LINGUISTIC DESCRIPTIONS OF ACTION AND OBJECT PERCEPTION: THE ROLE OF ACTION READINESS

by

TOMMY ANDERSON

Submitted to the Psychology Department
School of Natural and Social Sciences
in partial fulfillment of the requirements
for the degree of Bachelor of Arts

Purchase College
State University of New York

May 2019

Sponsor: Alexia Toskos, Ph.D.

Second Reader: Lauren Harburger, Ph.D.

Abstract

Research assessing the relationship between action and object perception is far-reaching. However, most of this research uses tangible objects or photos of objects. Less research has focused on the abstract ways in which we interact with objects, namely linguistically. This experiment focuses on how linguistic descriptions of action can influence object perception. Participants were tasked with listening to a game show scenario that had them toss water balloons into buckets using two different grasp types: precision grip (pinch) or power grasp (whole-hand). Participants then identified an ambiguous object, which could be interpreted as a cherry or apple. We also manipulated action readiness by describing participants as either about to pick up the water balloon (hand available) or having already picked up the water balloon (hand unavailable) at the end of the story. This was done because research suggests that if the hand is unavailable for action, the effect goes away. Results for this experiment revealed a congruency effect, which means that the grasp type described determined the interpretation of the ambiguous object, but only when the hand was available to perform action.

Linguistic Descriptions of Action and Object Perception: The Role of Action Readiness

Action and object perception have historically been treated as separate from one another. One set of representations is thought to underlie action while a different set of representations is thought to underlie perception. But even the most mundane task, like opening a door, requires that we integrate what we see in the world with what we're doing. This experiment examines whether there is any overlap between the representations that underlie action and perception.

Several theories describing the relationship between the action and object perception representations have been proposed. The two-visual pathways theory posits that information is processed in two parallel processing streams: what and where / how (see Fischer & Zwaan, 2008 for a review). For example, the ventral stream contains mechanisms that underlie object identification (the what pathway), and the dorsal stream has mechanisms for action control (the where pathway, Ungerleider & Mishkin, 1982). Behavioral evidence suggests that perception of objects can be distorted (e.g., a person's ability to identify the size of an object) without a corresponding difficulty in creating the right sized grasp aperture for picking up that same object (Glover, 2004), suggesting a divide between the two systems.

However, there is also support that perception and action have a common mechanism, particularly in the earlier planning phases of the action. The theory of event-coding (Hommel, 2004), suggests that viewing an action and planning to perform that same action use a common set of mechanisms (Wolpert, Doya & Kawato, 2003). However, control of an ongoing action involves a separate set of motor-specific representations.

But what are the specific underlying biological mechanisms where there is overlap? Some research suggests that mirror neurons, which are neurons attuned to action and perception of action, may be responsible. Mirror neurons were originally discovered in the macaque monkey, specifically in the ventral premotor area F5 (Di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992). These neurons not only fire when action is being performed, but when action is being viewed, which suggests that they are involved in action-planning. Due to this difference in function, researchers hypothesized that there are two types of mirror neurons: broadly congruent (achieving overall goal) and strictly congruent (how to achieve goal) (for review, see Rizolatti & Craighero, 2004). Some research even shows visual action does not have to be present, and that the sound of action i.e. crushing peanuts, can elicit the firing of mirror neurons (Keysers et al., 2003; Kohler et al., 2002). Current research indicates that mirror neurons are present in humans as well (Koski et al., 2002), specifically in Broca's area, which is the human equivalent of the F5 area in macaque monkey. Like F5, neurons in this area fire when action is being viewed. However, research has demonstrated that there are some differences in the mirror systems between the macaque monkey and humans (Nishitani & Hari, 2002). Unlike macaque monkeys, static images of action being performed are enough to excite the mirror neurons in humans (Buccino et al., 2005), suggesting that even more abstract forms of action can elicit mirror neuron response.

More broadly, researchers have tested the extent to which brain areas involved with action are also involved with object perception. In one study, researchers had participants view a variety of images – both graspable objects and objects or scenes that did not afford any type of manual action. Using functional magnetic resonance imaging (fMRI), graspable objects, like tools, elicited a stronger motor cortex response than non-graspable objects, like a house (Chao & Martin, 2000). Similarly, in another study, researchers used positron emission tomography (PET) to measure brain activity when participants viewed a variety of graspable objects (Grèzes & Decety, 2002). Participants had to either silently name the object or imagine themselves using the object. Both conditions revealed that there was motor activation when viewing and naming the object, as well as imagining interacting with the object, with specific activation in the middle frontal gyrus.

Though there is support that the motor cortex is involved object perception, does this mean that motor representations affect object perception in a causal way? There seems to be a behavioral element when it comes to motor representations of object perception. In a study conducted by Tucker and Ellis (2001), participants were tasked with responding to a variety of stimuli with two types of hand actions: precision grip or a power grasp. Participants were more likely to respond quicker if the item viewed matched the particular hand action. This demonstrates a congruency affect, meaning the hand action determines the interpretation of the stimuli. Similarly, another study found that participants respond more quickly to objects that match with a specific hand action (Bub, Masson, & Lin, 2013). Participants were tasked with learning a variety of different hand action primes of different orientations (horizontal or vertical) and handedness (right or left), and then presented with manipulable objects. If the hand action image matched with the manipulable object (i.e., a hand action that could be used to hold a flashlight, and a flashlight) this would be considered a congruent response, or a match between hand action prime and manipulable object.

Not only does object perception depend on motor representations, but so does perception of the hands themselves (Coslett, Medina & Burkey, 2010). Participants were tasked with identifying if a series of images of hands were either right hands or left hands. The task was made difficult by the hands being rotated up to 180 degrees either rightward or leftward. What researchers found was that the more rotated the hand, the slower people were to mentally rotate it in order to determine if it was a right or left hand. Furthermore, people were slower to mentally rotate hands in the medial direction than the lateral direction because that direction of rotation is physically more difficult, suggesting that even our perception of hands can change based off of how our motor experience with them. To further test this idea, the researchers found that people with chronic pain in one arm and not the other were slower to mentally rotate pictures of the affected hand.

Given that viewing a picture of a hand activates the motor system, might viewing a hand influence perception of a subsequent, ambiguous object? To test this question, researchers used different hand primes to see how they can influence the perception of an object (Dils, et al. under revision). Participants were presented with a precision grip or a power grasp and then asked to identify an image. Many iterations of this study yielded the same results – that hand primes influenced how an object was perceived. The initial study utilized an image that could be interpreted as a football or a coffee bean; a subsequent study utilized an image that could be interpreted as a nut or a flashlight; and finally, an image that could be interpreted as a cherry or an apple. If participants were to see a precision grip, followed by the ambiguous image, they were more likely to interpret the image as the smaller object (i.e. coffee bean, nut, cherry); however, if they saw the power grasp, they were more likely to interpret the image as the larger items (i.e. football, flashlight, apple). These congruency effects show that the hand prime can influence the perception of an ambiguous image. In a fourth study, researchers showed hand primes that already had an object being held. These occupied hand primes, or hands unavailable to perform action, did not yield the same results of the former three studies. Participants in the hands unavailable condition were unaffected by the hand primes, suggesting that the availability of the hand to perform action is essential for action readiness.

As mentioned previously, there has been support that more abstract forms of motor action, like static images, can affect our perception (Buccino et al, 2005). Another abstract form of motor action is the linguistic form. Researchers have found that even words that describe action can elicit a motor response (Hauk, Johnsrude, Pulvermüller, 2004). Participants would hear a variety of actions words that referred to either the face, arms, or legs (e.g. to lick, pick, or kick), and then asked on a seven-point scale how much each word reminded them of their respective area. fMRI brain scans revealed that action words activated the same brain areas that were activated when performing those action movements. This overlap suggests that even mere description of action can have effects on motor activation.

Past research indicates that there is a relationship between motor activity, object perception, and linguistic descriptions. The present study tested whether linguistic descriptions of action affect object perception in the same way that viewing action affects perception. Specifically, the study tested whether the action readiness of the action prime modulates the degree to which linguistic action affects perception. Participants heard a story about participating in water balloon toss. One version of the story mentioned holding the balloon by the knot, simulating a precision grip; and the other mentioned holding the balloon without touching the knot, simulating a power grasp. Action readiness of the hands was also manipulated in two conditions: hands either available or unavailable for subsequent action. That is, the story described the person as already holding the balloon (hands unavailable) or preparing to pick up the balloon (hands available). At that moment, the ambiguous cherry/apple image used by Dils et al. (under revision) appeared on the screen and participants had to name the object. We hypothesized that in the hands available condition, the effect would be present like in the Dils et al. (under revision). Similarly, we expect that the effect would not be present in the hands unavailable condition.

Method

Subjects

Two-thousand four hundred and fifty-nine participants were recruited utilizing Amazon's Mechanical Turk service. Basic demographics information, like age (37.41) and gender (M = 1,106, F = 1,253, NB = 17, Prefer not to answer = 6), were collected post-experiment, as well as the handedness of the participants.

Materials

Auditory Stimuli. A story describing a Game Show scenario, in which contestants are tasked with bringing water balloons to a bin, was presented. We manipulated two factors: grasp type (whole hand vs precision) and action readiness (hands empty vs hands full), to create four conditions. There were four endings of the story dependent on the condition. The first set of endings described the participants as having just grabbed water balloons (hands full) by the knot (precision) or not the knot (whole hand). The second set of endings described participants preparing to grab the balloon (hands empty) by the knot (precision) or not the knot (whole hand). The story was presented in three pieces so that pictures can be shown intermittently.

Visual Stimuli. There were four images, three non-graspable filler images (cactus, house, beach) and an ambiguous image that could be interpreted as a cherry or an apple (see Figure 1).

Design

This study was conducted using a between-subjects design. The two variables were hand grasp (precision or power) and hand availability (available hand or unavailable hand). The dependent variable was the interpretation of the ambiguous image.

Procedure

An audio test to assure participants have their volume raised was presented. The word "nine" was said and participants had to type out the word before the proceeding to the next screen. Once completed, participants had to listen to a story describing a Game Show scenario. In this scenario, contestants were tasked with bringing water balloons to a bin. Depending on the condition, participants either heard a version that had them grab the balloons by the knot (precision grip) or grab them without touching the knot (power grasp). Also, depending on the condition, participants either had just grabbed a pair of balloons when the story cut off (unavailable hand) or that they were preparing to grab a pair of balloons when the story cut off (available hand). As participants were listening to the story, three filler images (cactus, sea, house) were presented throughout the story to assure participants are paying attention to the visual stimuli on the screen. Once the story was completed, participants were asked to identify an ambiguous image that could be interpreted as either a cherry or an apple (i.e., You just saw a picture of a red object. What was it?). The following question inquired if they had any additional interpretations of the image. This was then followed by two tester questions, to assure that participants were paying attention throughout the study. The first question was open-ended, and asked participants how they were instructed to hold the balloon; whereas the second question was multiple choice and asked participants how many times they were told to pick up the balloon (i.e., never, once, repeatedly). Finally, participants were asked their age, gender, race, history of injury and chronic pain, and handedness in a post-experiment demographics questionnaire.

Results

All of the participants' responses to the cherry / apple image were first coded as either requiring a precision grip or whole hand grasp. If participants said that it was both it was coded as both, and if participants had unclear or incoherent answers it was coded as none (i.e., ball). Responses to the first comprehension question were coded as anything that could be interpreted as the knot or not the knot (i.e. the top of the balloon), generously construed. We counted the number of congruent responses, or responses where the action that was described in the story matched with the response that participants gave for the ambiguous image.

We tested whether participants were more likely to perceive the ambiguous object in a manner congruent to the action they just heard about. In the available hand condition, or preparing for action condition, 627 participants made the congruent condition while 552 participants had an incongruent response, meaning 53.18% of responses were congruent. This percentage was different from chance. A chi-square goodness of fit test revealed this relationship was significant, $X^2(1, N=1,179) = 4.64, p<.05$.

The same was done for the unavailable hand condition, or performing action condition. In this condition, 595 participants made a congruent response while 603 participants had an incongruent response, meaning 49.67% of responses were congruent. This percentage was not different from chance. A chi-square goodness of fit test revealed the relationship was not significant, $X^2(1, N=1,198) = .04, p=.842$. (see Figure 2 for graph of results).

Due to a large number of participants answering one or both of the comprehension questions incorrectly, we also ran the same tests including only those who got both of the comprehension questions correct. This dataset also excluded those who had any additional interpretations of the ambiguous image. In the available hand condition, or preparing action condition, 152 participants made a congruent response, while 158 participants had an incongruent response, meaning 49.03% of responses were congruent. This percentage was not different from chance, $X^2(1, N=310) = .08, p=.777$.

Once again, the same was done for the unavailable hand condition, or performing action condition. In this condition, 210 participants made a congruent response while 196 participants had an incongruent response, meaning 51.72% of responses were congruent. This percentage was not different from chance. A chi-square goodness of fit test revealed the relationship was not significant, $X^2(1, N=406) = .42$, $p=.517$. (see Figure 3 for graph of results).

Discussion

This study added to a body of research that seeks to define the relationship between action and perception. Though in the past action and perception have been treated as separate from one another, there has been increasing support that the two interact. Results for this experiment were in line with past research, specifically in regards to hand availability. Like the Dils et al. (under revision) study, when the hand was available to perform an action (i.e. it was empty), an effect of the hand prime, or in this study, linguistic description, was present. However, when action was described as already being performed, and the hand was unavailable for action, the effect was no longer present. This supports the theory that hand availability is essential in planning action (Proffitt, 2006), and that in order for someone to imagine themselves performing an action, they have to know they are capable of doing it, even abstractly.

However, there were several limitations of this study. The biggest limitations were the comprehension questions. A large number of participants answered either one or both of them incorrectly, suggesting that they were either too hard or were not clear enough for participants. The first comprehension question, which asked participants how they were instructed to hold the balloon, yielded many answers that could be interpreted in a variety of ways. Recurring answers were “the top” or “the bottom” of the balloon. This made coding difficult, as we were unsure if this meant the knot of the balloon or not the knot of the balloon. A second iteration of this study may not have an open-ended question, but rather a multiple-choice option that participants could pick from. However, this may not be as effective at weeding out participants who were not paying attention. The second comprehension question, which asked participants how many times the participants picked up the water balloon, may have also been problematic as it may have been too specific for participants to have remembered, leading to very poor performance.

Another limitation was participants' inclination to name the ambiguous object an apple. In both conditions of the experiment, participants were more likely to identify the image as an apple (1,271 responses), as opposed to a cherry (1,106 responses). The presence of this apple bias might be due to the water balloon being the primary object in the story. Water balloons, despite being able to be manipulated in both a precision grip and power grasp manner, more closley resemble an apple in their size and weight. Due to this similarity, participants may have been more inclined to identify the ambiguous image as an apple because they were unintentionally primed to be holding an object that can be more closely identified with an apple.

Despite these limitations, the implications of this experiment yield important information about the abstract ways perception can be influenced. In a study conducted by Fausey and Boroditsky (2010), researchers found that linguistic cues can shape how participants placed blame and financial liability in accident scenarios. Even when participants viewed a video of accident, the language that was used to describe the action influenced how participants viewed who blame was placed on. The real-life ramifications of this are staggering, as linguistic cues can be used to skew the testimonies of children (Carter, Bottoms & Levine, 1996), and adults (Fausey & Boroditsky, 2010), alike in eyewitness testimonies. Though this study did not explore these topics, it serves as an example of how one does not even have to view action, and that the mere description of action, can be enough to influence how our perception can be influenced.

Therefore, testimony involving what type of object a person saw under poor viewing conditions might be biased by what the eyewitness was doing with their hands in the moment.

References

Bub, D. N., Masson, M. E. J., & Lin, T. (2013). Features of planned hand actions influence identification of graspable objects. *Psychological Science, 24(7)*, 1269-1276.

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action related sentences modulates the activity of the motor system: A combined TMS and behavioral study. Cognitive Brain Research, *24, 355 – 363.*

Carter, C. A., Bottoms, B. L., & Levine, M. (1996). "Linguistic and socioemotional influences on the accuracy of children's reports": Erratum. *Law and Human Behavior*, *20*(3), 335-358.

Chao, L.L., Martin, A. (2000) Representation of manipulable man-made objects in the dorsal stream. Neuroimage, *12, 478–484*

Coslett, H. B., Medina, J., Kliot, D., & Burkey, A. R. (2010). Mental motor imagery indexes pain: The hand laterality task. *European Journal of Pain*, *14*(10), 1007-1013.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. Experimental Brain Research, 91, 176–180.

Dils, A., Flusberg, S.J., Boroditsky, L., (under revision)

Fausey, C. M., & Boroditsky, L. (2010). Subtle linguistic cues influence perceived blame and financial liability. Psychonomic Bulletin & Review, 17(5), 644–650.

Fausey, C. M., & Boroditsky, L. (2010). Who dunnit? Cross-linguistic differences in eye-witness memory. *Psychonomic Bulletin & Review, 18* (1), 150-157.

Fischer, M. H., & Zwaan, R. A. (2008). Embodied Language: A Review of the Role of the Motor System in Language Comprehension. *Quarterly Journal of Experimental Psychology*,*61*(6), 825-850.

Glover, S. (2004). Separate visual representations in the planning and control of action. Behavioral and Brain Sciences, *27*, *3–78*.

Grèzes, J., & Decety, J. (2002). Does visual perception of object afford action? Evidence from a neuroimaging study. *Neuropsychologia, 40*(2), 212-222.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic Representation of Action Words in Human Motor and Premotor Cortex. *Neuron, 41*(2), 301-307.

Hommel, B. (2004). Event files: Feature binding in and across perception and action. Trends in Cognitive Sciences, *8*, *494–500*.

Keysers, C., Kohler,E., Umilta, M.A., Fogassi,L., Nanetti, L., & Gallese, V. (2003). Audio-visual mirror neurons and action recognition. Experimental Brain Research, *153, 628–636*.

Kohler,E., Keysers, C., Umilta, M.A., Fogassi,L., Gallese, V., & Rizzolatti, G. (2002, August 2). Hearing sounds, understanding actions: Action representation in mirror neurons. Science, *297*, *846 – 848*.

Koski,L.,Wohlschla¨ger,A.,Bekkering,H.,Woods, R. P., Dubeau, M.-C., Mazziotta, J. C. et al. (2002). Modulation of motor and premotor activity during imitation of target-directed actions. Cerebral Cortex, *12, 847–855*.

Nishitani, N., & Hari, R. (2002). Viewing lip forms: Cortical dynamics. Neuron, 36, 1211–1220.

Proffitt, D. R. (2006). Embodied perception and the economy of action. *Perspectives on Psychological Science. 1*(2), 110-122.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. Annual Reviews Neuroscience, *27, 169–192*.

Tucker, M. & Ellis, R. (2001). The potentiation of grasp types during visual object categorization. *Visual Cognition*, *8(6)*, 769-800.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In M. Ingle, M. Goodale, & R. Mansfield (Eds.), Analysis of visual behavior (pp. 549–586). Cambridge, MA: MIT Press.

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. Philosophical Transactions of the Royal Society, *358, 593–602*.

Figures

Figure 1. The ambiguous image, which could be interpreted as a cherry or an apple, that served as the stimuli for the experiment.

Figure 2. Effect of grasp type on object perception across action readiness conditions. Error bars represent 1 SEM.

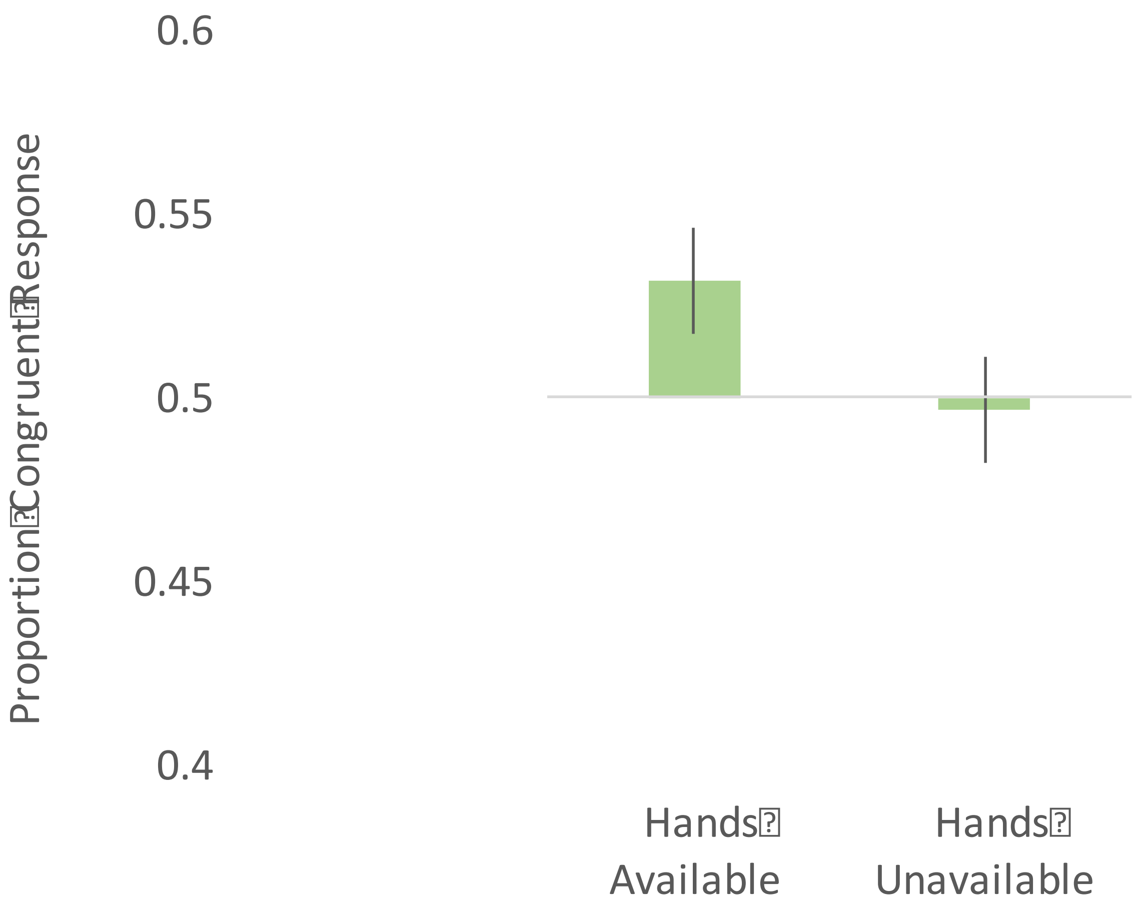

Figure 3. Results from the correct comprehension questions and no additional interpretations dataset. Error bars represent 1 SEM.