



The influence of sentence focus on mental simulation: A possible cause of ACE instability

Hua Jin^{1,2,3} · Guangfang Zhou² · Xiang Li⁴

Accepted: 5 March 2024 / Published online: 1 April 2024
© The Psychonomic Society, Inc. 2024

Abstract

Recent studies have revealed the instability of the action–sentence compatibility effect (ACE). The current study was designed to demonstrate the hypothesis that the instability of the ACE may be attributed to the instability of focused information in a sentence. A pilot study indicated that the focused information of sentences was relatively stable in the sentence–picture verification task but exhibited significant interindividual variability in the action–sentence compatibility paradigm in previous studies. Experiments 1 and 2 examined the effect of sentence focus on the shape match effect and the ACE by manipulating the focused information of sentences using the focus marker word “是” (is). Experiment 1 found that the shape match effect occurred in the original sentence, while it disappeared when the word “是” (is) was used to make an object noun no longer the focus of a sentence. Experiment 2 failed to observe the ACE regardless of whether the sentence focus was on the action information. Experiment 3 modified the focus manipulation to observe its impact on the ACE using different fonts and underlines to highlight the focused information. The results indicated that the ACE only occurred when the action information was the sentence focus. These findings suggest that sentence focus influences mental simulation, and the instability of the ACE is likely to be associated with the instability of sentence focus in previous studies. This outcome highlights the crucial role of identifying specific information as the critical element expressed in the current linguistic context for successful simulation.

Keywords Sentence focus · Action–sentence compatibility effect · Shape match effect · Mental simulation · Embodied cognition

A critical distinction between humans and other organisms is that humans can understand language and use it to communicate thoughts and information. How do humans understand language? How is the meaning of language represented in the brain? These questions have not been clearly explained.

Amodal system theory suggests that language processing is based on abstract, amodal symbols that are arbitrarily linked to the objects they represent. The brain is only a system that manipulates these symbols (Newell & Simon, 1976). Although this theory has produced rich research achievements and prompted the development of computational models of language comprehension, such as latent semantic analysis (LSA; Landauer & Dumais, 1997), it cannot explain how abstract symbols are associated with objects in the real world, which was referred to as the symbol grounding problem (Harnad, 1990). Therefore, researchers have sought to understand language function through the lens of embodied cognition.

In the early stages of the embodied language theory, researchers considered that sensorimotor experiences are essential for language comprehension. They have proposed that language concepts are rooted in the sensorimotor system and that language comprehension involves reactivating sensorimotor experiences (Barsalou, 1999; Qu et al., 2012). In this view, the sensorimotor experiences generated

Hua Jin and Guangfang Zhou contributed equally and should be considered co-first authors.

✉ Hua Jin
jinhua@mail.tjnu.edu.cn

¹ Key Research Base of Humanities and Social Sciences of the Ministry of Education, Academy of Psychology and Behavior, Tianjin Normal University, Tianjin, China

² Faculty of Psychology, Tianjin Normal University, Tianjin, China

³ Tianjin Social Science Laboratory of Students' Mental Development and Learning, Tianjin, China

⁴ School of Psychology, Xinxiang Medical University, Xinxiang, China

through interaction with the external world are extracted and transformed into analogue, modal symbols, which are stored in long-term memory. In language comprehension, the brain systems recruited to process embodied symbols are similar to actual perceptual and motor processes. Based on this process, language symbols are linked to objects in the real world, providing a reasonable explanation for the symbol grounding problem. This viewpoint was represented by the perceptual symbol system (Barsalou, 1999) and the immersive experienter theory (Zwaan, 2004), and they are collectively referred to as “strong embodiment” theories. For example, the perceptual symbol system proposed that accumulated sensorimotor experiences from interaction with the external world are reactivated to simulate events, scenes, and actions described by language in language comprehension, thereby leading to understanding of the meaning of sentences. Some researchers call this process “mental simulation” (Zhang et al., 2015). However, as the research further develops, it was revealed that embodied representations are flexible and not always observed, while symbolic representations may be more prevalent in language comprehension. Based on this perspective, some researchers proposed the language and situation simulation theory (Barsalou et al., 2008) and the symbol interdependency hypothesis (Louwerse & Jeuniaux, 2008), they consistently proposed that mental simulation occurs only when deep processing language and that symbolic representations are sufficient when roughly processing meets current situational or task demands. These theories are collectively referred to as “weak embodiment” theories.

Researchers have devoted efforts to finding empirical evidence for the activation of sensorimotor experiences during language comprehension. The two most representative behavioral paradigms are the sentence–picture verification task (SPVT; Stanfield & Zwaan, 2001) and the action–sentence compatibility paradigm (ASCP; Glenberg & Kaschak, 2002), which have been used to demonstrate the activation of perceptual and motor experiences, respectively, during language comprehension.

In Stanfield and Zwaan’s (2001) study, participants were required to read sentences that implied a specific orientation of an object (e.g., “John put the pencil in the drawer.” implying horizontal orientation; “John put the pencil in the cup.” implying vertical orientation). After reading the sentences, participants viewed pictures (a pencil oriented horizontally or vertically) and judged whether the objects in the pictures were mentioned in the preceding sentences. Match and mismatch conditions were constituted based on the congruence between the orientation of the objects in the pictures and the orientation implied in the sentences. The results showed that response times were significantly shorter in the match condition than in the mismatch condition (the match effect). The researchers argued that this difference was due to the

activation of object orientation features during the reading process, indirectly supporting the activation of perceptual experiences during language comprehension. After this paradigm was proposed, a series of studies using this paradigm demonstrated the activation of other perceptual features (e.g., shape, size, color, visibility) during language comprehension (Connell & Lynott, 2009; de Koning et al., 2017a; Horchak & Garrido, 2021; Speed & Majid, 2020; Yaxley & Zwaan, 2007; Zwaan et al., 2002). Moreover, several studies have conducted replication experiments on various perceptual features to explore their stability and reliability. Despite finding inconsistent results in the direction or color dimensions, these studies consistently demonstrated the stability of the shape match effect (de Koning et al., 2017b; Rommers et al., 2013; Zwaan & Pecher, 2012). Furthermore, the shape match effect has been observed in children and the elderly (Madden & Dijkstra, 2009; Wassenburg et al., 2017). Based on these research findings, while the reasons for the instability of the direction or color match effects are not yet conclusive, there is substantial evidence supporting the stability of the shape match effect.

Glenberg and Kaschak’s (2002) study aimed to demonstrate the simulation of motor experiences during sentence comprehension. In this study, participants were required to read sentences that implied a specific motion direction (toward or away from the body) and to judge the sensibility of the sentences by pressing the “YES” or “NO” button. The critical manipulation in this task was the button configuration. In half of the trials, participants had to move toward their own body when pressing the “YES” button and away from their own body when pressing the “NO” button. In the other half of the trials, the button configuration was reversed. The button direction and the motion direction implied in the sentences constituted congruent and incongruent conditions. Similar to the results of Stanfield and Zwaan (2001), this study found that reading times were significantly shorter when the button direction was congruent with the motion direction implied in the sentences compared with the incongruent condition (referred to as the action–sentence compatibility effect [ACE]), indirectly supporting the activation of motor experiences during language comprehension. Following this study, some researchers replicated and extended the ACE (Bergen & Wheeler, 2010; De Scalzi et al., 2015; Glenberg et al., 2008). However, in recent years, the stability and reliability of the ACE have been questioned. Initially, two studies attempted to replicate the ACE (Díez-Álamo et al., 2020; Papesh, 2015); although they tried to maintain consistency with the original studies with regard to sentence materials, experimental manipulations, and procedures, both obtained null results. Papesh (2015) conducted Bayesian analyses on previous data and found little robust evidence to support the alternative hypothesis, indicating the instability of the ACE. Building on these two studies, researchers

sought to provide more robust evidence for the stability and reliability of the ACE through preregistered multilab replications. However, none of the laboratory data supported the ACE (Morey et al., 2022). Furthermore, a meta-analysis revealed that although the effect size of the ACE was significant, its value was small (Cohen's $d = 0.129$). More importantly, this meta-analysis revealed a severe publication bias in ACE studies, suggesting that many studies that failed to find the ACE were not published (Winter et al., 2022).

In conclusion, (1) the SPVT and the ASCP have similar research logic that aims to demonstrate the mutual influence between language and actual perceptual (or motor) processes, supporting the activation of perceptual or motor experiences during sentence comprehension, and (2) the two paradigms differ in terms of replicability. For SPVT, although some of its results have shown instability, at least the shape match effect it revealed has been successfully replicated in subsequent studies. In contrast, the ACE has not been supported by recent studies, and the results of the meta-analysis and Bayesian analysis suggest the instability of this effect.

Before considering the reasons for the ACE instability, it is necessary to introduce the phrase “sentence focus” and its related research. Sentence focus refers to the most emphasized or relevant information in a sentence, while other information serves as the background to the focus (Birch & Rayner, 1997; Káldi & Babarczy, 2021). Some studies have found that readers allocate more attentional resources and engage in deeper processing for focused information during language processing, leading to more detailed semantic representations of the focused information. In contrast, non-focused information in a sentence, which occupies a secondary position, usually receives fewer attentional resources, resulting in shallower processing and coarse representation (Birch & Rayner, 1997; Lowder & Gordon, 2015; Yang et al., 2018). For example, in a change detection task, it was found that when a change occurs in nonfocused information during the second reading compared with when a change occurs in focused information, participants have significantly lower detection rates (Sanford et al., 2006, 2009; Sturt et al., 2004). This finding indicates that the representation of non-focused information is relatively coarse to the extent that participants may not notice the change when it occurs.

Thus, we can conclude that (1) according to the “weak embodiment” theory, the simulation system is necessary only when relevant information is deeply processed, and (2) the depth of information processing is influenced by sentence focus, with only the focused information receiving deep processing. Based on this, we can speculate that only the focused information would recruit mental simulation processes in language comprehension. Therefore, the instability of ACE may be related to the instability of the sentence focus (i.e., interindividual variability in the focused information). Specifically, the focus of sentences

used in the ASCP is not explicit. When participants read a sentence such as “*You kicked the football to Jack,*” some participants identify the action information (“*kicked something to Jack*”) as the sentence focus, leading to deep processing of the action information and the activation of motion direction feature. Conversely, other participants may identify the object noun (“*football,*” emphasizing what was kicked to Jack) as the sentence focus. Although previous research has indicated that noun processing may also activate related motor experiences (Carota et al., 2012; Gough et al., 2012). However, nouns do not imply any specific directionality, while the sentences in the ASCP do. Considering that ACE primarily arises from the consistency between the movement direction being activated in sentence comprehension and the actual action direction, when the action information cannot consistently become the sentence focus, the movement-direction features of the sentence cannot be steadily activated, resulting in ACE not being steadily observed. However, the focus of sentences used in the SPVT is more explicit, and the object in the picture always corresponds to the object noun in the sentence. This leads to consistency among participants in identifying the sentence focus and more stable results for the shape match effect.

The present study used a pretest and three formal experiments to validate our speculation that sentence focus influences the mental simulation process in language comprehension and that the instability of sentence focus may be one of the reasons for the instability of the ACE. The pretest first evaluated the focused information of the sentences used in previous studies through a questionnaire, which served as the foundation for the entire study. Experiments 1, 2, and 3 examined the effect of sentence focus on the shape match effect and ACE by manipulating the focused information of sentences. We hypothesized that (1) in the SPVT, the object nouns in the sentences (e.g., “*banana*” in “*Jerome saw the banana on the dessert.*”) would be stably recognized as the focus, while the focus of sentences used in the ASCP would show high variability between individuals; (2) in the SPVT, the shape match effect would disappear when the subject noun was manipulated to be the focus; and (3) in the ASCP, the ACE would appear when the action information was manipulated to be the focus.

Pretest: Evaluation of the focused information of the sentence used in the previous literature

Purpose

Through a questionnaire, this pretest aimed to demonstrate the stability of object nouns as the focused information of

sentences used in the SPVT and the instability of action-related words as the focused information of sentences used in the ASCP. Previous research has considered it-cleft sentences as sentence structures with highly explicit focused information (Birch & Rayner, 1997). Therefore, it-cleft sentences were included in the questionnaire to examine the validity of the focus evaluation method used in the present experiment.

Method

Participants

The sentences in the questionnaire were initially presented in English. Therefore, only proficient English speakers were recruited. Thirty female participants majoring in English were recruited from a university in Tianjin, China. Their ages ranged from 22 to 27 years, and all participants had an English proficiency level of English Majors-Band 8 or higher.

Research instrument

A self-designed sentence focus evaluation questionnaire was used that consisted of 63 sentences. Twenty-eight sentences were derived from studies using the SPVT (Zwaan et al., 2002), such as “*Jerome saw the banana on the dessert.*” Another 28 sentences were derived from studies using the ASCP (Glenberg & Kaschak, 2002; Papesh, 2015), such as “*You kicked the football to Jack.*” Additionally, the questionnaire included seven it-cleft sentences, such as “*It is the note that you slipped Heather.*” (with the focus on “*note*”). In the questionnaire, sentences from studies using the SPVT and studies using the ASCP alternated, and it-cleft sentences were inserted into the questionnaire at intervals of eight sentences. Furthermore, two questionnaire versions were created with the same sentences but with reversed presentation orders to balance order effects.

Procedure and scoring

The questionnaire was administered online. Participants were instructed to read each sentence carefully and select the part of the sentence that they believed emphasized or was most important for understanding the sentence (Osaka et al., 2002). Completing the questionnaire took approximately 5–10 minutes. Scoring followed the method used in previous research (Osaka et al., 2002). For each sentence, the percentage of participants who selected the action-related word or object noun as the focus was calculated. For the sentences used in the ASCP, such as “*You kicked the football to*

Jack.”, the proportion of participants who selected “*kicked*” and “*football*” as the focus was calculated. For the sentences used in the SPVT, such as “*Jerome saw the banana on the dessert.*”, the proportion of participants who selected “*banana*” as the focus was calculated. For it-cleft sentences, such as “*It is the note that you slipped Heather.*”, the proportion of participants who selected “*note*” as the focus was calculated. Then, the average selection rates for different types of sentences were calculated. Because previous research has considered 70% the threshold for determining sentence focus (Osaka et al., 2002), chi-square tests were conducted to statistically compare the times participants selected each type of word with the times corresponding to 70% of the threshold for each type of sentence.

Results and discussion

For it-cleft sentences, the average selection rate for the focused noun was 86.19%. Chi-square tests revealed a significant difference, $\chi^2(1) = 15.16, p < .001$. This result indicates that participants selected the focused noun significantly more often than the 70% threshold, supporting the validity of the focus evaluation method used in the present experiment.

For the sentences used in the SPVT, the average selection rate for the object noun was 80.60%. Chi-square tests revealed a significant difference, $\chi^2(1) = 24.78, p < .001$. This result indicates that participants selected the object noun significantly more often than the 70% threshold, suggesting that the focused information in these sentences was explicit and consistent among the participants.

For the sentences used in the ASCP, the average selection rate for the action-related word was 24.29%, while the average selection rate for the object noun was 41.55%. Chi-square tests revealed that the selection rates for both the action-related word and the object noun were significantly lower than 70%—action-related word: $\chi^2(1) = 350.4, p < .001$; object noun: $\chi^2(1) = 136.69, p < .001$. This result indicates that the focused information in these sentences was vague and varied between individuals.

In summary, this pretest first demonstrated the effectiveness of the focus evaluation method by it-cleft sentences. Similar to it-cleft sentences, the sentences used in the SPVT had explicit focused information, and participants consistently identified their focused information as object nouns. In contrast, the sentences used in the ASCP exhibited inconsistency in focus among participants. Some participants recognized the action-related word for the same sentences as the focus, while others identified other parts of the sentences, mostly the object noun, as the focus. These findings support the hypothesis proposed in the introduction, suggesting that the sentences used in the ASCP do not exhibit explicit and stable focused information. Instead, the focused information of these sentences

shows significant interindividual variability. Is the instability of the ACE related to the instability of sentence focus found in the present experiment? We hypothesize that if the instability of the ACE is due to the instability of sentence focus, then manipulating the action information as the focus should result in the ACE in the ASCP, and manipulating the subject noun as the focus should eliminate the shape match effect in the SPVT.

Experiment 1: The influence of sentence focus on the shape match effect—Manipulating by focus marker word “是” (*is*)

Purpose

The SPVT was employed to investigate the influence of sentence focus on the shape match effect. Using the focus-marking word “是” (*is*), the sentence focus was shifted from the object noun to the subject noun to explore whether the shape match effect disappeared under this condition.

Method

Participants

The prior sample size was determined using G*Power software, with statistical significance at $\alpha = 0.05$, a power of $1 - \beta = 0.95$, and a moderate effect size of $f = 0.25$. The result indicated that 36 participants would be required. Finally, 41 participants were recruited from a university in Tianjin, China. One participant reported being affected by external noise during the experiment, and her data were excluded. The final analysis included data from forty participants (32 females, age range: 18–25 years).

Experimental design

A 2 (sentence type: object noun-focused sentence, subject noun-focused sentence) \times 2 (match: match, mismatch) within-subject design was used with the dependent variable of response time for picture judgment.

Materials

There were 40 sets of experimental sentence–picture pairs and 160 filler sentence–picture pairs. The experimental sentence–picture pairs were partly from previous studies and partly self-created. Each set of experimental sentence–picture pairs consisted of four types of sentences (two shapes



\times two focus conditions) and two pictures corresponding to the implied shape of objects in the sentences. Therefore, there were eight combinations of sentence–picture pairs (see Table 1). Previous studies consistently used personal names (e.g., John, Tom) as the subjects of the sentences. However, the subjects were modified to be occupational nouns (e.g., police officer, manager) in the current experiment. Furthermore, in previous studies, the objects presented in the pictures always corresponded to the object noun of the experimental sentences (e.g., “banana” in “Jerome saw the banana on the dessert.”), which attracted more attention from the participants and made it the focus of the sentence. This would have interfered with our focus manipulation. Therefore, in the present experiment, additional filler materials (40 sentence–picture pairs) were added so that the subject noun could also be the content of the pictures. However, because personal names are difficult to represent with pictures while occupational nouns are relatively feasible, the subjects of the experimental sentences were modified in the present experiment.

The sentence focus was manipulated by the focus marker word “是” (*is*). Previous linguistic studies have suggested that focus marker words (or focus-sensitive particles) are an essential means to mark focused information, such as “有” (*have*), “是” (*is*), and “连” (*even*) in Chinese (Tang, 2020). Among them, “是” (*is*) is the most prominent focus marker word in Modern Chinese, and its essential role in marking the focus cannot be replaced by any other word (Liu, 2013). In Chinese, if it is necessary to emphasize a particular element in a sentence, the focus marker word “是” (*is*) can be added before it (Jiang, 2012). Therefore, in the present experiment, the subject noun of the original sentence was placed at the end of the sentence, and the focus marker word “是” (*is*) was placed before the subject noun to make it the focus of the sentence (similar to it-cleft sentence in English). Because the pretest results showed that the participants could stably identify the object nouns as the focus for the original sentence, the original sentences were used for the object noun-focused condition.

Among the 160 filler sentence–picture pairs, 40 consisted of sentences with the same structure as the experimental sentences, but with the subject noun as the picture content. In addition, 20 sentences with different structures from the experimental sentences were added to prevent participants from perceiving the manipulation of the sentence focus. In the 40 sets of experimental sentence–picture pairs and 60 filler sentence–picture pairs, the objects in the pictures were mentioned in the sentences, requiring participants to provide a “YES” response. To balance the “YES” and “NO” responses, 100 sentence–picture pairs were included where the objects in the pictures were not mentioned in the sentences.

Based on the above materials, eight experimental lists were created, each containing 40 experimental

Table 1 Example of sentence–picture pair for each combination (Experiment 1)

Match	Sentence type	Sentence	Picture
Match	Object noun-focused	牧民看到天空中有一只鹰 The herdsman saw the eagle in the sky	
	Subject noun-focused	看到天空中有一只鹰的是牧民 It's the herdsman who saw the eagle in the sky	
Mismatch	Object noun-focused	牧民看到巢穴里有一只鹰 The herdsman saw the eagle in the nest	
	Subject noun-focused	看到巢穴里有一只鹰的是牧民 It's the herdsman who saw the eagle in the nest	
Match	Object noun-focused	牧民看到巢穴里有一只鹰 The herdsman saw the eagle in the nest	
	Subject noun-focused	看到巢穴里有一只鹰的是牧民 It's the herdsman who saw the eagle in the nest	
Mismatch	Object noun-focused	牧民看到天空中有一只鹰 The herdsman saw the eagle in the sky	
	Subject noun-focused	看到天空中有一只鹰的是牧民 It's the herdsman who saw the eagle in the sky	

The sentences were presented in Chinese in the experiment

sentence–picture pairs and 160 filler sentence–picture pairs. For the experimental sentence–picture pairs, each list included only one type of combination from each set, and the types of combinations were counterbalanced in the experimental lists. The filler sentence–picture pairs were consistent across different lists.

Procedure

The experiment was programmed using E-Prime 3.0. In the formal experiment, a fixation was presented; the fixation disappeared after 500 ms, followed by the automatic presentation of a sentence. Participants were instructed to carefully read the sentence and press the space bar when they understood the meaning of the sentence. Then, another fixation cross was presented for 500 ms, followed by the automatic presentation of a picture. Participants were instructed to judge whether the object in the picture had been mentioned in the sentence they had just read as quickly and accurately as possible. They were instructed to press the “YES” button if the object was mentioned and the “NO” button if it was not. After providing a response, the participants proceeded to the subsequent trial until the end of the experiment. The formal experiment consisted of 200 trials presented in a pseudorandom order to ensure that the same conditions did not appear consecutively. The experiment was divided into four blocks, each consisting of 50 trials. The participants

could take self-paced breaks between blocks. To ensure that the participants read the sentences carefully, they were informed that a recall test would be conducted after the formal experiment. The recall test consisted of 48 sentences, with 24 sentences appearing in the formal experiment and 24 not appearing. The participants were asked to judge whether each sentence had been read in the formal experiment. The experiment took 20~25 minutes.

Before the formal experiment, the participants completed six practice trials. The practice procedure was identical to the formal experiment, but the materials used did not appear in the formal experiment.

Data analysis

Consistent with previous studies, the dependent variable in the present experiment was the response time (RT) for picture judgment. First, the response times were log-transformed, and then a linear mixed model (LMM) was established using the *lme4* package (Bates et al., 2015). The LMM included the sentence type, match, and their interaction as fixed effects and the participant and item as random effects, with the RT for picture judgment as the dependent variable. In model construction, the model fitting started from the maximal random effects, and if the complex model failed to converge, it was gradually simplified until convergence was achieved. If the interaction effect was significant, simple

effects analysis was conducted using the *lsmeans* package (Lenth, 2016). This analysis method was chosen because LMM can simultaneously consider the variation of participants and items, resulting in more accurate and stable results than an analysis of variance (ANOVA; Solana & Santiago, 2022).

Results and discussion

The average accuracy of the participants in the picture judgment was 97.34%, and the average accuracy in sentence recall was 80.73%, indicating that the participants performed the task according to the instructions. Before the statistical analysis, the following trials were removed: (1) trials with incorrect responses (accounting for 5.44% of all trials); (2) trials with RTs exceeding 2.5 standard deviations from the mean (accounting for 1.56% of all trials); and (3) trials with response times greater than 3,000 ms or less than 300 ms (accounting for 0.94% of all trials; Engelen et al., 2011).

The LMM¹ for RT revealed a significant interaction between sentence type and match, $b = 0.019$, $SE = 0.008$, $t = 2.28$, $p = .023$, 95% CI [0.003, 0.034]. Simple effects analysis revealed that in the object noun-focused condition, participants responded significantly faster when the picture matched than when it did not match the shape implied by the sentence, $b = -0.095$, $SE = 0.023$, $t = -4.19$, $p < .001$, 95% CI [-0.140, -0.051]. However, there was no significant difference between the match and mismatch conditions in the subject noun-focused sentence, $b = -0.021$, $SE = 0.023$, $t = -0.92$, $p = .359$, 95% CI [-0.067, 0.024]. In addition, the main effect of the match was significant, $b = -0.029$, $SE = 0.008$, $t = -3.59$, $p < .001$, 95% CI [-0.045, -0.013], but this effect was mediated by sentence type. The main effect of sentence type was not significant, $b = 0.006$, $SE = 0.008$, $t = 0.75$, $p = .461$, 95% CI [-0.010, 0.023]. The descriptive results for the match and sentence conditions are illustrated in Fig. 1.

In general, the interaction between the sentence types and the match supports the hypothesis proposed in the introduction, providing evidence for the influence of sentence focus on the shape match effect and mental simulation. Specifically, the shape match effect disappears when the object noun is not the sentence focus. Furthermore, the present experiment replicates previous findings in the object noun-focused sentence (original sentence), demonstrating the stability and replicability of this effect. Based on the results of Experiment 1, we speculate that if the instability of the ACE arises from the instability of the sentence focus in the ASCP, then the ACE can be found by marking the action information as the focus using the focus marker word “是” (*is*). We conducted Experiment 2 to test this hypothesis.

¹ The final model, in lme4's model notation, was $\text{prt} \sim \text{focus} * \text{match} + (1 + \text{focus} | \text{participant}) + (1 | \text{item})$.

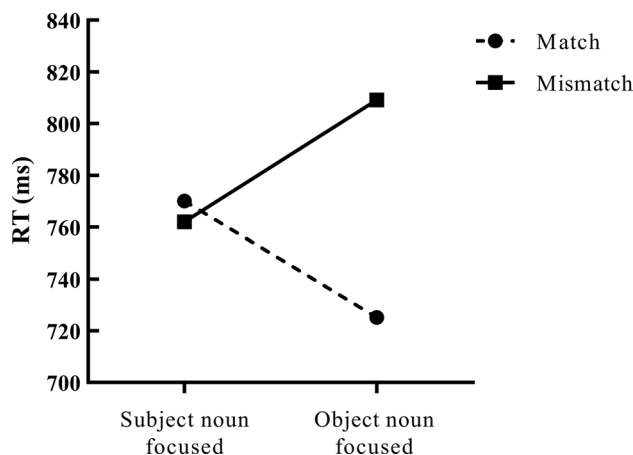


Fig. 1 The descriptive results for each experimental condition in the Experiment 1

Experiment 2: The influence of sentence focus on the ACE—Manipulating by focus marker word “是” (*is*)

Purpose

The impact of sentence focus on motor simulation was examined using the ASCP by manipulating the sentence focus to either action or nonaction information using focus-marking words “是” (*is*). This investigation explored whether the ACE was observed when action information became the sentence focus.

Method

Participants

Similar to Experiment 1, at least 36 participants would be required in Experiment 2. Finally, 48 participants were recruited from a university in Tianjin, China. Among them, two participants did not complete the entire experiment due to technical issues, three had an accuracy rate lower than 80% in the sensibility judgment task, and three did not follow the experimental instructions. The data of these participants were discarded. Therefore, the final analysis included data from forty participants (32 females, between 17 and 25 years of age).

Experimental design

A 2 (sentence type: action-focused sentence, nonaction-focused sentence) \times 2 (consistency between direction implied by sentence and response direction: consistent, inconsistent) within-subject design was employed, and the dependent variable was the reading time for the sentence.

Table 2 Example of sentence for each combination (Experiment 2)

Sentence type × Direction	Example
Action-focused sentence—toward	彤彤把一块披萨是送给你了 A pizza was given to you by Tongtong
Action-focused sentence—away	你把一块披萨是送给彤彤了 A pizza was given to Tongtong by you
Nonaction-focused sentence—toward	彤彤送给你的是一块披萨 What was given to you by Tongtong is a piece of pizza
Nonaction-focused sentence—away	你送给彤彤的是一块披萨 What was given to Tongtong by you is a piece of pizza

The sentences were presented in Chinese in the experiment

Materials

There were 48 sets of experimental sentences and 48 filler sentences. The experimental sentences were partly derived from previous studies (Glenberg & Kaschak, 2002; Papesh, 2015) and partly self-created. All experimental sentences were sensible, and there were four versions for each set of experimental sentences (2 sentence type × 2 direction): action-focused—toward sentence, action-focused—away sentence, nonaction-focused—toward sentence, and nonaction-focused—away sentence. The direction of the sentences was implied by the combination of the subject, predicate verb, and object complement. Moreover, the sentence focus was manipulated by the focus marker word “是” (*is*; examples of each condition are shown in Table 2).

The filler sentences consisted of three types of nonsense sentences: inappropriate verb-object collocations (e.g., *Tom scrubs an honor to you*), inappropriate subjects (e.g., *The sun throws a bottle to you*), and inappropriate quantifiers (e.g., *Tony gives you a meal of roses*). This was intended to encourage the participants to read the sentence carefully. Additionally, to ensure consistency between the experimental and filler sentences, the same experimental manipulations were applied to the filler sentences, resulting in four types of filler sentences corresponding to the experimental sentences.

Four lists were created based on the materials mentioned above, each containing 48 experimental sentences and 48 filler sentences. For the experimental sentences, each list included only one version of each set of sentences. The sentence versions were counterbalanced in the lists, with 12 action-focused—toward sentences, 12 action-focused—away sentences, 12 nonaction-focused—toward sentences, and 12 nonaction-focused—away sentences. The filler sentences in the different lists were identical.

Procedure

The experiment was programmed using E-Prime 3.0. In the formal experiment, a fixation point was first presented in the center of the screen. After the participants pressed the “START”

button, the fixation disappeared, and the sentence was presented. The participants were required to carefully read the sentence and judge its sensibility as quickly as possible. The participants needed to hold the “START” button during this process; otherwise, the sentence would disappear. After reading the sentence, the participants released the “START” button and then pressed either the “YES” or “NO” button to respond (“YES” if they thought the sentence was sensible and “NO” otherwise). After providing a response, the participants proceeded to the subsequent trial until the end of the experiment. The formal experiment consisted of 96 trials, with all sentences presented randomly. The entire experiment took approximately 10 minutes.

The participants responded with an external keyboard. The keyboard was placed next to the laptop with a 90-degree counterclockwise rotation, and the keys “A,” “G,” and “L” were arranged vertically. The vertical distances between “A” and “G” and between “G” and “L” were equal, with “A” closer to the participants and “L” farther away. Additionally, stickers were used to replace “G” with “START,” and “A” and “L” were replaced with “YES” or “NO.” The positions of the “YES” and “NO” keys were counterbalanced between the participants. For example, for half of the participants the “A” key was “YES” and the “L” key was “NO,” while these were reversed for the other half of the participants (see Fig. 2).

Before the formal experiment, the participants completed two practice sessions. The first session included 20 trials and aimed to familiarize the participants with the button configuration. The word “YES” or “NO” was presented in this stage, and the participants were asked to press the corresponding button. After each response, feedback was presented to the participants. The second practice session consisted of eight trials and aimed to familiarize the participants with the procedure of the formal experiment. The second practice followed the same procedure as the formal experiment, but the materials used in this practice did not appear in the formal experiment.

Data analysis

Consistent with previous research, the dependent variable in the present experiment was the sentence reading time,

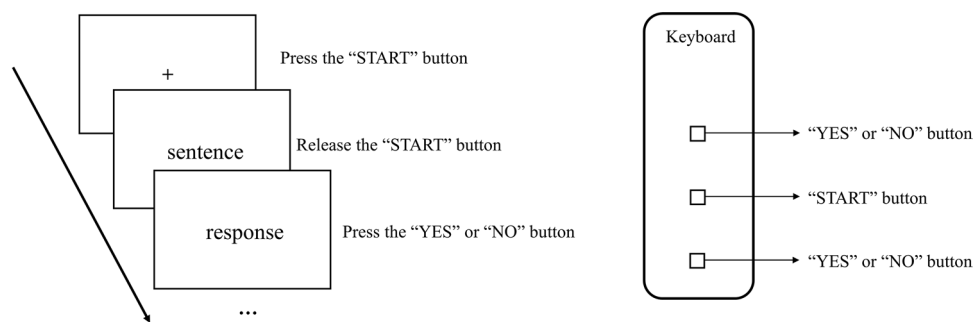


Fig. 2 The general structure of a trial (left panel) and the button configuration (right panel) for Experiment 2

which referred to the latency between the sentence onset and the participant releasing the “START” button (Glenberg & Kaschak, 2002). First, the reading times were log-transformed, and then a linear mixed model (LMM) was established using the *lme4* package (Bates et al., 2015). The LMM included the sentence type, consistency, and their interaction as fixed effects, the participant and item as random effects, and the reading time as the dependent variable. In model construction, the model fitting started from the maximal random effects, and if the complex model failed to converge, it was gradually simplified until convergence was achieved. If the interaction was significant, a simple effects analysis was conducted using the *lsmeans* package (Lenth, 2016). In addition, the current experiment did not observe ACE in both sentence types. In order to exclude the possibility that the null results stemmed from low statistical power, an additional Bayesian analysis was conducted. This analysis was performed using the *BayesFactor* package (Morey et al., 2015). Referring to the approach of Yao et al. (2022), we calculated the Bayesian factor of the interaction between sentence direction and response direction in action-focused sentences and nonaction-focused sentences, respectively. The Bayesian factor of the interaction was the ratio of the model “including sentence direction, response direction, interaction, participant intercept, and item intercept” to the model “only including sentence direction, response direction, participant intercept, and item intercepts.”

Results and discussion

The average accuracy of the sensibility judgment task was 91.61%, indicating that participants performed the task according to the instructions. Before statistical analysis, the following trials were removed: (1) trials with incorrect responses (accounting for 1.77% of all trials) and (2) trials exceeding 2.5 standard deviations from the mean reading time (accounting for 1.41% of all trials). Furthermore, (3) in previous research, participants were required to read the sentences and respond as quickly as possible, and the maximum presentation time for each sentence was set at 3,000 ms

(Glenberg & Kaschak, 2002). Therefore, trials with reading times exceeding 3,000 ms were also excluded (accounting for 4.79% of all trials).

The LMM² for reading time showed that the interaction between sentence type and consistency was not significant, $b = -0.002$, $SE = 0.006$, $t = -0.39$, $p = .699$, 95% CI $[-0.013, 0.009]$. Additionally, the main effect of consistency was not significant, $b = -0.001$, $SE = 0.006$, $t = -0.25$, $p = .807$, 95% CI $[-0.013, 0.010]$. This result indicates that no significant ACE was observed in either the action-focused or nonaction-focused sentences. However, the main effect of sentence type was significant, $b = 0.016$, $SE = 0.006$, $t = 2.81$, $p = .005$, 95% CI $[0.005, 0.027]$, and the reading time was significantly longer for action-focused sentences than for nonaction-focused sentences.

Furthermore, in previous literature, ACE has been defined as the interaction between sentence direction and response direction (Glenberg & Kaschak, 2002). In order to maintain consistency with the previous studies, an LMM analysis was conducted with sentence type, sentence direction, response direction, and their interactions as predictor variables. The results³ revealed that the interaction of the three factors was not significant, $b = -0.002$, $SE = 0.006$, $t = -0.33$, $p = .746$, 95% CI $[-0.013, 0.009]$. More importantly, the interaction between sentence direction and response direction was also not significant, $b = -0.002$, $SE = 0.006$, $t = -0.38$, $p = .708$, 95% CI $[-0.014, 0.009]$, once again indicating no significant ACE was observed in either the action-focused or nonaction-focused sentences. Additionally, the main effect of sentence type was observed again, $b = 0.015$, $SE = 0.006$, $t = 2.54$, $p = .011$, 95% CI $[0.003, 0.026]$, with reading times for action-focused sentences significantly longer than for nonaction-focused sentences. No other effects were significant (all $ps > .190$). The descriptive results for each condition are presented in Fig. 3.

² The final model, in *lme4*’s model notation, was $rt \sim \text{focus} * \text{consistency} + (1 + \text{consistency} | \text{participant}) + (1 | \text{item})$.

³ The final model, in *lme4*’s model notation, was $rt \sim \text{focus} * \text{sdirection} * \text{rdirection} + (1 + \text{sdirection} | \text{participant}) + (1 | \text{item})$.

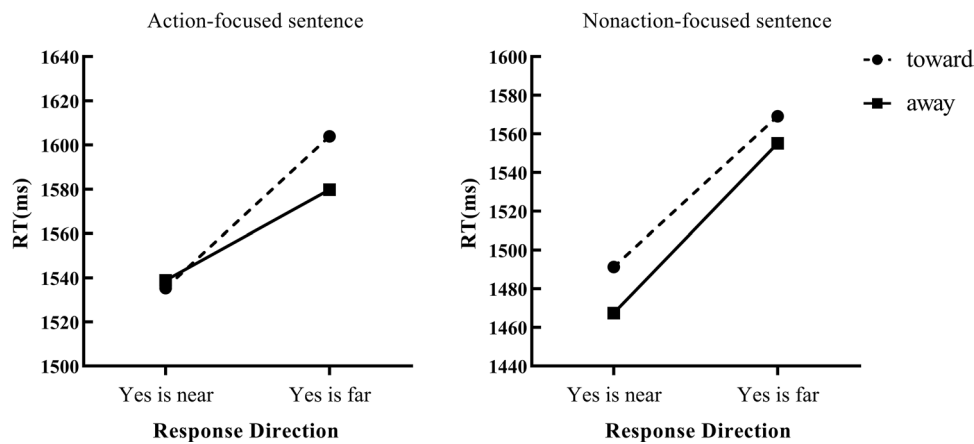


Fig. 3 The descriptive results for each experimental condition in the Experiment 2

Experiment 2 found that the sentence reading time was significantly longer for action-focused sentences. This may be because action-focused sentences generally have more words than nonaction-focused sentences (e.g., “彤彤送给你的是一块披萨” vs. “彤彤把一块披萨是送给你了”). Furthermore, the ACE was not found in either condition. We consider that this negative result can be explained in two ways. Firstly, it may stem from low statistical power (Button et al., 2013). Despite the priori sample size calculation conducted in the present study, Kumle et al. (2021) suggested that G*Power is not suitable for the LMM analysis employed in the present study, as it does not allow us to consider the variability attributable to the random factors of the design. Additionally, because the present study is novel (no previous research has focused on the interaction between sentence focus and ACE), the effect size was estimated to be moderate ($f = 0.25$), as in some previous studies (e.g., Matsumoto et al., 2022). These two factors may have contributed to insufficient power to detect minor effects in the current experiment, resulting in a null result. Secondly, it may be attributed to an inappropriate focus manipulation method, specifically that the focus manipulation method used in the current experiment was not suitable for the sentence used in the ASCP in Chinese. There is evidence to support this hypothesis. First, in the second practice session, most participants specifically noticed the presence of the focus marker word “是” (*is*) and asked the experimenter whether these sentences were considered nonsense because they thought that “是” (*is*) made the sentence less fluent. We explained to the participants that this task required them to judge whether the meaning expressed by the sentence was sensible without considering the fluency of the sentence. However, some participants still reported being disturbed by the fluency of the sentences during the experiment. This indicates some

issues with the focus manipulation method used in the present experiment. Second, we conducted a postexperiment assessment of the sentence focus. We selected 12 action-focused sentences and 12 nonaction-focused sentences and recruited 33 participants. The results showed that an average of 18.43% of the participants identified the focused information as the verb for the nonaction-focused sentences, and 46.72% identified the focused information as the action-focused sentences. Although the latter percentage was relatively higher, it was significantly lower than the 70% threshold in previous studies, $\chi^2(1) = 7.92, p = .005$. Taken together, these results suggest that the focus marker word “是” (*is*) may not be suitable for sentences used in the ASCP. It did not achieve the desired effect of manipulating the focus, that is, consistently leading the participants to identify the action information as the sentence focus.

To investigate whether the null result of Experiment 2 stemmed from low statistical power, we conducted a Bayesian analysis. Some researchers have suggested that Bayesian analysis can be used to explore the degree to which the data supports the null hypothesis, thereby overcoming the limitation of null hypothesis significance testing (NHST), which cannot distinguish whether the null result comes from the absence of the effect or the low statistical power (Wagenmakers et al., 2011, 2018). The Bayesian analysis results revealed that the Bayesian factor (BF_{10}) for the interaction between sentence direction and response direction was 0.12 in action-focused sentences and it was 0.08 in nonaction-focused sentences. According to the categorization of BF_{10} by Wagenmakers et al. (2018), these results indicate moderate evidence to support the null hypothesis under the action-focused condition and strong evidence under the nonaction-focused condition. Based on these results, we are inclined to infer that the null results of Experiment 2 might have stemmed from inappropriate

focus manipulation. Therefore, in Experiment 3, we altered the focus manipulation method to investigate the impact of sentence focus on the ACE.

Experiment 3: The influence of sentence focus on the ACE—Manipulation by external markers

Purpose

Experiment 1 revealed that the focus marker word “是” (*is*) influenced the shape match effect. That is, marking the subject noun as the sentence focus led to the disappearance of the shape match effect. However, in Experiment 2, using the focus marker word “是” (*is*) to mark the action information failed to elicit the ACE. Based on the participants' performance during the experiment and the post hoc questionnaire, we speculate that the negative results in Experiment 2 were due to the ineffectiveness of the focus manipulation. Therefore, in Experiment 3, we manipulated sentence focus using external markers (italics, underlining, etc.) to explore the influence of sentence focus on the ACE. Previous research has demonstrated the effectiveness of this approach in manipulating sentence focus (Sanford et al., 2006; Káldi & Babarczy, 2021). Additionally, this manipulation method does not alter word order, fluency, or sentence length, ensuring consistency between the current experiment and the original paradigm and overcoming the shortcoming of manipulating sentence focus using the focus marker word “是” (*is*).

Method

Participants

Similar to Experiment 1, G*Power analysis indicated that 36 participants would be required. However, considering the requirement to balance four material lists, four presentation orders (see Materials for more details), and two response configurations across participants, the present experiment planned to recruit 32 participants ($1 * 32$). Finally, 35 participants were recruited from a university in Tianjin, China. One participant had an accuracy rate below 80% in the sensibility judgment task, and two participants did not follow the task instructions. The data from these participants were discarded. Therefore, the final analysis included data from 32 participants (27 females, between 18 and 27 years of age).

Experimental design

A 2 (sentence type: action-focused sentence, nonaction-focused sentence) \times 2 (consistency between sentence

direction and response direction: consistent, inconsistent) within-subject design was employed. The dependent variable was the reading time for the sentence.

Materials

The experimental sentences in Experiment 3 were similar to those in Experiment 2, with the only difference being the focused manipulation. Referring to the micromarker that emphasizes specific text content in Chinese reading research (He & Mo, 2002; Song et al., 2011), the current experiment marked the focused information by changing the font, font size, bolding, and underlining. Specifically, the focused information was presented in 12 pt SimHei bold and underlined font, while nonfocused information was presented in 10 pt SimSun font. Previous eye-tracking studies have demonstrated that marked content is processed more deeply by participants, as evidenced by longer gaze durations, higher fixation frequency, larger pupil diameters, and shorter saccade distances on marked information (Wang et al., 2004), consistent with the function of sentence focus. Therefore, the subject (e.g., “you” in “you kicked the football to Jack.”), predicate verb, and object complement (e.g., “to Jack”) of action-focused sentences were all marked by the above marking mean. This is because we believe that the direction of the sentence in the ASCP is not determined solely by the predicate verb but by the combination of these three components. In contrast, no content was marked in nonaction-focused sentences, as the pilot study demonstrated that the focus of unmarked sentences was not action-related information. Examples of sentences for each condition are shown in Table 3.

To prevent the participants from detecting the experimental purpose, the present experiment increased the number of filler sentences to 102. There were 27 sensible sentences, all with different structures from the experimental sentences. There were nine sentences with external markers in the middle of the sentence, nine with external markers at the end of the sentence, and nine without markers. The remaining 75 filler sentences were all nonsense sentences. Among them, 24 had the same structure as the experimental sentences: six had external markers at the beginning of the sentence, six had external markers at the end, and 12 did not have markers. Additionally, there were 51 nonsense sentences with different structures from the experimental sentences: 17 had external markers in the middle of the sentence, 17 had external markers at the end, and 17 did not have markers.

Similar to Experiment 2, four lists were created based on the above materials. Each list comprised 48 experimental and 102 filler sentences that were evenly distributed between sensible and nonsense sentences. In addition, each list contains four different presentation orders. For each order, the condition to which the first experimental sentence read by

Table 3 Example of sentence for each combination (Experiment 3)

Sentence type × Direction	Example
Action-focused sentence—toward	彤彤送给你一块披萨 Tongtong gave a pizza to you
Action-focused sentence—away	你送给彤彤一块披萨 You gave a pizza to Tongtong
Nonaction-focused sentence—toward	彤彤送给你一块披萨 Tongtong gave a pizza to you
Nonaction-focused sentence—away	你送给彤彤一块披萨 You gave a pizza to Tongtong

The sentences were presented in Chinese in the experiment. Furthermore, due to the grammatical differences between Chinese and English, the three sentence components marked as focused are separated in English, whereas they are continuous in Chinese

the participant belongs was manipulated, while the other sentences were presented randomly (e.g., Participant 1 first read the action-focus—away sentence, Participant 2 first read the action-focus—toward sentence, Participant 3 first read the nonaction-focus—away sentence, Participant 4 first read the nonaction-focus—toward sentence).

Procedure

The procedure was similar to Experiment 2, except the sentences were presented in a pseudorandom order. To avoid the influence of order effects, the conditions of the first experimental sentence were counterbalanced across participants.

Data analysis

The same analysis plan used for Experiment 2 was used for Experiment 3.

Results and discussion

The average accuracy of the sensibility judgment task was 90.7%, indicating that participants performed the task according to the instructions. Before statistical analysis, the following trials were removed: (1) trials that were affected by external factors during the experiment (due to keying errors from the previous trial, which disrupted the reading of subsequent trials, accounting for 0.13% of all trials); (2) trials with incorrect responses (accounting for 8.46% of all trials); (3) trials exceeding 2.5 standard deviations from the mean reading time (accounting for 1.30% of all trials); and (4) trials with reading times exceeding 3,000 ms (accounting for 3.91% of all trials).

The LMM⁴ for reading time revealed a significant interaction between sentence type and consistency, $b = -0.021$,

⁴ The final model, in lme4's model notation, was $rt \sim \text{focus} * \text{consistency} + (1 | \text{participant}) + (1 | \text{item})$.

$SE = 0.008$, $t = -2.65$, $p = .008$, 95% CI $[-0.037, -0.005]$. Simple effects analysis revealed that in the action-focused sentence, the reading time of participants was significantly shorter in the condition where the direction of the response was consistent with the direction implied by the sentence compared with the condition where they were inconsistent, $b = -0.053$, $SE = 0.023$, $t = -2.32$, $p = .021$, 95% CI $[-0.097, -0.008]$. However, there was no significant difference in the nonaction-focused sentence, $b = 0.032$, $SE = 0.022$, $t = 1.42$, $p = .156$, 95% CI $[-0.012, 0.076]$. In addition, the main effect of the sentence type was significant, $b = 0.031$, $SE = 0.008$, $t = 3.89$, $p < .001$, 95% CI $[0.015, 0.047]$, and the reading time for action-focused sentences was significantly longer than for nonfocused sentences. The main effect of consistency was not significant, $b = -0.005$, $SE = 0.008$, $t = -0.65$, $p = .515$, 95% CI $[-0.021, 0.010]$.

Similar to Experiment 2, a further LMM analysis⁵ was conducted with sentence type, sentence direction, response direction, and their interactions as predictor variables. The results revealed that all main effects were non-significant (all $ps > .060$), but the interaction of the three factors was significant, $b = -0.021$, $SE = 0.008$, $t = -2.60$, $p = .009$, 95% CI $[-0.036, -0.005]$. ACE was examined separately in action-focused and nonaction-focused sentences to better understand the nature of the three-way interaction. For action-focused sentences, the LMM analysis⁶ revealed that the interaction between sentence direction and response direction was significant, $b = -0.027$, $SE = 0.009$, $t = -2.91$, $p = .007$, 95% CI $[-0.045, -0.009]$, indicating that when the sentence direction and response direction were consistent, participants' reading times were significantly shorter than in the inconsistent condition (for "yes is near" condition, toward sentences vs. away sentences: $b = -0.058$, $SE = 0.026$, $t = -2.26$, $p = .032$, 95% CI $[-0.111, -0.006]$; for "yes is far" condition, toward sentences vs. away sentences: $b = 0.050$, $SE = 0.027$, $t = 1.85$, $p = .075$, 95% CI $[-0.005, 0.105]$). However, for nonaction-focused sentences, the interaction between sentence direction and response direction was not significant ($b = 0.016$, $SE = 0.015$, $t = 1.07$, $p = 0.294$, 95% CI $[-0.013, 0.045]$).⁷ These results indicate that ACE can be observed only in action-focused sentences. The descriptive results for each condition are presented in Fig. 4.

In conclusion, the present study demonstrated the impact of sentence focus on ACE, supporting the hypothesis proposed in the introduction. Specifically, the ACE was

⁵ The final model, in lme4's model notation, was $rt \sim \text{focus} * \text{sdirection} * \text{rdirection} + (1 | \text{participant}) + (1 | \text{item})$.

⁶ The final model, in lme4's model notation, was $rt \sim \text{sdirection} * \text{rdirection} + (1 + \text{sdirection} | \text{participant}) + (1 | \text{item})$.

⁷ The final model, in lme4's model notation, was $rt \sim \text{sdirection} * \text{rdirection} + (1 + \text{sdirection} | \text{participant}) + (1 | \text{item})$.

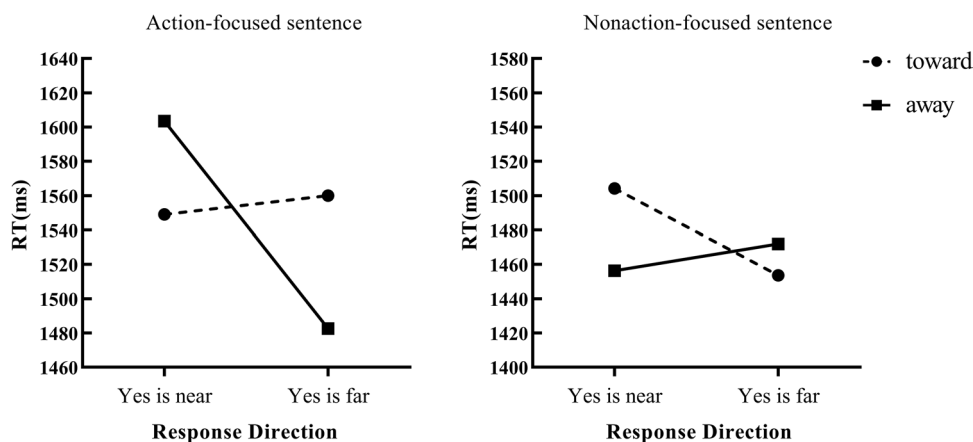


Fig. 4 The descriptive results for each experimental condition in the Experiment 3

observed only when action-related information was the sentence focus. In contrast, when there were no external markers in the sentence, the action-related information could not stably become the focus, making it difficult to find the ACE.

General discussion

The present study conducted a series of experiments to investigate the influence of sentence focus on mental simulation with the aim of uncovering the reasons for the replication failures of the ACE reported in recent studies. A pilot experiment revealed more significant interindividual variability for the focus of sentences used in the ASCP compared with it-cleft sentences and those used in the SPVT. Based upon this finding, Experiments 1 and 2 manipulated the sentence focus using the focus marker word “是” (*is*) to further explore how the shape match effect and the ACE are influenced by sentence focus. Experiment 1 found that the shape match effect occurred only in object noun-focused sentences (original sentences), while it disappeared when the sentence focus was changed to the subject noun. However, Experiment 2 found no ACE regardless of whether the action information was the sentence focus. Experiment 3 used an external marker (the font and font size of the focused information were modified, and the focused information was bolded and underlined) to manipulate sentence focus. The results revealed that the ACE was observed only when the action-related information was the focus. Taken together, these findings suggest that sentence focus may influence the generation of mental simulation. In language comprehension, only the focused information is mentally simulated. The instability of the ACE reported in previous studies is likely because action information does not usually become the sentence focus during reading comprehension.

Sentence focus affects mental simulation

The influence of sentence focus on the ACE or shape match effect in sentence comprehension may be related to its association with deeper semantic processing. The generation of the ACE or shape match effect is attributed to the activation of motion direction or object perceptual features during sentence comprehension. This simulation interacts with actual motion or perception processes, influencing participants' reaction times in the task (Glenberg & Kaschak, 2002; Stanfield & Zwaan, 2001). Moreover, as mentioned in the introduction, deep information processing is likely necessary for simulation generation. Therefore, the influence of sentence focus on the ACE and shape match effect may be explained by its impact on the depth of information processing, which further affects the occurrence of action/perception simulation. This assumption is consistent with the findings of studies on focused information processing. Qi (2012) proposed that focused information is the most prominent part of the discourse. He used a “*theater metaphor*” to illustrate the role of focus in the cognitive structure: “*From the linguistic perspective, verbal activities are like a play, and focus is the spotlight on the stage. Its function is to illuminate a certain part of the stage. The speaker is the director of this play, deciding where the light shines.*” In other words, the function of focus is to draw readers' attention to the most essential information in the language and to engage in deeper processing of focused information. Empirical studies on focused information processing also support this view. For example, eye-tracking and event-related potentials (ERPs) studies have found that people allocate more attention to focused information during reading and that focused information is encoded and integrated at a deeper level (Birch & Rayner, 1997; Káldi & Babarczy, 2021; Lowder & Gordon, 2015; Yang et al., 2018).

Another line of research that is highly relevant to our findings is the investigation of sentence focus using change detection tasks. A key finding from these studies is that under conditions of near semantic distance (i.e., when the information is replaced with semantically similar information), changes are more easily detected when the altered information is the sentence focus (Sanford et al., 2006, 2009; Sturt et al., 2004). Sturt et al. (2004) argued that this finding supports the concept of “good enough representation” proposed by Ferreira et al. (2002), which suggests that language processing does not always construct a rich and detailed semantic representation. Instead, people often rely on a set of rapid and heuristic strategies to understand the meaning of sentences, creating a “good enough representation” corresponding to the task demands. Based on their research, Sturt et al. (2004) proposed that the focus is also an important influencing factor that modulates the level of semantic representation. During language processing, focused information is more likely to be engaged in deeper processing, resulting in a more detailed semantic representation. Therefore, changes in the focused information are more easily detected. Importantly, Kaschak and Madden (2021) suggested that if readers employ only a “good enough representation” processing strategy during language comprehension, there may be no need to engage in mental simulation. Simulation processes are only engaged when a more detailed representation is required. Additionally, studies have found that mental simulation in sentence comprehension could facilitate the detection of picture changes (Holman & Gîrbă, 2019). Based on these research findings, we propose that the difference in the level of semantic representation between focused and nonfocused information, as observed by Sturt et al. (2004), may be attributed to the differences in mental simulation processes. The richer and more detailed semantic representation of focused information may correspond to the sensorimotor experiences activated during mental simulation.

In summary, the present study suggests that readers need to engage in deep processing and more detailed representation of action/perception information only when it is the sentence focus. This processing and representation require support from the simulation system, leading to the activation of motion direction or object shape features and the emergence of the ACE or shape match effect. Conversely, when the action/perception information is not the sentence focus, it may be shallowly processed and coarsely represented without the involvement of the simulation system. As a result, the motion direction features/object shape features would not be activated, and the ACE or shape match effect would not occur. Combined with the pilot experiment results, it can be inferred that the instability of the sentence focus is one of the reasons for the instability of ACE in the previous study. However, it is essential to claim that we did not consider the sentence focus to be

the exclusive cause, as replicability crises are not limited to the ACE but appear to affect the embodiment literature in general. In fact, some studies have not even used sentences as research material, and thus, the unstable results cannot be attributed to the sentence focus. For instance, in recent years, researchers have attempted to replicate important effects in the field of embodied cognition, such as the attentional spatial–numerical association of response codes (SNARC) effect (Fischer et al., 2003) and the motor interference effect on action verb memory (Shebani & Pulvermüller, 2013). Despite using larger sample sizes, similar to the replication studies of ACE, they encountered replication failures (Colling et al., 2020; Montero-Melis et al., 2022). Additionally, Saccone et al. (2021) also failed to replicate the motor interference effect in tool naming discovered by Witt et al. (2010). More importantly, Witt et al. (2020) reanalyzed their data after learning that their results could not be replicated, finding that their previous results were merely the by-product of the analysis pipeline. Furthermore, meta-analysis works on motor cortex stimulation studies during the comprehension of action language found that these studies are quite underpowered (20%–30%), suggesting that a large percentage of studies included could not be replicated, and these studies also exhibited publication bias (Solana & Santiago, 2022, 2023). These results mean that the replicability crisis is widespread in the field of embodied cognition, and nonoptimal research practices should also contribute to the low reproducibility, such as (1) the flexibility in data collection, analysis, and reporting; (2) the use of insufficient sample sizes; and (3) the selective publication of positive results. Additionally, Kaschak and Madden (2021) proposed that the ambiguity of embodied cognitive theories in language comprehension may cause the low replicability of studies. In conclusion, the instability of research results may stem from the combined effects of multiple factors. More crucially, all the factors mentioned above may also be present in ACE studies. Therefore, it must be emphasized that the low reliability of ACE is likely derived from multiple factors working together, and the instability of the sentence focus is just one of the reasons.

The present study aligns with previous research on the flexibility of mental simulation during language comprehension. To our knowledge, research on mental simulation flexibility primarily focuses on tense and task factors. First, the influence of tense on the match effect and the ACE has been revealed by several studies. For example, Kang et al. (2020) found that the match effect disappears when sentences are presented in the future tense. Other studies have found the ACE only in present progressive tense sentences but not when the sentence is in perfect tense (Bergen & Wheeler, 2010; Liu & Bergen, 2016). Second, the influence of task factors on mental simulation has been supported by numerous studies (Gilead et al., 2016; Hoedemaker & Gordon, 2014; Huettig et al., 2020; Lebois et al., 2015; Van Dam et al., 2012). For instance, Van Dam et al. (2012) found a significant increase in motor cortical activation

only when participants were required to consider the action-related features of the concept. Similarly, Gilead et al. (2016) found that reading action phrases increased excitability in the sensorimotor area when participants were asked to consider “*how to perform this action*” but activated the default network when participants were asked to consider “*why to perform this action*.” More surprisingly, the “indirect action demand” implied by the current context can activate the motor cortex, even in the absence of action information in the sentence (Van Ackeren et al., 2012). Integrating our findings with existing research on the flexibility of mental simulation, we posit that the impact of tense and task factors on mental simulation may share a common mechanism with sentence focus. These factors highlight specific aspects of language, directing readers to allocate more attention to it and resulting in deeper processing and more refined representations of the relevant information. The corresponding sensorimotor experiences may thus be activated. For instance, the present progressive tense emphasizes the process of motion, leading to the activation of relevant motor experiences and the occurrence of the ACE. Conversely, the present (or past) perfect tense highlights the result of motion, which likely prompts perceptual representations of motion outcomes; thus, the ACE would not occur. Likewise, studies on task factors indicate that the sensorimotor cortex is activated only when the task demand or context explicitly emphasizes action-related information. However, this hypothesis and the more precise mechanisms of this process must be explored with new experimental designs in the future.

Theoretical implication

According to different views on the importance of embodied representation in language understanding, Meteyard et al. (2012) divided embodied theory into “strong embodied” theory and “weak embodied” theory (as elaborated in the introduction, the “strong embodiment” theory posits that language comprehension relies entirely on sensorimotor experience, while the “weak embodiment” theory suggests that language processes are flexible, with sensorimotor experience conditionally involved in language comprehension). So far, a substantial body of behavioral and neuroimaging research has provided evidence for the flexibility of sensorimotor representation in language comprehension, such as the findings that tense and task factors may affect mental simulation. These findings indicate that sensorimotor representations are not necessary for language comprehension, thus supporting the “weak embodiment” theory.

Our finding that sentence focus influences the ACE or shape match effects extends the view of the flexibility of sensorimotor representations in language comprehension and the content of the “weak embodiment” theory. The sentence focus differs fundamentally from tense and task factors. First, task demand or context refers to factors beyond

language and emphasizes the impact of external factors on the current process of language comprehension. Second, while tense is a linguistic feature, it is only a grammatical aspect. In contrast, the focus represents the information structure of the sentence and highlights a specific component, directing readers’ attention to that component (Wang et al., 2014). Therefore, the two factors are fundamentally distinct. Based on the above, the present study provides the first evidence that the linguistic structure is also an important factor that needs to be considered in the flexibility of mental simulation, offering a significant complement to the “weak embodiment” theory.

Limitations and future research

Firstly, it is essential to note some limitations of the SPVT and ASCP. (1) Neither task directly reflects the activity of the sensorimotor system in language comprehension. Therefore, some researchers have questioned whether the match effect and ACE can serve as evidence for embodied language processing because they can also be explained from amodal, nonembodied views. For example, Ostarek and Huettig (2019) suggested that participants may extract abstract shape information about objects during sentence–picture verification tasks, leading to the observed match effect, which is unrelated to the sensorimotor system. Mahon and Caramazza (2008) proposed that both the match effect and ACE are more related to decision mechanisms rather than directly representing the activation of the sensorimotor system. Even Ostarek et al. (2019) found that there is no causal relationship between perceptual simulation and the match effect in the sentence–picture verification paradigm. Taking these considerations into account, the present study seems to only conclusively establish the influence of sentence focus on the match effect and ACE. Still, more objective and direct evidence is needed to determine whether sentence focus influences the activity of the sensorimotor system in language comprehension. (2) In both tasks, participants engaged in picture judgment or action execution after the sentence processing, thus they cannot reveal when the language system and the sensorimotor system interact, making it difficult to confirm whether sensorimotor recruitment is a necessary component or merely a byproduct in the language comprehension (Mahon & Caramazza, 2008). Considering these two issues, future research using high temporal resolution techniques such as EEG (de Vega et al., 2021; van Elk et al., 2010) and sEMG (Fino et al., 2016; Gálvez-García et al., 2020) to investigate the influence of sentence focus on the activity of the sensorimotor system is highly appropriate. However, it should be noted that although these techniques can provide direct and objective metrics of sensorimotor system activation and

answer questions about when the language and sensorimotor systems interact, they neither provide strong evidence for embodiment. This is because their results are merely correlational, meaning that one cannot assert that sensorimotor activation is caused by language comprehension. Therefore, in future research, employing causal methods such as brain lesions or brain stimulation techniques (for a review, see Ostarek & Bottini, 2021) to explore the causal relationship of the sensorimotor system with language comprehension is essential. This approach can provide decisive evidence for the embodiment view.

Secondly, the ACE observed under the action-focused sentences can also be interpreted from the perspective of motor imagery in Experiment 3. According to the perspective of motor imagery, external marking action information may encourage the participants to explicitly imagine the actions, thereby facilitating them to respond in the imagined direction. Based on this interpretation, the null results of Experiment 2 may also be related to the focus marker word “是” (*is*) failing to prompt participants to imagine action content actively. If so, the instability of ACE may be related to whether the participant voluntarily imagines the actions described in language. According to this explanation, the activation of the motor system should involve active imagery after language processing is completed, not as part of the language comprehension process; thereby, the ACE cannot be regarded as evidence supporting the embodiment view (Mahon & Caramazza, 2008). In fact, the match effect can also be explained from the perspective of mental imagery, but to some extent, Pecher et al. (2009) ruled out this possibility by employing a delayed SPVT. While no direct evidence can rule out the mental imagery explanation of ACE, the ASCP is less susceptible to task settings compared with the SPVT (in the SPVT, participants are required to directly compare the content of the sentence with that of the picture, which may prompt them to engage in mental imagery, whereas in the ASCP, participants are only required to judge the sensibility of the sentence, which is entirely unrelated to the action direction), and some previous studies have considered ACE as having evidence supporting action simulation (Bergen & Wheeler, 2010; Borreggine & Kaschak, 2006; Glenberg & Kaschak, 2002; Glenberg et al., 2008; Van Dam & Desai, 2017), so we are more inclined to interpret the ACE in action-focused sentences from the perspective of mental simulation. Meanwhile, we believe it is essential to compare the cortical motor areas involved in action-focused sentence comprehension and motor imagery tasks in future research, as it can distinguish between the mental simulation and mental imagery explanations of ACE.

Conclusion

Mental simulation may be influenced by sentence focus, in which only the focused information has the potential to be simulated in language comprehension. This finding not only suggests that the instability of the ACE in previous studies may be attributed to the instability of sentence focus but also highlights the flexibility of embodied representations, supporting the “weak embodiment” theory. Moreover, the results of the present study demonstrate the impact of the hierarchical structure of language itself on mental simulation, indicating that the key to mentally simulating specific information may lie in whether readers identify it as the sentence focus. In conclusion, the results of the present study extend the “weak embodiment” theory.

Acknowledgments We thank Shuo Tang for his help in the data collection and the participants for their support of the present study.

Author note Hua JIN contributed to the study design and the manuscript writing. Guangfang ZHOU completed the data collection, analysis and the manuscript writing under the supervision of Hua JIN. Xiang LI provided critical and constructive advice on the writing of the manuscript.

Funding No funding was received for conducting this study.

Data Availability The datasets analysed during the current study are available in the Open Science Framework (OSF) at <https://osf.io/gaued>.

Declarations

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

Ethics approval This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of Tianjin Normal University (No. 2023021801).

Consent to participate Informed consent was obtained from all individual participants included in the study.

Consent to publication Not applicable.

References

- Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and Brain Sciences*, 22(4), 637–660. <https://doi.org/10.1017/S0140525X99532147>
- Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. In M. de Vega, A. Glenberg, & A. Graesser (Eds.), *Symbols, embodiment, and meaning* (pp. 245–283). Oxford Academic. <https://doi.org/10.1093/acprof:oso/9780199217274.003.0013>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>

- Bergen, B., & Wheeler, K. (2010). Grammatical aspect and mental simulation. *Brain and Language*, 112(3), 150–158. <https://doi.org/10.1016/j.bandl.2009.07.002>
- Birch, S., & Rayner, K. (1997). Linguistic focus affects eye movements during reading. *Memory & Cognition*, 25(5), 653–660. <https://doi.org/10.3758/BF03211306>
- Borreggine, K. L., & Kaschak, M. P. (2006). The action–sentence compatibility effect: It’s all in the timing. *Cognitive Science*, 30(6), 1097–1112. https://doi.org/10.1207/s15516709cog0000_91
- Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5), 365–376. <https://doi.org/10.1038/nrn3475>
- Carota, F., Moseley, R., & Pulvermüller, F. (2012). Body-part-specific representations of semantic noun categories. *Journal of Cognitive Neuroscience*, 24(6), 1492–1509. https://doi.org/10.1162/jocn_a_00219
- Colling, L. J., Szűcs, D., De Marco, D., Cipora, K., Ulrich, R., Nuerk, H. C., ..., McShane, B. B. (2020). Registered replication report on Fischer, Castel, Dodd, and Pratt (2003). *Advances in Methods and Practices in Psychological Science*, 3(2), 143–162. <https://doi.org/10.1177/2515245920903079>
- Connell, L., & Lynott, D. (2009). Is a bear white in the woods? Parallel representation of implied object color during language comprehension. *Psychonomic Bulletin & Review*, 16, 573–577. <https://doi.org/10.3758/PBR.16.3.573>
- de Koning, B. B., Wassenburg, S. I., Bos, L. T., & Van der Schoot, M. (2017). Size does matter: Implied object size is mentally simulated during language comprehension. *Discourse Processes*, 54(7), 493–503. <https://doi.org/10.1080/0163853X.2015.1119604>
- de Koning, B. B., Wassenburg, S. I., Bos, L. T., & van der Schoot, M. (2017). Mental simulation of four visual object properties: Similarities and differences as assessed by the sentence–picture verification task. *Journal of Cognitive Psychology*, 29(4), 420–432. <https://doi.org/10.1080/20445911.2017.1281283>
- de Vega, M., Dutriaux, L., Moreno, I. Z., García-Marco, E., Seigneuric, A., & Gyselinck, V. (2021). Crossing hands behind your back reduces recall of manual action sentences and alters brain dynamics. *Cortex*, 140, 51–65. <https://doi.org/10.1016/j.cortex.2021.03.016>
- De Scalzi, M., Rusted, J., & Oakhill, J. (2015). Embodiment effects and language comprehension in Alzheimer’s disease. *Cognitive Science*, 39(5), 890–917. <https://doi.org/10.1111/cogs.12187>
- Díez-Álamo, A. M., Glenberg, A. M., Díez, E., Alonso, M. A., & Fernandez, A. (2020). The linguistic looming effect. *Journal of Memory and Language*, 114, 104147. <https://doi.org/10.1016/j.jml.2020.104147>
- Engelen, J. A., Bouwmeester, S., de Bruin, A. B., & Zwaan, R. A. (2011). Perceptual simulation in developing language comprehension. *Journal of Experimental Child Psychology*, 110(4), 659–675. <https://doi.org/10.1016/j.jecp.2011.06.009>
- Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, 11(1), 11–15. <https://doi.org/10.1111/1467-8721.00158>
- Fino, E., Menegatti, M., Avenanti, A., & Rubini, M. (2016). Enjoying vs. smiling: Facial muscular activation in response to emotional language. *Biological Psychology*, 118, 126–135. <https://doi.org/10.1016/j.biopsycho.2016.04.069>
- Fischer, M. H., Castel, A. D., Dodd, M. D., & Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nature Neuroscience*, 6(6), 555–556. <https://doi.org/10.1038/nm1066>
- Gálvez-García, G., Aldunate, N., Bascour-Sandoval, C., Martínez-Molina, A., Peña, J., & Barramuño, M. (2020). Muscle activation in semantic processing: An electromyography approach. *Biological Psychology*, 152, 107881. <https://doi.org/10.1016/j.biopsycho.2020.107881>
- Gilead, M., Liberman, N., & Maril, A. (2016). The effects of an action’s “age-of-acquisition” on action–sentence processing. *NeuroImage*, 141, 341–349. <https://doi.org/10.1016/j.neuroimage.2016.07.034>
- Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, 9(3), 558–565. <https://doi.org/10.3758/BF03196313>
- Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., & Buccino, G. (2008). Processing abstract language modulates motor system activity. *Quarterly Journal of Experimental Psychology*, 61(6), 905–919. <https://doi.org/10.1080/17470210701625550>
- Gough, P. M., Riggio, L., Chersi, F., Sato, M., Fogassi, L., & Buccino, G. (2012). Nouns referring to tools and natural objects differentially modulate the motor system. *Neuropsychologia*, 50(1), 19–25. <https://doi.org/10.1016/j.neuropsychologia.2011.10.017>
- Harnad, S. (1990). The symbol grounding problem. *Physica D: Non-linear Phenomena*, 42(1/3), 335–346. [https://doi.org/10.1016/0167-2789\(90\)90087-6](https://doi.org/10.1016/0167-2789(90)90087-6)
- He, X. Y., & Mo, L. (2002). A preliminary study of the effects of text signals. *Journal of Psychological Science*, (03), 279–282+381. <https://doi.org/10.16719/j.cnki.1671-6981.2002.03.006>
- Hoedemaker, R. S., & Gordon, P. C. (2014). Embodied language comprehension: Encoding-based and goal-driven processes. *Journal of Experimental Psychology: General*, 143(2), 914–929. <https://doi.org/10.1037/a0032348>
- Holman, A. C., & Gírbă, A. (2019). The match in orientation between verbal context and object accelerates change detection. *Psihologija*, 52(1), 93–105. <https://doi.org/10.2298/PSI180412033H>
- Horchak, O. V., & Garrido, M. V. (2021). Dropping bowling balls on tomatoes: Representations of object state-changes during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 47(5), 838–857. <https://doi.org/10.1037/xlm0000980>
- Huettig, F., Guerra, E., & Helo, A. (2020). Towards understanding the task dependency of embodied language processing: The influence of colour during language–vision interactions. *Journal of Cognition*, 3(1), 41. <https://doi.org/10.5334/joc.135>
- Jiang, L. (2012). On the focus marker “shi.” *The Guide of Science & Education*, 16, 102–143. <https://doi.org/10.16400/j.cnki.kjkdz.2012.08.024>
- Káldi, T., & Babarczy, A. (2021). Linguistic focus guides attention during the encoding and refreshing of working memory content. *Journal of Memory and Language*, 116, 104187. <https://doi.org/10.1016/j.jml.2020.104187>
- Kang, X., Eerland, A., Joergensen, G. H., Zwaan, R. A., & Altmann, G. T. M. (2020). The influence of state change on object representations in language comprehension. *Memory & Cognition*, 48(3), 390–399. <https://doi.org/10.3758/s13421-019-00977-7>
- Kaschak, M. P., & Madden, J. (2021). Embodiment in the lab: Theory, measurement, and reproducibility. In M. D. Robinson & L. E. Thomas (Eds.), *Handbook of embodied psychology: Thinking, feeling, and acting* (pp. 619–635). Springer.
- Kumle, L., Vö, M. L. H., & Draschkow, D. (2021). Estimating power in (generalized) linear mixed models: An open introduction and tutorial in R. *Behavior Research Methods*, 53(6), 2528–2543. <https://doi.org/10.3758/s13428-021-01546-0>
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato’s problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2), 211–240. <https://doi.org/10.1037/0033-295X.104.2.211>
- Lebois, L. A., Wilson-Mendenhall, C. D., & Barsalou, L. W. (2015). Are automatic conceptual cores the gold standard of

- semantic processing? The context-dependence of spatial meaning in grounded congruency effects. *Cognitive Science*, 39(8), 1764–1801. <https://doi.org/10.1111/cogs.12174>
- Lenth, R. V. (2016). Least-squares means: The R package *lsmeans*. *Journal of Statistical Software*, 69, 1–33. <https://doi.org/10.18637/jss.v069.i01>
- Liu, L. (2013). *The study on the focus marker in modern Chinese (Unpublished doctoral dissertation)*. Fudan University.
- Liu, N., & Bergen, B. (2016). When do language comprehenders mentally simulate locations? *Cognitive Linguistics*, 27(2), 181–203. <https://doi.org/10.1515/cog-2015-0123>
- Louwerse, M. M., & Jeuniaux, P. (2008). Language comprehension is both embodied and symbolic. In M. de Vega, A. Glenberg, & A. Graesser (Eds.), *Symbols and embodiment: Debates on meaning and cognition* (pp. 309–326). Oxford Academic. <https://doi.org/10.1093/acprof:oso/9780199217274.003.0015>
- Lowder, M. W., & Gordon, P. C. (2015). Focus takes time: Structural effects on reading. *Psychonomic Bulletin & Review*, 22(6), 1733–1738. <https://doi.org/10.3758/s13423-015-0843-2>
- Madden, C. J., & Dijkstra, K. (2009). Contextual constraints in situation model construction: An investigation of age and reading span. *Aging, Neuropsychology, and Cognition*, 17(1), 19–34. <https://doi.org/10.1080/13825580902927604>
- Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*, 102(1/3), 59–70. <https://doi.org/10.1016/j.jphysparis.2008.03.004>
- Matsumoto, N., Watson, L. A., Fujino, M., Ito, Y., & Kobayashi, M. (2022). Subjective judgments on direct and generative retrieval of autobiographical memory: The role of interoceptive sensibility and emotion. *Memory & Cognition*, 50(8), 1644–1663. <https://doi.org/10.3758/s13421-022-01280-8>
- Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex*, 48(7), 788–804. <https://doi.org/10.1016/j.cortex.2010.11.002>
- Montero-Melis, G., Van Paridon, J., Ostarek, M., & Bylund, E. (2022). No evidence for embodiment: The motor system is not needed to keep action verbs in working memory. *Cortex*, 150, 108–125. <https://doi.org/10.1016/j.cortex.2022.02.006>
- Morey, R. D., Rouder, J. N., & Jamil, X. (2015). *BayesFactor: Computation of Bayes factors for common designs* (Version 0.9.12) [Computer software]. <https://bayesfactorppl.r-forge.rproject.org>. Accessed 28 Feb 2024
- Morey, R. D., Kaschak, M. P., Diez-Alamo, A. M., Glenberg, A. M., Zwaan, R. A., Lakens, D., ..., Ziv-Crispel, N. (2022). A pre-registered, multi-lab non-replication of the action–sentence compatibility effect (ACE). *Psychonomic Bulletin & Review*, 29(2), 613–626. <https://doi.org/10.3758/s13423-021-01927-8>
- Newell, A., & Simon, H. A. (1976). Computer science as empirical inquiry. *Communications of the ACM*, 19(3), 113–126. <https://doi.org/10.1145/360018.360022>
- Osaka, M., Nishizaki, Y., Komori, M., & Osaka, N. (2002). Effect of focus on verbal working memory: Critical role of the focus word in reading. *Memory & Cognition*, 30(4), 562–571. <https://doi.org/10.3758/BF03194957>
- Ostarek, M., & Bottini, R. (2021). Towards strong inference in research on embodiment—Possibilities and limitations of causal paradigms. *Journal of Cognition*, 4(1), 5. <https://doi.org/10.5334/joc.139>
- Ostarek, M., & Huettig, F. (2019). Six challenges for embodiment research. *Current Directions in Psychological Science*, 28(6), 593–599. <https://doi.org/10.1177/0963721419866441>
- Ostarek, M., Joosen, D., Ishag, A., De Nijs, M., & Huettig, F. (2019). Are visual processes causally involved in “perceptual simulation” effects in the sentence–picture verification task? *Cognition*, 182, 84–94. <https://doi.org/10.1016/j.cognition.2018.08.017>
- Papesh, M. H. (2015). Just out of reach: On the reliability of the action–sentence compatibility effect. *Journal of Experimental Psychology: General*, 144(6), e116–e141. <https://doi.org/10.1037/xge000125>
- Pecher, D., van Dantzig, S., Zwaan, R. A., & Zeelenberg, R. (2009). Short article: Language comprehenders retain implied shape and orientation of objects. *Quarterly Journal of Experimental Psychology*, 62(6), 1108–1114. <https://doi.org/10.1080/17470210802633255>
- Qi, F. (2012). *The study on focus in modern Chinese (Unpublished doctoral dissertation)*. Fudan University.
- Qu, F. B., Yin, R., Zhong, Y., & Ye, H. S. (2012). Motor perception in language comprehension: Perspective from embodied cognition. *Advances in Psychological Science*, 20(6), 834–842.
- Rommers, J., Meyer, A. S., & Huettig, F. (2013). Object shape and orientation do not routinely influence performance during language processing. *Psychological Science*, 24(11), 2218–2225. <https://doi.org/10.1177/0956797613490746>
- Saccone, E. J., Thomas, N. A., & Nicholls, M. E. R. (2021). One-handed motor activity does not interfere with naming lateralized pictures of tools. *Journal of Experimental Psychology: Human Perception and Performance*, 47(4), 529–544. <https://doi.org/10.1037/xhp0000863>
- Sanford, A. J. S., Sanford, A. J., Molle, J., & Emmott, C. (2006). Shallow processing and attention capture in written and spoken discourse. *Discourse Processes*, 42(2), 109–130. https://doi.org/10.1207/s15326950dp4202_2
- Sanford, A. J., Price, J., & Sanford, A. J. (2009). Enhancement and suppression effects resulting from information structuring in sentences. *Memory & Cognition*, 37(6), 880–888. <https://doi.org/10.3758/MC.37.6.880>
- Shebani, Z., & Pulvermüller, F. (2013). Moving the hands and feet specifically impairs working memory for arm- and leg-related action words. *Cortex*, 49(1), 222–231. <https://doi.org/10.1016/j.cortex.2011.10.005>
- Solana, P., & Santiago, J. (2022). Does the involvement of motor cortex in embodied language comprehension stand on solid ground? A p-curve analysis and test for excess significance of the TMS and tDCS evidence. *Neuroscience & Biobehavioral Reviews*, 41, 104834. <https://doi.org/10.1016/j.neubiorev.2022.104834>
- Solana, P., & Santiago, J. (2023). Worse than expected: A z-curve reanalysis of motor cortex stimulation studies of embodied language comprehension. *Psicológica*, 44(2), e15661. <https://doi.org/10.20350/digitalCSIC/15661>
- Song, G. W., Xing, H. J., & Ma, Y. H. (2011). Influences of levels of problem posing and structure training on the mechanism of text signals. *Journal of Psychological Science*, 34(03), 588–592. <https://doi.org/10.16719/j.cnki.1671-6981.2011.03.043>
- Speed, L. J., & Majid, A. (2020). Grounding language in the neglected senses of touch, taste, and smell. *Cognitive Neuropsychology*, 37(5/6), 363–392. <https://doi.org/10.1080/02643294.2019.1623188>
- Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, 12(2), 153–156. <https://doi.org/10.1111/1467-9280.00326>
- Sturt, P., Sanford, A. J., Stewart, A., & Dawydiak, E. (2004). Linguistic focus and good-enough representations: An application of the change-detection paradigm. *Psychonomic Bulletin & Review*, 11(5), 882–888. <https://doi.org/10.3758/BF03196716>
- Tang, Y. (2020). *Focus markers promote Chinese reading: Evidence from eye movements (Unpublished master's thesis)*. Tianjin Normal University.

- Van Dam, W. O., & Desai, R. H. (2017). Embodied simulations are modulated by sentential perspective. *Cognitive Science*, 41(6), 1613–1628. <https://doi.org/10.1111/cogs.12449>
- Van Ackeren, M. J., Casasanto, D., Bekkering, H., Hagoort, P., & Rueschemeyer, S. A. (2012). Pragmatics in action: Indirect requests engage theory of mind areas and the cortical motor network. *Journal of Cognitive Neuroscience*, 24(11), 2237–2247. https://doi.org/10.1162/jocn_a_00274
- Van Dam, W. O., Van Dijk, M., Bekkering, H., & Rueschemeyer, S. A. (2012). Flexibility in embodied lexical-semantic representations. *Human Brain Mapping*, 33(10), 2322–2333. <https://doi.org/10.1002/hbm.21365>
- van Elk, M., van Schie, H. T., Zwaan, R. A., & Bekkering, H. (2010). The functional role of motor activation in language processing: Motor cortical oscillations support lexical-semantic retrieval. *NeuroImage*, 50(2), 665–677. <https://doi.org/10.1016/j.neuroimage.2009.12.123>
- Wagenmakers, E. J., Wetzels, R., Borsboom, D., & van der Maas, H. L. (2011). Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). *Journal of Personality and Social Psychology*, 100(3), 426–432. <https://doi.org/10.1037/a0022790>
- Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ..., Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, 25, 58–76. <https://doi.org/10.3758/s13423-017-1323-7>
- Wang, R., Yan, G. L., & Bai, X. J. (2004). Effect of signals on text memory and its mechanism. *Studies of Psychology and Behavior*, 03, 549–554.
- Wang, L., Li, X., & Yang, Y. (2014). A review on the cognitive function of information structure during language comprehension. *Cognitive Neurodynamics*, 8, 353–361. <https://doi.org/10.1007/s11571-014-9305-1>
- Wassenburg, S. I., de Koning, B. B., de Vries, M. H., Boonstra, A. M., & van der Schoot, M. (2017). Gender differences in mental simulation during sentence and word processing. *Journal of Research in Reading*, 40(3), 274–296. <https://doi.org/10.1111/1467-9817.12066>
- Winter, A., Dudschig, C., Miller, J., Ulrich, R., & Kaup, B. (2022). The action–sentence compatibility effect (ACE): Meta-analysis of a benchmark finding for embodiment. *Acta Psychologica*, 230, 103712. <https://doi.org/10.1016/j.actpsy.2022.103712>
- Witt, J. K., Kemmerer, D., Linkenauger, S. A., & Culham, J. (2010). A functional role for motor simulation in identifying tools. *Psychological Science*, 21, 1215–1219. <https://doi.org/10.1177/0956797610378307>
- Witt, J. K., Kemmerer, D., Linkenauger, S. A., & Culham, J. C. (2020). Reanalysis suggests evidence for motor simulation in naming tools is limited: A commentary on Witt, Kemmerer, Linkenauger, and Culham (2010). *Psychological Science*, 31(8), 1036–1039. <https://doi.org/10.1177/0956797620940555>
- Yang, C. L., Zhang, H., Duan, H., & Pan, H. (2018). Linguistic focus promotes the ease of discourse integration processes in reading comprehension: Evidence from event-related potentials. *Frontiers in Psychology*, 9, 2718. <https://doi.org/10.3389/fpsyg.2018.02718>
- Yao, P., Staub, A., & Li, X. (2022). Predictability eliminates neighborhood effects during Chinese sentence reading. *Psychonomic Bulletin & Review*, 29(2), 243–252. <https://doi.org/10.3758/s13423-021-01966-1>
- Yaxley, R. H., & Zwaan, R. A. (2007). Simulating visibility during language comprehension. *Cognition*, 105(1), 229–236. <https://doi.org/10.1016/j.cognition.2006.09.003>
- Zhang, J. J., Wang, X. X., & He, X. M. (2015). Effects of agent's motivation on mental simulation during sentence comprehension. *Acta Psychologica Sinica*, 47(10), 1247–1259. <https://doi.org/10.3724/SP.J.1041.2015.01247>
- Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. *Psychology of Learning and Motivation*, 44, 35–62. [https://doi.org/10.1016/S0079-7421\(03\)44002-4](https://doi.org/10.1016/S0079-7421(03)44002-4)
- Zwaan, R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension: Six replication attempts. *PLOS ONE*, 7(12), e51382. <https://doi.org/10.1371/journal.pone.0051382>
- Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, 13(2), 168–171. <https://doi.org/10.1111/1467-9280.00430>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open practices statement The data and code for all experiments are available online (<https://osf.io/gaued>), and none of the experiments was preregistered.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.