Extending Gesture’s Impact on Word Learning to Reading: A Self-Paced Reading Study

Sarah S. Hughes-Berheim, Spyridoula Cheimariou, John F. Shelley-Tremblay, Margaret M. Doheny & Laura M. Morett

To cite this article: Sarah S. Hughes-Berheim, Spyridoula Cheimariou, John F. Shelley-Tremblay, Margaret M. Doheny & Laura M. Morett (2022): Extending Gesture’s Impact on Word Learning to Reading: A Self-Paced Reading Study, Discourse Processes, DOI: 10.1080/0163853X.2022.2132080

To link to this article: https://doi.org/10.1080/0163853X.2022.2132080
Extending Gesture’s Impact on Word Learning to Reading: A Self-Paced Reading Study

Sarah S. Hughes-Berheim, Spyridoula Cheimariou, John F. Shelley-Tremblay, Margaret M. Doheny, and Laura M. Morett

ABSTRACT
Taken together, the Coherence Principle of Multimedia Learning Theory and the Integrated Systems Hypothesis propose that co-occurring and semantically congruent verbal and visual information should be integrated into one mental representation that enhances memory. The purpose of this paper was to examine how learning pseudowords with matching versus mismatching gestures affects subsequent identification and integration of these newly learned pseudowords into read sentential contexts. Additionally, the pseudowords were manipulated to occur in either semantically congruent or semantically incongruent read sentential contexts, based on the pseudowords’ learned definition. To investigate the research question, two experiments utilizing self-paced reading paradigms were employed. Results of Experiment 1 indicated partial support for the Integrated Systems Hypothesis. In Experiment 2, results indicated that pseudowords learned with matching gestures were identified more quickly and accurately after being read in semantically congruent sentences compared to semantically incongruent sentences, as was expected based on the Integrated Systems Hypothesis as well as the Coherence Principle of Multimedia Learning Theory. Additional results and implications are reported.

Introduction
Vocabulary is an important component of reading comprehension ability, with individuals with greater vocabulary skills demonstrating increased reading comprehension (Landi, 2010) and academic achievement (Berger et al., 2017; Jackson, 2005). Although the bulk of vocabulary learning occurs during childhood, the average adult is exposed to around 2,000 new words each year (Nation & Waring, 1997). For individuals at the university level, this number can only be expected to increase as they are introduced to increasingly complex content in their lectures and textbooks. Despite the importance of vocabulary for reading comprehension and academic success, more information about how college students learn vocabulary words and subsequently integrate them into larger contexts is needed.

Previous research has indicated that observing semantically congruent gestures concurrently with words aids their learning (e.g., Kelly et al., 2009; Morett, 2014). Studies that have presented related material across modalities (e.g., verbal and visual) have been shown to enhance encoding and memory, providing support for Dual Coding and Multimedia Learning Theories (Clark & Paivio, 1991; Mayer, 2002). Additionally, the increased learning that occurs when words are learned with semantically congruent gestures compared to semantically incongruent gestures suggests these two modalities may
be integrated into one neural representation during processing, a major claim of the Integrated Systems Hypothesis (Kelly et al., 2010). Building upon the insight that these theories have provided into gesture and speech processing, which will be discussed subsequently in further detail, the purpose of this study was to extend their findings by investigating whether presenting vocabulary words with semantically congruent gestures, compared to semantically incongruent gestures, aids processing when newly learned words are read in context.

Multimedia learning theory & vocabulary learning

Vocabulary learning involves associating novel words with known meanings during memory encoding and subsequently drawing upon these associations during memory retrieval. Multimedia vocabulary learning occurs when a mental representation of a word comprises both words and images. Dual Coding Theory (Clark & Paivio, 1991) suggests that the combination of verbal and non-verbal (i.e., visual) representations during vocabulary encoding enhances memory. For example, pairing the word *dactylion* with a picture of the tip of the middle finger (its referent) results in visual and verbal representations for the word, improving encoding and recollection. Indeed, pairing novel vocabulary words with images of their referents enhances memory for them (Akbulut, 2007; Chun & Plass, 1996). Specifically, research on second language acquisition indicates increased vocabulary learning occurs when novel words (e.g., *pájaro*) are paired with images of their referents (e.g., a picture of a bird) compared to when words are learned by pairing them with their native language translations (bird; Morett, 2019).

In addition to images aiding vocabulary learning, pairing novel vocabulary with representational gestures may also enhance vocabulary encoding. Representational (i.e., metaphoric and iconic) gestures convey meanings relevant to co-occurring speech via their shape and motion (McNeill, 1992, 2005). They can be used to describe actions (e.g., swinging a bat), to depict spatial properties (e.g., describing a ring as round), or to refer to concrete entities associated with abstract ideas (e.g., putting a hand over one’s heart to convey love; Hostetter, 2011). These gestures, when co-occurring with speech, aid vocabulary encoding as they display information in the visual modality that complements verbal information conveyed via the auditory or verbal modality.

The Coherence Principle of Multimedia Learning Theory posits that both visual (e.g., gesture) and verbal (e.g., speech) stimuli must be relevant to each other and related in order to enhance memory (Mayer, 2002). This principle is supported by studies that have manipulated the use of media during learning. For example, one study found evidence that including visual information, such as animations, with verbal instructions is helpful in enhancing memory retention and transfer (Moreno & Mayer, 2000). Conversely, including irrelevant media (i.e., words, pictures, sounds) along with verbal information hinders rather than benefits recall (Harp & Mayer, 1998; Moreno & Mayer, 2000).

Following this logic, representational gestures should enhance learning of corresponding vocabulary words as long as the meanings of the gestures and words are contextually related. Indeed, studies that have manipulated the co-occurrence of novel vocabulary words and gestures have found gestures that are contextually related to, or semantically congruent with, vocabulary enhance recall (Garcia-Gamez & Macizo, 2019; Morett, 2014). Conversely, representational gestures that depict referents that are unrelated to the corresponding vocabulary words have been found to hinder both word learning and subsequent word identification, since the gesture and word are irrelevant to each other in meaning. These effects of gesture impacting vocabulary learning have occurred even when vocabulary words are presented as text, forcing both the novel words and gestures to be processed via the visual modality (Garcia-Gamez & Macizo, 2019). This is a crucial point, as the Coherence Principle of Multimedia Learning Theory suggests beneficial multimedia learning effects can only occur when stimuli are presented across multiple modalities (i.e., visual and verbal), in order to avoid overloading one system (Mayer, 2002).
**Gesture & speech processing**

Beyond Multimedia Learning Theory, additional research points to an interaction between gesture and speech in language. For example, the Gesture as Simulated Action framework (Hostetter & Alibali, 2008) proposes gestures arise due to mental simulations of actions referred to during speech (e.g., mentally simulating the action of drinking while talking about drinking). Additionally, processing written words, such as drink, activates the motor and premotor areas of the brain that correspond with completing those actions (Hauk et al., 2004), suggesting a link between language and gesture conceptually. Findings such as these have led researchers to believe language is grounded in our sensorimotor experiences (Glenberg & Robertson, 2000), a key claim of embodied cognition.

Embodied cognition proposes that cognitive information, such as language, is learned through bodily interactions with the world (Wilson, 2002). As such, this framework suggests word knowledge should contain semantic information that is gained through one’s experiences (Vigliocco et al., 2009). This interpretation is supported by findings that indicate individuals use the affordances of an object (i.e., how one can interact with said object) when creating meaning (Kaschak & Glenberg, 2000), and by evidence that individuals create mental simulations of the sentences they are reading that aid processing (Stanfield & Zwaan, 2001; Zwaan et al., 2002). During speech, co-occurring gestures are thought to be physical manifestations of the speakers’ mental imagery that facilitate speech production and comprehension (Hostetter & Alibali, 2008; Kelly et al., 2010).

The impact of the relationship between gesture and speech on cognition has been described via the Integrated Systems Hypothesis, which posits that when gesture and speech occur conjointly, both are processed bi-directionally and obligatorily (Kelly et al., 2010). Thus, language comprehension occurs in such a way that information from one modality (e.g., speech) cannot be processed without being influenced by information from the other modality (e.g., gesture) when both are present. This bi-directional and obligatory processing of gesture and speech affects comprehension. For example, an action in a prime video (e.g., chopping) is identified more quickly and accurately when it is followed by a target video that displays semantically congruent representational gesture and speech (e.g., gesture: chopping; speech: “chop”) related to the prime compared to when the prime is followed by a target video containing semantically incongruent gesture, speech, or both (e.g., gesture: twisting; speech: “twist”). Notably, this is the case even when instructions are given to only attend to the speech and ignore the accompanying gesture (Kelly et al., 2010). These results suggest gesture and speech are not only integrated during language comprehension, but that this integration actually enhances comprehension when the information conveyed via both modalities matches.

**Current study**

The purpose of this paper was to further investigate the bi-directional and obligatory relationship between gesture and language by determining whether semantically related representational gestures aid (1) pseudoword learning and (2) pseudoword processing during reading. In alignment with previous studies (Kelly et al., 2007, 2015), the effect of gesture on word learning was explored by manipulating pseudoword-gesture congruency during word learning. Our first research question was whether matching gestures (i.e., gestures depicting the pseudoword’s referent), compared to mismatching gestures (i.e., gestures depicting another pseudoword’s referent), aided pseudoword learning. In alignment with the Coherence Principle of Multimedia Learning Theory, it was hypothesized that pseudowords learned with matching representational gestures would be identified more accurately and quickly than pseudowords learned with mismatching gestures.

Additionally, this study aimed to extend findings from the Integrated Systems Hypothesis by investigating whether gesture-definition congruency at word learning influences subsequent processing of newly-learned words within read sentential contexts. This study is the first of its kind, to our knowledge, to identify whether presenting congruent information during word learning enhances subsequent processing when words are read in sentential contexts. Investigating how learned words
are processed in sentential contexts is important as reading sentences is a more complex process than processing words in isolation. During sentence reading, word-to-text integration – the process of coupling the meaning of the word with one’s representation of the meaning of the sentence – must occur for comprehension to occur (Stafura & Perfetti, 2017). This process requires the reader to continually build and update their understanding of the sentence (Zwaan & Madden, 2004). As a more complex process, it was unclear whether the beneficial effect of learning words via visual input would occur when newly learned words were read in sentences.

In order to investigate the Integrated Systems Hypothesis, participants read newly learned words in sentences that matched some, part, or all of the material presented at word learning. As is explained further in the methods, sentential contexts could be semantically congruent with the gesture the word was learned with, the definition the word was learned with, both the gesture and definition, or neither the gesture nor definition. Because the Integrated Systems Hypothesis proposes gesture and speech are both integrated into one representation, we expected coherent representations to aid processing. Thus, we expected words learned with matching gestures to be read quicker in semantically congruent sentences than words learned with mismatching gestures read in semantically congruent sentences. Additionally, we predicted that if words learned with mismatching gestures were subsequently read in mismatching contexts, the match of the gesture the word was learned with to the context of the sentence would result in processing delays.

Experiment 1 investigated research questions 1 and 2 by having participants learn pseudowords and subsequently read them in sentences via a self-paced reading paradigm. After reading the sentences, participants were asked whether the learned word made sense in the sentence. Participant accuracy on this question was coded as pseudoword-sentential congruency probe accuracy and used as a measure of pseudoword identification. Participants’ reaction time on the pseudoword-sentential congruency probe reflects word meaning identification speed after the word has been read. Finally, time participants spent reading the pseudoword in its sentential context, or self-paced reading latency, reflects on-line pseudoword processing speed during reading. In Experiment 1, all pseudowords occurred in the sentence-final position. Experiment 2 investigated the same two research questions with a self-paced reading paradigm but had participants read sentences where pseudowords occurred in the middle of sentences to account for possible sentence wrap-up effects, which are linguistic processes that take place at the end of a sentence (e.g., semantic integration), that may have affected the results of Experiment 1 (Rayner et al., 1995). Similar measures of pseudoword identification and processing are included.

**Experiment 1**

**Methods**

**Participants**
Thirty-two adult native English speakers (age range: 18–23 yrs.; M = 19.59, SD = 2.07; 23 females, 9 males) were recruited from a university in the Southeastern U.S. to participate in the current study for course credit. All participants had normal hearing and normal or corrected-to-normal vision and were screened for documented speech, language, or learning disabilities. Two participants reported being bilingual in a language other than English and were removed from the data analysis. An additional two participants were removed from data analysis because of technical issues. Participants were briefed on the study and gave written informed consent before participating. All procedures were approved by the institution’s Institutional Review Board.

**Materials**

**Pseudowords**. All pseudowords (e.g., *kroosk*) were generated using the application *Wuggy* (Keuleers & Brysbaert, 2010). Pseudowords were consistent with English phonotactic and graphemic rules, contained between 3–6 letters, had a maximum of two syllables, and had a maximum neighborhood
density of three words. Each pseudoword was randomly assigned a concrete English verb as its referent (e.g., to drink). English verb referents were able to be conveyed transparently via a gesture. All English verbs were taught in the infinitive form and definition assignment was consistent throughout each paradigm.

**Gestures.** Each pseudoword was accompanied by a representational gesture during word learning. All gestures were conveyed via videos that contained a single talker (i.e., a White adult female) producing representational gestures enacting the referents of each verb (e.g., drinking). Videos of all gestures can be found in a publicly accessible database (Hughes-Berheim et al., 2020). Because some referents were facial expressions (e.g., to smile), the talker’s face was visible in all videos.

**Pseudoword-gesture pairings.** During word learning, pseudoword definitions occurred with either a gesture video that matched (i.e., was semantically congruent) or mismatched (i.e., was semantically incongruent) its English definition. For counterbalancing purposes, pseudowords always occurred in pairs (e.g., kroosk – to drink and jesp – to sweep). These word learning pairs were created based on sentence reading trials, such that sentences were as similar as possible (e.g., “She used the broom to sweep” vs. “She used the cup to drink”) to minimize expectancy effects. In the matching learning condition, the pseudoword occurred with the gesture video that matched its English definition. In the mismatching learning condition, the pseudoword occurred with the alternate gesture video within the pseudoword pair, which mismatched its English definition. Pseudoword-gesture pairings were normed by a separate group of 33 native English speakers. Results indicated matching gesture-definition pairs were rated as more semantically similar than mismatching gesture-definition pairs (match: $M = 5.95$, $SD = 1.35$; mismatch: $M = 2.20$, $SD = 1.64$; $t = -3.76$, $t < .001$; Hughes-Berheim et al., 2020).

Two pairs of pseudowords were presented in each learning trial block. Therefore, word learning trials consisted of both pairs of pseudowords learned with a matching gesture, both pairs of pseudowords learned with a mismatching gesture, or one matching and one mismatching pair. In order to determine whether the gestures accurately depicted their matching English definitions, another norming study was completed with an additional set of 30 adult native English speakers. Norming participants selected English definitions representing the meaning of the gestures from among those in the same block with 77–100% accuracy ($M = 94\%$, $SD = 4.5\%$).

**Sentences.** Ninety-six context and critical sentences (one of each per pseudoword) were created for use in self-paced reading blocks. Context sentences, which varied in length, provided semantic context for subsequent critical sentences (e.g., *She was thirsty*). Critical sentences consisted of 4–8 words (e.g., *She used the cup to* . . . ) and always ended with one of the four pseudowords introduced in the previous learning block (e.g., kroosk). A cloze norming study, done with a final group of 30 adult native English speakers, indicated that critical sentences were completed with the expected English definition 73–100% of the time ($M = 88\%$; $SD = 7.4\%$).

All pseudowords, definitions, context sentences, and semantically congruent and semantically incongruent sentences are available in Appendix A. All stimuli are available on OSF via DOI 10.17605/OSF.IO/T7Y3 G.

**Procedure**

Through a succession of interleaved word learning and self-paced reading trials (see below), participants learned a total of 96 pseudowords and subsequently read them in sentences containing them. Pseudoword learning and self-paced reading trials were distributed across 24 blocks, each of which featured four pseudowords. Having participants learn four pseudowords per block before being tested was chosen for counterbalancing purposes and to avoid overloading working memory.
Word learning blocks. During word learning blocks, participants viewed the following sequence twice in succession: pseudoword as text; video of representational gesture conveying either the pseudoword’s referent (matching; see Figure 1a) or a dissimilar referent of a paired pseudoword (mismatching; see Figure 1b); English definition as text; inter-stimulus interval (ISI; 1000 ms). Sequences were repeated to mitigate sporadic inattention and aid memory for pseudowords. Representational gestures were shown after pseudowords and before English definitions rather than concurrently with either of these words to avoid the split attention effect, in which attention is divided between two visual stimuli, hindering learning (Moussa-Inaty et al., 2012). Participants were instructed to remember the English definition shown in each word learning trial. Participants were not given any explicit instructions about the gestures.

For counterbalancing purposes, mismatching and matching pseudoword-gesture pairings were randomized, creating four lists (see Table 1 for example). In this way, each pseudoword was learned in each of the four conditions. The presentation order of learning trials within each list was randomized for each participant.

Self-paced reading blocks. After each set of four word learning trials, participants completed a corresponding self-paced reading block. Participants were instructed to read all stimuli as quickly as possible, pressing a button to proceed. These steps were taken to prepare the stimuli for

Table 1. Pseudoword pairs and gesture-definition pairings for Experiment 1 and 2.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Pair</th>
<th>Word</th>
<th>Gesture</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match/Match</td>
<td>1</td>
<td>Kroosk</td>
<td>Drinking</td>
<td>To drink</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Clarg</td>
<td>Eating</td>
<td>To eat</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Fesp</td>
<td>Sweeping</td>
<td>To sweep</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Prilp</td>
<td>Writing</td>
<td>To write</td>
</tr>
<tr>
<td>Mismatch/Mismatch</td>
<td>1</td>
<td>Kroosk</td>
<td>Sweeping</td>
<td>To drink</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Clarg</td>
<td>Drinking</td>
<td>To sweep</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Fesp</td>
<td>Eating</td>
<td>To eat</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Prilp</td>
<td>Writing</td>
<td>To write</td>
</tr>
<tr>
<td>Match/Mismatch</td>
<td>1</td>
<td>Kroosk</td>
<td>Drinking</td>
<td>To drink</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Clarg</td>
<td>Sweeping</td>
<td>To sweep</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Fesp</td>
<td>Eating</td>
<td>To eat</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Prilp</td>
<td>Writing</td>
<td>To write</td>
</tr>
<tr>
<td>Mismatch/Match</td>
<td>1</td>
<td>Kroosk</td>
<td>Sweeping</td>
<td>To drink</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Clarg</td>
<td>Drinking</td>
<td>To sweep</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Fesp</td>
<td>Eating</td>
<td>To eat</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Prilp</td>
<td>Writing</td>
<td>To write</td>
</tr>
</tbody>
</table>
a subsequent ERP study, which, in turn, is based on the design of previous ERP studies (e.g., Kutas & Federmeier, 1999).

During self-paced reading blocks (Figure 2), participants were instructed to proceed through the following sequence as quickly as possible by pressing a button: context sentence (presented wholesale); critical sentence ending with pseudoword (presented one word at a time); pseudoword-sentential congruency probe (presented wholesale). All stimuli were presented centrally, such that words in the critical sentence appeared to replace one another in the center of the screen upon button presses. The pseudoword-sentential congruency probe, which was the same in all trials, read as follows: Did the last word make sense in the sentence? Participants responded to this probe as quickly as possible by pressing a button to indicate yes or no. Self-paced reading conditions were counterbalanced within each block across participants. Additionally, the presentation order of trials within each self-paced reading block was randomized separately from the presentation order of trials within the preceding word learning block.

Critical sentences belonged to one of four possible conditions based on the semantic congruency of both definitions and gestures with which pseudowords were learned (see Table 2 for examples):

C/C: Definition semantically congruent within the sentential context learned with a matching gesture semantically congruent with the sentential context (definition-congruent, gesture-congruent);
IC/IC: Definition semantically incongruent within the sentential context learned with a matching gesture semantically incongruent with the sentential context (definition-incongruent, gesture-incongruent);
C/IC: Definition semantically congruent within the sentential context learned with a mismatching gesture semantically incongruent with the sentential context (definition-congruent, gesture-incongruent);
IC/C: Definition semantically incongruent within the sentential context learned with a mismatching gesture semantically congruent with the sentential context (definition-incongruent, gesture-congruent).

### Table 2. Semantic congruency of definitions and gestures relative to critical sentences for Experiment 1.

<table>
<thead>
<tr>
<th>Semantic Congruency (Definition/Gesture)</th>
<th>Context Sentence</th>
<th>Critical Sentence Pseudoword</th>
<th>Definition (Gesture Learned With)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C, C</td>
<td>She was thirsty.</td>
<td>She used the cup to kroosk.</td>
<td>Drink (Drinking)</td>
</tr>
<tr>
<td>IC, IC</td>
<td>She was thirsty.</td>
<td>She used the cup to fesp.</td>
<td>Sweep (Sweeping)</td>
</tr>
<tr>
<td>C, IC</td>
<td>She was thirsty.</td>
<td>She used the cup to kroosk.</td>
<td>Drink (Drinking)</td>
</tr>
<tr>
<td>IC, C</td>
<td>She was thirsty.</td>
<td>She used the cup to fesp.</td>
<td>Sweep (Sweeping)</td>
</tr>
</tbody>
</table>

C = Congruent, IC = Incongruent
Results

Self-paced reading latency, or the amount of time participants spent reading pseudowords in sentences, was the main outcome of interest. Additionally, accuracy and reaction time on the pseudoword-sentential congruency probe, hereafter referred to as congruency probe, were analyzed. Prior to the analysis, congruency probe latency data for trials with incorrect responses (21.75%) were excluded. Additionally, congruency probe (3.21%) and SPR latency (15.8%) values 3 standard deviations above the mean were excluded.

Crucially, the design of the study was not completely crossed. Due to the nature of the paradigm, pseudowords learned with matching definitions and gestures could only be read in conditions in which sentential contexts were either semantically congruent with both the definition and gesture of the pseudoword (C/C) or semantically incongruent with both the definition and gesture of the pseudoword (IC/IC). Further, pseudowords learned with mismatching definitions and gestures could only be read in conditions in which sentential contexts were either semantically congruent with the definition of the pseudoword and semantically incongruent with the gesture the pseudoword was learned with (C/IC) or semantically incongruent with the definition of the pseudoword and semantically congruent with the gesture the pseudoword was learned with (IC/C).

Due to the design of the study, self-paced reading latency for the pseudoword and congruency probe accuracy were analyzed via two models, described below. Self-paced reading latency and congruency probe accuracy were analyzed using $2 \times 2$ linear mixed effect models with gesture-definition congruency (matching vs. mismatching gesture) at learning and pseudoword definition congruency in critical sentences (semantically congruent vs. semantically incongruent definition) as fixed effects. In all models, fixed effects were coded using weighted mean centered (Helmert) contrast coding. Random intercepts for participants and items were entered into the models to account for different baseline levels of performance on these factors.

We employed a data-driven approach to model selection, such that the maximal random effect structure permitting convergence was used to reduce type I error rates while maintaining power (Matuschek et al., 2017). When multiple models converged, a significance test was run to determine whether the more complicated models, with random slopes and intercepts, fit the data significantly better than the simpler models, with only random intercepts. The model comparison procedure as well as the complete data analysis scripts can be referenced via the R code located in the supplementary materials on OSF via DOI 10.17605/OSF.IO/T7Y3G.

Additionally, all model results can be found in Appendix B. Self-paced reading latency for pseudowords and congruency probe latency were fit using the lmer() function of the lme4 package (Bates et al., 2015). Congruency probe accuracy was analyzed using the glmer() function of the lme4 package.

Self-paced reading latency

The main effect of learning congruency failed to reach significance, indicating that pseudowords learned with matching gestures were read with a latency similar to pseudowords learned with mismatching gestures ($\beta = -3.56, SE = 2.44, t (2143.65) = -1.46, p = .146$; see Table S1). Additionally, there was a marginal main effect of the semantic congruency of the definitions of pseudowords within sentences. Pseudowords were read slightly more quickly in sentences in which their definitions were semantically congruent than in which they were semantically incongruent ($\beta = 4.48, SE = 2.45, t (2162.44) = 1.83, p = .067$). Finally, the interaction between learning congruency and pseudoword definition congruency failed to reach significance ($\beta = -2.68, SE = 4.91, t (2176.87) = -0.55, p = .585$). See Figure 3 for illustration with means and significant contrasts.
The main effect of learning congruency on congruency probe accuracy failed to reach significance, indicating that pseudowords learned with matching and mismatching gestures were identified with similar accuracy ($\beta = -0.11$, $SE = 0.099$, $z = -1.08$, $p = .279$; see Table S2). By contrast, the main effect of pseudoword definition congruency reached significance, indicating that identification accuracy was greater for pseudowords read in sentences semantically congruent rather than semantically incongruent with their learned definitions ($\beta = -0.59$, $SE = 0.099$, $z = -5.99$, $p < .001$). Finally, the interaction between learning and pseudoword definition congruency reached significance ($\beta = 0.48$, $SE = 0.197$, $z = 2.43$, $p = .015$). See Figure 4 for illustration and means with significant contrasts.

In alignment with the Coherence Principle of Multimedia Learning Theory, post-hoc tests (see Table 3) revealed that pseudowords learned with matching definitions and gestures were identified more accurately when they appeared in semantically congruent sentential contexts (C/C) compared to semantically incongruent sentential contexts (IC/IC; $\beta = 0.83$, $SE = 0.14$, $z = 5.83$, $p < .001$). Moreover, pseudowords learned with matching gestures appearing in semantically congruent contexts (C/C) were identified more accurately than pseudowords learned with mismatching gestures appearing in semantically incongruent contexts (IC/C; $\beta = 0.70$, $SE = 0.15$, $z = 4.81$, $p < .001$). Finally, pseudowords learned with mismatching gestures appearing in semantically congruent contexts (C/IC) were identified with higher accuracy than pseudowords learned with matching gestures appearing in semantically incongruent contexts (IC/IC; $\beta = 0.49$, $SE = 0.13$, $z = 3.62$, $p < .002$).
Congruency probe latency

Results revealed that the main effect of learning congruency failed to reach significance, indicating that responses to congruency probes had similar latencies for pseudowords learned with matching gestures than mismatching gestures ($\beta = 27.60$, $SE = 37.93$, $t (1907.14) = 0.73$, $p = .467$; see Table S3). By contrast, the main effect of pseudoword definition congruency reached significance, indicating that responses to congruency probes were quicker when pseudowords were read in semantically congruent rather than semantically incongruent sentential contexts ($\beta = 263.89$, $SE = 38.08$, $t (1936.19) = 6.93$, $p < .001$). Finally, the interaction between learning and pseudoword definition congruency was marginal ($\beta = -146.98$, $SE = 75.94$, $t (1925.43) = -1.94$, $p = .053$). See Figure 5 for illustration and means with significant contrasts. More information about the results and how they were derived are available on OSF via DOI 10.17605/OSF.IO/T7Y3G.

Table 3. Post-Hoc contrasts for congruency probe accuracy in Experiment 1.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean</th>
<th>SD</th>
<th>Tukey’s HSD Contrasts</th>
</tr>
</thead>
<tbody>
<tr>
<td>C/C</td>
<td>84.7%</td>
<td>36.5%</td>
<td>C/C 0.093</td>
</tr>
<tr>
<td>C/IC</td>
<td>80.0%</td>
<td>40.1%</td>
<td>C/IC &lt;.001 .05</td>
</tr>
<tr>
<td>IC/C</td>
<td>74.5%</td>
<td>43.6%</td>
<td>IC/C &lt;.001 .002 .729</td>
</tr>
<tr>
<td>IC/IC</td>
<td>72.0%</td>
<td>45.0%</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4. Accuracy on congruency probe across conditions in Experiment 1. * $p < .05$; ** $p < .01$; *** $p < .005$
Contrary to our hypothesis, pseudowords learned with matching gestures were not identified more accurately or quickly than pseudowords learned with mismatching gestures. However, pseudowords learned with matching gestures were identified more quickly and accurately after being read in sentential contexts in which both the definition and gesture of the pseudoword were congruent (C/C condition) compared to sentences with which both the definition and gesture of the pseudoword were semantically incongruent (IC/IC). This finding indicates partial support for the Integrated Systems Hypothesis. The Integrated Systems Hypothesis postulates that when semantically congruent information is displayed across multiple modalities (e.g., matching gesture and definition during word learning), the information is integrated into a unitary mental representation that enhances subsequent word processing (Kelly et al., 2010). The result of identification accuracy increasing when both parts (i.e., gesture and definition) of a word are congruent rather than incongruent with the sentential context suggests the congruent and unitary mental representation aided subsequent word processing.

Contrary to the hypothesis, no differences in self-paced reading latencies were detected when pseudowords learned with matching gestures were read in sentences in which both the definition and gesture of the pseudoword were semantically congruent (C/C) compared to both being semantically incongruent (IC/IC). This finding contradicts the Integrated Systems Hypothesis. A possible reason for this discrepancy may be because critical words occurred at the end of a sentence when wrap-up processes are taking place. During self-paced reading trials in Experiment 1, single words were presented centrally on the screen consecutively, replacing one another, rather than revealing masked words consecutively across the screen in the more typical moving window configuration. Although

**Figure 5.** Latency on congruency probe across conditions in Experiment 1. *p < .05; **p < .01; ***p < .005
these trials were consistent in structure with ERP trials, self-paced reading latencies may have encompassed wrap-up effects due to placement of pseudowords at the end of sentences (Rayner et al., 1995). Wrap-up effects include specific linguistic processes that occur at the end of a sentence, such as assigning referents to pronouns, connecting clauses, and determining semantic properties of the sentence (Stowe et al., 2018). As such, presenting critical words in the sentence-final position may have affected self-paced reading latencies via simultaneous semantic integration processes (i.e., determining whether the word was semantically congruent with the context of the sentence). In Experiment 2, we addressed these procedural differences and determined whether the results were replicable in a more typical moving window self-paced reading paradigm in which pseudowords appeared in the middle, rather than at the end, of sentences.

**Experiment 2**

**Methods**

**Participants**

Participants of Experiment 2 were 46 adult native English speakers (age range: 18–35 yrs.; M = 20.51, SD = 3.18; 37 females, 8 males) recruited from the same university in the Southeastern U.S. who had not participated in Experiment 1. Exclusion criteria were the same for Experiment 2 as they were for Experiment 1. One participant reported being bilingual in a language other than English and was removed from the data analysis.

**Procedure**

The procedure of Experiment 2 was similar to that of Experiment 1. The major changes in Experiment 2 compared to Experiment 1 were that self-paced reading trials were adjusted such that participants read critical sentences via a noncumulative moving window self-paced reading paradigm (Mitchell, 1984) and the pseudoword occurred in the middle of the sentence instead of at the end. The noncumulative moving-window paradigm (see Figure 6) was introduced so that the procedure was more typical of other self-paced reading studies (Jegerski & VanPatten, 2014). Additionally, the pseudoword was moved to the middle of the sentence rather than the end to eliminate any wrap-up effects (Vasishth & Lewis, 2006) that may have occurred in Experiment 1. In order to keep Experiment 2 as similar as possible to Experiment 1, target sentences were minimally adjusted (e.g., She took the cup to *kroosk* vs. She went to *kroosk* from the cup). Critical sentences used in Experiment 2 can be referenced via Appendix A. All stimuli are available on OSF via DOI 10.17605/OSF.IO/T7Y3 G.

**Figure 6.** Self-paced reading trial in Experiment 2. Figure 3 displays the pseudoword *kroosk* occurring in a semantically-congruent sentence reading trial in which the definition of *kroosk* makes sense within the sentence. In a semantically-incongruent sentence reading trial, participants would read the pseudoword *fosp* in place of *kroosk*.
Additional changes to the paradigm in Experiment 2 were implemented to improve participant attention and comprehension. First, Experiment 2 was split into two halves with 12 blocks each that participants completed on separate days 48 hours apart. This was done because the paradigm in full may have been too taxing for participants and may have resulted in inattention based on participant observation through live video feed in Experiment 1. During self-paced reading blocks, participants were not told to proceed as quickly as possible, as they were in Experiment 1, as we were concerned that these instructions may have hindered full comprehension of critical sentences. Moreover, the comprehension question was changed so that it didn’t require pseudoword recall, such that participants were asked “Was the word [insert pseudoword read in sentence] used correctly in the sentence?”

Finally, catch trials were introduced at random following word learning and self-paced reading blocks. These catch trials asked participants whether a certain pseudoword was learned in the previous block. Eight catch trials were included in order to gauge participant attention across the two paradigms. On average, participants answered the catch trials correctly 86.14% of the time (SD = 17.12%). Participants who scored below 50% accuracy on these catch trials were removed (n = 4). The complete data analysis scripts for Experiment 2 can be referenced on OSF via DOI 10.17605/OSF.IO/T7Y3G.

Results

Self-paced reading latency
The main effect of learning congruency failed to reach significance, indicating that pseudowords learned with matching and mismatching definitions and gestures were read with similar latencies ($\beta =$

![Figure 7. SPR latency for pseudowords in sentences across conditions in Experiment 2. * $p < .05$; ** $p < .01$; *** $p < .005$](image-url)
−9.91, SE = 19.90, t(39.16) = −0.50, p = .62; see Table S4). Additionally, there was a marginal main effect of definition congruency on self-paced reading latency; pseudowords were read slightly faster in sentences in which their definitions were semantically congruent than in which they were semantically incongruent (β = 25.52, SE = 14.94, t(3610.52) = 1.71, p = .09). Finally, the interaction between learning and definition congruency reached significance (β = −86.22, SE = 30.11, t(3608.06) = −2.86, p = .004). See Figure 7 for illustration and means with significant contrasts. In alignment with Experiment 1, post-hoc tests (see Table 4) revealed only one significant contrast. Pseudowords learned with matching gestures appearing in semantically congruent contexts were read quicker than pseudowords learned with matching gestures appearing in semantically incongruent contexts (β = −68.6, SE = 21.2, t = −3.24, p = .007).

### Congruency probe accuracy

The main effect of learning congruency reached significance, indicating that pseudowords learned with matching definitions and gestures were identified more accurately than pseudowords learned...

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean</th>
<th>SD</th>
<th>C/C</th>
<th>C/IC</th>
<th>IC/C</th>
</tr>
</thead>
<tbody>
<tr>
<td>C/C</td>
<td>641.02</td>
<td>538.22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C/IC</td>
<td>673.21</td>
<td>595.56</td>
<td>.543</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IC/C</td>
<td>663.18</td>
<td>597.44</td>
<td>.924</td>
<td>.841</td>
<td></td>
</tr>
<tr>
<td>IC/IC</td>
<td>715.53</td>
<td>689.48</td>
<td>.007</td>
<td>.481</td>
<td>.145</td>
</tr>
</tbody>
</table>

Figure 8. Accuracy on congruency probe across conditions for Experiment 2. * p < .05; ** p < .01; *** p < .005
with mismatching definitions and gestures ($\beta = -0.28, SE = 0.09, z = -3.23, p = .001$; see Table S5). Additionally, the main effect of definition congruency reached significance, indicating that pseudowords were more accurately identified when they appeared in sentential contexts semantically congruent than semantically incongruent with their learned definitions ($\beta = -1.11, SE = 0.089, z = -12.42, p < .001$). Finally, the interaction between learning and definition congruency failed to reach significance ($\beta = 0.03, SE = 0.18, z = 0.19, p = .85$). See Figure 8 for illustration and means with significant contrasts.

**Congruency probe latency**

The main effect of learning congruency was significant, indicating that response latencies to congruency probes were shorter for pseudowords learned with matching gestures than mismatching gestures ($\beta = 124.89, SE = 49.86, t (34.29) = 2.51, p = .017$; see Table S6). Additionally, the main effect of definition congruency reached significance, indicating that response latencies to congruency probes were shorter when pseudowords appeared in sentences in which their definitions were semantically congruent than in which they were semantically incongruent ($\beta = 376.73, SE = 32.95, t (2914.02) = 11.43, p < .001$). Finally, the interaction between learning and definition congruency reached significance ($\beta = 347.20, SE = 65.98, t (2936.82) = 5.26, p < .001$). See Figure 9 for illustration and means with significant contrasts.

Post-hoc analyses (see Table 5) revealed congruency probe latencies were significantly shorter when pseudowords appeared in semantically congruent rather than semantically incongruent contexts, regardless of whether they were learned with matching or mismatching gestures ($p = .83$) compared to in all other conditions ($p \leq .001$). Congruency probe latencies for pseudowords learned with

![Figure 9](image_url)  
**Figure 9.** Latency on congruency probe across conditions in Experiment 2. * $p < .05$; ** $p < .01$; *** $p < .005$
matching gestures occurring in sentences semantically incongruent with their learned definitions (IC/IC) were longer than those for such pseudowords occurring in semantically congruent sentential contexts (C/C; p = .001) but were shorter than those for pseudowords learned with mismatching gestures occurring in semantically incongruent sentential contexts (IC/C; p < .001). Finally, congruency probe latencies were significantly longer for pseudowords leaned with mismatching gestures occurring in semantically incongruent sentential contexts compared to all other conditions (IC/C; p < .001). More information about the results and how they were derived are available on OSF via DOI 10.17605/OSF.IO/T7Y3 G.

**Discussion**

Results of Experiment 2, in general, were similar to those of Experiment 1. For self-paced reading latency, there was no main effect of learning congruency. Contrary to Experiment 1, there was a main effect of sentential congruency, such that participants read pseudowords quicker when they occurred in semantically congruent rather than semantically congruent contexts. Finally, in Experiment 2, pseudowords learned with matching gestures were read quicker in sentences in which the definition and gesture were semantically congruent (C/C) compared to sentences in which the definition and gesture were semantically incongruent (IC/IC). This result supports the Integrated Systems Hypothesis’s claim that co-occurring, semantically congruent inputs that occur across modalities are integrated into a single mental representation (Kelly et al., 2010). Additionally, this integration seems to improve subsequent processing of pseudowords. As in Experiment 1, there were no differences in self-paced reading latency for pseudowords learned with mismatching gestures that were read in sentences.

Contrary to Experiment 1, there was a main effect of learning congruency on congruency probe accuracy and latency, such that pseudowords learned with matching gestures were identified more quickly and accurately, in general, compared to pseudowords learned with mismatching gestures. These results were expected based on previous literature and indicate support for the Coherence Principle of Multimedia Learning Theory (Mayer, 2002) by indicating that increased learning occurs when semantically-congruent and co-occurring input occurs across modalities. Similar to Experiment 1, results from Experiment 2 indicate a main effect of sentential congruency on both accuracy and reaction time on the congruency question, such that there was improved accuracy and reduced reaction times for words read in semantically congruent rather than semantically incongruent contexts. In general, these results support the interpretation that the sentential context a word occurs in influences word processing.

Unlike Experiment 1, there was no significant interaction between learning and sentential congruency on congruency probe accuracy; however, there was a significant interaction between the two factors on reaction time to the congruency probe (the interaction was marginal in Experiment 1). Post-hoc contrasts revealed pseudowords learned with matching gestures were identified more quickly after being read in semantically congruent sentences (C/C) compared to semantically incongruent sentences (IC/IC). Additionally, reaction times on the congruency probes for words learned with matching gestures and read in mismatching contexts (IC/IC) were shorter than the reaction times for words
learned with mismatching gestures and read in contexts that are semantically incongruent with the definition of the word but semantically congruent with the gesture the word was learned with (IC/C).

These results indicate some support for the Integrated Systems Hypothesis in two ways. First, the reduced processing times for words that are completely congruent rather than completely incongruent with sentential context suggest easier processing occurs when words are read in semantically congruent rather than semantically incongruent contexts. Additionally, the increase in processing times for words that were partially congruent with the sentential context rather than completely incongruent, suggest that the partially-congruent lexical representation may have increased the difficulty of the task. Thus, this result supports the idea that co-occurring gestures and speech during word learning may be integrated into a single lexical representation that influences subsequent word processing. Contrary to our predictions, there were no differences in congruency probe reaction time results between words learned with matching gestures that were read in semantically congruent contexts (C/C) and words learned with mismatching gestures that were read in semantically congruent contexts (C/IC), which does not support the above interpretation or the Integrated Systems Hypothesis.

**General discussion**

Representational gestures co-occurring with speech can facilitate novel word learning by enhancing encoding. Co-occurring gestures and words must be relevant to each other and related in order to enhance memory encoding. However, no studies to date have investigated how words learned with semantically congruent versus semantically incongruent gestures are recognized and identified in sentential contexts. The purpose of the present study was to examine how learning words via text and with representational gestures that either match or mismatch their referents affects subsequent identification and integration of these newly learned words in read sentential contexts. We investigated this research question across two experiments using a self-paced reading paradigm. In these experiments, participants read newly learned words in sentences in which the definitions of newly learned words were either semantically congruent or semantically incongruent within the context of the sentence. Crucially, these words were learned accompanied either by matching or mismatching gestures. We measured self-paced reading latencies for pseudowords and identification speed and accuracy, as indicated by a probe question. Importantly, button presses for the probe question were not counterbalanced across keys, and yes responses to the congruency probe were always made with the right hand. For this reason, accuracy and latency for the congruency probe should be interpreted with caution. We hypothesized that words learned with matching gestures would be processed more quickly than words learned with mismatching gestures when they are read in semantically congruent sentential contexts. We also hypothesized that words learned with matching representational gestures would be identified more accurately and quickly than words learned with mismatching gestures when these words are read in semantically congruent sentential contexts.

Three main results were found, partially confirming our hypotheses. First, in Experiment 2, although not in Experiment 1, words learned with matching definitions and gestures were read faster when they appeared in semantically congruent contexts compared to semantically incongruent contexts. This effect replicates past research on contextual effects in reading documenting contextual facilitation effects for words preceded by semantically congruent contexts (Becker, 1980; Ehrlich & Rayner, 1981; Federmeier, 2007; Kutas & Hillyard, 1984; Stanovich & West, 1979, 1983). Most importantly, in Experiment 2, words learned with matching gestures were read quicker in sentences in which representations of words were semantically congruent compared to sentences in which representations of words were semantically incongruent. This finding largely supports the Integrated Systems Hypothesis, which posits that when semantically congruent information is presented (e.g., matching gesture and definition during word learning), this information is integrated into a unitary mental representation that enhances subsequent word processing (Kelly et al., 2010).
Related to this, when processing mismatching gestures, participants’ self-paced reading latency did not differ by sentential context. That is, reading times of the pseudowords were similar in sentential contexts semantically incongruent with either the definition or the gesture. This lack of difference in mismatching gestures’ self-paced reading latency may indicate that the gesture had an effect on word learning, such that representations of words mismatched independently of whether the incongruency stemmed from the gesture or the definition. This offers evidence supporting the Integrated Systems Hypothesis as it suggests that a unitary representation was formed and that one learning modality (e.g., gesture) was not more important than the other (e.g., definition) for learning.

Third, unlike Experiment 1, Experiment 2 showed effects of learning congruency on accuracy and latency for the congruency probe, consistent with our hypotheses. That is, words were identified more accurately and more quickly when learned with matching definitions and gestures than when learned with mismatching definitions and gestures. This is in alignment with the Coherence Principle of Multimedia Learning Theory, which posits that co-occurring stimuli must be contextually relevant to enhance memory encoding and to benefit recall and comprehension (Kelly et al., 2010). Importantly, our results provide further evidence consistent with the idea that these beneficial effects can occur even when the learned stimuli pairs appear in the same modality (text and gesture).

Contrary to our hypothesis, the effect of learning congruency was not observed in any of the measures in Experiment 1. Notably, neither experiment showed learning congruency effects in self-paced reading latency. Methodological differences between the two experiments may shed light on these results. In Experiment 1, stimuli were presented in an artificial reading mode, i.e., one word at a time in the center of the screen, which may have introduced additional processing difficulties. The fact that we found accuracy and latency differences on the congruency probe when a more realistic presentation format was used in Experiment 2 (the moving-window technique) may indicate that presentation format may have affected processing efficiency. When comparing different methods of self-paced reading paradigms, (Just et al. (1982) showed that the moving-window technique most closely resembles naturalistic reading as it approximates gaze duration in eye-fixation studies, whereas the stationary method (as used here in Experiment 1) came in a close second.

Another explanation for the differences between the two experiments may have to do with the presence of wrap-up effects. Wrap-up effects, such as determining the semantic properties of a sentence, have been frequently documented in eye-tracking reading studies (Asahara, 2018; Hirotani et al., 2006; Just & Carpenter, 1980; Rayner et al., 2000; Warren et al., 2009), and are known to influence self-paced reading latencies of sentence-final words. One of the changes introduced in Experiment 2 was the addition of additional words at the end of the sentences in order to avoid these wrap-up effects. Given that the results in self-paced reading latency were roughly similar in the two experiments, wrap-up effects may not have influenced reading times. However, congruency probe measures differed between the two experiments, suggesting that this difference may have been induced by the change in the probe question. The probe question in Experiment 1 required recall of the pseudoword, which may have increased cognitive load, compared to in Experiment 2 where the pseudoword was repeated in the probe question, avoiding additional cognitive load. Increased cognitive load is associated with increased task difficulty, which could have made the task in Experiment 1 more difficult than Experiment 2, resulting in reduced accuracy and increased reaction time on the congruency probe, although there is no way to know this for sure.

In general, the main effect of sentential congruency across almost all comparisons in Experiment 1 and 2 (except for SPR latency in Experiment 1) indicates that participants relied more on definitions of pseudowords than on gestures associated with them when responding to probe questions. When investigating the interaction effects in Experiment 1 and 2 for words learned with mismatching definitions and gestures, congruency probes were generally processed more accurately (in Experiment 1) and quickly (in Experiment 2) in sentential contexts semantically congruent with the definition rather than the gesture. This finding may be related to the obligatory relationship between gesture and speech in the Integrated Systems Hypothesis. According to this idea, when speech and
gesture co-occur, information from one modality cannot be processed without being influenced by information from the other modality. It may be the case that obligatory conjunctive processing of language and gesture does not apply in the written modality.

As already mentioned, the present study focused on written language, whereas most of the research regarding gesture has focused on spoken language. The focus of previous research on gesture processing in conjunction with spoken language is unsurprising given that gestures are typically produced in conjunction with spoken language. However, one of the major contributions of the present study is that it offers insight into the way that different modalities contribute to encoding and integration of learned words with text. Thus, since gestures are not an integral part of written language, it may have been easier for participants in the present study to isolate gestures from their referents when processing written text. As such, the results may indicate a boundary condition for obligatory integration of gesture with language. Importantly, pseudoword definitions were always learned in the infinitive verb form, which is not always the most common way these words are read. This discrepancy may have affected how learned words are subsequently processed. Future studies investigating similar research questions may define pseudowords as nouns or examine learning of rare rather than fake words to enhance ecological validity.

To conclude, this study is the first to investigate how words learned with matching and mismatching definitions and gestures are integrated into written sentential contexts. Many key results were replicated across two experiments. Participants were able to form a unitary integrated mental representation of a word that was subsequently integrated into a sentential context, replicating standard contextual facilitation effects. Our results support the Coherence Principle of Multimedia Learning Theory as well as the Integrated Systems Hypothesis, although they offer insight into the boundaries of obligatory integration of gesture and language. Results of the study suggest that when professors are teaching college students new terminology during lectures, they may want to consistently use representational gestures that match the meanings of the new words to aid word learning. By including definitions and semantically-congruent gestures during spoken word learning, the students may have a stronger mental representation of the word that will aid subsequent word processing when they subsequently read the words in context.

Future research should seek to verify and replicate these results using similar or related methods such as event-related potentials (ERPs) and eye-tracking. For example, the N400 ERP, which reflects semantic integration, or the P600 ERP, which reflects misprediction (DeLong & Kutas, 2020; Thornhill & Van Petten, 2012; Van Petten & Luka, 2012), may shed more light on different aspects of processing words learned with matching versus mismatching representational gestures within sentential contexts. Additionally, future studies should consider including a measure of word learning in addition to the dependent variables to verify the extent to which words were learned and recalled accurately. Nevertheless, the results of the current research suggest that representational gestures can facilitate the learning and subsequent processing of novel vocabulary presented in the written modality in college students, providing further insight into how gesture facilitates learning in this population and context.

**Acknowledgements**

We would like to thank Mary Bray, Raymond Bulger, Jake Feiler, Cailee Nelson, and Gracie Springer for their help on this project.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**ORCID**

Sarah S. Hughes-Berheim (http://orcid.org/0000-0002-5661-5499)
References


