

## Body part representations in verbal semantics

**BENJAMIN BERGEN**

*University of California, San Diego, La Jolla, California*

**TING-TING CHAN LAU**

*University of Hawaii, Manoa, Hawaii*

**SHWETA NARAYAN**

*University of California, Berkeley, California*

AND

**DIANA STOJANOVIC AND KATHRYN WHEELER**

*University of Hawaii, Manoa, Hawaii*

Embodied theories of language propose that word meaning is inextricably tied to—grounded in—mental representations of perceptual, motor, and affective experiences of the world. The four experiments described in this article demonstrate that accessing the meanings of action verbs like *smile*, *punch*, and *kick* requires language understanders to activate modality-specific cognitive representations responsible for performing and perceiving those same actions. The main task used is a word–image matching task, where participants see an action verb and an image depicting an action. Their task is to decide as quickly as possible whether the verb and the image depict the same action. Of critical interest is participants' behavior when the verb and image do not match, in which case the two actions can use the same effector or different effectors. In Experiment 1, we found that participants took significantly longer to reject a verb–image pair when the actions depicted by the image and denoted by the verb used the same effector than when they used different effectors. Experiment 2 yielded the same result when the order of presentation was reversed, replicating the effect in Cantonese. Experiment 3 replicated the effect in English with a verb–verb near-synonym task, and in Experiment 4, we once again replicated the effect with learners of English as a second language. This robust interference effect, whereby a shared effector slows discrimination, shows that language understanders activate effector-specific neurocognitive representations during both picture perception and action word understanding.

Across the cognitive sciences, researchers are finding that aspects of higher cognition, such as language, memory, and concept representation, are shaped by, grounded in, or reducible to elements of lower cognition, such as perception and action. In this article, we focus on how language is embodied in these low-level cognitive systems. A series of experiments tested whether language users accessing the meanings of action verbs like *kick*, *grab*, and *smile* make use of perceptual and motor systems dedicated to recognizing and performing those same actions. The findings we report support the claim (Feldman, 2006; Feldman & Narayanan, 2004; Gallese & Lakoff, 2005) that concepts in general, and action concepts in particular, are grounded in perceptuomotor circuitry dedicated to multimodal interactions with the world.

There is substantial evidence from different empirical methods that processing language denoting actions performed with particular body parts activates areas of motor and premotor cortex involved in producing motor actions using those same body parts. Using behavioral and

neurophysiological evidence, Pulvermüller, Härle, and Hummel (2001) and Hauk, Johnsrude, and Pulvermüller (2004) found that verbs denoting actions performed with different effectors were processed in different regions of motor cortex; the areas of motor cortex involved in leg motion, for instance, displayed more activation during the processing of leg-related words than during processing of mouth-related words and vice versa. More recently, Tettamanti et al. (2005) have shown, through functional brain imaging, that passive listening to sentences describing mouth versus leg versus hand actions activates different parts of premotor cortex (as well as other areas—specifically, BA 6, BA 40, and BA 44). Similar findings have been reported in a number of other studies (such as Buccino et al., 2005), although in at least one case, they were not replicated (Postle, McMahon, Ashton, Meredith, & de Zubicaray, 2008). Behavioral studies have convergently shown that language describing motor actions activates motor systems (Bergen & Wheeler, 2005; Glenberg & Kaschak, 2002; Helbig, Graf, & Kiefer, 2006; Masson,

---

B. Bergen, bkbergen@ucsd.edu

---

Bub, & Warren, 2008; Tseng & Bergen, 2005; Zwaan & Taylor, 2006).

The finding that motor systems are engaged during the processing of language about action is not surprising, since motor systems are in fact used for multifarious purposes. Motor control is supported by a network of brain structures that includes primary motor cortex, premotor area, and supplementary motor area, as well as the cerebellum and basal ganglia (Winter, 1990). Recently, it has been shown that some of these motor areas fire not only when individuals execute actions, but also when they perceive the same actions (Rizzolatti & Craighero, 2004). Neurons that display the property of firing during both execution and perception of action, *mirror neurons*, have been extensively documented in single-cell recording studies in monkeys (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). Although there are no equivalent single-unit studies in humans, comparable *mirror activity* patterns have been demonstrated through brain imaging studies. These human studies show that the mototopic organization of motor cortex—with the various effectors arrayed so that the legs and feet are in superior motor cortex, below which are the hands, arms, and mouth—extends to the perception of actions as well (Buccino et al., 2001). The execution or observation of actions produced by the mouth, leg, and hand activate distinct parts of premotor cortex, found in ventral sites, dorsal sites, and intermediate foci, respectively. When appropriate target objects are present, there is also activation in a somatotopic activity map in parietal cortex (Buccino et al., 2001). Similar findings have shown activation in motor areas during motor imagery (Kosslyn, Ganis, & Thompson, 2001) and during recall of previously performed actions (Nyberg et al., 2001). In sum, performing, perceiving, and imagining actions all make use of common systems dedicated to specific effectors.

Existing work thus shows that motor areas dedicated to particular effectors become active when participants perform a variety of higher cognitive functions, including processing linguistic input. But the spatial resolution of current brain imaging techniques does not allow for inspection of individual neurons. Consequently, the presence of neural activation in the same region for both action and language understanding, as determined by fMRI, does not necessarily imply that the very same neural circuits are used for the two functions; different circuits in close proximity to each other could perform different functions. So this question remains open: Are the same systems used for performing, perceiving, and processing language about actions?

What's more, even if the very same neural circuits are active during both motor control and language understanding, this does not imply that such activation constitutes a functional component of language understanding. It could be the case, for example, that the activation found by Hauk et al. (2004), Pulvermüller et al. (2001), and Tettamanti et al. (2005) is simply a downstream, collateral pattern of activation that plays no functional role in language understanding. Thus, although imaging studies suggest a role

for effector-specific motor systems in processing action language, it remains to be determined whether and to what extent the same neural resources dedicated to motor control are recruited for the purpose of extracting meaning from linguistic input.

We designed a series of experiments to address these two issues left open by previous research: (1) whether action and language understanding use the same motor circuitry, and (2) whether this activation of motor areas is a functional part of language understanding. The paradigm we used investigates whether there is interference between visual and linguistic input during the process of matching images and verbs that depict related actions.

This basic logic was the same for all four experiments we conducted. Experiment 1 used an image–verb matching paradigm, in which participants were presented with a stick figure image for 1 sec, followed by a written verb, and were asked to decide as quickly as possible whether the verb was a good description of the image or not. The images and verbs all depicted actions primarily associated with one body area—hands/arms, mouth/face, or feet/legs. In half of the trials, the verb was a good description of the image (the “matching” condition). The other half of the trials, in which the verb was not a good description of the image, was split into the two conditions of interest (Figure 1). In one of these nonmatching conditions, the same-effector condition, the different actions depicted by the image and described by the verb made use of the same major body area, or effector. In the other half of the nonmatching trials, the different-effector condition, the image and verb depicted actions performed predominantly using different effectors. We hypothesized that there would be greater interference between the neurocognitive structures responsible for perceiving and understanding language about actions using the same effector than between those responsible for actions using different effectors. We therefore predicted slower responses in the same-effector condition than in the different-effector condition.

The reasoning behind this prediction of interference goes as follows. The brain imaging and behavioral studies cited above suggest that understanding action verbs makes use of the same neurocognitive resources used for performing and recognizing motor actions. If this is the case, having participants match an action verb with an action image should activate the circuits responsible for perceiving and performing the actions associated with both the word and the image. When the image and verb depict the same motor action (in the matching condition), only the one matching perceptuomotor circuit becomes active. But when a verb and image do not match, the neural structures responsible for both should become active.

Critically, this simultaneous activation of two distinct motor circuits ought to have different effects, depending on the relation between the circuits. Motor and perceptual systems display lateral inhibition among structures responsible for related but incompatible functions (Windhorst, 1996); that is, neural structures representing similar actions must inhibit each other more strongly than those representing less closely related actions. The consequences of

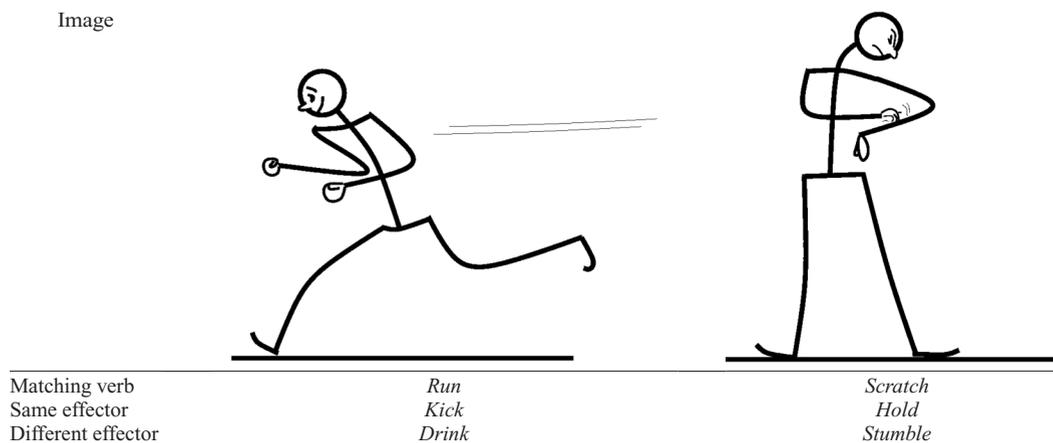


Figure 1. Verbs in the three conditions for the images *run* and *scratch*.

trying to walk and chew gum simultaneously are less dire than those of trying to walk and waltz simultaneously.

As a result, when two distinct active circuits encode motor actions using the same effector, they will be in direct competition with each other, which will produce interference and longer response times (RTs). By contrast, when the two distinct active circuits encode actions using different effectors, which are thus not laterally inhibitory, they should compete less with each other, leading to faster responses. The logic is thus that if language and visual images activate perceptuomotor circuits of the same type, then participants should display greater interference in identifying a mismatched verb and image when they use the same effector (the same-effector condition) than when an image and verb using different effectors are compared.

There are accounts in the literature (see, e.g., Bergen, 2008; Kaschak et al., 2005) of when to expect facilitation (as in Glenberg & Kaschak, 2002, and Stanfield & Zwaan, 2001, among many others) and when to expect inhibition effects (as in Bergen, Lindsay, Matlock, & Narayanan, 2007; Estes, Verges, & Barsalou, 2008; and Richardson, Spivey, Barsalou, & McRae, 2003, among others) in behavioral research like this. Interference has the potential to be a more compelling finding than facilitation, simply because it can allow for stronger inferences about the causal role of the systems involved. Imagine, for instance, that we observe that processing a sentence about action (say, “You punched the wall”) using a particular effector facilitates subsequent actions using that same effector (such as pushing a button in front of you). Although this facilitation would suggest that the motor system is engaged during language processing, it would not tell us whether that motor system activity plays a functional role in language processing. However, if performing a task that engages the motor system, like perceiving motor actions, selectively interferes with subsequent processing of language describing similar but incompatible actions, as we expect to find in the work described here, this suggests that activating the motor system is a functional part of accessing an action word’s meaning. The reasoning behind our expectation of interference, as well as other potential explanations for the observed effect, are fleshed out

in more detail in the Discussion section following Experiment 1, as well as in the General Discussion section.

## EXPERIMENT 1 Image–Verb Matching

### Method

**Participants.** Participants were 39 students at the University of California (UC) at Berkeley, all native English speakers with normal hearing and language competence and with normal or corrected-to-normal vision. They received course credit in exchange for participation in the experiment.

**Procedure.** Each trial consisted of a visual stimulus, like the images shown in Figure 1, presented for 1,000 msec, followed by a 500-msec interstimulus interval (ISI). During the first 450 msec of this interval, participants saw a visual mask covering the whole screen. This was meant to reduce potential priming effects resulting from visual perception. In order to avoid forward masking, a blank screen was shown for the final 50 msec of this ISI. An English verb in written form was then presented until the participant pressed a button indicating that the verb was or was not a good description for the action depicted in the image. The verb fell into one of the three conditions described above: (1) matching, (2) same effector, and (3) different effector.

**Materials.** A total of 48 black-and-white stick figure images were hand-created by one of the experimenters, using a graphics editor. One third of these images were intended to specifically depict particular motor actions performed primarily by each of the three effectors: mouth/face, hand/arm, and foot/leg. In order to evoke specific actions like *trip* and *scream*, we manipulated posture, occasional movement lines, and head and eye positions, as well as overall body shape. Many of the actions depicted by these images thus also included characteristic positions or movements of body parts other than the effector centrally involved in the action.

Verbs that appropriately described these images were selected through a norming study in which 13 native speakers of American English were presented with each image and were asked to provide the verb they thought best described the action depicted by the image. The most frequent response to each image (always provided by more than half of the participants in the norming) was taken as the matching verb. To produce the nonmatching verb–image pairs, each image was randomly assigned one of the verbs that matched another verb with the same effector and one of the verbs with another effector; these were the same-effector and different-effector verbs for that image. The same verbs and images were thus used in the same-effector and different-effector conditions. We verified that, by contrast with the matching pairs, none of the same-effector or different-effector verbs

had been produced in the norming for the images they were paired with. Participants were given the following instructions:

This experiment tests how people relate words and images. You will first see an image of a person performing an action. Then you will see a verb that is either a good description of the action or not. Your job is to decide as quickly as possible whether or not the word is a good description of the action. If the word is a good description of the action, press the “Yes” button. If it is not, press the “No” button.

**Design.** Each participant was randomly placed in one of two groups. After a practice session with 14 image–word trials, each group was presented with each of the 48 images twice, once in each half of the experiment, in a pseudorandom order. Overall, each participant saw (1) each image followed by a matching verb once, (2) half of the images followed by a (nonmatching) different-effector verb, and (3) the other half of the images followed by a (nonmatching) same-effector verb. The verbs were distributed such that each image that was shown to one group in the same-effector condition was in the different-effector condition for the other group.

**Results**

There were fewer than 2% incorrect responses, overall, and no significant differences in error frequency among the conditions. In order to ensure that any significant differences in RT across conditions did not result from a small set of outliers, we removed all trials that deviated more than two standard deviations from the mean for that item. This resulted in the removal of less than 4% of the remaining data. In addition, six items were responded to correctly by fewer than 80% of participants; in order to avoid any potentially ambiguous images, we removed these items in all their conditions from analysis. In what follows, we consider only correct “no” responses to trials in the same-effector and different-effector conditions.

Of those data that remained, the means were different in the same-effector and different-effector conditions, as shown in Table 1. We can see that the mean RT to nonmatching verbs is, on average, 48 msec longer when the verb and image use the same effector than when they use different effectors. This critical difference was significant in one-way repeated measures ANOVAs by both participants [ $F_1(1,38) = 11.47, p < .01, \eta_p^2 = .24$ ] and items (where images are taken as items) [ $F_2(1,41) = 4.37, p < .05, \eta_p^2 = .15$ ].

**Analysis of Global Semantic Similarity**

The direction of the significant difference between the two nonmatching conditions conforms to the hypothesis that rejecting a verb as the description of an image takes longer when the same effector is involved in the actions depicted

by both. An alternative explanation for this behavior is that it is semantic similarity in general, rather than the sharing of an effector in particular, that gives rise to this effect. Actions that share an effector might, in general, be similar to each other in dimensions other than the identity of the effector; if so, participants might simply have taken longer to reject verbs that described actions that were, in some way, more similar to the actions depicted by the images they followed.

We addressed this concern by investigating whether similarity between the two presented actions was a good predictor of RT. The metric we used was a similarity rating produced by latent semantic analysis (LSA; Landauer, Foltz, & Laham, 1998). LSA, among other things, is a statistical method for extracting and representing the similarity between words or texts on the basis of the contexts they appear or do not appear in. Two words or texts will be rated as more similar the more alike their distributions are in a given corpus. Of relevance to the present discussion is the pairwise comparison function, which produces a similarity rating from  $-1$  to  $1$  for any pair of texts. Identical texts would have a rating of  $1$ , whereas completely dissimilar ones would have a rating of  $-1$ . Because LSA similarity scores rely exclusively on distributional similarity of the words in a corpus, there has been extensive criticism of their use as a representation of actual meaning (e.g., Glenberg & Robertson, 2000). However, LSA does provide a stable measure of semantic similarity that aligns relatively well with human judgments of similarity. LSA has been shown to perform quite like humans in a range of behaviors, including synonym (Landauer & Dumais, 1997) and multiple-choice (Landauer et al., 1998) tasks. We are using LSA only as a measure of similarity, not as a theory of meaning.

LSA’s pairwise comparison function requires input in the form of pairs of words. We constructed these pairs by taking the matching verb for each image to represent that image and pairing it with each of the verbs presented for that image in the three conditions. For instance, for the *run* image (Figure 1), we determined a semantic similarity score between *run* and *run* (matching), between *run* and *kick* (nonmatching, same effector), and between *run* and *drink* (nonmatching, different effector). This is an indirect way of evaluating the similarity between an image and a verb, since it is mediated by a verb describing the image but provides a rough estimate of semantic proximity.

To determine the degree to which semantic similarity accounted for the results reported above, we regressed the average RT per item (i.e., per image–verb pair) on the LSA rating for that item. This regression included only the nonmatching conditions, because the distribution of LSA

**Table 1**  
**Mean Response Times and Standard Deviations (in Milliseconds) in**  
**Same-Effector and Different-Effector Conditions in Experiments 1, 2, and 3**

	Experiment 1		Experiment 2				Experiment 3	
	Image–Verb		English Verb–Image		Cantonese Verb–Image		Verb–Verb	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Same effector	798	251	783	115	847	184	1,030	300
Different effector	750	204	732	104	801	156	930	240

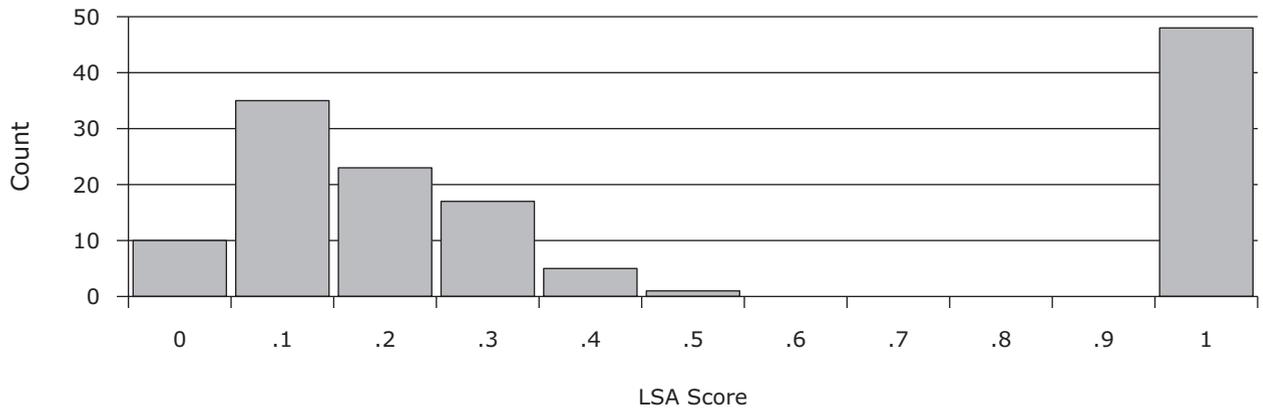


Figure 2. Histogram depicting number of word pairs with ranges of latent semantic analysis (LSA) ratings.

ratings across items is abnormal if it includes the matching condition, as can be seen in Figure 2.

Considering only the two nonmatching conditions, there was a very weak correlation between LSA rating and RT ( $R = .094, p = .38$ ). As can be seen from the regression graph in Figure 3 (and from the positive value of the regression coefficient  $R$ ), the relation is in the predicted direction; the trend is for participants to take longer to reject more similar pairs of words and pictures than less similar ones. However, whereas the similarity between a nonmatching verb and image as measured by LSA qualitatively might account for a small amount of the variance in RT, it does not do so significantly. Of course, this does not prove that sharing an effector (and not other sorts of similarity) is responsible for the RT effects we observed. The LSA rating might be a flawed measure of similarity in general or with respect to verb–image similarity. For this reason, further studies like the ones described below are needed to test the source of the effect. The absence of a significant relation between LSA rating and RT shown by the regression above does, however, suggest that an ex-

planation based on similarity alone does not transparently account for the interference behavior we found.

**Discussion**

Participants took significantly longer to decide that an image and verb did not match when the depicted and described actions shared an effector than when they did not. This result provides evidence that, when understanding motor action verbs, language users recruit some resources normally used for perceiving or performing those same actions. Given findings that motor actions are perceived in part through activating some brain circuits involved in motor control, this finding is also compatible with the idea that understanding action verbs depends on the active simulation or imagination of motor control (Bergen & Chang, 2005; Bergen & Wheeler, 2005; Feldman & Narayanan, 2004; Gallese & Lakoff, 2005; Glenberg & Kaschak, 2002; Tseng & Bergen, 2005; Zwaan & Taylor, 2006; etc.).

We have loosely suggested above that the observed interference caused by processing images and same-effector verbs could arise from lateral inhibition. A more detailed

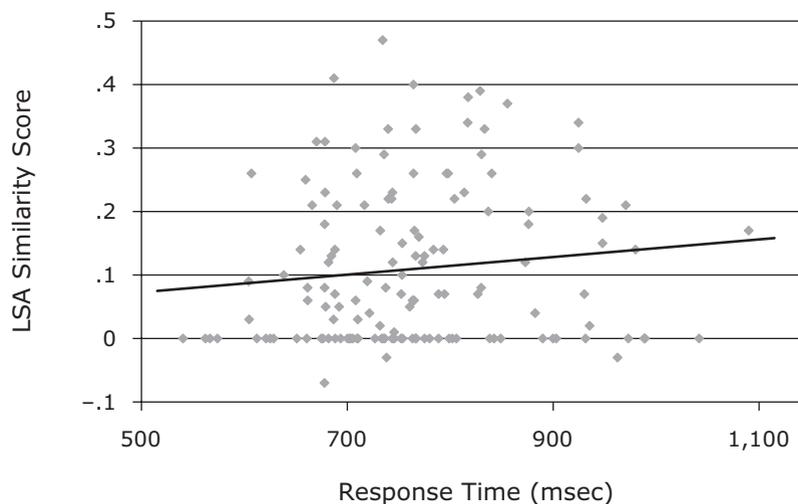


Figure 3. Regression of response time by latent semantic analysis (LSA) rating per item.

account is possible, however. We know that that the mirror circuits involved in the recognition of verbs and images are specific to particular types of action; this specificity has been seen in monkeys (Gallese et al., 1996), where a given mirror neuron may code a specific type of gesture, such as a precision grasp. We have already argued that when two actions use the same effector, it is necessary for mirror circuits that encode those actions to mutually inhibit each other. This lateral inhibition may give rise to the sort of delay in response evidenced in the experiment described above in the following way.

Processing words pertaining to motor actions yields activation of the motor structures responsible for performing those specific actions (Hauk et al., 2004; Pulvermüller et al., 2001; Tettamanti et al., 2005). There are convincing arguments that mirror circuits are among the motor structures activated by motor action language (Feldman & Narayanan, 2004; Gallese & Lakoff, 2005). In terms of the task at hand, then, neural representations for actions are activated by both the image and the verb. Participants take these two inputs and determine whether they match. A crucial part of this process must, on any account, be the determination by the participant of having perceived one action or two. At the neural level, this can be translated into the strong activation of one (in the matching condition) versus two (in the two nonmatching conditions) motor action circuits. In the same-effector condition, the two motor structures are responsible for actions sharing an effector, and as such, they will strongly inhibit each other. It will thus take longer, in such a trial, for the two distinct motor circuits to become fully active and longer for a participant to arrive at two distinct action perceptions and a “no” response. By comparison, when the actions do not make use of the same effector—for example, *jog* and *laugh*—less mutual inhibition of the motor circuits leads to a faster time course to coactivation of the two separate circuits and a “no” response.

Another possible account of the observed interference effect, suggested by an anonymous reviewer, is not lateral inhibition but, instead, lateral priming. Helbig et al. (2006) have shown that motor affordances for objects that are different but use the same effector prime each other. It could be that recognizing the initial image in each trial activated not only the motor representation of the depicted action, but also representations of other actions using the same effector. In this case, a subsequent word describing a nonmatching action using the same effector might take longer to reject, since its motor representation is already somewhat active.

Regardless of the ultimate neural explanation, the findings above show that word processing depends on perceptuomotor processes akin to those used in action perception. Neural imaging studies show that understanding motor action language yields activation of motor areas but leave open the possibility that different circuits located in close physical proximity are responsible for the different behaviors—actions versus language understanding. However, slower responses to language as a result of activating perceptuomotor systems, as shown in this experiment, confirm that the perceptuomotor structures activated by language are of the same type as those activated by action

perception. Moreover, these results demonstrate that this activation of perceptuomotor systems plays a functional role in language processing. If it did not, then the normal mechanisms of verb processing should be able to go through without interference, whether or not the actions described and depicted used the same effector.

There is one other alternative explanation for the observed effect that is worth addressing here. It might have been that seeing the images led participants to activate representations not only for the depicted action, but also for other actions using the same effector. Although these would eventually be rejected as nonmatching, it could be that a same-effector verb that was among this list of candidates, and thus previously activated, took longer to reject than an unrelated, and thus previously unactivated, different-effector verb.

Experiment 2, below, addressed this concern by reversing the order of presentation of the verb and image. The results provide further confirmation of the robustness and reliability of the same-effector interference effect. We also replicated the study in Cantonese, showing that the effect is not specific to English or, by extension, to alphabetic writing systems.

## EXPERIMENT 2

### Verb–Image Matching in English and Cantonese

This experiment replicated Experiment 1 with two modifications. First, the order of stimuli presentation was reversed, so that participants saw first the verb, then the image. We predicted that if the previous results were due to the activation of action-specific neural structures, it should not matter what order the verb and image were presented in; we should still find the same slower responses to same-effector mismatches. Experiment 2 thus tested whether a delay in matching is a generalized, rather than a directional, interference effect.

The second modification was a change in the language of presentation and the native language of participants. Experiment 1, like other studies reported above, provided evidence that motor systems are activated during language understanding. But these studies use mostly English and exclusively Indo-European languages as the native language of participants and the language of experimental materials. Experiment 2 included both a group of English speakers performing the task in English and a group of native Cantonese speakers presented with Cantonese stimuli. We chose to investigate Cantonese, in particular, because of potential differences caused by moving from the alphabetic writing system used for English to the logographic writing system that Cantonese employs. Some Cantonese action verbs include components—radicals—that indicate the effector used to perform the described action. The presence of an explicit marker of effector on a verb might augment the size of the observed effect.

We hypothesized that we would replicate the image–verb interference effect with a verb–image interference effect for both English and Cantonese speakers. This rep-

lication thus provided a platform for evaluating the robustness of the effect observed in Experiment 1, across both variations in presentation order and languages with different writing systems.

## Method

**Participants.** Participants were 21 native English speakers and 27 native Cantonese speakers. Participants received a small gift in exchange for participation.

**Procedure.** The procedure replicated that in Experiment 1, the only modification being reversal of presentation order so that the word was presented first and the image second. The ISI still contained a visual mask in order to inhibit any influence of visual memory. Participants saw the instructions and experimental items in their native language: English or Cantonese.

**Materials.** A total of 36 words (12 each of mouth, hand, and leg verbs) were paired with images. Most were a subset of those used in Experiment 1. Some of the original images were deemed by one of the authors (a native Cantonese speaker) to be impossible to represent with a single character in Cantonese, and, thus, were not included in either the English or the Cantonese version of this experiment. Some of the deleted images were replaced and represented by only one word or character in each language. New images were created by an experimenter and were drawn as consistently as possible with the original stimuli.

English words paired with images were taken from Experiment 1, when possible. Cantonese words were chosen on the basis of a norming study with 20 native Hong Kong Cantonese speakers who did not subsequently participate in the main experiment. For the norming, participants were presented with an image along with 3 single-character Cantonese words and were asked to choose the word that best described the image. From 104 words describing 36 images, the descriptor most frequently selected was then paired with its image as the matching verb for the main experiment.

**Design.** All 36 images were presented twice to each participant, once in the matching and once in one of the two nonmatching conditions. Instructions, training, and all other protocols were the same as in Experiment 1.

## Results

Accuracy was relatively high in both language groups, with fewer than 4% incorrect responses among English speakers and fewer than 5% among Cantonese speakers. There were no significant differences in accuracy across conditions. Outlying responses were removed from analysis if they fell more than two standard deviations from the mean for that item. This resulted in the removal of less than 4% of the data. No participants or items had outlying mean RTs or mean accuracy lower than 80%.

We investigated only the nonmatching conditions, because a comparison with the matching condition bore no theoretical interest. The results in Table 1 show the same effect as in the previous experiment: faster reactions to different-effector than to same-effector pairs, in both languages. We performed a single omnibus analysis with language (English or Cantonese) as a between-participants factor and condition (same or different) as a within-participants factor. The effect of interest was the difference in RTs between the two nonmatching conditions. There was a large, significant main effect of condition in a repeated measures ANOVA by participants [ $F_1(1,46) = 22.63, p < .001$ ], but not by items ( $F_2 < 1$ ). There was no main effect of language or interaction between language and condition. The critical interaction between different-

effector and same-effector responses was significant in participant analyses across both English and Cantonese language groups, as shown by planned repeated measures ANOVAs for each language group. For English speakers, there was a significant difference between same and different responses by participants [ $F_1(1,20) = 12.68, p < .01, \eta_p^2 = .28$ ], as was the case for Cantonese [ $F_1(1,26) = 10.64, p < .01, \eta_p^2 = .22$ ]. However, item analyses showed no significant differences for either (both  $F_2$ s  $< 1$ ).

## Discussion

As in Experiment 1, both groups in Experiment 2 took longer to respond when a verb and image encoded different actions that shared an effector (*kick* and *jump*) than when they encoded different actions using different main effectors (*kick* and *grab*).

The results from Experiment 2 provide two new pieces of information about the activation of effector-specific motor knowledge during word processing. First, the effect is robust, even given different orders of image and verb presentation. This means that the results from Experiment 2 could not be due to the proposed alternative explanation for Experiment 1: that participants actively came up with verbs to describe the picture and were more likely to come up with a nonmatching same-effector one than a nonmatching different-effector one, and thus took longer to reject same-effector verbs. Instead, finding the same effect when the verb was presented first supports the initial explanation offered in the discussion for Experiment 1: that it is competition among motor representations activated by the verb and image that produces the interference in the same-effector condition. A second outcome of the findings reported here is that the effect is not language specific, since it is demonstrated here with both English and Cantonese speakers. This is an important point, since any psychologically plausible theory of language understanding gains validity if it is observed across languages.

A potential confound in both Experiment 1 and Experiment 2 is that participants may have been responding to visual properties of the images, rather than to motor properties of the actions depicted. For instance, a picture of one action might, in some way, have resembled a participant's mental image of another action involving the same effector, whether or not the actions themselves were similar. It is plausible to believe that there is more visual similarity between actions that use the same effector than between those using different effectors. This would have yielded greater visual interference (and thus the longer RTs) when nonmatching items used the same effector than when they used different ones. Or there might have been a difference in the visual attention required for the two conditions. In the same-effector condition, participants might have needed to perform more detailed inspection of a particular part of the image, the part representing the effector used by both actions. However, in the different-effector condition, the participant needed only to visually determine that a different part of the body was involved. Either of these alternate accounts could explain the observed effect in Experiment 2.

### EXPERIMENT 3 Verb–Verb Matching

In order to exclude the possibility that it was the visual properties of the presented images that yielded the effects observed in Experiments 1 and 2, and not the identity of the effector used, we conducted a third experiment in which all images were replaced by verbs. Participants in Experiment 3 were thus shown a verb followed by a near-synonym or a mismatching verb. The mismatches again fell into two different categories; the actions being compared used either the same effector or different effectors. Participants were asked to decide whether the two verbs had approximately the same meaning. The absence of images in this design removed the visual confound. We hypothesized that as in the previous experiments, when presented with verb–verb pairs and asked to judge whether they matched, participants would respond more slowly to mismatches when the verbs shared an effector than when they did not.

#### Method

**Participants.** Fifty-seven undergraduate cognitive science students at UC Berkeley participated in the study in exchange for either extra credit or \$5. All were native English speakers. The classes for which students could receive extra credit by participating in this experiment had not, at the time of the experiment, covered material that would help them guess the purpose of the experiment.

**Procedure.** Participants performed a task similar to those in Experiments 1 and 2, with some necessary differences. Participants were informed that they would see two verbs in succession and then would be prompted to indicate as quickly as possible whether they considered the two to mean “approximately the same thing.” This phrasing was necessary because there are very few exact synonyms for common English action verbs, if any.

The experiment began with a practice session consisting of eight randomly chosen filler trials, during which participants were given displayed feedback. This served to demonstrate that “approximately the same thing” meant very similar, although not necessarily identical, verbs. In each trial of the main experiment, participants were presented with a fixation cross in the center of the screen for 2 sec. A longer fixation cross time was used in this experiment because piloting showed that a shorter presentation made the study undesirably stressful for participants. The fixation cross was followed by an English action verb presented for 1 sec, then a visual mask for 450 msec and a blank screen for 50 msec. (The visual mask from the previous experiment was kept to maintain the parallel structure of the experiments.) Then the second verb was displayed; it stayed on the screen until the participant pressed “yes” or “no.” All verbs were presented in capital letters in the center of the screen.

**Materials.** Experimental materials were constructed from 42 pairs of hand, foot, or mouth verbs and constituted 50% of the trials. The remaining trials were fillers included to obfuscate the aim of the experiment. For the fillers, we picked verb pairs that did not obviously involve hand, foot, or mouth actions. In critical trials, the verbs were related in one of three ways. Half were near-synonyms—the matching condition. The other half of the critical trials were non-matching; verbs had clearly different meanings. Half of these pairs (same-effector cases) included two verbs denoting actions using the same primary effector (hand, foot, or mouth). Different-effector cases constituted the other half of the nonmatching trials; in them, the two verbs did not share a primary effector. Examples of the different conditions are shown in Table 2.

There are not many English bodily action verbs that are short, commonly used, and unambiguous in a neutral context. In order to ensure consistent stimulus sets across experiments, the list of verbs came from those produced by participants in the pretest to Experiment 1, in

**Table 2**  
Examples of Verbs in Each of the Three Conditions

Initial Verb	Match	Same Effector	Different Effector
scream	shriek	lick	step
tie	knot	clap	run
dance	waltz	limp	yell

which they named a verb associated with each visual image. Choosing true synonyms was difficult because, for the most part, there are no commonly occurring synonyms for simple action verbs. Some of the synonyms given were thus arguably mismatches (for instance, “wink” and “blink”). No trials in which participants disagreed with our classification were included in the analysis, however, because we omitted all “incorrect” answers from our analysis (e.g., trials where participants judged “wink” and “blink” to be nonmatching).

**Design.** Each of the 42 critical initial verbs was presented twice to each participant: once followed by a near-synonym, once by a mismatch. As in Experiments 1 and 2, each participant saw each verb with only one mismatch: either the same-effector case or the different-effector case. Mismatches were chosen out of the same pool of verbs as initial verbs. Each participant saw any given non-matching verb in the second position only once, as either a same-effector mismatch or a different-effector mismatch. Given these constraints, nonmatching verbs were assigned randomly to initial verbs. This divided participants into two groups, depending on which set of initial verbs they saw in the same-effector condition. The experimental software alternated between groups.

#### Results

One participant had an outlying mean RT (greater than three standard deviations from the global mean) and was removed from analysis. There was also 1 participant whose accuracy rate was lower than 80% overall and who was removed from analysis. Together, this resulted in the removal of 2 participants from the analysis. Items would have been removed if they had had outlying mean responses (greater than two standard deviations from the overall mean) or low overall accuracy (lower than 80%), but none met these criteria, and as a consequence, results from all items were analyzed.

By comparison with the first two experiments, there were far more errors in this study. Participants provided the correct answer for 99.1% of the different-effector mismatches, but only 93.5% of the same-effector mismatches and 87.8% of the near-synonyms. Performance on the different-effector mismatches was thus significantly more accurate than performance on either the near-synonyms ( $p < .001$ ) or the same-effector mismatches ( $p < .001$ ), and performance on same-effector mismatches was significantly more accurate than performance in the near-synonym condition ( $p < .001$ ).

Including only correct responses, the mean RT was 100 msec faster in the different-effector condition than in the same-effector condition (Table 1). This effect was significant in repeated measures ANOVAs, both by participants [ $F_1(1,54) = 30.27, p < .001, \eta_p^2 = .29$ ] and by items [ $F_2(1,41) = 10.80, p < .01, \eta_p^2 = .18$ ].

#### Discussion

Participants took significantly longer to reject a verb pair as near-synonyms when the two verbs shared an effector

than when they did not. This replicates the finding from Experiments 1 and 2, where participants took longer to reject a verb–image pair that did not match when the action referenced by the verb and the action depicted in the image shared an effector. These results provide evidence that the meanings of action verbs include in their representations information about bodily action—in particular, the body part used to perform the action. Moreover, the outcomes from these experiments cannot result from visual ambiguity of picture stimuli, because the present experiment showed the effect in a task where no images were presented.

#### EXPERIMENT 4 Image–Verb Matching in Second-Language Learning

The preceding experiments showed that language understanders engage motor representations while processing words and that this motor activation is effector-specific. We observed this effect regardless of the order of presentation of the image and verb stimuli, and even with purely linguistic stimuli. It appears when native English speakers process words in English and when native Cantonese speakers process words in Cantonese. But we have not investigated the extent to which native speakers of one language access perceptuomotor knowledge when processing words from another language.

Our final experiment investigated the development and use of motor knowledge by second-language learners. We ask whether nonnative English speakers display perceptuomotor activation effects parallel to those of native English speakers. We pursue two angles on this question below. First is the question of whether it is ever possible for nonnative speakers to activate perceptuomotor representations during language processing; it is quite plausible that they never do. For instance, fluent activation of embodied representations may be something that only native speakers do—that this is part of what distinguishes native from nonnative speakers of a language. On the other hand, highly proficient speakers of second languages often display great fluency in producing and understanding meaningful speech, and if activating perceptuomotor systems plays a part in these processes, highly proficient second-language users should show effects qualitatively similar to those of native speakers.

This leads to the second question under consideration. Gaining competence in a second language involves achieving increasingly native-like language processing capacities. But we currently do not know whether nonnative speakers with greater fluency in a second language come to more fluently access perceptuomotor representations. If so, more fluent nonnative language users should display larger (i.e., more native-like) effects when we test them using the paradigm from the three experiments above. So, we ask whether proficiency in a second language correlates with the size of a participant's image–verb matching effect.

#### Method

This final experiment used the design from Experiment 1 with nonnative (second-language) English speakers as the participants.

**Participants.** Participants were 40 right-handed, nonnative speakers of English enrolled in mainstream classes (i.e., not special classes dedicated to second-language learners of English) at the University of Hawai'i at Manoa. They participated in exchange for either course credit or \$5.

Participants varied in two important ways: native language and English proficiency. Participants for this experiment were drawn from a variety of native language groups, predominantly Japanese and other Asian languages. We did not aim to study speakers of only a single native language, because we saw no clear way that native language would influence the presence or magnitude of a same-versus different-effector effect.

The second dimension of variation across participants, English proficiency, is difficult to measure. All participants showed a high level of English proficiency; enrolling in mainstream classes at the University of Hawai'i Manoa requires a minimum TOEFL composite score of 173 (computer), 61 (Internet), or 500 (paper). In order to test for a relation between proficiency and the size of the image–verb matching effect, we needed an independent measure of English proficiency for each participant. We therefore included a word meaning test at the end of the experiment as a measure of their test-specific lexical knowledge. We used this information as described below to investigate whether participants' English proficiency correlated with the size of the interference effect they displayed. The idea that language acquisition in part involves learning to pair words with perceptuomotorically grounded representations of their meanings predicts such a correlation. The word meaning test was administered after the main experiment and was explained by the following instructions:

You will see an action verb. Please decide what body part the action uses:

- mouth
- hand/arm
- foot/leg

For example, "jump" uses the foot/leg.

Participants saw all 48 verbs used in the study in a random order and pressed buttons labeled *mouth*, *hand*, or *foot*. Response accuracy was measured, but latency was not.

**Procedure.** For the main experiment, the protocol was identical to that in Experiment 1, with the exception of the instructions, which were presented in simpler terms more accessible to nonnative speakers. Participants performed a forced choice task in which they decided whether images and written verbs depicted the same action or different actions. When the actions were different, the body part involved in the action (mouth, arm, or leg) was either "same" or "different." We reasoned that if nonnative speakers activate motor structures during understanding, they should display the same effect described for the preceding experiments: RTs should be longer when the actions are different but the involved effectors are the "same," as compared with the condition in which the effectors are "different."

**Materials.** The set of items used in the main experiment and the word meaning test included the same items as in Experiment 1, although we suspected that several would be unknown to even advanced nonnative English learners. In order to preserve the original design, these items were included with the intention of excluding outlying items prior to analysis.

#### Results

As was expected, responses were less accurate for this population than for the native English speakers who participated in the preceding experiments. Overall, 11% of responses were incorrect: 14% in the matching condition and 12% in the nonmatching same-effector condition, but only 3% in the nonmatching different-effector condition. Paired *t* tests showed that accuracy was significantly different be-

**Table 3**  
**Mean Participant Response Times and Standard Deviations (in Milliseconds)**  
**in Nonmatching Conditions by Effector and Combined for All Effectors**

Condition	Foot		Hand		Mouth		All Effectors	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Same effector	1,276	293	1,218	269	1,124	236	1,158	199
Different effector	1,039	210	1,113	212	1,056	257	1,043	205

tween the two nonmatching conditions ( $t = 8.67, p < .0001$ ) and between the matching and different-effector conditions ( $t = 10.53, p < .0001$ ), whereas the difference between the matching and same-effector conditions approached significance ( $t = 1.82, p = .07$ ). Four participants were excluded for performing the experimental task with less than 80% accuracy in any condition. One of the remaining participants had fewer than 10 correct items/responses in all conditions and was removed from further analysis. Results from the remaining 35 participants were analyzed.

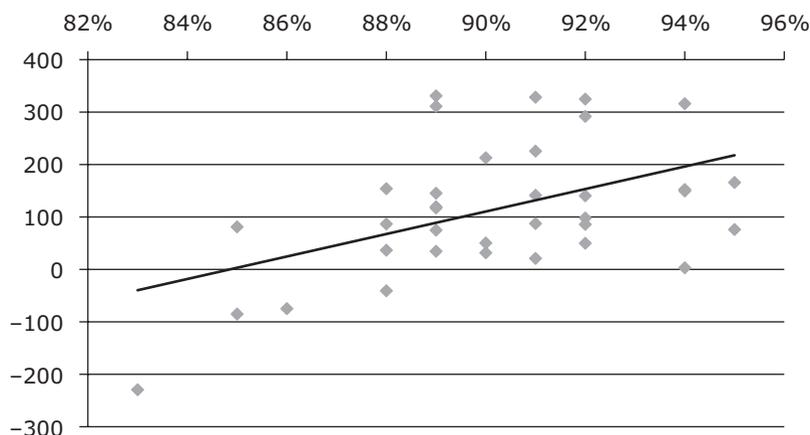
As shown by the low accuracy rates reported above, some words were unfamiliar, and we excluded from analysis 10 items (i.e., 10 images, in all three conditions) that had mean accuracy rates lower than 80%. Excluded verbs came relatively evenly from each of the effector groups: three for mouth, two for hand, and five for foot. We also removed all trials with RTs more than two standard deviations from the mean for that item. This resulted in the removal of less than 2.7% of the remaining data. For each participant, outlying RTs (those more than two standard deviations from the participant’s mean in a particular condition) were also eliminated.

The critical measure was the difference between the nonmatching conditions. Same-effector responses were 115 msec slower than different-effector responses (Table 3), a difference significant in a repeated measures ANOVA by participants [ $F_1(1,34) = 28.33, p < .0001, \eta_p^2 = .45$ ]. The items analysis of this same difference was marginally significant [ $F_2(1,37) = 3.89, p = .06, \eta_p^2 = .09$ ].

Turning now to the correlation between language proficiency and the size of the same–different difference, we had two measures of proficiency at hand. The first was the

word meaning test, which we conducted after the end of the main experiment for each participant. Participants performed well, overall: Accuracy rates varied between 75% and 100%, with a median score of 92%. This yielded sufficient variation for a regression analysis correlating the word meaning test accuracy rate for each participant with the size of their interference effect in the main experiment. The interference effect size was calculated as the difference for each participant between the mean RTs in the same-effector condition and the different-effector condition. Larger interaction effect sizes, reflecting greater degrees of interference, were expected to correlate positively with accuracy on the vocabulary test. The better the participant’s mastery of English, the more he or she should be activating effector-specific motor structures in performing the task. However, posttest accuracy did not correlate significantly with the size of the effect:  $\beta = .13, p = .46$ . Accuracy on the vocabulary test was not significantly correlated with the size of the effect in the main experiment. This lack of a correlation does not appear to have been due to the vocabulary test being insufficiently reliable; it, in fact, showed moderately high reliability (Cronbach’s  $\alpha = .72$ ).

We did, however, have a second available measure of proficiency: participants’ accuracy in the main experiment itself. Whereas the vocabulary test was untimed, allowing for slow, reflective responses, in the main experiment timed responses might have led to accuracy numbers that could have been a better measure of each participant’s proficiency in accessing word meanings online. Consistent with this reasoning, accuracy in the main experiment correlated significantly with the size of the effect participants displayed in it ( $\beta = .482, p = .003$ ), as shown in Figure 4, below.



**Figure 4. Linear regression correlating main experiment accuracy (on the x-axis) with same–different effect (on the y-axis).**

## Discussion

Although they unsurprisingly produced slower RTs than native speakers did in the previous experiments and were less accurate overall, nonnative English speakers performed like natives in that responses when the image and verb shared an effector were slower than when they did not share an effector. We have argued that this effect derives from inhibition between competing neural structures active when the participant must simultaneously process competing inputs. We can conclude that like native speakers, these nonnative speakers relied on motor structure activation to understand words.

The other result of this experiment is the finding that language proficiency (as measured by overall accuracy in the main experiment) correlated positively with effect size. The more successful participants were on the task (which most likely correlates with proficiency in the language in general), the more native-like their perceptuomotor interference effects were. One interpretation of this result is that learners progressively develop the ability to activate neural circuitry that supports grounded, modal representations when understanding language. It could be that early or less proficient learners of a language process words using mechanisms that are not grounded in perception and action—for instance, through lexical association or through shallow translation into their native language. As they become more fully proficient users of their nonnative language, the connections between words and the modal representations of their meanings become stronger and more easily activated, such that they perform more like native speakers in behavioral language tasks. In other words, the more deeply linguistic representations are grounded, the more language users understand, or, the more they understand, the deeper their linguistic representations reach.

There is, however, an alternative explanation for why participants who were less accurate showed a smaller interference effect. It could be that they were still accessing modal systems, just like native speakers, but that they were only able to do so for the words they knew, which were fewer in number. The fewer verbs they knew, the more frequently they were guessing in the main experiment, and as a result, the smaller the effect size was. The power of this explanation is tempered somewhat because all incorrect responses in the main experiment were excluded prior to calculation of the effect; these could, therefore, not have influenced its size, but correctly guessed (but unknown) words could have. There would have been relatively few of these, since all participants included in the analysis performed at more than 80% accuracy in the main experiment, which is much higher than chance (50%). If they were successfully guessing the meanings of the words in the main experiment—assuming they were at chance accuracy on those verbs they did not know—they could not have been guessing on more than 40% of the items overall (half of which would be wrong, yielding the minimum 80% overall accuracy). This would mean that up to a quarter of the actually included responses might have been the results of correct guesses. This could, in principle, have resulted in a dilution of the effect in less proficient participants by a

minority of unknown verbs. This explanation for the findings would lead us to reason that second-language learners can perform in native-like ways on some tasks (like the image–verb matching task) without necessarily engaging in the same understanding processes that native speakers rely upon—in this case, activating the perceptuomotor system. If this is indeed the case, the effect size gives us an indirect measure of whether nonnative speakers who are behaving like natives are using native-like response strategies or alternative ones. We leave open an exploration of potentially useful applications of such measures to second-language proficiency testing.

To sum up the results from this last experiment, as was expected, nonnative English-speaking participants were slower to respond when the image action and verb action did not match but the effector (mouth, hand, foot) was the “same,” rather than when it was “different.” We have reasoned that in order to compare visual and linguistic input, understanders engage neural structures related to action execution. This is the same effect that was found in the preceding experiments. The first contribution of this experiment is thus to show that advanced nonnative learners display the same interference effect. The second is to show a positive correlation between language proficiency and the interference effect size. We identified two possible explanations for this relation. The first proposes that mastering the vocabulary of a target language involves learning to directly access perceptuomotor systems during language processing. The second argues that second-language users access modal systems just as native speakers do and that the correlation was an artifact of the number of words whose meaning participants had to guess. In either case, we have seen that the greater a second-language learner’s mastery of the target language, the more of a native-like interference effect that learner displays.

## GENERAL DISCUSSION

The present set of experiments has shown that, when presented with images and verbs depicting different actions, participants are slower to decide that they do not match when the two actions make primary use of the same effector. Several competing explanations for the observed effect were considered and rejected above: visual ambiguity of the images and global semantic similarity between the denoted actions. The most plausible explanation is that our findings result from interference between perceptuomotor circuits representing distinct actions that use the same effector. This mutual inhibition account argues that both images and verbs activate neural structures that laterally inhibit each other; as was mentioned above, mirror circuits primarily used to perform or perceive depicted actions are prime candidates for such structures. The work reported here provides evidence for activation of modality-specific neurocognitive structures while language users process the meanings of verbs. It does so using a completely novel method—distinct from priming (Glenberg & Kaschak, 2002; Zwaan, Stanfield, & Yaxley, 2002) or previously observed interference effects (Bergen et al., 2007; Kaschak et al., 2005; Richardson et al., 2003)—and thus provides critical convergent evidence on

the function of motor and perceptual systems in language understanding.

A key feature of the methodology employed in the experiments reported on here is how observing an interference effect, rather than a facilitatory priming effect, for example, allows us to make strong conclusions about the function of perceptuomotor activation during the performance of language tasks. In each experiment, it took participants longer to process the meanings of action verbs when comparing them with nonmatching actions performed with the same effector than when comparing them with actions using a different effector. Since there is no reason to believe that the different-effector condition causes facilitation of responses, we have reasoned that the slower RTs in the same-effector condition are the result of interference. This finding complements those of neural imaging research showing that processing action verbs yields activation in effector-specific parts of motor cortex (Hauk et al., 2004; Pulvermüller et al., 2001; Tettamanti et al., 2005). Not only are these brain regions active, but the experiments above show that activated perceptuomotor circuits interfere with a range of lexical tasks. In its broadest interpretation, this suggests that accessing the meanings of action verbs includes activation of grounded, modal knowledge as a critical component. If activation of perceptuomotor circuitry in verb processing were only a downstream effect, unimportant to accessing word meaning, there would be no reason for participants to take longer to make matching decisions in the same-effector condition, as they were shown to above.

The neural mechanism we proposed to be responsible for the observed interference effect affords a clear, testable prediction. The mutual inhibition account should extend to tasks other than image-verb matching. If action verb processing activates motor representations for described actions, it should also delay performing another, competing action using the same effector, but not another action using a different effector. Whether or not this is the case remains to be empirically determined.

The activation of perceptuomotor circuits during language processing is consistent with the hypothesis that accessing the meaning of language about action in general makes use of the internal enactment of mental simulations of described action scenes, as various scholars have argued (Bailey, Chang, Feldman, & Narayanan, 1998; Barsalou, 1999; Glenberg & Robertson, 2000; MacWhinney, 2005; Zwaan, 1999). For instance, in order to understand an utterance like "John threw the water balloon," language users might activate a subset of the motor structures responsible for that particular type of throwing. Or they might engage their visual system to mentally simulate what it would be like to perceive someone throwing a water balloon. Other systems might plausibly be evoked: haptic knowledge of the feel of the rubber balloon in the thrower's hand or the water on the receiver's body, auditory knowledge of the sound of impact, or even affective knowledge of joy at a throw well aimed or of anger and embarrassment at having been targeted. Language users have the remarkable ability to make immediate inferences about actions they have heard described. For example, upon hearing "John

threw the water balloon," the understander can immediately answer questions about the hand shape used, the amount of pressure applied to the balloon, the trajectory of the arm, and so on, or can respond appropriately, with congratulations or empathy, among other things (Bergen & Chang, 2005).

There is substantial experimental evidence for this view, suggesting that both motor simulation (Bergen & Wheeler, 2005, 2010; Borreggine & Kaschak, 2006; Glenberg & Kaschak, 2002; Zwaan & Taylor, 2006) and perceptual simulation (Bergen et al., 2007; Kiefer, Sim, Herrnberger, Grothe, & Hoenig, 2008; Matlock, 2004; Richardson et al., 2003; Stanfield & Zwaan, 2001; Zwaan et al., 2002) are unconsciously and automatically engaged during language understanding. As was noted at the beginning of this article, evidence from brain imaging studies also shows that processing language associated with particular perceptual or motor functions activates the neural areas responsible for those same functions (Pulvermüller et al., 2001; Tettamanti et al., 2005). Modal brain areas in fact appear to serve a number of different uses, including recalling actions (Nyberg et al., 2001; Wheeler, Petersen, & Buckner, 2000) and modal imagery (Decety, Sjöholm, Ryding, Stenberg, & Ingvar, 1990; Ehrsson, Geyer, & Naito, 2003; Lotze et al., 1999; Porro et al., 1996; Roland, Larsen, Lassen, & Skinhøj, 1980; Sirigu et al., 1996). It is therefore no surprise that perceptual and motor systems are also engaged in the process of language understanding.

Together, the demonstrated use of low-level motor and perceptual structures for imagined action, recalled action, and, now, language processing lend credence to a view of meaning and thought tightly grounded in the experiences individuals have interacting with the world around them. The findings reported here support an embodied view of the meanings of linguistic units and the utterances they appear in: one in which motor language has meaning through reference to the modal experiences that the individual can evoke. It seems that the particularities of human neural circuitry are deeply explanatory of properties of the language understanding process.

#### AUTHOR NOTE

Our thanks to Jerome Feldman for helping to conceive of this line of experiments; to Nathaniel Smith, who provided a great deal of analytical and technical assistance; to Arthur Glenberg for invaluable consultation on the analysis; to Zachary Weinberg for his collaboration on Experiment 3; to Lori Yancura and Ashley Maynard for helpful comments on an early draft; and to reviewers for and audiences at the Berkeley Linguistics Society and Cognitive Science Society conferences, where early versions of parts of this work were presented. Correspondence concerning this article should be addressed to B. Bergen, Cognitive Science Department, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0515 (e-mail: bkbergen@ucsd.edu).

#### REFERENCES

- BAILEY, D., CHANG, N., FELDMAN, J., & NARAYANAN, S. (1998). Extending embodied lexical development. In M. A. Gernsbacher & S. J. Derry (Eds.), *Proceedings of the 20th Annual Conference of the Cognitive Science Society* (pp. 84-89). Mahwah, NJ: Erlbaum.
- BARSALOU, L. W. (1999). Perceptual symbol systems. *Behavioral & Brain Sciences*, *22*, 577-609.
- BERGEN, B. (2008). Experimental methods for simulation semantics. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson, & M. Spivey

- (Eds.), *Methods in cognitive linguistics* (pp. 277-301). Amsterdam: Benjamins.
- BERGEN, B., & CHANG, N. (2005). Embodied construction grammar in simulation-based language understanding. In J.-O. Östman & M. Fried (Eds.), *Construction grammars: Cognitive grounding and theoretical extensions* (pp. 147-190). Amsterdam: Benjamins.
- BERGEN, B., LINDSAY, S., MATLOCK, T., & NARAYANAN, S. (2007). Spatial and linguistic aspects of mental imagery in language comprehension. *Cognitive Science*, **31**, 733-764.
- BERGEN, B., & WHEELER, K. (2005). Sentence understanding engages motor processes. In B. G. Bara, L. W. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 238-243). Mahwah, NJ: Erlbaum.
- BERGEN, B., & WHEELER, K. (2010). Grammatical aspect and mental simulation. *Brain & Language*, **112**, 150-158.
- BORREGGINE, K. L., & KASCHAK, M. P. (2006). The action-sentence compatibility effect: It's all in the timing. *Cognitive Science*, **30**, 1097-1112.
- BUCCINO, G., BINKOFKI, F., FINK, G. R., FADIGA, L., FOGASSI, L., GALLESE, V., ET AL. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience*, **13**, 400-404.
- BUCCINO, G., RIGGIO, L., MELLI, G., BINKOFKI, F., GALLESE, V., & RIZZOLATTI, G. (2005). Listening to action-related sentences modulates the activity of the motor system: A combined TMS and behavioral study. *Cognitive Brain Research*, **24**, 355-363.
- DECETY, J., SJÖHOLM, H., RYDING, E., STENBERG, G., & INGVAR, D. H. (1990). The cerebellum participates in mental activity: Tomographic measurements of regional cerebral blood flow. *Brain Research*, **535**, 313-317.
- EHRSSON, H. H., GEYER, S., & NAITO, E. (2003). Imagery of voluntary movement of fingers, toes, and tongue activates corresponding body-part-specific motor representations. *Journal of Neurophysiology*, **90**, 3304-3316.
- ESTES, Z., VERGES, M., & BARSALOU, L. W. (2008). Head up, foot down: Object words orient attention to the objects' typical location. *Psychological Science*, **19**, 93-97.
- FELDMAN, J. (2006). *From molecule to metaphor: A neural theory of language*. Cambridge: MIT Press.
- FELDMAN, J., & NARAYANAN, S. (2004). Embodiment in a neural theory of language. *Brain & Language*, **89**, 385-392.
- GALLESE, V., FADIGA, L., FOGASSI, L., & RIZZOLATTI, G. (1996). Action recognition in the premotor cortex. *Brain*, **119**, 593-609.
- GALLESE, V., & LAKOFF, G. (2005). The brain's concepts: The role of the sensory-motor system in reason and language. *Cognitive Neuropsychology*, **22**, 455-479.
- GLENBERG, A., & KASCHAK, M. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, **9**, 558-565.
- GLENBERG, A., & ROBERTSON, D. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory & Language*, **43**, 379-401.
- HAUK, O., JOHNSRUDE, I., & PULVERMÜLLER, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, **41**, 301-307.
- HELBIG, H. B., GRAF, M., & KIEFER, M. (2006). The role of action representations in visual object recognition. *Experimental Brain Research*, **174**, 221-228.
- KASCHAK, M. P., MADDEN, C. J., THERRIAULT, D. J., YAXLEY, R. H., AVEYARD, M., BLANCHARD, A. A., & ZWAAN, R. A. (2005). Perception of motion affects language processing. *Cognition*, **94**, B79-B89.
- KIEFER, M., SIM, E.-J., HERRNBERGER, B., GROTHE, J., & HOENIG, K. (2008). The sound of concepts: Four markers for a link between auditory and conceptual brain systems. *Journal of Neuroscience*, **28**, 12224-12230.
- KOSSLYN, S. M., GANIS, G., & THOMPSON, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, **2**, 635-642.
- LANDAUER, T. K., & DUMAIS, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, **104**, 211-240.
- LANDAUER, T. K., FOLTZ, P. W., & LAHAM, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, **25**, 259-284.
- LOTZE, M., MONTOYA, P., ERB, M., HÜLSMANN, E., FLOR, H., KLOSE, U., ET AL. (1999). Activation of cortical and cerebellar motor areas during executed and imagined hand movements: An fMRI study. *Journal of Cognitive Neuroscience*, **11**, 491-501.
- MACWHINNEY, B. (2005). The emergence of grammar from perspective taking. In D. Pecher & R. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thinking* (pp. 198-223). Cambridge: Cambridge University Press.
- MASSON, M. E. J., BUB, D. N., & WARREN, C. M. (2008). Kicking calculators: Contribution of embodied representations to sentence comprehension. *Journal of Memory & Language*, **59**, 256-265.
- MATLOCK, T. (2004). Fictive motion as cognitive simulation. *Memory & Cognition*, **32**, 1389-1400.
- NYBERG, L., PETERSSON, K. M., NILSSON, L. G., SANDBLOM, J., ÅBERG, C., & INGVAR, M. (2001). Reactivation of motor brain areas during explicit memory for actions. *NeuroImage*, **14**, 521-528.
- PORRO, C. A., FRANCESCATO, M. P., CETTOLO, V., DIAMOND, M. E., BARALDI, P., ZUIANI, C., ET AL. (1996). Primary motor and sensory cortex activation during motor performance and motor imagery: A functional magnetic resonance imaging study. *Journal of Neuroscience*, **16**, 7688-7698.
- POSTLE, N., MCMAHON, K. L., ASHTON, R., MEREDITH, M., & DE ZUBICARAY, G. I. (2008). Action word meaning representations in cytoarchitecturally defined primary and premotor cortices. *NeuroImage*, **43**, 634-644.
- PULVERMÜLLER, F., HÄRLE, M., & HUMMEL, F. (2001). Walking or talking? Behavioral and neurophysiological correlates of action verb processing. *Brain & Language*, **78**, 143-168.
- RICHARDSON, D. C., SPIVEY, M. J., BARSALOU, L. W., & McRAE, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science*, **27**, 767-780.
- RIZZOLATTI, G., & CRAIGHERO, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, **27**, 169-192.
- RIZZOLATTI, G., FADIGA, L., GALLESE, V., & FOGASSI, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, **3**, 131-141.
- ROLAND, P. E., LARSEN, B., LASSEN, N. A., & SKINHØI, E. (1980). Supplementary motor area and other cortical areas in organization of voluntary movements in man. *Journal of Neurophysiology*, **43**, 118-136.
- SIRIGU, A., DUHAMEL, J.-R., COHEN, L., PILLON, B., DUBOIS, B., & AGID, Y. (1996). The mental representation of hand movements after parietal cortex damage. *Science*, **273**, 1564-1568.
- STANFIELD, R. A., & ZWAAN, R. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, **12**, 153-156.
- TETTAMANTI, M., BUCCINO, G., SACCOMAN, M. C., GALLESE, V., DANNA, M., SCIFO, P., ET AL. (2005). Sentences describing actions activate visuomotor execution and observation systems. *Journal of Cognitive Neuroscience*, **17**, 273-281.
- TSENG, M., & BERGEN, B. (2005). Lexical processing drives motor simulation. In B. G. Bara, L. W. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 2206-2212). Mahwah, NJ: Erlbaum.
- WHEELER, M. E., PETERSEN, S. E., & BUCKNER, R. L. (2000). Memory's echo: Vivid remembering reactivates sensory specific cortex. *Proceedings of the National Academy of Sciences*, **97**, 11125-11129.
- WINDHORST, U. (1996). On the role of recurrent inhibitory feedback in motor control. *Progress in Neurobiology*, **49**, 517-587.
- WINTER, D. A. (1990). *Biomechanics and motor control of human movement*. New York: Wiley.
- ZWAAN, R. A. (1999). Embodied cognition, perceptual symbols, and situation models. *Discourse Processes*, **28**, 81-88.
- ZWAAN, R. A., STANFIELD, R. A., & YAXLEY, R. H. (2002). Do language comprehenders routinely represent the shapes of objects? *Psychological Science*, **13**, 168-171.
- ZWAAN, R. A., & TAYLOR, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General*, **135**, 1-11.