Summer 7-22-2014

Language Through the Body: The Grounding of Motor Language Processing

Mariah Kristine Hamang
University of Colorado Boulder, mariah.hamang@colorado.edu

Follow this and additional works at: http://scholar.colorado.edu/ling_gradetds
Part of the Cognitive Psychology Commons, and the Linguistics Commons

Recommended Citation

This Thesis is brought to you for free and open access by Linguistics at CU Scholar. It has been accepted for inclusion in Linguistics Graduate Theses & Dissertations by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
LANGUAGE THROUGH THE BODY:
THE GROUNDING OF MOTOR LANGUAGE PROCESSING

by

MARIAH K. HAMANG

B.A., Indiana University Northwest, 2012

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirement for the degree of
Master of Arts
Department of Linguistics
2014
This thesis entitled:
Language Through the Body: The Grounding of Motor Language Processing
written by Mariah K. Hamang
has been approved for the Department of Linguistics

___________________________________
(Dr. Bhuvana Narasimhan)

___________________________________
(Dr. Eliana Colunga)

Date: July 14, 2014

The final copy of this thesis has been examined by the signatories, and we
Find that both the content and the form meet acceptable presentation standards
Of scholarly work in the above mentioned discipline.

IRB protocol # 130555
Abstract

Hamang, Mariah K. (M.A., Linguistics)

Language Through the Body: The Grounding of Motor Language Processing

Thesis directed by Associate Professor Dr. Bhuvana Narasimhan

In contrast to theories of language which conceive of linguistic representations in the mind as amodal symbols that are divorced from the rest of cognition, embodied theories of language contend that knowledge of semantic concepts, relations, and linguistic cues are grounded within the holistic cognitive lives of language users, giving rise to linguistic knowledge that is fundamentally connected to experiences of perception, sensation, action, and other worldly encounters. Empirical evidence, both neurological and behavioral, has been emerging in support of an embodied view of language processing, which encompasses the notion of mental simulation and imagery in response to linguistic stimuli. This paper presents original experimental research that investigates a) how an ongoing mental simulation can affect the processing of a sentence which evokes a new mental simulation, b) the degree of detail in mental simulations of the body that are activated by language of motion, and c) whether the subject pronoun of a sentence can encourage a reader to adopt a particular perspective in their simulation—either as a participant or as an observer—and how this perspective affects their processing speeds. Although statistically inconclusive, the results of the experiment suggest that mental simulations are involved during the processing of language and that when two closely related mental simulations are evoked nearly simultaneously, there occurs a slight processing interference. There was no statistically appreciable effect of the perspective of the subject pronoun on processing.
Dedicated to Thelma J. “Meemee” Ritchie,
The earliest inspiration for my goals and accomplishments in education,
Who motivated me to the end by her own example.

I am forever grateful.
Acknowledgements

First and foremost, I must offer my deepest appreciation to my advisor Dr. Bhuvana Narasimhan, whose feedback and insight have been invaluable throughout this experience, especially through the rough patches. I am profoundly grateful to have had such a patient, knowledgeable, and attentive mentor to hear my ideas and break ground with me. She has a remarkable ability to challenge and encourage simultaneously, and I have learned and accomplished more than I ever could have without her supportive spirit.

In addition, I would like to thank the members of the committee, Dr. Laura Michaelis-Cummings and Dr. Eliana Colunga, whose assistance throughout the research process has kept me critical of my own research. The support of my committee and of the Department of Linguistics has been comprehensive, ranging from community to financial sponsorship, and I am honored to be in a position to acknowledge them.

Gratitude is also in order for the Professional Research Assistants of Dr. Tim Curran’s lab, Chris Bird and Krystin Corby, who made hassle-free arrangements to lend me a millisecond-accurate keyboard that was crucial to the precision of the experimental data. Also fundamental to the experiment was Kraig Ewert, who hand-drew all the line drawings that served as the stimuli. I am grateful for his artistic skills, for his friendship, for his offer to assist.

I must also recognize my friends and colleagues, both within the Department of Linguistics and without, whose advice and open ears create an unparalleled sense of solidarity. I would also like to express my thanks and admiration for my family and the many shapes and sizes of their lifelong support. Finally, I owe a special form of gratitude to Joe Bartholomew for enduring life with me behind the scenes.

Thank you.
Contents

Abstract .......................................................................................................................................... iii
Dedication...................................................................................................................................... iv
Acknowledgements ......................................................................................................................... v
Contents ......................................................................................................................................... vi
List of Tables ................................................................................................................................ vii
List of Figures .............................................................................................................................. viii
Introduction..................................................................................................................................... 1
Empirical Background: Theories and Findings ................................................................................ 8
    Brain Activation During Language of Motion and the Perception of Action ......................... 8
    Behavioral Motor Responses to Processing Motion Language .............................................. 15
Methods......................................................................................................................................... 25
    Participants.............................................................................................................................. 25
    Procedure ................................................................................................................................ 25
    Stimuli ..................................................................................................................................... 28
    Design ..................................................................................................................................... 30
Results ........................................................................................................................................... 33
Discussion ..................................................................................................................................... 40
References ..................................................................................................................................... 46
Appendices.................................................................................................................................... 51
    Appendix A – List of Sentences (Example of One Pronoun Set in “I” Condition) ............... 51
    Appendix B – Complete Distribution of Match, Mismatch, and Neutral Conditions Across All Stimulus Lists (Latin-Square) ..................................................................................... 53
    Appendix C – Distribution of Subject Pronouns Across Six Participants .............................. 54
    Appendix D – List of Images .................................................................................................. 55
List of Tables

Table 1: Distribution of Handshape Images and Sentences in Target Conditions..........................28
List of Figures

Figure 1: Average Response Time in Match, Mismatch, and Neutral Conditions....................... 34
Figure 2: Average RT in Match, Mismatch, and Neutral Conditions (with Error Bars)...............34
Figure 3: Average Response Time by Handshape Type in Target Conditions............................. 35
Figure 4: Average RT by Handshape Type in Target Conditions (with Error Bars).....................36
Figure 5: Average Response Time per Subject Pronoun in Target Conditions............................ 37
Figure 6: Average RT per Subject Pronoun in Target Conditions (with Error Bars)....................37
Introduction

Imagine you are sitting on a park bench. It is a warm, sunny day, and your eyes are closed; you hear music from the apartment above you, and you smile. When you open your eyes, you see a big red playground and a woman going for a run along a trail. Or, imagine you are a professional baseball player, and you’re up to bat. The ball flies at you at 80 mph, and you swing your arms and send the ball to the outfield. Or, you’re getting out of your car and you slam your finger in the car door. These scenarios surely create some degree of mental imagery as we picture ourselves in these situations, some more vivid than others. But where does mental imagery come from? Can it tell us more about how we think, or how we develop? Is it possible that we’re using our prior knowledge of worldly movements and experiences in constructing these mental scenes? And even more provoking, could we be activating the same brain structures we use to physically smile or see a playground or feel pain from errant car doors when we imagine these actions?

Evidence has gradually been emerging to support the idea that indeed, the accumulation of our bodily interactions with the world may underlie, to a degree of uncertain functional significance, our ability to reason about the world (or imaginary worlds), to understand and produce language about the world (or imaginary worlds), and to imagine ourselves doing things in the world that we may or may not be able to do in reality, such as hitting a fly ball from an 80 mph pitch. To argue that memories of sensorimotor experiences are functional to any degree in cognitive acts such as language use or considering hypothetical consequences to arrive at a decision is to claim that these cognitive functions would be different, impoverished, or even impossible if they were not informed by these prior bodily experiences and encounters. This notion has come to be known as the theory of embodied, or grounded, cognition (Jeannerod
One of its main tenets is the concept of simulation, the process of mentally re-creating the scenario, the procedures, and the motions of some event or process (Kahneman & Tversky 1982; Grèzes & Decety 2001; Zwaan 2009). This representation would be what we may experience as mental imagery. Although mechanisms of mental simulation may have originally evolved to aid in action control and practice (perhaps as a form of action visualization) (Pezzulo 2011), evidence suggests that simulation has been exapted for other extended cognitive functions, such as empathizing with others and social cognition (Jackson et al. 2005; Goldman 2006), understanding the goals and intentions of others (Blakemore & Decety 2001), and language use and comprehension (Glenberg & Robertson 2000). Constructing mental scenes, even in the form of rudimentary fragments and loose associations, helps us to better understand the stories and experiences of others as they unravel. This is perhaps because our own physical and social interactions with our environment help to mediate meaning between the shared, arbitrary symbols of language and their users.

Embodied cognition as it relates to language processing assumes that words and sentences have meaning determined at least partially by connections that have developed between prior bodily experience (seeing, hearing, feeling, moving, breathing, having emotions, etc.) and language associated with such experience. Furthermore, according to the theory, encountering such language activates these associations as the utterance is perceived and processed. Sensory or perceptual information related to the words is then available to use in constructing the simulation that aids and enhances understanding, within milliseconds’ time of encountering the stimulus and often well below our conscious awareness.

Theories of language processing that embrace embodiment stand in contrast to amodal theories of linguistic cognition in which linguistic representations and capacities—semantic,
syntactic, and phonological—exist within a cognitive system which works according to principles that are independent of perceptual and sensory systems. Within the field of cognitive science, these amodal approaches have been intimately tied to the emergence of computer programming, in which a pre-specified code carries out the functions of the linguistic expressions, and indeed, the analogy of the mind as a computer aptly describes amodal theories of language such as those put forth by Fodor & Pylyshyn (1988) and Noam Chomsky (1957). Other amodal theories propose that knowledge of semantic items and relations are contained within lexical entries that specify lists of characteristic features and syntactic roles to guide understanding (Jackendoff 1990), but these entries are not embedded in the experiences or wider cognitive life of the language user beyond the insulated system of their language. Instead of these computational and amodal representations, grounded theories of language integrate knowledge of semantic concepts and linguistic cues into the broader sensory, perceptual, and physical experiences of the individual.

The experiment performed for this thesis extends current behavioral findings that support the notion of the functional involvement of motor simulation during the comprehension of linguistic input—that is, mental simulation that has a physical, embodied foundation in neural circuits related to those that perform actual movement. It relies crucially on the claim that seeing functional actions, or body parts arranged in functional ways, induces motor simulation that would theoretically compete for the same neurological areas and resources that become active during semantic and syntactic processing of motion or spatial language that depict these actions. This experiment exploits this postulate of embodiment theory, using images of functional handshapes (in this case, an open palm, a pinch, and a fist) to induce a mental simulation as a prime for the processing of a sentence that expresses motion using a functional handshape, such
as clapping (open palm), sewing (pinch), or boxing (fist). The hypothesis is that people processing a sentence that denotes a particular hand action will be slower to do so when they have been immediately primed with an image of a handshape that is incongruent with that expressed in the sentence (interference), and will be faster to do so when they have been primed with a congruent handshape (facilitation), as compared to the neutral condition of the handshape sentence being primed with a non-handshape image.

This experiment addresses mental simulation at a fine-grained and nuanced level. Prior research, to be discussed in more depth below, has been concerned with interference or facilitation of the participants’ actual physical movements in response to a stimulus (as when a response is faster when the handshape required by the participants’ response is congruent with that involved in the stimulus sentence [Wheeler 2010; Masson 2008]). Other work has considered interference and facilitation at the level of the effector (that is, the whole body part area that responds to the same stimulus in the same area of motor cortex) between two actions that are being intentionally compared (such as a slower response time to judge that the word “kick” is not the action depicted in a line drawing of a figure running, an interference at the “leg” effector [Bergen et al. 2010]). The goal of the present experiment is to investigate these interference and facilitation effects between two mental simulations induced by an image and a sentence alone without any conscious awareness or deliberation about them by the participant. It also aims to explore more deeply the level of detail and robustness involved in bodily mental simulations that are triggered by stimuli and an experimental task that involve, in theory, activation in the brain not from significant bodily movement but from representational and semantic content alone. The experiment evaluates the effects as they occur between three different types of functional handshapes, the hand being a very particular area of the hand and
arm effector, and asks a) whether congruent, matching handshapes are processed as an advantage to understanding in tasks where the simulations overlap and b) whether mismatching depictions of handshapes are processed as incongruent, impeding performance in such tasks.

Should the expected interference and facilitation results occur, the findings would suggest that the mental simulations resulting from both the image, a visual depiction, and the sentence, a closely related linguistic input, were somehow important or operational in processing both stimuli. If the simulations were not functional, we would not expect to see any difference in response time, as both simulations, if they even occurred, could both exist in parallel and not affect the processing of either stimulus. Interference might occur because the image makes one handshape more salient in working memory, suppressing the brain’s ability to process the other handshapes as quickly because of its preoccupation with the first stimulus. It is perhaps a matter of neurological resources in competition to process related but incompatible stimuli. On the other hand, facilitation might occur because the saliency of the handshape representation resulting from the first stimulus is reinforced by or prepared for the subsequent stimulus without any conflict between the active simulation from the prime and the incoming sentence, accelerating sentence processing. The results would also indicate that motor simulation may not only activate the level of the whole effector, or body part area, but may be rather fine-grained and specific in making distinctions between different handshapes associated with different functions. The further implication is that the neurological circuitry of bodily movement and of the physical or spatial relations that are abstracted from embodied interactions throughout development may be involved in the semantic and conceptual representations of the actions and of language about the actions.
An additional consideration of the present experiment is that of the perspective conveyed by the sentence as encoded by its personal subject pronoun. Prior work has shown that a sentence that encourages its reader to adopt a participant perspective (with second-person “You” or first-person “I” as the subject) is responded to faster than one which promotes an observer’s perspective (with third-person “He” or “She” as the subject) (Brunyé 2009). Brunyé’s research, discussed below in more detail, also found, however, that a sentence with “I” as the subject can be manipulated to induce an observer’s perspective. Given the flexibility of the first-person subject pronoun in its evocation of the participant perspective, the current research hypothesizes that the use of the second-person pronoun “You” as the subject in sentences such as “You stitched the dress” will in general result in faster processing of the sentence than either the first-person or third-person subjects.

This thesis begins with an overview of the relevant neurological, behavioral, and linguistic evidence and research related to simulation and simulation semantics, the particular type of simulation that is brought about by encounters with linguistic utterances. The methods, procedures, and design of the experiment are then described in detail, followed by a discussion of the results and their potential implications for cognition and language comprehension. To preview, while this behavioral experiment does not directly address the neurological underpinnings of simulation effects, it does appear to find indications of simulation effects. Unfortunately, the trends yielded from the data are not statistically significant. However, they suggest that motor simulation, grounded in the brain’s motion circuitry, is relevant and meaningful to understanding language of motion and action in some potentially surprising ways. The data also offer some preliminary evidence that these motor simulations may be strikingly detailed, such that simulations potentially distinguish between different functional configurations.
of body part effectors and the actions they afford. The role of the subject pronoun and the perspective it evokes does not seem to affect processing speeds, at least within the current empirical design.
Empirical Background: Theories and Findings

Brain Activation During Language of Motion and the Perception of Action

It is important to note that mental simulation is not always a conscious process, and its functional significance remains a point of debate. The actual mechanisms responsible for simulation are also ill-defined, although several essential neurological findings have advanced the discussion in rather exciting directions. For instance, fMRI evidence shows that the same areas of motor cortex that are activated when performing an action are active when a person perceives someone else performing the action (Hari et al. 1998), suggesting that the motor areas may contribute to imitation in early development and beyond. These same areas of motor movement also become active when a person reads about the particular action (van Elk et al. 2010). These responses are effector-specific, which means that seeing someone smile activates the mouth area of the motor cortex, reading about someone kicking a ball activates the foot and leg area of the motor cortex, etc. (Carillo-de-la-Peña 2006). The activation is also more pronounced when the perceived action is purposeful and goal-directed, as opposed to a meaningless, non-functional movement (Decety et al. 1997).

This finding of motor activation in the presence of, not physical motor movement itself, but of language about motion or by perceiving someone else’s motion, is intimately tied to the notion of a mirroring system, or mirror neurons, which is, in theory, common to the primates and results in the activation of brain areas responsible for motor movement when seeing the action performed (Rizzolatti & Craighero 2004; Cattaneo & Rizzolatti 2009). Evidence of mirror neurons was first discovered in macaque monkeys, who were found to demonstrate activation of particular neurons in the frontoparietal lobe and specifically in the inferior parietal lobule—roughly equivalent to Broca’s area in the human brain, known for its role in language production.
and comprehension—upon viewing a human researcher interact meaningfully, or intentionally, with a food item, such as grasping it, manipulating it, and placing it down (Gallese et al. 1996). These same neurons had been shown to fire during the physical execution of “goal-directed hand and mouth movements” by the monkeys (Gallese et al. 1996). Later research showed that a small population of the neurons (15%) even responded to meaningful action-related sounds in isolation, such as the cracking of peanut shells (Kohler et al. 2002).

The claim that analogous mirror neurons are present in humans has remained unfalsifiable because the invasive brain procedures used to measure the macaques’ neurons are an ethical concern for human research. Another restriction to consider is the poor resolution of currently available human brain imaging techniques—fMRIs, PET scans and their ilk can show activation in a particular area, but are not specific enough to show individual neurons within areas of activation. It is therefore problematic to claim with confidence that the precise neural circuits of the motor cortex or of the mirror neurons are what we see activated on fMRI imagery. Even if it is the exact same circuitry that activates when performing an action and when thinking about or perceiving an action, it is still unclear how executive control mechanisms subdue the compulsion to act when simulating the action (Jeannerod 1994; for selective prefrontal cortex activity during imitation tasks, see Decety et al. 1997).

Although the direct parallel between the differing species’ neural physiology remains untestable, it seems increasingly likely that mirror neurons are shared amongst the primates and are similarly manifest and functional in human cognition (Gallese 2003; Rizzolatti & Craighero 2005). The experimental findings above—that brain areas corresponding to a motor action are active when perceiving or reading about the action, and that their activation is in part dependent upon the characteristics of the action (e.g. whether it is associated with a goal-directed or
meaningful function)—do seem to indicate that a type of mirror neuron system may contribute to such activation patterns. A mirroring mechanism could very well underlie processes of mental imagery, motor simulation, and the development of the intersubjective experience which factors into empathy and social cognition. The available evidence seems to support the theory of mirror neurons, and indeed, initiatives in the study of embodied cognition are afforded a much deeper explanatory power if we accept the presence of a mirror neuron type of system, at least until a better understanding of the parallel between macaque mirror neurons and human neurology is achieved.

Even if we assume the existence of mirror neurons, there remains a major unresolved issue in regards to the functionality or necessity of such a system and of the simulation it enables. Would our understanding of a description of a fight scene in a story be different, or under-informed, if these brain areas were not active? Of course visualizing ourselves performing some type of intentional action, such as learning a particular dance move from watching and imitating someone else’s demonstration, involves a more deliberate simulation, but the activation of these effector-specific brain areas in e.g. the processing of language that denotes using those areas seems questionable and could be potentially unnecessary to understand language. It could be the case that this sensorimotor activation in the cortex is an artifact of the semantics of the language, a non-operational “downstream” effect that serves no useful function in the processing of the sentence (Bergen et al. 2010).

Certainly another concern raised by the theory is the questions it may or may not imply for situations of people with congenital disability, such as deafness or blindness. While the area remains undoubtedly open to investigation, the idea of grounded cognition does bring to bear the potentially divisive question of whether or not e.g. a congenitally deaf individual exhibits
patterns of cortex activation or simulation that differ from a hearing individual in processing
language related to hearing experiences. Would this imply that people with physical disabilities
from birth may have different mental representations associated with their affected physical
impairment, and that they therefore may even understand language about particular bodily
experiences with a conceptualization that differs from people without the disability? Again, the
possibility that the activation of motor cortex during the processing of motion language could be
a non-functional artifact of language comprehension cannot be ignored. Still, as will be discussed
in more detail below, experimental evidence is beginning to suggest that perhaps these patterns
of brain activation are, in fact, somehow meaningful in comprehending sensorimotor language.

The most crucial aspect of embodied cognition for the present research is its potential
role in language processing. It has already been mentioned that brain imaging studies reveal
effector-specific activation in motor cortex when reading about a particular action (van Elk et al.
2010), giving a neurological foundation to the possibility of simulation during language
comprehension. But this is not the only research that finds a relation between language
processing and brain activity in motor areas. Jirak et al. (2010) give an extensive meta-analysis
of research that corroborates these findings, including seminal work by Pulvermüller (2005)
which discovered unique activation patterns when mouth, hand, and foot actions, such as “lick”,
“pick”, and “kick”, were respectively displayed during a lexical decision task. The patterns of
brain activity found in this experiment correspond to the somatotopic organization of the motor
cortex, which refers to its configuration as a strip with areas that are identifiably associated with
the various parts of the body (Knierim 2010). Thus, in Pulvermüller’s research (2005), reading
“lick” resulted in activation in the mouth area of the cortex, “pick” in the hand area, and “kick”
in the foot area. Tettamanti (2005) found equivalent results when people read full sentences denoting particular mouth, hand, and foot actions.

Also mentioned above was the research of Hari et al., who found somatotopically organized brain responses in participants who watched the performance of an action by someone else (1998). We related this finding to the theory of mirror neurons, and indeed, the motor cortex shares the same general vicinity of that expected by the human mirror neuron system—the responses from the macaque monkeys were measured in the ventral premotor cortex (Rizzolatti 1996), a widely functional area of behavioral motor planning and execution in non-human primates. The patterns of brain activation in humans occur in the premotor cortex (an area important to the planning of action), the primary somatosensory cortex (the crucial receptor of haptic, or touch, stimuli), the supplementary motor area (which contributes to the control of physical movement), and the inferior parietal cortex (a major sensory receptor) (Molenberghs 2009). These areas have also been found to be active during both mental motor imagery (Kosslyn et al. 2007) and the recall of previously performed actions (Nyberg et al. 2001). The activation of these important motor areas in such a wide array of perceptual, sensory, and mental tasks suggests that the same neurological structures responsible for physically controlling and executing actions may play a functional role in a number of facets of general cognition, including language use and comprehension.

What would it mean for language processing to involve the same circuits used to move around when we hear or read language about moving around? Such a model would ascribe some degree of groundedness and experience to linguistic meaning and representation that has been largely dismissed because of issues such as those that arise when considering abstract and metaphorical language or congenitally disabled yet cognitively and linguistically competent
individuals, a criticism which was debated above. It seems that mental simulation is a relatively well-defended mechanism of cognitive reasoning at large, insofar as it is involved in spatial reasoning (Sima et al. 2013), categorization (Barsalou 1999), analogy (Geminiani et al. 1995), and other components of cognition. Mental imagery also underpins a number of influential cognitive theories concerned with the encoding and processing of language and action, such as dual coding theory (DCT), which hypothesizes that visual information is coupled with verbal associations as individuals learn and give meaning to environmental information (Paivio 1991), or the common coding theory—an alternative to DCT which claims that action and perception share a common representation, such that seeing and performing an action both activate concepts, linguistic or otherwise, of associated perceptual events that constitute the prior knowledge relied on by both functions to provide understanding (Prinz 1990; van der Wel et al. 2013).

While in the present work these theories and others like them will not be dealt with in much critical detail, the point remains that linguists and also other interdisciplinary researchers, such as psychologists and neuroscientists, have begun to consider how the perceptual and sensory experiences of the body during development may impact the cognitive and linguistic associations we use to guide our understanding. It is apparent by now that language is a primary stimulus that triggers mental imagery and mental simulation. If it is the case that mental simulations are evoked as we process language, then it should also hold true that a preexisting mental simulation should encourage processing of compatible linguistic input, and that to interrupt or manipulate the mental simulation should disrupt the language processing to a measurable degree.
These principles of behavioral interference and facilitation underlie the following experiment, which intends to empirically investigate the extent to which the activation of a motor representation in response to a stimulus image of a body part interferes with or facilitates processing language about action related to that body part. Specifically, it predicts that after seeing an image of a particular functional handshape (i.e., a fist, an open palm, or a pinch), a person should be slower to process a sentence that denotes an action requiring a different handshape as compared to conditions in which the sentence is preceded by a neutral, non-bodily image, such as a tree or a stoplight. After a person perceives the image of the hand, their active motor cortex effectors are theoretically simulating the experience of making, using, and seeing the handshape, given what we know about mirror neurons and visual observation. When the subject is immediately presented with the task of processing a sentence denoting a hand action which requires a different, mismatching handshape (e.g. “I slapped the liar” after seeing an image of a fist), the simulation should cause interference, i.e. a slower response time, in processing the sentence than if the handshape of the image matches that denoted in the sentence, or if a neutral (non-body part) image is shown. (Processing the sentence is operationalized here as a response to a judgment of sensibility.) It is possible that this interference occurs because the handshape simulation from the image is active and salient in working memory, creating neurological and representational competition and, therefore, suppression between closely related pathways that try to immediately process other handshapes before the first simulation has decayed.

Likewise, when the subject is required to immediately process a sentence denoting a hand action which requires a compatible handshape that matches that depicted in the prime, we predict that facilitation will occur, as the active simulation from the prime corresponds to the simulation
evoked by the following sentence. Even if the same neural resources were being used for both simulations, this would not create competition that impedes processing, but would instead allow the sentence to be processed faster, since the simulation from the new stimulus is already active from the older stimulus.

This experiment empirically addresses the extent to which a theory of fine-grained and highly-specified mental simulation is appropriate and defensible. Should an interference effect occur, this would indicate that the mental simulation triggered by the handshape image is detailed enough to include the shape the hand was making when it was perceived, and to cause a functional interference when the motor effector “re-activates” itself to process and conceptualize language about an action with a different handshape. In this case, the conflict results from the simulation of both the image and the sentence as being fine-grained. If the simulation were not fine-grained, activating the hand area of the motor cortex in general would likely not affect or may even facilitate, or accelerate response time, even if the handshapes from the image and from the sentence were different. Assuming the presence of the expected effects, we could then speculate that the same neural resources are being used to create simulations for both functions—both the processing of the image and the processing of the sentence.

**Behavioral Motor Responses to Processing Motion Language**

Although we have seen that brain imaging has revealed some intriguing patterns of activity in motor cortex in response to language of motion, we have also seen that researchers have hesitated to draw hasty conclusions from the discovery of this increased localized brain activity. The activation of these respective areas does not necessarily indicate that the motor cortex is serving a functional purpose in the processing of perceived linguistic or physical stimuli (Bergen 2010). The areas along the motor cortex that appear to be firing in these imaging studies
may be non-operatively activated as a downstream, non-functional effect. They may also simply be nearby circuits that are performing different, unrelated functions. Given the limitations of this technology, it is currently impossible to know the true role of these activated perceptuomotor areas in abstracting meaning from language if we rely exclusively on fMRI or EEG results.

Therefore, it has been necessary to examine via more indirect, experimental means the role of motor simulation and the motor cortex in processing language and deriving meaning. A number of studies have cleverly shown that the sensorimotor system appears to be inherently functional in the execution of linguistic operations, including response to linguistic stimuli. Bergen and colleagues have provided convincing evidence that embodied mental simulation plays an operative role in language processing, using empirical paradigms such as Action Compatibility Effect (ACE) experiments (Bergen 2007). One such study (Bergen 2010) presented participants with sentences that were syntactically comparable but which semantically indicated a particular orientation of an object, such as “John hammered the nail into the wall,” (horizontal orientation) or, “John hammered the nail into the floor” (vertical orientation). Subjects were then shown an image of a nail and told to identify whether the depiction was, in fact, a nail. Of course, every image of the nail is a nail, but participants were quicker to respond affirmatively when the image corresponded to the orientation of the object described in the sentence (a vertical nail for “into the floor,” and a horizontal nail for “into the wall”). The implication is that upon reading the sentence, the subjects were, in milliseconds’ time, re-creating a scene in which a nail is being hammered in a particular direction, thusly embodying the motor experience and projecting themselves into the described scene, as either an observer or a participant. Furthermore, it seems that the representation they simulated in response to the linguistic stimuli was salient enough to expedite congruent perceptual processing and to disrupt
incongruent perceptual processing, which is suggestive of a functionally relevant operation. At
the semantic and syntactic level, language can evoke richer simulations than individual lexemes
in isolation, as the words combine to denote details of particular actions, orientations, spatial
relationships, and inner states, all of which we are capable of simulating as we decipher the
scenes and events described to us.

Another compatibility effect was found when subjects read a sentence denoting an action
that would be performed with the handshape of a flat palm (e.g. carrying a watermelon) or of a
closed fist (e.g. carrying a marble). In the target events, the participant indicated the
grammaticalcy of the sentence by pressing a button with either an open palm or a closed fist. The
experiment (Wheeler 2010) found that respondents were faster to judge grammaticality when the
shape they were physically making with their hand corresponded to the handshape required of
the action denoted in the sentence. This result suggests that the content of the sentence prompts
the participant to construct a mental scene that simulates the action described at the sentence
level, not at the level of verbal semantics or argument structure, as the two representations
evoked by carry are only differentiated by the nature of the direct object. The evidence that the
simulation is manifested by sentence-level constructions fuels questions regarding when
simulation starts, what drives the simulation through the hearer or speaker’s encounter with the
sentence, and whether or not pieces of the simulation recruited from cognitive schemata begin
developing loosely before the end of the sentence fills in the remaining gaps with details relevant
for the representational scene. The simulations induced in this experimental task are compatible
with the real-life experiences of the body performing the actions, and this cognitive re-creation
of engaged motor performance expedites the processing of the congruent sentence.
A similar result was found in a comparable study by researchers at the University of Victoria in which participants listened to a sentence denoting people engaging with objects such as calculators and toothbrushes, but in an attentive, non-manual way (looking at, approaching, etc.) (Masson et al. 2008). They were then cued to grab one of the multiple attachments atop a contraption called the Graspasaurus, which each required a different handshape to grasp. Again, reaction times were faster when the handshape required by the cued attachment corresponded to the handshape that would be required to manually use the mentioned object, not look at it, which interestingly abets previous findings by providing evidence that functional simulations are automatically set in motion even if the subject is not prompted with a sentence containing an event that functionally involves the object.

A number of other experiments have additionally contributed to the mounting evidence that perceptuomotor information is an integral operative component of sentence comprehension, contrary to the view of traditional cognitive science in which the sentence activates the long-term memory store to computationally retrieve abstract representations that construct the meaning. Consider the following study (Zwaan & Taylor 2006) in which participants listened to a sentence implying clockwise or counterclockwise manual rotation (i.e. “Jane started the car” or “Eric turned down the volume”, respectively) and judged whether or not the sentence made sense. Participants indicated their response by turning a knob either clockwise for “yes” or counterclockwise for “no” (reversed and counterbalanced across participants). Zwaan and Taylor found that subjects’ response times were in general significantly faster (by 38 ms) when the direction of rotation denoted in the sentence corresponded to the direction that the subjects turned their hand in order to respond to the sensibility judgment. This effect is due to what the authors term “motor resonance,” in which performing an action, watching others perform an
action, and mentally representing or processing language about performing an action all recruit the same neural substrates, a phenomenon we have seen to relate to mirror neurons and semantic or conceptual simulation.

Previous research into simulation semantics, including the studies described above, has largely considered facilitation effects in which, e.g. language processing is expedited (with a faster response time) when the action depicted in a sentence is compatible with the physical action being performed by the body. As suggestive as the facilitation effects discussed so far may be, a far more compelling and comprehensive testament to the functionality of motor circuits in making meaning of language is offered if inhibition effects are considered as well. Lateral inhibition refers to the process whereby an activated neural structure that represents a perceived action is interfered with or inhibited by a neural structure responsible for representing a similar but incompatible action (Bergen 2010). That is, if the nearly simultaneous processing of two different but related actions activates the same body part, or modality effector, the two neural structures are in close competition to produce the accurate, appropriate representation in response to either given stimulus. This cortical clashing should result in a noticeably longer response time to process the second incoming stimulus, as the neurons compete for one representation to successfully win out.

In fact, this is exactly what has been found in empirical studies performed by Ben Bergen in his lab. When participants see a black-and-white line image of someone performing an action (e.g. run) and then immediately judge whether a particular verb (e.g. kick) depicts the action in the image, there is an interference effect (an average slower response time of 48 ms) if the presented verb uses the same body part effector (e.g. leg) but is not a match for the image (Bergen 2010). This suggests that the mental simulation of the action upon perceiving it visually
in a sketch, which incorporates the role of mirror neurons in simulation activation, is performing a functional role in the processing of linguistic stimuli similar to that of performing the action oneself or listening to a sentence depicting the action. The fact that the same effect was found in a comparable cross-linguistic experiment design using Cantonese stimuli for native Cantonese speakers (Bergen 2010) suggests that the implied effect is psychologically and cross-linguistically plausible.

Interestingly, in a continuation of this experiment series, a new group of participants were not shown an image of an action but were instead shown a modality-specific verb (e.g. tie) and asked to judge whether a following verb was nearly synonymous to the first verb (e.g. knot, or clap). An interference effect of approximately 100 ms occurred when the verbs required the same active effector but were not near-synonyms (Bergen 2010), as in the tie vs. clap example above. This design further supports the functional inhibition effects seen in other similar studies by demonstrating the constancy of the interference effect even when the visual image of the action is eliminated as a stimulus. Furthermore, such findings indicate that representations induced by exposure to perceptual stimuli are indeed conceptual and are initiated not just by physical perception or execution of the action itself, but also by exposure to words, at the very least to those with rich semantic and functional associations. It was mentioned above that simulations aren’t completely determined until the end of the sentence, but this does not rule out the very likely possibility that simulations begin at the beginning of the sentence or with small pieces of language and adapt accordingly to further input and context. Indeed, at the level of discourse, it is viable that simulations change and re-structure themselves far beyond the scope of any individual sentence.
Under different conditions, another experiment discovered a lateral inhibition effect similar to those in Bergen’s studies described above (Nyberg et. al. 2001, cited in Bergen 2012). In this study, “a small black rectangle on a computer screen that moved upward, downward, or horizontally for 4 seconds” (Bergen 2012, p. 228) was shown to participants. After seeing a blank screen for 1.5 seconds after the rectangle disappeared, participants saw a sentence appear on the screen that described upward or downward motion (e.g. “The diver swims to the surface” or “The meteorite crashes into the ground”, respectively). Participants read the sentence and then pushed a button as soon as they understood it. Like the language processing interference described in Bergen’s experiment above, this simple experimental task revealed that participants were approximately 100 ms slower to read and comprehend a sentence denoting either downward motion when the black rectangle had been moving downward or upward motion when the rectangle had been moving upward. This finding implies that the interference of activated competing motor circuits occurs when the cortex is called upon to process not only physical, manual actions but also motion orientations. However, it is striking that the interference occurred when the orientations of the triangle and of the sentence matched, as opposed to when they mismatched.

In attempting to offer an explanation that accounts for the results of this experiment, Bergen (2012) makes a number of suggestions. The first is that of a motion aftereffect in which motion-detecting neurons are fatigued after watching the rectangle move in a particular direction; therefore, automatic processing of direction-identical linguistic input is prevented because the neurons are less able to simulate motion in that direction on immediate subsequent command. The other potential account for the effect is that participants stored the rectangle’s motion in their memory systems and, maintaining the memory throughout sentence processing, have interfering
difficulty in comprehending a very similar denoted event over the maintenance of the present event in short-term memory recall. The role of the memory system in mediating between event perception and the construction of sentence meaning would offer an explanation for processing interference that does not rely on the neural pathways of the motor cortex specifically, and this possibility cannot be ruled out even as a motor cortical and neural explanation is developed in this essay. Evidence for either interpretation is more challenging to come by than merely observing the effect; this is a prevalent and inherent problem in the domain of simulation semantics, as it is difficult to attribute an effect to any single given cognitive process or localized brain region, as seen in the above discussion of fMRI technology. Nevertheless, under either explanatory account, when embodied simulations cannot be engaged instantaneously by the comprehender, language understanding is inhibited and slowed. The implication is that the cognitive architectures responsible for encoding and recognizing motion, action, and orientation are functional in the perception and comprehension of motion-conscious linguistic stimuli.

Separately, facilitation and inhibition effects could be attributed to downstream or nearby, unrelated effects, as discussed above. However, when taken together, these two sides of the same coin contribute to a more complete illustration of simulation semantics that attests to the functionality of motor circuits in understanding language and extracting meaning. Not only does the motor cortex aid and accelerate language processing in compatible situations, but it also interferes with itself in processing incompatible, related situations.

Within the ACE paradigm, other findings seminal to the present research have been discovered in regards to the effect of subject pronouns on directing the mental simulation of the sentence. Brunyé et al. (2009) found that the use of second-person “you” or first-person “I” pronouns in a sentence such as “I am slicing the tomato” resulted in faster verification that a
picture depicted the action when the picture was from the participant’s internal perspective. The use of the third-person “he” pronoun resulted in faster verification when the picture was from the observer’s external perspective. However, first person “I” sentences could be manipulated to match faster with the external observer’s perspective when the sentence is preceded by a short description of the narrator that uniquely identifies the narrator as someone other than the participant, such as “I am a 30-year-old deli employee. I’m making a vegetable wrap. I’m slicing a tomato.” This result implies that the use of pronouns in narrative can induce a particular perspective in a mental simulation, and that the perspective evoked by the pronouns can vary with context. The role of a sentence’s denotation of participant or observer perspective in the construction of mental simulations is discussed more extensively in Zwaan (2004).

Similarly, Bergen and colleagues (2012) discovered that when participants listened to a sentence such as “You threw the baseball to the catcher” and were asked at the same time to identify whether a picture, shown once, was the same as another picture shown right after it, participants were faster to verify that the two pictures of a baseball were the same when the second picture appeared smaller than the first one, giving the impression that the baseball was moving away from the participant, as would be the case if the participant were the immersed agent throwing the baseball. It seems as though the perspective indicated by the subject of a sentence has bearing on how the motion scene is simulated by the listener or reader.

These findings motivate the present experiment’s consideration, introduced above, of whether or not the use of a first-, second-, or third-person subject pronoun will affect the speed with which participants process a hand action sentence. Given the apparent narrative-dependent flexibility of the first person “I” subject pronoun and the external observer’s perspective associated with the third person “he/she” subject pronoun, it is hypothesized that the second-
person “you” subjects, being more reliably associated with a participant perspective, would correlate with a faster response time.
Methods

Participants

A total of 79 right-handed undergraduate students from the University of Colorado at Boulder participated in the following experiment in return for either course credit or extra credit dependent upon the course, as approved through the Internal Review Board at the University of Colorado at Boulder, (Protocol #13-0555). Twenty-nine (29) were male, and all but one (a 25- to 35-year-old male) were 18 to 25 years of age. According to an individually self-reported Edinburgh Handedness Inventory, the mean handedness score of participants was 0.74, between “right handed” at 0.5 and “purely right handed” at 1.0, with no scores below 0.5.

Procedure

After providing consent, demographic information, and a handedness evaluation, participants were seated at a computer where they read further instructions and performed the experiment on Psychopy, which was the software used for the experiment design and interface (Pierce 2007). Each trial began with a two-second blank screen, followed by a fixation cross which displayed on the screen for one second. Then, a 500 ms blank screen was shown before the subsequent presentation of a hand-drawn black-and-white line drawing of an object1 which displayed in the center of the screen for 500 ms. After 350 ms of a blank Inter-Stimulus Interval (ISI), the subject of the stimulus sentence, which was always “I”, “You”, “He”, or “She”, displayed in the center of the screen and progressed through the sentence one word at a time at a rate of 350 ms for the subject, 500 ms for the verb, and 800 ms for the direct object. This method of presentation ensured that reading time remained uniform across participants. All three target

\[1\] I owe much gratitude to Kraig Ewert for his artistic skills in drawing the image set, for his enthusiasm to sketch the line drawings, and for his patience as I adjusted the objects commissioned by the task.
handshape drawings (i.e. a fist, an open palm, and a pinch) were drawn from the perspective of a right-handed person looking at their own right hand making the shape. This was done in order to remain consistent between the three images, but also to control for and harmonize with the simulations expected when exclusively right-handed participants simulate themselves performing hand actions.

Displaying the priming image for 500 ms followed by an ISI of 350 ms is rather short—at least, shorter than the 1000 ms priming display and 500 ms ISI between image and verbal stimulus in a related experiment by Bergen et al (2010). However, as mentioned above, the intention of Bergen et al.’s study was for participants to make a conscious judgment of similarity between the priming image and the verbal stimulus (e.g. how well the image matched the following verb), so more time would allow the participants to retain a better representation of the motion depicted by the image in working memory. Contrarily, in the present experiment, participants were told merely to attend to the image and that it was unrelated to the following sentence. The shorter display of the image in this experiment reflects the proportional unimportance of the image in the participants’ conscious judgment, and the shorter ISI between prime and critical stimulus ensures that the superficially attended to image still has an active representation when the participants encounter the target sentence.²

In regards to the display durations of the various sentence components, the subject of the sentence was given less screen time (350 ms), as it was always of a recognizable closed class and

² An attempt was made to add to the experiment a component which would ask the participant to identify the image they had seen on the just-completed trial. This would provide an additional measure of accuracy that would indicate whether or not the participant was, in fact, attending to the image. Unfortunately, this feature proved unworkable with the software at the time, so the fixation cross, the explicit instructions to attend to the “visual distractions” (i.e., the images), and the short ISI between priming image and sentence were thought to sufficiently induce participants to direct attention to the image.
was therefore to some degree predictable, or expected. The verb was given 500 ms, a slightly longer display, because the verb is more semantically rich in this situation and is more critical to determining the sensibility of the sentence. The direct object displayed quite a bit longer not only because it is during the display of the direct object that participants are to indicate their response with a key press, a motor response which requires time to execute, but also because the direct object is the place where the sentence is disambiguated. Since semantically implausible or nonsensical sentences may require reflection and therefore take longer to process, enough time was allotted so as to allow for full comprehension of the sentence and immediate motor response to its sensibility. A limit was put on the amount of time allotted for the participants’ response, as opposed to allowing an open-ended response, so as to discourage extended deliberation of the sentence such that the participants’ initial instinct and, presumably, the influence of their active sensorimotor representation were most accurately captured by their response. If the participant did not respond within 800 ms, a screen appeared to inform them that they had taken too long to respond and to press the space bar to continue with the next trial. No response was recorded for these trials, and they were subsequently deleted as error trials in later data analysis. Based on feedback from initial participants, the durations of the sentence component displays were felt to be the optimal speeds at which to capture the various abilities of college-level readers while still allowing for a swift, controlled unveiling of the simulation-inducing stimulus. These durations, along with the decision to display the sentences one word at a time as opposed to showing the full sentence on the screen in its entirety, were confirmed as effective by feedback from colleagues who offered to assist with informal piloting. Displaying the sentences one word at a time also helped to control variability in reading time that might affect response time (RT), the critical measure. In initial post-experiment debriefings, no initial participants expressed
dissatisfaction or difficulty with the reading speed afforded by experiment trials, which corroborated with the researcher’s intuitions, so no changes were made to the reading speed in subsequent sessions.

**Stimuli**

All sentences were of the form “A verbed B”, where A is a singular first, second, or third person human (e.g. “he” or “she”) pronoun and B is a direct object whose function in the argument structure as PATIENT or RECIPIENT is either semantically sensible or non-sensible with the verb, as confirmed by 45 native English speakers who independently judged the sensibility of the sentences with an average 94% agreement. (The lowest agreement in the target, non-filler sentences occurred only with “You petted the dog” at 86%. However, given the limited availability of verbs that satisfy the handshape motion verbs required by the target condition sentences, this sentence was kept in the stimulus set for the main experiment, as the sensibility agreement was felt to be adequate.) This independent norming study was also approved under IRB Protocol #13-0555 and was performed by participants who did not participate in the main experiment.

All direct objects contained an article determiner (e.g. “the”, “a”) which was displayed as the complete noun phrase when the direct object displayed (e.g. “I applauded the show”). In nearly every sentence, this determiner was “the,” but there were a small number of sentences in which a different determiner such as “my” or “his” was felt to be more natural than “the” (e.g. “I zipped my jacket” vs. “I zipped the jacket”, or “She pinched his sleeve” vs. “She pinched the sleeve”), so the more natural determiner was chosen in these cases. See Appendix A for a complete list of stimulus sentences.) Target sentences were also chosen with an attempt to control for syllable length in mind, although given the limited possibilities of sensible direct
objects to pair with transitive motion verbs such as “smooth”, “grip”, or “zip”, syllable length was not 100% balanced ($M = 5, SD = 0.8$).

Subjects in the main experiment were seated with their hands on the computer keyboard in the standard “home key” typing position. If they considered a sentence to be sensible, they were instructed to indicate this by pressing the dented “J” key with their right index finger, and if they considered the sentence to be non-sensible, they pressed the dented “F” key with their left index finger. A green star sticker was present on “J” to emphasize the “Yes, sensible” nature of the key, while a red star sticker emphasized the “No, non-sensible” nature of the “F” key. All participants were right-handed so as to control for variability in motor experience between the left hand and the right hand. This is an important choice, as Borghi & Scorolli (2009) have found that people do in fact respond faster to sensible sentences with their dominant hand than their non-dominant hand in an experiment with a motion simulation paradigm such as this one. As all target sentences are sensible sentences and therefore require a right-handed response, any potential disparity between the motor response of a given individual’s dominant right hand versus their non-dominant left hand does not factor into the critical response times in which we are interested.

While subjects were led to believe that the experiment was concerned with their judgments of sentence sensibility and whether they would be affected in some way by the visual distraction of the image, the crucial variable was in fact the amount of time it took them to respond to the sentence. They were all told to indicate their answers as soon as they understood the sentence any time after the direct object (or “B” in our schematic representation) appeared on the screen, as the direct object is the disambiguating piece of information needed to determine the sensibility of the sentence, as discussed in the empirical evidence above—and indeed,
response time was measured from the onset of the direct object. This was felt to measure the speed with which they actually processed the complete sentence as precisely as possible. All participants completed 15 unrecorded practice trials before beginning the main experiment so as to familiarize themselves with the reading pace and the experimental task.

Design

There are three target conditions in this experiment: match, mismatch, and neutral. The match condition designates when the handshape in the image corresponds to the handshape denoted by the action of the sentence. The mismatch condition is the situation in which the handshape in the image does not correspond to that denoted by the action of the sentence and instead corresponds to another handshape. The neutral condition refers to the trials in which a filler, non-handshape image (e.g. a tree, a stoplight, a lamp, etc.—See Appendix D for the full list of images) is paired with a target handshape sentence. This is the control condition, as no motor simulation is theoretically evoked in the absence of an image depicting body parts or motion.

Each condition has three manifestations, depicted in Table 1 below. A FIST sentence matches with a FIST sentence and mismatches with either of the other two handshapes, PALM or PINCH, and the other two handshapes parallel this distribution.

<table>
<thead>
<tr>
<th>Sentence</th>
<th>FIST</th>
<th>PALM</th>
<th>PINCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match Image</td>
<td>FIST</td>
<td>PALM</td>
<td>PINCH</td>
</tr>
<tr>
<td>Mismatch Image</td>
<td>PALM PINCH</td>
<td>FIST PINCH</td>
<td>FIST PALM</td>
</tr>
<tr>
<td>Neutral Image</td>
<td>filler object</td>
<td>filler object</td>
<td>filler object</td>
</tr>
</tbody>
</table>

Table 2: Distribution of Handshape Images and Sentences in Target Conditions
There are 9 stimulus sentences used for each handshape for a total of 27 target sentences. Although all participants viewed an equal number of stimuli in the match, mismatch, and neutral conditions, the stimuli were arranged such that any one participant encountered each sentence only once; this was a tactical decision intended to decrease the likelihood of the participant cluing into the true purpose of the experiment and perhaps compromising their unbiased approach to all stimuli. It also avoided the prospective risk of the participant being unduly influenced by previous exposures to the same sentences. A Latin-square design was used to distribute the stimuli across participants. This resulted in a total of 36 stimulus lists, each of which was completed by two separate participants (72 out of 79 participants). See Appendix B for a representation of the Latin-square distribution. The other 7 participants completed lists to replace those that had been discarded due to participant error, judged to be an overall inaccuracy of sensibility judgments below 85%. This parameter resulted in the exclusion of data from 11 participants.

Each participant additionally made judgments on 60 filler sentences, 45 of which were non-sensible and 15 of which were sensible, on account of the fact that all 27 target sentences are sensible and there is a desire to balance sensibility evenly across the stimulus set. Ultimately, therefore, each participant responded to 87 sentences, plus 15 initial practice trials.

A further factor considered in this experiment, as discussed previously, is whether a person’s perceptuomotor simulations might be affected by the subject of the sentence being in the first, second, or third person (i.e. I, you, or he and she). (See above for more details.) To evaluate this hypothesis, the three person conditions were distributed such that each participant saw all 27 target sentences with the same person as the subject (i.e. first, second, or third). The filler sentences were balanced such that they contained more of the other two types of subjects.
than the target condition and an even number of each type total, such that the use of the pronouns seemed random and irrelevant. For every list in the third person condition, half of the sentences had “He” as the subject and the other half had “She”. The half with “He” and the half with “She” were reversed and balanced across subjects to appropriately account for the odd halving of 27 (13 and 14 sentences for each gender). Thus, there were 12 lists in each condition, which were seen by two separate participants, making for a total of 24 participants to each person condition (excluding error trials, as above). A schematic representation of the pronoun distribution is provided in Appendix C.
Results

The data were analyzed using a mixed effect linear regression analysis with pronoun type (i.e. first-, second-, or third-person), sentence type (i.e. fist, palm, pinch), and condition (i.e. match, mismatch, neutral) as predictor variables, and with gender as a control variable. The model included participant and item as crossed random effects with varying slopes and intercepts. There were no significant main or interaction effects of the predictor variables. However, a likelihood ratio test comparing the intercept-only model with a model that included condition as the predictor variable approached significance, $\chi^2 (2) = 5.14, p = 0.077$.

Although there were no statistically significant main effects or interactions, there are a number of consistent and noteworthy trends to report, as represented in Figures 1-3. Response times in the mismatch condition, when the handshape in the image was incongruent with that expressed by the sentence, were the longest ($M = 0.73, SD = 0.17$). Response times in the match condition, when the handshape from the image was the same as that denoted by the action in the sentence, were the next longest ($M = 0.71, SD = 0.16$). The fastest condition was the neutral condition ($M = 0.70, SD = 0.16$). These values are represented in Figure 1 below. Figure 2 shows the values to scale such that the standard deviations can be represented. All graphs with error bars have been provided to demonstrate that the distinctions that are seen on a smaller scale without the standard deviations are not reliable.

Assuming the trends are significant in a larger sample population, the results indicate a longer response time (interference) in the mismatch condition, as hypothesized. However, the response time in the matching condition is also longer in relation to the neutral condition, against prediction. This unexpected finding suggests either the presence of interference in the matching condition as well as in the mismatching condition (although perhaps not as strongly), or that the
simulations activated at the hand effector were not sufficiently detailed to distinguish between the different handshape types in the various conditions. These possibilities and their potential implications will be discussed in more depth in the following section.

**Figure 1: Average Response Time in Match, Mismatch, and Neutral Conditions**

**Figure 2: Average RT in Match, Mismatch, and Neutral Conditions (with Error Bars)**
Within the matching condition, *pinch* sentences were responded to fastest ($M = 0.71$, $SD = 0.15$), followed closely by *palm* sentences ($M = 0.71$, $SD = 0.18$). *Fist* sentences exhibited the longest response time ($M = 0.72$, $SD = 0.16$), although, again, the differences between the response times are marginal and are not significant. Within the mismatching condition, *palm* sentences preceded by *fist* or *pinch* images were responded to fastest (respectively, $M = 0.71$, $SD = 0.15$; $M = 0.69$, $SD = 0.16$). *Fist* sentences preceded by *palm* or *pinch* images were responded to slowest (respectively, $M = 0.76$, $SD = 0.18$; $M = 0.76$, $SD = 0.17$), as they were in the matching condition as well. Response times for *pinch* sentences that were mismatched with *fist* or *palm* images fell in between the other two sentence types (respectively, $M = 0.74$, $SD = 0.19$; $M = 0.71$, $SD = 0.18$). The relative patterns of the values in the mismatching condition resemble that of the neutral condition, with *palm* sentences being responded to fastest, followed by *pinch* sentences, with *fist* sentences eliciting the slowest response time. The average response times for sentences of each handshape type in the three conditions are represented in Figure 3 below. Figure 4 shows the values to scale such that the standard deviations can be represented.

![Figure 3: Average Response Time per Handshape Type in Target Conditions](image-url)
Figure 4: Average RT per Handshape Type in Target Conditions (with Error Bars)

While it is important to remember that the differences in response times between the various combinations of matching and mismatching handshape types are statistically inconclusive and may therefore be coincidental, the trends at least bear speculation. The tendency for the \textit{fist} sentences to consistently result in longer response times than either the \textit{palm} or the \textit{pinch} sentences in both matching and mismatching conditions suggests that the \textit{fist} sentences were potentially problematic. It is possible that the fist and fist-associated actions are for some reason more difficult to simulate than actions associated with the palm or with pinching, although this possibility cannot be explained or substantiated by present research.

There were no main effects of the subject pronoun on response time. However, as with the main conditions, there are suggestive trends. Participants took the longest amount of time to respond to sentences with a second-person “\textit{You}” subject ($M = 0.73, SD = 0.18$). Sentences with a third-person “\textit{He}” or “\textit{She}” subject pronoun were responded to slightly faster ($M = 0.71, SD =$
0.16), and first-person “I” subject sentences demonstrated the fastest response by a narrow margin ($M = 0.71, SD = 0.16$). These values are represented in Figure 5 below.

**Figure 5: Average Response Time per Subject Pronoun in Target Conditions**

**Figure 6: Average RT per Subject Pronoun in Target Conditions**
Given the statistical invalidity of these findings, no reliable conclusions can be drawn from the data. Nonetheless, the results do not support the original hypothesis that a second-person “You” subject would demonstrate the fastest response time. Instead, the findings suggest that the first-person subject is more likely to evoke faster processing.

The overall empirical results, although not statistically conclusive, are characterized by quantitative trends that seem to indicate the presence of a behavioral interference between visual and verbal primes of incongruent handshapes when they are presented in immediate succession, resulting in a longer response time of 300 milliseconds (ms) to process the sentence, as compared to the neutral condition. Additionally, however, there is also the indication of an interference effect between visual and verbal primes of congruent handshapes, resulting in a 100 ms delay in comparison to the neutral condition. This is less of an impediment to response time than in the mismatching condition, but it is still a surprising and unexpected result that urges exploration. Although, again, given the lack of statistical evidence to support a meaningful difference between the response times in the two conditions, any inferences about the implications of these trends are made with caution.

The null statistical findings are not sufficiently significant to support the hypothesis that mental simulations are a functional or necessary component of language processing, nor do they conclusively support the hypothesis that the participant perspective, encoded by the second-person “You” subject pronoun, induces faster processing than the observer’s perspective, exemplified by third-person subjects. However, there are trends in the predicted direction that suggest that stimulus-induced mental simulations may play a role in sentence comprehension and that they are potentially occurring to a degree detailed and vivid enough to include not just the “hand” body part effector, but the shape of the hand and the functional concepts and actions
associated with the handshape as well. They also suggest that, in the absence of a larger
explanatory discourse or narrative, sentences beginning with the first-person “I” pronoun are
likely to be processed faster by means of evoking the participant’s perspective more strongly
than second-person subjects. Perhaps the trends indicated by the present data would reach
statistical significance in a larger sample population, thereby allowing more conclusive findings
to be drawn.
Discussion

Although the results of the present experiment are statistically inconclusive, the quantitative trends of the data are informative and, assuming their statistical significance in a sufficiently larger sample population, are suggestive of implications relevant to the discussion of simulation semantics in the current literature.

The prediction that a processing delay results from incongruence between the handshape depicted in an image and the handshape required to carry out an action expressed in a sentence, shown nearly simultaneously, does not appear contradicted in the data. This effect of inhibition on sentence comprehension could be explained in one of several ways. It could be the case that visually perceiving the handshape activates representational or semantic concepts associated with that certain handshape in long-term memory, concepts that, while enjoying strong activation in working memory, suppress other nearby concepts that are different but fundamentally related. Another possibility is that the image triggers a motor simulation of the handshape, including the physical form and function of the handshape, a simulation that conflicts with the input from the sentence as the motor cortex attempts to activate a new configuration of neurons in the area of the same effector. Both of these possibilities implicate language into a broad system of cognitive functioning that necessarily incorporates detailed mental imagery or highly-specified schemata that have been embedded in prior experience in its response to stimuli.

However, the data also indicate an unexpected result—that a processing delay also occurs when the handshape depicted in an image is the same as the handshape required to perform an action expressed by a sentence shown at nearly the same time. While the interference is not as long in this case as in the mismatching condition (100 ms and 300 ms, respectively) and, again, is not statistically significant, the effect is reminiscent of the results found in Nyberg et al.
(2001), discussed above. In that experiment, inhibition occurred when processing a sentence describing e.g. upwards motion when the participant had been primed with a triangle moving upwards, resulting in an interference effect between stimuli with congruent directions of motion. Earlier it was suggested that such inhibition in conditions of compatible stimuli may be caused by a motion fatigue aftereffect, in which motion-detecting neurons are depleted after watching movement in a particular direction. This fatigue then creates interference when the motion detectors are called upon immediately afterwards to process simulated motion in the same direction. An analogy between motion-detecting neurons and effector-simulating neurons may appropriately account for the effects demonstrated by the current empirical data.

Another potential explanation for the delay during processing of a sentence that is compatible with its prime is that of the mediation of the memory system during the task. It might be the case that limited working memory, fixated on the form and orientation of the handshape in the initial prime, inhibits the threshold at which that same form or orientation can be registered from a different type of stimulus (visual/perceptual vs. semantic/conceptual). The design of the present experiment did encourage participants, with both visual and verbal instructions, to attend to the image preceding the sentences, so this is potentially a valid interpretation of the effect. Alternatively, it is possible that participants did not register the difference between matching and mismatching handshapes, as the difference between the average response times in these two conditions were inconclusive. This would suggest that their mental simulations were not sufficiently detailed to distinguish between different forms of the same effector.

The data do not support the predicted hypothesis that sentences beginning with a second-person “You” subject would, in general, provoke a faster response time than sentences beginning with first-person “I” or third-person “He/She” subjects. This hypothesis reflected the original
expectation that the second-person subject pronoun would induce the reader to adopt a participant perspective, in which they project themselves into the simulated scene as the agent of the action. This perspective encourages the participant to more readily identify with the action as the performer, instead of as the observer, which theoretically expedites processing of the stimulus. It appears that, in contrast, the first-person “I” subjects tended toward a faster response, although this was not a statistically conclusive finding. Nevertheless, the inability of the present experiment to establish sound evidence in this regard should not be taken as evidence against the participant’s vs. the observer’s perspective in mental simulation in principle. It could be the case that an unforeseen shortcoming in the experimental design restricted the opportunity to accurately examine this variable, such as the presentation of the sentence one word at a time, a factor worth manipulating in future studies. The question remains open, however, as to how the perspective of a sentence or a narrative as encoded by, for example, its pronouns might bear on the relationship the reader or listener develops with the text and therefore the types of simulations they create. It seems a worthwhile hypothesis to pursue, especially in light of the evidence that does exist in favor of the effect of perspective on mental simulation, as discussed above (Brunyé et al. 2009; Bergen 2012).

Assuming the current experimental findings are replicable and appropriately significant in a larger sample of participants, the interference effects indicated by the data would seem to support the role of semantic simulation as a functional component of language processing, even if the findings do complicate current groundwork in the field. The trends in the data also offer some preliminary, although unconfirmed, evidence that the simulations evoked by linguistic input are detailed and vivid, not underspecified and abstract. Language of motion has provided a convenient window into the groundedness of semantic meaning, as we perhaps naturally
anticipate that if linguistic meaning is embedded in knowledge of prior bodily encounters, that physical language about the body is an ideal place to investigate this connection. Of course, if language and other facets of cognition are truly embodied and do rely on situated associations between our worldly interactions and our inner mental lives, we should expect to find effects of interference and facilitation in other behaviors and in more abstract types of language. The investigation of both interference and facilitation effects is crucial, as the two effects combined (expediency in some situations and disruption in others) indicates stronger, more comprehensive evidence for the functional reality of simulation semantics across a variety of situations. There is also a great need to investigate the level of detail of mental simulations and mental imagery (Wheeler 2010) and the degree to which such detail is operational, an endeavor which has been attempted here but which remains an open question.

In order to continue this line of inquiry, it is important to continue modifying empirical approaches and paradigms. The present experiment carries the assumption inherent in its design that, given the supposed role of mirror neurons in activating the motor cortex, merely seeing an image of a handshape and not physically making the handshape should serve as a sufficient prime for simulation and the activation of motor cortex pathways. However, it is possible that the trends seen in the data were not as robust as they might have been had the stimuli been of a different nature—perhaps a moving video for the prime, or the hand drawn from a different perspective, or incorporating a more pronounced component of cross-modality (such as listening to the sentence instead of reading it). The effects from such modifications may bear significantly on the discussion of motor simulation and embodied cognition.

Moreover, the implications of the current findings in this domain can be informative for future research. The role of cognitive simulations in the processing of metaphorical language
remains controversial and under-investigated. Bergen and Lindsay et. al. (2007) have found empirical evidence that processing motion- and action-denoting sentences used figuratively in metaphorical language (e.g. “Stock prices climbed”) in the presence of supposedly interfering upward or downward primes does not significantly yield the slower, inhibited response times seen in identical conditions with sentences whose meaning denotes the comparative literal counterparts (e.g. “The ant climbed”). However, consider Desai et. al. (2011) whose fMRI studies have shown similar neurological activation in the left anterior inferior parietal lobule during the comprehension of sentences with the same verb used in literal and metaphorical contexts (e.g. “The daughter grasped the flowers” and “The public grasped the idea,” respectively). To move beyond the literal and assess whether mental representations and simulations function in the processing and usage of metaphorical or polysemous language and to what degree is a promising route of future inquiry.

Perceptual, sensorimotor, and cognitive processes overlap to such a great degree that it is inherently difficult to attribute behavioral effects seen in experimental tasks to any one specific cognitive cause or cortical location. The question endures, therefore, as to how individual perceptions and interactions with the world become “wired” such that their influence on linguistic, spatial, and motoric reasoning becomes an inextricable component of our worldly and semantic concepts. Perhaps it is a matter of what the human “hardware” (i.e. brains and sensory organs) makes of the input from “software” (i.e. social development, cognitive capacities, and experience, linguistic or otherwise) over the course of the lifetime—perhaps the two accumulate holistically such that our cognition is always embedded in prior experience. Research into the functions of mental simulation and concept formation potentially challenge theoretic frameworks which conceive of language and grammar as a computationally abstract, instinctual knowledge
base that one must merely tap into. Grounded theories of language instead purport a fully integrated and multimodal theory of general cognition that takes into account all types of perceptual, sensorimotor, and therefore perhaps even cultural experience as the roots of linguistic meaning.

Embodied, or grounded, cognition is a domain of interdisciplinary research where exciting new observations, analyses, and technologies will undoubtedly continue to unravel. The very real potential that semantic simulation is integral to our linguistic usage is currently an area of promising investigation by contemporary neurobiological, psychological, and linguistic (and particularly cross-linguistic) research. The activation of perceptuomotor circuitry either by visual or physical stimuli just before or during language processing is increasingly being proven to have a functional influence on language processing. If it is the case that our sensorimotor and perceptual systems are responsible, even in part, for the development of our linguistic comprehension and conceptual representation, then our linguistic behavior ought to reflect the connections and associations that have been integrated into the holistic cognitive system. Although many details remain unresolved, the search for meaning—why it varies, where we find it, and how it works—continues to advance as we investigate further the process of extracting meaning from linguistic encoding.
References


Carillo-de-la-Peña, M. T., Lastra-Barreira, C., & Galdo-Álvarez, S. (2006). Limb (hand vs. foot) and response conflict have similar effects on event-related potentials (ERPs) recorded during motor imagery and overt execution. European Journal of Neuroscience, 24, 635-643.


Appendices

Appendix A – List of Sentences (Example of One Pronoun Set in “I” Condition)

Open Palm

1. I slapped the liar.
2. I high-fived the winner.
3. I patted the pillow.
4. I petted the dog.
5. I smoothed the tablecloth.
6. I spanked the boy.
7. I stroked the horse.
8. I smacked the countertop.
9. I applauded the show.
34. You tied the treehouse.
35. You shopped the whiskers.
36. You strolled the mother.
37. You blistered the justice.
38. You congratulated the lampshade.
39. You angered the fragrance.
40. You brewed the statue.
41. You grew the furnace.
42. You squatted the dictionary.
43. She knelt the mirror.
44. She misbehaved the yoga.
45. She baked the nostril.
46. She extinguished the water.
47. She hung the temperature.
48. She enamored the ladder.
49. She vacationed the pebble.
50. She rusted the breath mints.
51. He hummed the shield.
52. He perched the rainforest.
53. He hiccupped the marriage.
54. He parked the jingle.
55. He snorkeled the hairbrush.
56. He shaved the television.
57. He weeded the porkchops.
58. I condescended the ball.
59. I microwaved the discussion.
60. I offended the chainsaw.
61. I arrested the bedroom.
62. I danced the ashtray.
63. I excused the toolbox.
64. I governed the steak.
65. I persuaded the briefcase.
66. I believed the berries.
67. I gargled the motorcycle.
68. I forgave the kitchen.
69. I shampooed the reason.
70. I planted the oven.
71. I profited the ocean.
72. I sneezed the meeting.

Filler – Non-sensible

28. You assumed the toilet.
29. You swam the doctor.
30. You chewed the ceiling.
31. You swore the vegetable.
32. You knotted the doorframe.
33. You waddled the tattoo.

Fist

10. I punched the wall.
11. I gripped the microphone.
12. I pounded the bully.
13. I boxed the fighter.
15. I sluged the thief.
16. I squeezed the baseball.
17. I wrung the dishrag.
18. I grabbed the rope.

Pinch

19. I pinched his sleeve.
20. I inserted the coins.
21. I held the chopsticks.
22. I fastened the button.
23. I tweezed the splinter.
24. I stitched the dress.
25. I plucked my eyebrows.
26. I zipped my jacket.
27. I picked the lock.

Filler – Sensible

73. You asked a question.
74. You sang a hymn.
75. You entered the room.
76. You tasted the champagne.
77. You read the headlines.
78. She admired the basket.
79. She considered the offer.
80. He gathered the documents.
81. He designed the graphics.
82. He forfeited the trophy.
83. I finished the puzzle.
84. I followed the trail.
85. I dated an attorney.
86. I recorded the concert.
87. I explained the options.
Appendix B – Complete Distribution of Match, Mismatch, and Neutral Conditions Across All Stimulus Lists (Latin-Square)

### Participant Number (All First-Person Subject Pronouns)

<table>
<thead>
<tr>
<th></th>
<th>1 (ABC₁)</th>
<th>2 (ACB₁)</th>
<th>3 (BAC₁)</th>
<th>4 (BCA₂)</th>
<th>5 (CAB₂)</th>
<th>6 (CBA₂)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentences 1-9 (A)*</td>
<td>Match</td>
<td>Neutral</td>
<td>Mismatch¹</td>
<td>Match</td>
<td>Neutral</td>
<td>Mismatch²</td>
</tr>
<tr>
<td>Sentences 10-18 (B)*</td>
<td>Mismatch¹</td>
<td>Match</td>
<td>Neutral</td>
<td>Mismatch²</td>
<td>Match</td>
<td>Neutral</td>
</tr>
<tr>
<td>Sentences 19-27 (C)*</td>
<td>Neutral</td>
<td>Mismatch¹</td>
<td>Match</td>
<td>Neutral</td>
<td>Mismatch²</td>
<td>Match</td>
</tr>
</tbody>
</table>

### Participant Number (All Second-Person Subject Pronouns)

<table>
<thead>
<tr>
<th></th>
<th>7 (ABC₁)</th>
<th>8 (ACB₁)</th>
<th>9 (BAC₁)</th>
<th>10 (BCA₂)</th>
<th>11 (CAB₂)</th>
<th>12 (CBA₂)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentences 1-9 (A)*</td>
<td>Match</td>
<td>Neutral</td>
<td>Mismatch¹</td>
<td>Match</td>
<td>Neutral</td>
<td>Mismatch²</td>
</tr>
<tr>
<td>Sentences 10-18 (B)*</td>
<td>Mismatch¹</td>
<td>Match</td>
<td>Neutral</td>
<td>Mismatch²</td>
<td>Match</td>
<td>Neutral</td>
</tr>
<tr>
<td>Sentences 19-27 (C)*</td>
<td>Neutral</td>
<td>Mismatch¹</td>
<td>Match</td>
<td>Neutral</td>
<td>Mismatch²</td>
<td>Match</td>
</tr>
</tbody>
</table>

### Participant Number (All Third-Person Subject Pronouns)

<table>
<thead>
<tr>
<th></th>
<th>1 (ABC₁)</th>
<th>2 (ACB₁)</th>
<th>3 (BAC₁)</th>
<th>4 (BCA₂)</th>
<th>5 (CAB₂)</th>
<th>6 (CBA₂)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentences 1-9 (A)*</td>
<td>Match</td>
<td>Neutral</td>
<td>Mismatch¹</td>
<td>Match</td>
<td>Neutral</td>
<td>Mismatch²</td>
</tr>
<tr>
<td>Sentences 10-18 (B)*</td>
<td>Mismatch¹</td>
<td>Match</td>
<td>Neutral</td>
<td>Mismatch²</td>
<td>Match</td>
<td>Neutral</td>
</tr>
<tr>
<td>Sentences 19-27 (C)*</td>
<td>Neutral</td>
<td>Mismatch¹</td>
<td>Match</td>
<td>Neutral</td>
<td>Mismatch²</td>
<td>Match</td>
</tr>
</tbody>
</table>

### Key

- Mismatch¹ = Mismatched pairing Type 1 (e.g. Fist sentence with Palm image)
- Mismatch² = Mismatched pairing Type 2 (e.g. Fist sentence with Pinch image)
- *Each chunk of 9 sentences (A; B; C) includes the same 3 fist sentences, 3 palm sentences, and 3 pinch sentences.
Appendix C – Distribution of Subject Pronouns Across Six Participants

<table>
<thead>
<tr>
<th>Participant 1</th>
<th>Participant 2</th>
<th>Participant 3</th>
<th>Participant 4</th>
<th>Participant 5</th>
<th>Participant 6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Open Palm</strong></td>
<td><strong>Open Palm</strong></td>
<td><strong>Open Palm</strong></td>
<td><strong>Open Palm</strong></td>
<td><strong>Open Palm</strong></td>
<td><strong>Open Palm</strong></td>
</tr>
<tr>
<td>1. I</td>
<td>1. You</td>
<td>1. He</td>
<td>1. I</td>
<td>1. You</td>
<td>1. She</td>
</tr>
<tr>
<td>5. I</td>
<td>5. You</td>
<td>5. He</td>
<td>5. I</td>
<td>5. You</td>
<td>5. She</td>
</tr>
<tr>
<td><strong>Fist</strong></td>
<td><strong>Fist</strong></td>
<td><strong>Fist</strong></td>
<td><strong>Fist</strong></td>
<td><strong>Fist</strong></td>
<td><strong>Fist</strong></td>
</tr>
<tr>
<td><strong>Pinch</strong></td>
<td><strong>Pinch</strong></td>
<td><strong>Pinch</strong></td>
<td><strong>Pinch</strong></td>
<td><strong>Pinch</strong></td>
<td><strong>Pinch</strong></td>
</tr>
<tr>
<td>27. I</td>
<td>27. You</td>
<td>27. He</td>
<td>27. I</td>
<td>27. You</td>
<td>27. She</td>
</tr>
</tbody>
</table>
Appendix D – List of Images

Target
1. OPEN PALM
2. FIST
3. PINCH

Neutral/Filler
4. fish
5. necklace
6. house
7. boat
8. mountains
9. river
10. candle
11. chair
12. tree
13. fruit basket
14. lamp
15. soccer ball
16. dress
17. flowers
18. bonfire
19. cat
20. sun
21. giraffe
22. boots
23. airplane
24. stoplight
25. eye (only filler, not used in neutral condition)