can shape language processing across different settings and tasks.

The significant interaction in the reaction time model indicates that response directions that matched the learning locations yielded faster trials than when they mismatched. In other words, participants who learned artificial words as labels for novel objects in a multisensory environment showed a compatibility effect between the spatial location of the objects and the direction of the response in a Stroop-like task. This suggests that the sensorimotor information that was available during word learning was reactivated when the words were processed in a situation where no referent was present. Our finding provides direct support for the assumption underlying embodied models of language comprehension - that words become associated with sensorimotor experiences during initial word learning, and that this influential information can become encoded rapidly, not just after years of co-occurrence.

This finding has several implications for the cognitive processes involved in word learning and embodied accounts of language. First, it shows that sensorimotor experiences are not only relevant for the processing of words that have been acquired over many years of language use, but also for words that have been learned in a short period of time. This supports the idea that sensorimotor information is a fundamental component of word meaning representations, rather than a by-product of long-term co-occurrence. Second, it provides additional evidence that sensorimotor experiences are automatically reactivated during word processing, even when the semantic content word is not relevant for the task at hand. Third, it shows that sensorimotor experiences influence word processing beyond the influence of language statistics. In this study, participants did not learn the words in a linguistic context that could have established associations between the words and spatial terms, demonstrating that
sensorimotor information can affect word processing based on the direct experience with the word’s referent, rather than on the co-occurrence with other words, especially in adults. The first research question outlined in this study asks how robust the effects of embodiment are during new word learning across new contexts. The results of this experiment suggest that they are at least somewhat robust in computerized settings.

However, while the study designs employed in our research may approximate some aspects of language learning, they may not generally reflect language learning in adults, especially second language (L2) learning. In the context of L2 acquisition, adults are typically engaged in a process of mapping new lexical information onto existing concepts, a process that is fundamentally different from the initial concept formation that occurs in early L1 learning. This difference may greatly impact the effects of embodiment in more naturalistic novel word learning studies and subsequent educational applications. For example, embodied information from L1 semantic representations may either enhance or interfere with sensorimotor trace patterns when learning new linguistic labels for known concepts. New experiments that specifically investigate the role of L1 interference in embodied language processing must be conducted in order to determine how these effects may differ when integrated with L1 representations. This line of research could provide valuable insights into the cognitive processes underlying L2 learning and the extent to which these processes are influenced by sensorimotor experiences. It could also shed light on the potential benefits and challenges of using embodied learning approaches in L2 instruction. With these results in mind, we can turn to new questions regarding L1 interference in embodied language processing during new word learning.
Experiment 2

While this artificial language paradigm appears to replicate in computerized settings, many questions remain about the nature of embodiment in more complex language learning contexts. For example, empirical evidence suggests that the semantic representations formed during L2 word learning are reliant on similar L1 concepts, though others suggest that different semantic networks are built over time (Foroni, 2015; Monaco et al., 2019). However, studies in this domain usually involve L2 bilinguals with relatively high degrees of proficiency, making it difficult to determine the mechanisms and constraints of L1 influence on novel word learning. Further, L2 vocabulary learning usually involves acquiring alternative labels for familiar concepts, so L1 and L2 representations can be expected to align in many sensorimotor dimensions. To examine the role that L1 representations might play in new word learning, experimental designs must either find or create scenarios in which they are not aligned. Observing direct competition between L1 and L2 embodied representations could reveal more information on this relationship.

Accordingly, a second experiment was designed to investigate whether features of L1 concepts which imply vertical spatial associations interfere with or enhance newly built embodied semantic representations. This experiment featured an almost identical procedure to Experiment 1, except the images of the objects were modified. To elicit vertical spatial associations, half of the images were edited to appear as though the objects have wings. Specifically, the wings were designed to appear as though the object is in flight – a feature used in previous embodied language processing research (Zwaan et al., 2002). The objects which were given wings also had any features representing legs removed, while the objects which are not given wings retain their “legs” (see Table 4). Wings were chosen because participants are likely
to recognize and associate them with other, previously known winged animals (or the general concept of flight), thereby potentially recruiting upward spatial associations when forming representations of these words. The objects that have legs but not wings may be associated with ground-dwelling animals and may elicit downward spatial associations, or at least would not featurally elicit upward associations.

Half of the objects in the high and low presentation zones were given wings. This allows for the observance of feature based L1 facilitation or interference. If L1 sensorimotor information is involved when building new word representations, then the combined effects of immediate spatial information during learning (e.g., Experiment 1) and the vertical spatial associations from familiar featural concepts should accumulate. For example, winged objects which are learned in a high position that also require upward responses movements (or legged objects learned low with downward response movements) should elicit faster reaction times than when those factors are misaligned. If this pattern is not observed, it could be argued that new word groundings of this kind are based more strongly on immediate experience, and not as reliant on L1 representations as some accounts may suggest (Monaco et al., 2019).

Method

Participants

Sixty-eight participants ($M_{age} = 40.0$, $SD_{age} = 11.5$, 66% female) were recruited from the Prolific recruitment platform for financial compensation of $5 USD to participate in this study. All participants were native English speakers. No participants exhibited outliers of average reaction time or error rates during any part of the study, so no exclusion criteria were imposed. Participants completed the task online using their own computers.
Materials

The same eight words from Experiment 1 were used in this experiment. To establish vertical spatial associations from familiar concepts, half of the images were digitally edited to appear as though the objects have naturalistic outspread wings, which were designed to convey the impression that the object is in flight. The objects with wings had any leg-like features removed, while those without wings retained their "legs". Two of the four objects in each of the upper and lower zones were given wings. For the purposes of this experiment, the objects “Joster” and “Elat” were swapped before editing to maintain an equal distribution of winged and legged creatures in each zone (as “Joster” was the only creature that had no discernable legs to begin with).
Table 4

*Modified object images.*

<table>
<thead>
<tr>
<th>Lower spatial location</th>
<th>Upper spatial location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word</td>
<td>Image</td>
</tr>
<tr>
<td>MORA</td>
<td><img src="image1.png" alt="Image" /></td>
</tr>
<tr>
<td>JOSTER</td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td>KAFOR</td>
<td><img src="image5.png" alt="Image" /></td>
</tr>
<tr>
<td>RESSEL</td>
<td><img src="image7.png" alt="Image" /></td>
</tr>
</tbody>
</table>

*Procedure*

The procedure of this experiment was identical to Experiment 1.

*Design*

This experiment employed a 2x2x2 within-subjects design. The factors include object location during learning (2 levels: up and down), response direction (2 levels: upward and downward), and features (2 levels: wings or legs).
Results

In the learning phase, a paired samples t-test indicated that participants correctly identified more objects during the fourth free response round than the first round, $t(67) = 9.42, p < .001$. This result indicates that participants were generally able to learn the objects’ names before the test phase. The learning data over successive free response rounds generally reflects the observations of Öttl et al. (2017) and Experiment 1 (see Figure 7). An independent samples t-test between accuracy at the end the fourth learning cycle between Experiments 1 and 2 indicates no significant difference between learning performance, $t(93.71) = 1.53, p = .13$.

![Figure 7](image)

**Figure 7.** Free response accuracy over successive learning cycles through Experiment 2.

As in Experiment 1, reaction times were analyzed using a linear mixed-effects model. Incorrect trials (1.3% of the data) were excluded from this analysis. Reaction times below 200ms and greater than 2.5 standard deviations more than each participant’s average were also excluded from this analysis (2.4% of the remaining data). Raw reaction times were positively skewed (skewness = 17.8) so they were log transformed (skewness = 1.1), and then entered into the
model as the dependent variable. Object location during learning ("location"), the direction of response during the test phase task ("response"), and wing/leg featural condition ("feature") were entered as fixed effects, as were the three-way interaction and all two-way interactions. The random effect structure was determined by entering the maximal structure (including subject intercepts) and systematically removing slope terms until convergence was achieved. If models with equally complex structures converged, AIC was used to select between them. This model selection procedure was designed to comply with the recommendations of Barr et al. (2013) to use the maximal random effects structure supported by the data, as well as guidelines from Meteyard and Davies (2020) for best practices in reporting.

This analysis revealed a significant interaction between response and location (\(B = -.05, p = .004\)). No other interactions or main effects were statistically significant predictors in the model. Figure 8 graphically represents these group means, using raw RT rather than log transformed RT for interpretability; Table 5 contains the model output. Since no effects involving the wing/leg predictor were significant in the linear model, post hoc contrasts were performed collapsing across this condition using the Scheffé adjustment. These analyses indicated that high position/upward responses were faster than low position/downward responses (\(p = .03, \text{Cohen’s } d = .10\)), low position/upward responses were faster than high position/downward responses (\(p = .003, \text{Cohen’s } d = .12\)), and high position/upward responses were faster than low position/upward responses (\(p < .001, \text{Cohen’s } d = .16\)). No other contrasts were statistically significant.

A logistic mixed-effects model was constructed to analyze error rate. The model construction and selection processes were identical to that of the reaction time model. This analysis yielded no significant interactions or main effects. Table 6 and Figure 9 represent this analysis.
Figure 8. Reaction time means grouped by location, response direction, and feature. Figure represents group means and standard errors rather than marginal means of linear model.
Figure 9. Error rate means grouped by location, response direction, and feature. Figure represents group means and standard errors rather than marginal means of linear model.
Table 5
Mixed effects model predicting reaction times in Experiment 2.

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Estimates</th>
<th>std. Error</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>6.69</td>
<td>0.05</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Response (0 = down)</td>
<td>-0.00</td>
<td>0.01</td>
<td>0.955</td>
</tr>
<tr>
<td>Location (0 = down)</td>
<td>0.02</td>
<td>0.01</td>
<td>0.134</td>
</tr>
<tr>
<td>Feature (0 = legs)</td>
<td>0.00</td>
<td>0.01</td>
<td>0.726</td>
</tr>
<tr>
<td>Response × Location</td>
<td>-0.05</td>
<td>0.02</td>
<td>0.004</td>
</tr>
<tr>
<td>Response × Feature</td>
<td>-0.03</td>
<td>0.02</td>
<td>0.122</td>
</tr>
<tr>
<td>Location × Feature</td>
<td>0.00</td>
<td>0.02</td>
<td>0.943</td>
</tr>
<tr>
<td>Response × Location × Feature</td>
<td>0.04</td>
<td>0.03</td>
<td>0.183</td>
</tr>
</tbody>
</table>

**Random Effects**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2$</td>
<td>0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau_{00 \text{ subject_id}}$</td>
<td>0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICC</td>
<td>0.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N subject_id</td>
<td>68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>8371</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marginal R² / Conditional R²</td>
<td>0.002 / 0.610</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Model equation: logRT ~ response * location * feature + (1|subject)

$\tau =$ estimated variance of random effects
Table 6

Mixed effects model predicting error rates in Experiment 2.

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>Predictors</th>
<th>Odds Ratios</th>
<th>CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Intercept)</td>
<td>0.01</td>
<td>0.01 − 0.02</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Response (0 = down)</td>
<td>1.12</td>
<td>0.59 − 2.14</td>
<td>0.728</td>
</tr>
<tr>
<td></td>
<td>Location (0 = down)</td>
<td>0.65</td>
<td>0.33 − 1.31</td>
<td>0.229</td>
</tr>
<tr>
<td></td>
<td>Feature (0 = legs)</td>
<td>0.60</td>
<td>0.28 − 1.29</td>
<td>0.192</td>
</tr>
<tr>
<td></td>
<td>Response × Location</td>
<td>0.72</td>
<td>0.26 − 1.95</td>
<td>0.512</td>
</tr>
<tr>
<td></td>
<td>Response × Feature</td>
<td>0.97</td>
<td>0.34 − 2.78</td>
<td>0.961</td>
</tr>
<tr>
<td></td>
<td>Location × Feature</td>
<td>1.52</td>
<td>0.48 − 4.82</td>
<td>0.474</td>
</tr>
<tr>
<td></td>
<td>Response × Location × Feature</td>
<td>3.43</td>
<td>0.74 − 15.95</td>
<td>0.116</td>
</tr>
</tbody>
</table>

Random Effects

<table>
<thead>
<tr>
<th></th>
<th>σ²</th>
<th>3.29</th>
</tr>
</thead>
<tbody>
<tr>
<td>τ₀₀ subject_id</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td>ICC</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>N subject_id</td>
<td>68</td>
<td></td>
</tr>
</tbody>
</table>

Observations 8489

Marginal R² / Conditional R² 0.030 / 0.182

Note. Model equation: error ~ response * location * feature + (1|subject), family = binomial
τ = estimated variance of random effects

Discussion

The results of this experiment provide insight into the role of sensorimotor grounding
during novel word learning, and distinct from Experiment 1, the role that immediate grounding
might play while “new” concepts exhibit recognizable features that either align or contrast with
the sensory context.
The performance curve during the free-response trials of the learning phase is very similar to Experiment 1. This is to say that participants seemed to successfully learn the names of the objects over the course of the task. Each point along the curve appears to be slightly lower than the free response data reported in the model study, but not so much lower that the validity of the task is compromised, as learning performance right before the test phase was not significantly different between Experiments 1 and 2. These data suggest that the wing/leg image manipulation neither enhanced nor detracted from participants’ ability to learn the stimuli.

The results of the mixed-effects model predicting reaction time during the test phase is also similar to those of Experiment 1. Specifically, there was a significant interaction between response direction and learned location such that that response directions that matched the learning locations yielded faster trials than when they did not match. This result matches that of Experiment 1, which also demonstrated a compatibility effect between the spatial location of the objects and the direction of the response, which supports the main theoretical inference from Experiment 1: words can become associated with sensorimotor experiences during initial word learning in very few exposures.

The patterns exhibited in the post hoc contrasts generally support this interpretation, with some nuance. The strongest contrast by effect size was between upward responses to objects learned high (in which vertical elements align) and upward responses to object learned low (in which vertical elements misalign). This result seems to indicate that the alignment of sign-referent spatial associations is a key factor in processing facilitation effects. However, this effect was only observed when the conditions aligned in an upward direction; downward responses to low stimuli were not significantly different than downward responses to high stimuli. Further, upward responses to high stimuli were faster than downward responses to low stimuli. The
vertical elements of these conditions align in both cases, which seems to suggest that processing facilitation of this nature might only (or predominantly) occur in the upward direction. This asymmetry is found throughout the literature of spatial processing effects in language (Dudschig et al., 2013; Lakens, 2012; Pecher et al., 2010), and has a few possible explanations. One account of these asymmetric effects is the principle of polarity correspondence (see Proctor & Xiong, 2015 for an overview). The simplest description of this principle is that “For a variety of binary classification tasks, people code the stimulus alternatives and the response alternatives as + polarity and - polarity, and response selection is faster when the polarities correspond than when they do not” (Proctor & Cho, 2006, p. 418). Further, Lakens (2012) offers descriptions of how +polar stimuli elicit processing benefits over −polar stimuli, which may result in all-positive polar stimuli (e.g., high stimuli and upward responses) receiving more processing benefits than all-negative polar stimuli (e.g., low stimuli and downward responses). This is partially consistent with the pattern observed in this experiment, but this pattern of effects was not observed in Experiment 1, and the wing/leg feature dimension resulted in a null effect that should have been significant according to polarity-based predictions (e.g., that winged object would be +polar and result in more processing facilitation when aligned with other +polar dimensions). While polarity correspondence theory may account for some of the results of this experiment, some of the results are incongruent, especially considering the results of Experiment 1, suggesting that polarity is a weak account of the observed effects.

Another possible explanation for asymmetric effect is that the study was conducted online without explicit instructions regarding viewing setup or posture. Participants likely completed the tasks in various conditions that do not reflect a typical laboratory environment. Since we cannot be sure if participants completed the study at a desk with a standard viewing
angle, the "high" and "low" object positions may have been perceived in a variety of angles and perspectives, resulting in noise or asymmetries in the observed effects. It may also be that these contrasts are over-powered. Further, the effect sizes of each of the significant contrasts is small (all Cohen’s $d$ are less than .2). These post-hoc contrasts are conducted on a version of the data that are collapsed across the wing/leg feature condition, and as such, treat the design as 2x2 factorial, similar to Experiment 1. However, this experiment used roughly twice as many participants as Experiment 1 in order to power the original 2x2x2 design. As a result, these relatively weak effects may have not been detectable in Experiment 1 but result in significance in this experiment when the post-hoc analyses on collapsed data are over-powered. However, we cannot be certain of this interpretation; it also could be that the differences in the stimuli between the two experiments cause different patterns of effects in each of the experiments beyond these design and analysis considerations.

The lack of any significant results regarding the wing/leg feature manipulation warrants focused examination. This manipulation was designed to create clear vertical spatial associations with the stimuli independent of their location during learning. In other words, objects with wings were intended to elicit embodied semantic activation of upwardness, and objects with legs were intended to elicit embodied semantic activation of downwardness. Previous studies have demonstrated that linguistically processing animal words which are associated with flight elicits embodied processing facilitation effects (Pecher et al., 2010; Šetić & Domijan, 2007), including images of winged animals (Zwaan et al., 2002), even during shallow tasks like the Stroop-like task (Lachmair et al., 2011; Schütt et al., 2022). However, in the reaction time model, the wing/leg feature manipulation predictor did not yield a significant interaction with response direction, which would have replicated these findings. The interaction between feature and
high/low learned location was also not statistically significant, nor was the three-way interaction between feature, location, and response direction. This likely indicates that the wing/leg features did not achieve their intended effect; that they did not elicit upward or downward semantic associations strongly enough to cause either facilitation or interference in combination with the other conditions of the experiment.

There are a few possible explanations for this null result. First, it could be that the edited images were simply not convincing to participants. While participants likely recognized the wings as such, transferring images of naturalistic bird wings onto cartoonish, static objects did not cause participants to engage in cognitive associations or simulations of these objects with sufficient depth to elicit embodied semantic facilitation. In simpler terms, participants may not have imagined these objects as flying or being high in the visual field beyond what they saw in the task. Relatedly, leaving non-winged objects with legs may not be sufficient to associate those objects with downwardness. A second explanation of these findings could be more directly related to the fact that the objects were unfamiliar to participants. While the previously mentioned studies show that words and images of flying animals can elicit spatial processing facilitation effects during language tasks, the stimuli in those studies included animal words which were likely highly familiar to participants, which capitalizes on the richness of the participants’ embodied semantic network and experiences with those animals and their features. Since the objects in this task were novel to participants, they may not have activated associative semantic networks as richly despite possessing some recognizable features, resulting in a lack of processing interactions.

This experiment was designed to determine whether features of L1 concepts which imply vertical spatial associations interfere with or enhance newly built representations. There is no
evidence of such an influence in this case. Future research should seek to determine other, stronger, and more naturalistic ways of controlling learned object features to address the uncertainty left by the limitations of this experiment. Examination of the findings and limitations of this experiment suggests that the novelty of the objects may have made it more difficult to find an effect where expected. This raises a crucial issue: if the objects were not novel, how might the results differ?
Experiment 3

While Experiment 2 examined featural additions that may elicit vertical spatial information, this design can be further adapted to reflect more naturalistic L2 learning. Specifically, rather than adding minor edits to the novel animal images to elicit spatial information, Experiment 3 involved replacing the novel images with images of objects that are recognizable and nameable to native English speakers. This design modification allows a stronger opportunity for L1 grounding to influence newly built word-space associations. In this way, the results of Experiment 3 should yield similar patterns of results to Experiment 2 but may be stronger in degree.

In this experiment, the same words were used as in the previous two experiments, but the images were replaced with icons of birds or marine animals (see Table 5). Participants are likely to recognize these animals, thereby potentially recruiting spatial associations when forming representations of these words. Flying and swimming animals were chosen in accordance with previous experiments in embodied language processing that require participants to judge animals as related to the sky or the ocean in order to elicit upward and downward spatial associations (Louwerse, 2011; Pecher et al., 2010).

Half of the objects in the high and low presentation zones were assigned as flying or swimming animals, which allows for the observance of either L1 facilitation or interference. As in Experiment 2, if L1 sensorimotor information is involved when building new word representations, then the combined effects of immediate spatial information during learning (e.g., Experiment 1) and the vertical spatial associations from familiar concepts should accumulate. For example, flying animals which are learned in a high position that also require upward responses movements (or ocean animals learned low with downward response movements)
should elicit faster reaction times than when those factors are misaligned. If this pattern is not observed, it could be argued that new word groundings of this kind are primed more strongly from immediate experience, and not as reliant on L1 representations as some accounts may suggest (Monaco et al., 2019), or that L1 representations supersede or outweigh any grounding that occurs during recent experiences.

**Method**

**Participants**

Seventy participants ($M_{age} = 37.2$, $SD_{age} = 12.1$, 40% female) were recruited from the Prolific recruitment platform for financial compensation of $5 USD to participate in this study. All participants were native English speakers. No participants exhibited outliers of average reaction time or error rates during any part of the study, so no exclusion criteria were imposed. Participants completed the task online using their own computers.

**Materials**

The same eight words from Experiments 1 and 2 were used in this experiment. In order to establish vertical spatial associations from familiar concepts, half of the images in each spatial zone were replaced with icons of flying animals, while the other half were replaced with icons of swimming animals (see Table 7).
Table 7

Modified object images for Experiment 3

<table>
<thead>
<tr>
<th>Lower spatial location</th>
<th>Upper spatial location</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Word</strong></td>
<td><strong>Image</strong></td>
</tr>
<tr>
<td>MORA</td>
<td><img src="image1.png" alt="Image" /></td>
</tr>
<tr>
<td>JOSTER</td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td>KAFOR</td>
<td><img src="image5.png" alt="Image" /></td>
</tr>
<tr>
<td>RESSEL</td>
<td><img src="image7.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Procedure

The procedure of this experiment is identical to Experiments 1 and 2.

Design

This experiment used a 2x2x2 within-subjects design. The factors include object location during learning (2 levels: up and down), response direction (2 levels: upward and downward), and animal type (2 levels: avian or marine).

Results

In the learning phase, a paired samples t-test indicated that participants correctly identified more objects during the fourth free response round than the first round, \( t(69) = 8.80, p < .001 \). This result indicates that participants were generally able to learn the objects’ names.
before the test phase. The learning data over successive free response rounds generally reflects the observations of Öttl et al. (2017) and the previous experiments (see Figure 10). A one-way ANOVA of accuracy at the end the fourth learning cycle between all experiments indicates no significant difference of learning performance, $F(2, 170) = 1.10, p = .34$.

![Figure 10](image.png)

*Figure 10.* Free response accuracy over successive learning cycles for all experiments.

As in the previous experiments, reaction times were analyzed using a linear mixed-effects model. Incorrect trials (1.6% of the data) were excluded from this analysis. Reaction times below 200ms and greater than 2.5 standard deviations more than each participant’s average were also excluded from this analysis (3.0% of the remaining data). Raw reaction times were positively skewed (skewness = 6.5) so they were log transformed (skewness = 1.2), and then entered into the model as the dependent variable. Object location during learning (“location”), the direction of response during the test phase task (“response”), and animal type (“animal”) were entered as fixed effects, as were the three-way interaction and all two-way interactions. The random effect structure was determined by entering the maximal structure (including subject intercepts) and
systematically removing slope terms until convergence was achieved. If models with equally complex structures converged, AIC was used to select between them. This model selection procedure was designed to comply with the recommendations of Barr et al. (2013) to use the maximal random effects structure supported by the data, as well as guidelines from Meteyard and Davies (2020) for best practices in reporting.

This analysis revealed a significant interaction between response and animal type ($B = -0.05$, $p < .05$) and a significant main effect of animal type ($B = .03$, $p < .05$). No other interactions or main effects were statistically significant predictors in the model. Notably, the key interaction between learned location and response direction is no longer significant, as it was in Experiments 1 and 2. Figure 11 graphically represents these group means, using raw RT rather than log transformed RT for interpretability; Table 8 contains the model output. Since no effects involving the learned location predictor were significant in the model, post hoc contrasts were performed collapsing across this condition using the Scheffé adjustment. These analyses indicated that upward responses for flying animals were significantly faster than upward responses for swimming animals ($p = .03$, Cohen’s $d = .10$). No other contrasts were statistically significant. A logistic mixed-effects model was constructed to analyze error rate. The model construction and selection processes were identical to that of the reaction time model. This analysis yielded no significant interactions or main effects. Table 9 and Figure 12 represent this analysis.
Figure 11. Reaction time means grouped by location, response direction, and animal type. Figure represents group means and standard errors rather than marginal means of linear model.
Figure 12. Error rates grouped by location, response direction, and animal type. Figure represents group means and standard errors rather than marginal means of linear model.
Mixed effects model predicting reaction times in Experiment 3

### Fixed effects

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Estimates</th>
<th>std. Error</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>6.67</td>
<td>0.04</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Response (0 = down)</td>
<td>0.02</td>
<td>0.01</td>
<td>0.087</td>
</tr>
<tr>
<td>Location (0 = down)</td>
<td>0.01</td>
<td>0.01</td>
<td>0.691</td>
</tr>
<tr>
<td>Animal (0 = swimming)</td>
<td>0.03</td>
<td>0.02</td>
<td>0.038</td>
</tr>
<tr>
<td>Response × Location</td>
<td>-0.02</td>
<td>0.02</td>
<td>0.313</td>
</tr>
<tr>
<td>Response × Animal</td>
<td>-0.05</td>
<td>0.02</td>
<td>0.024</td>
</tr>
<tr>
<td>Location × Animal</td>
<td>-0.01</td>
<td>0.02</td>
<td>0.583</td>
</tr>
<tr>
<td>Response × Location × Animal</td>
<td>0.00</td>
<td>0.03</td>
<td>0.888</td>
</tr>
</tbody>
</table>

### Random Effects

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2$</td>
<td>0.13</td>
</tr>
<tr>
<td>$\tau_{00 \text{ subject_id}}$</td>
<td>0.13</td>
</tr>
<tr>
<td>ICC</td>
<td>0.51</td>
</tr>
<tr>
<td>N subject_id</td>
<td>70</td>
</tr>
</tbody>
</table>

Observations 8552

Marginal $R^2$ / Conditional $R^2$ 0.001 / 0.510

Note. Model equation: logRT ~ response * location * animal + (1|subject)

$\tau$ = estimated variance of random effects
Table 9
Mixed effects model predicting error rates in Experiment 3

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>Odds Ratios</th>
<th>CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>0.01</td>
<td>0.00 – 0.01</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Response (0 = down)</td>
<td>1.67</td>
<td>0.81 – 3.47</td>
<td>0.166</td>
</tr>
<tr>
<td>Location (0 = down)</td>
<td>0.94</td>
<td>0.43 – 2.06</td>
<td>0.879</td>
</tr>
<tr>
<td>Animal (0 = swimming)</td>
<td>1.63</td>
<td>0.78 – 3.40</td>
<td>0.193</td>
</tr>
<tr>
<td>Response × Location</td>
<td>1.09</td>
<td>0.41 – 2.92</td>
<td>0.860</td>
</tr>
<tr>
<td>Response × Animal</td>
<td>0.86</td>
<td>0.34 – 2.21</td>
<td>0.754</td>
</tr>
<tr>
<td>Location × Animal</td>
<td>0.51</td>
<td>0.16 – 1.66</td>
<td>0.266</td>
</tr>
<tr>
<td>Response × Location × Animal</td>
<td>1.43</td>
<td>0.33 – 6.15</td>
<td>0.632</td>
</tr>
</tbody>
</table>

Random Effects

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>σ²</td>
<td>3.29</td>
</tr>
<tr>
<td>τ00 subject_id</td>
<td>1.25</td>
</tr>
<tr>
<td>ICC</td>
<td>0.28</td>
</tr>
<tr>
<td>N subject_id</td>
<td>70</td>
</tr>
</tbody>
</table>

Observations 8694
Marginal R² / Conditional R² 0.026 / 0.295

Note. Model equation: logRT ~ response * location * animal + (1|subject), family = binomial
τ = estimated variance of random effects

Discussion

The performance curve during the free-response trials of the learning phase of this experiment is similar to those of Experiments 1 and 2. Participants appeared to successfully learn the names of the objects over the course of the task. Again, each point along the curve appears to be slightly lower than the free response data reported in the model study, but not so much lower that the validity of the task is compromised. These data suggest that changing to clearly
recognizable and nameable animals neither enhanced nor detracted from participants’ ability to learn the stimuli.

In contrast to the previous experiments (and the model study), there was no significant interaction between response location and learned location in the reaction time model. However, this analysis did yield a significant interaction between animal type and response direction. While there were no significant contrasts between estimated marginal means, the pattern of the interaction is such that upward responses were faster for flying animals than downward responses, and the opposite was true for marine animals. This result closely aligns with similar studies involving the Stroop-like paradigm (Lachmair et al., 2011; Schütt et al., 2022) and other binary lexical classification tasks which feature including sky- and ocean-dwelling animals (Pecher et al., 2010). This finding supports an embodied account of language processing as discussed in those studies. Further, because this effect was observed while using recently acquired novel labels, it also supports an account of embodiment in bilingualism and second language acquisition such that semantic (and embodied) representations formed during L2 word learning strongly rely on mapping to L1 representations (Foroni, 2015; Monaco et al., 2019).

While some of the results of this study align with the predictions of embodied language processing literature in this way, certain elements may contrast with the literature on embodiment and novel word learning. Specifically, there was no significant interaction between response direction and learned location in this experiment, which does not align with the predictions of the model study (Öttl et al., 2017) or the results of Experiments 1 and 2. Using recognizable L1 concepts in the novel word learning paradigm seems to have shifted the processing emphasis during the test phase away from recently built spatial associations and towards embodied representations in L1. This shift seems quite strong; it could have been the case that both recent
spatial associations and L1 associations affect processing during the test phase, but the predictions of that account are not reflected in the data. To use a specific example, if both sources of spatial information contribute to semantic processing, then conditions in which all three dimensions align (e.g., words for flying animals learned in high positions requiring upward responses) should result in more processing facilitation than any other condition. However, this pattern was not observed in the data; there was no significant three-way interaction between these conditions and estimated marginal contrasts do not support that conclusion. In this context, L1 representations seem to outweigh or supersede representations based on recent experience.

Further, the only significant post hoc contrast showed that upward responses to sky-related animals were faster than upward responses to ocean-related animals. This result seems to indicate that the alignment of spatial associations is a key factor in processing facilitation effects for L1 concepts. However, this effect was only observed when the conditions aligned in an upward direction; downward responses to ocean animals were not significantly different than downward responses to sky animals. This asymmetrical effect is similar to one observed in Experiment 2. However, the corresponding asymmetry in Experiment 2 was found in learned position and not concept/featural associations. As in Experiment 2, this asymmetry may be attributable to polarity correspondence, but since the effect was not also observed in the learned position condition for this experiment and other key predictions fail, polarity offers only a weak account of this effect.

This experiment was designed to determine whether spatial associations in L1 interfere with or enhance newly built L2 embodied representations. These results suggest that L1 representations are dominant in embodied language processing when they are available, even when using non-L1 labels. Many overlapping explanations may account for this effect. Through
a distributional lens, it could be that L1 representations are much stronger than newer representations because they are built on a vastly greater amount of sensory experience. In this case, the size and attractor strength of L2 representations is not enough to contribute to semantic processing in comparison to the L1 network. Alternatively (though not necessarily to the contrary), it could be the case that “new” representations aren’t being built at all. Rather, the novel label may simply be integrated as a part of the L1 representation for a recognizable object.

Future research should seek to examine whether this pattern of interplay between L1 and L2 representations is robust to different task conditions and phases of learning. For example, it could be that L2 representations become more involved in processing as they become stronger in semantic memory with more repeated exposures than was allowed in this experiment.
General discussion

This study represents a bridge between two related but distinct bodies of literature. The first and more abundant of the two includes studies seeking to test whether (and under what circumstances) language comprehension is an embodied process. These types of experiments have shown that people respond faster to images that match the context of what they read and make judgments about words faster when they are displayed in positions that match spatial characteristics of their meanings (Lachmair et al., 2011; Louwerse & Jeuniaux, 2010; Pecher et al., 2010; Zwaan & Yaxley, 2003). The common interpretation of these findings is that comprehending language activates perceptual and motoric representations of their referents, which acts as preactivation or priming during tasks that recruit from those same sensorimotor neurocognitive sources (see Dreyer & Pulvermüller, 2018; Vigliocco et al., 2014 for examples of neurological evidence). These paradigms almost always rely on known words and stimuli, and do not offer direct examination of the cognitive mechanisms and constraints involved during word learning processes during which associative grounding takes place.

The other endpoint in the literature includes studies of embodiment during language acquisition. These paradigms typically involve artificial languages or bilingual participants. For example, Öttl et al. (2017) demonstrated that spatial associations between words and sensorimotor experiences can form almost immediately during word learning. Research on bilingual individuals shows L2 words also elicit embodied effects, but it remains unclear if these form new sensorimotor encodings or rely on L1 translations; most of these studies involve proficient bilinguals (Kühne & Gianelli, 2019; Monaco et al., 2019) making it difficult to observe the real-time influence of L1 groundings on new L2 words. The series of experiments described in this document addresses the gap between these two literatures by replicating and
extending methodologies borrowed from each. Experiment 1 was designed as a direct replication of an embodied word learning experiment using novel words and images (Öttl et al., 2017). Experiment 3 uses similar methods, but its stimuli more closely reflected those of other embodied language processing studies in that the images included recognizable semantic categories to participants, namely, animals with clear vertical spatial associations (as in Pecher et al., 2010, for example). Experiment 2 creates a midpoint between these two designs in which the object images were not known semantic icons but had visual features that imply vertical spatial associations. By incorporating methodological elements from across this spectrum of literature, the results of each experiment provide some new empirical foundations for our understanding of the role of embodiment in language learning.

The purpose of Experiment 1 was to replicate and extend Öttl et al.'s (2017) study, examining the connection between language and sensorimotor traces during novel word learning in a fully online, computerized setting. While the model study utilized stuffed animals, a large experimental space, and a specially constructed response box, we observed similar effects using images of the stuffed animals on participants’ computer screens using traditional interfaces as response instruments. The observed results largely reflect the original findings. Specifically, objects that were learned in vertical zones which matched the direction of hand movement during a Stroop-like judgment task showed semantic processing facilitation. This finding demonstrates that spatial information can become rapidly encoded in semantic memory, and can be effectively activated and measured in constrained, virtual environments. This also demonstrates that embodied language processing theories can be robust to some contextual changes, which is a concern in this domain (Dudschig & Kaup, 2017)
Experiment 2 aimed to investigate whether recognizable features that elicit spatial embodied associations in L1 vertical spatial associations would impact sensorimotor traces during novel word learning. To do so, the images of the stuffed animals were edited to appear as though they had wings or legs. This experiment did not yield significant results for the wing/leg feature manipulation, but again replicated the finding that the alignment of learned location and response direction results in processing facilitation. This experiment highlights the challenge of creating novel yet effective stimuli for embodied cognition research. Future studies should explore more naturalistic ways to present such features, possibly leveraging virtual reality to enhance sensorimotor engagement.

Experiment 3 introduced more complete L1 concepts to examine their interaction with newly learned L2 labels. This was achieved by replacing images of the stuffed animals with images of recognizable, namable animals that were either sky- or ocean-dwelling. Unlike the previous experiments, the spatial compatibility effect was overshadowed by strong L1 associations, highlighting the dominance of existing semantic networks. This finding suggests that L1 representations, built upon a lifetime of sensory experiences and distributional linguistic input, supersede or outweigh newly formed L2 representations. Experiment 3’s deviation in the pattern of results raises many questions regarding the specific mechanisms and constraints regarding the dominance of L1 representations. Figure 13 contains a graphical comparison of the results of all 3 experiments collapsed to the learning location/response direction interaction.
Figure 13. Reaction time results in Experiments 1 (left), 2 (middle), and 3 (right) collapsed to learning position and response direction. Figure represents group means and standard errors rather than marginal means of linear model.

Taken together, the results of these experiments contribute some insights and raise further questions in the scientific literature on the role embodied processing in language comprehension and learning. Most research on embodied language learning has focused on L1 acquisition in early childhood, while more advanced groups, such as adults and L2 learners, have been largely overlooked. As a result, the role of embodiment in learning may be underexamined in adults who are learning new labels for already known concepts or who have a larger lexicon that grounds existing concepts. Additionally, the specific mechanisms and contexts of encoding and activation during embodied learning tasks have not been examined in enough detail to draw firm conclusions about how and to what extent particular types of sensorimotor information facilitate learning.

While these interpretations are well supported by the data, other possible explanations of these effects may be considered. The most common alternative explanation to the effects found in this literature is the theory of polarity correspondence (Proctor & Cho, 2006). This principle posits that faster response selection occurs when the polarities of stimuli and responses correspond, with +polar stimuli (e.g., high presentations, upward responses) eliciting more
processing benefits than −polar stimuli (e.g., low presentations, downward responses) (Lakens, 2012). This account is somewhat supported by the asymmetric effects found in Experiments 2 and 3. More specifically, position-response alignment facilitation was only present for upward directions in Experiment 2, and animal type-response alignment facilitation was only present for sky-related animals in Experiment 3. While polarity correspondence partially explains the experiments’ asymmetrical results, it doesn't fully account for the results of all conditions in each experiment. Specifically, alignment between wing/leg feature conditions and the other dimensions did not yield facilitation in Experiment 2, and the position-response alignment facilitation effect seen in Experiments 1 and 2 was not observed in Experiment 3. Additionally, the position-response alignment facilitation effect observed in Experiment 1 was not asymmetrical; it was observed for upward and downward alignments equally. Since the predictions of polarity correspondence fail more often than they succeed, this theory offers only a weak account of these results at best, and the embodied explanation of neurocognitive simulation during language processing remains a favorable interpretation.

Several implications and challenges to a few domains of related literature are raised by the results of this study. In embodiment research, examinations of vertical spatial associations have been a reliable means to observe the effects of sensorimotor representation in language processing. The results of this study show that these effects may be dependent not only on task dynamics (Dudschig & Kaup, 2017), also on the recency of acquisition of the stimuli. Since a three-way interaction was not observed in Experiment 2, and the significant interaction with response direction shifted from learned location to animal type between Experiments 2 and 3, it stands to reason that strong, long-term, and perhaps prototypical representations are responsible for the effects found in that literature. Future research in that domain should seek to examine
how and whether these effects hold while systematically examining more recently acquired, less semantically prototypical concepts. With respect to embodied language learning research (Günther et al., 2020; Öttl et al., 2017), these results demonstrate that immediate sensorimotor grounding may be overpowered by stronger associations. Research in this domain should seek to determine what level of strength (e.g., how many learning phases, over how much time) new associations might need to compete with long-term associations. Relatedly, these experiments support the theory that L2 representations borrow or rely on sensorimotor information from L1 concepts when they are available (Kühne & Gianelli, 2019) rather than building separate representations. New research in this domain may consider examining similar effects at different levels of bilingual fluency to determine whether, how much, and when it is possible for L2 and L1 concepts to diverge or compete. Along these lines but in a broader scope, this research may have implications in linguistic relativity (Dyke, 2022). Specifically, since novel sensorimotor groundings can be built rapidly and reliably when no other semantic mappings are not available, separate cognitive representations of the similar concepts may be theoretically possible to construct across different languages if learned in systematically different contexts. However, since new representations appear to be overpowered by long-term associations, and since sensorimotor learning contexts only rarely systematically misaligned across languages (especially for concrete concepts), linguistic relatively may be only weakly supported by this data.

It is difficult to make many direct, practical inferences from this research into the domain of education. First, there was no difference in participants’ overall learning performance between any of the experiments, so no firm conclusions can be drawn about enhancing recognition or recall. The learning task in these experiments is highly decontextualized, and language learning
(and concept acquisition in general) usually takes place in much richer environments. Further, vertical spatial associations represent a very small fraction of the perceptual and motor information that may be associated with lexical concepts in semantic memory. Since sensorimotor groundings with vertical associations were able to be built in Experiments 1 and 2, it may be inferred that supplementing concept presentations with congruent sensorimotor information may increase the strength of those modal representations, but this is a weak inference because there was no control condition for comparison (e.g., no condition in which all stimuli were presented at the same height) and no learning differences between experiments. Educators and researchers seeking to determine how embodiment may enhance learning may focus on providing multimodal learning environments that are as rich as possible, but few recommendations can be made with any more specificity based on the results of these experiments.

These experiments are also constrained by certain limitations that warrant further investigation. First, these experiments were conducted online without instructions or checks regarding the physical context of the task outside of what was on the computer screen. Participants likely completed the tasks in various conditions that do not reflect a typical laboratory environment. As a result, viewing angles, motor movements, and other contextual effects may have had influential and/or noisy effects on the data. This is of direct concern to embodiment research, and many embodied effects are highly context-dependent (Dudschig & Kaup, 2017). Future research should seek to control (or systematically manipulate) these conditions. Second, the participants of this study were adult, monolingual English speakers. People who are multilingual or are more likely to have recent language learning experience (e.g., taking a college-level foreign language course) may be predisposed to alternative learning or
representational abilities and strategies. As a result, any interpretation of the results of this study should be generalized with strict caution. Further, children, especially those in critical language development periods, may perform quite differently on these tasks as well. Future studies should seek to examine similar effects in more diverse populations, especially with respect to age, language- and nation-of-origin.

The primary goal of this study was to investigate the role of embodied language processing when learning novel words and concepts. More specifically, this series of experiments was designed to extend recent empirical research in this domain to examine how recently built embodied associations interface with L1 embodied semantic representations. These experiments open new opportunities for research into the rapid formation and activation of sensorimotor traces in language learning. This study not only replicates previous findings in the field, but also offers potential insights into the mechanisms of embodied language processing and raises further questions for future research.
References


Schroer, S. E. (2023). *Embodied and environmental influences on early word learning.*

https://hdl.handle.net/2152/120462


