Title
Semantic Representation in the Mirror Neuron System

Permalink
https://escholarship.org/uc/item/3hb134q3

Journal
Berkeley Scientific Journal, 16(1)

ISSN
1097-0967

Authors
Leong, Josiah
Ivry, Richard

Publication Date
2011

Peer reviewed|Undergraduate
Semantic Representation in the Mirror Neuron System

Josiah Leong¹, Richard Ivry²

¹Psychology, College of Letters and Science, University of California, Berkeley
²Department of Psychology, University of California, Berkeley

Keywords: mirror neurons, symbolic gestures, reaction time, lexical decision, priming

ABSTRACT

The mirror neuron system is a fronto-parietal network of neurons that is activated both when a person performs an action and when he or she observes that action. The goal of this study was to investigate the semantic representation in this system during action language and gesture processing. This was done in a set of two behavioral experiments. Experiment 1 employed a simple priming paradigm—subjects viewed videos of symbolic gestures or landscapes, which served as the prime, followed by a word that was congruent or unrelated to the video prime, or a pseudo-word. The subject performed a lexical decision task on the target word. The study found a significant priming effect for semantically congruent target words, relative to semantically unrelated target words. However, this same priming effect was found for the video primes of landscapes. In experiment 2 we aimed to determine whether the videos were primed through effector-specific means, that is, whether hand and arm gestures would activate mirror neurons somatotopically and lead to different priming results when comparing hand responses to foot responses. We used the same priming study as experiment 1 except that subjects made their lexical decision responses on a foot pedal rather than a keyboard. Results suggest that symbolic gestures prime in a very diffuse way, such that semantic priming occurs independent of the effector being used to respond.

INTRODUCTION

Gestures are defined by Kendon (1996) as bodily actions that are recognized as communicative. For example, someone ushering another person to ‘come here’, or holding up both hands to say ‘stop’, or waving ‘hello’ are all symbolic gestures regularly used to communicate. These gestures convey meaning non-verbally and require some neural substrate to lead to semantic comprehension.

A clue to where the neural overlap between gestures and language may be was presented with the discovery of the mirror neuron system (Rizzolatti & Craighero, 2004). This frontal-parietal network of neurons fires both in the production of actions and the observation of those actions (Rizzolatti et al., 2001). In fact, Rizzolatti et al. (2002) pointed to Brodmann area 44 in the inferior frontal gyrus as a potential location for understanding action, as BA44 is the human homologue to the monkey’s F5 area which he studied. Furthermore, Gentilucci et al. (2006) used repetitive TMS to show that Broca’s area has access to the conceptual information being conveyed by gestures. Broca’s area has traditionally been viewed as an essential part of the language system, and Nishitani et al. (2005) found that Broca’s contributes to action planning, observation, understanding and imitation.

It has also been speculated that the neural substrates for language were preceded and subserved...
by the motor system. Gentilucci & Corballis (2006) suggest that language may have been formed from manual gestures rather than vocalizations due to the strong overlap between action and language in communication. Additionally, developmental studies show that children used gestures to communicate even before babbling and word production (Bates & Dick, 2002). This is phylogenetic evidence pointing to the motor system as a major contributor to language. Finally, Bernardis & Gentilucci (2006) provided direct evidence showing that speech and gestures enhanced each other when performed together. Subjects’ voicing parameters were enhanced when they performed a symbolic gesture simultaneously to pronouncing a word. Observing another person speak and gesture simultaneously also enhanced the subjects’ voicing parameters, as compared to when the actor only did one or the other.

Our current aims are to further explore the mirror neuron system as a possible shared link between the language and action systems. The following studies investigate the semantic relationship between the comprehension of symbolic gestures and the identification of words.

We designed a priming study to observe the relationship between the observation of symbolic gestures and of words that are semantically congruent to the gesture. In experiment 1, participants performed a lexical decision task—deciding whether the target is an actual word or not. This target word was either semantically congruent or unrelated to the symbolic gesture previously viewed in the trial. If there is indeed a relation between the neural system for comprehending gestures and language, then the congruent trials should have a quicker reaction time compared to baseline conditions.

The human mirror neuron system has been shown to have a somatotopic organization; leading to different activations if the action observed involves different body parts, and even if it is only an action word of different body parts that’s being processed (Buccino et al., 2001; Hauk et al., 2004). In experiment 2 we tested whether using a different effector than the one observed in the video would decrease or abolish the semantic priming effect. If so, this would suggest that there is a very specific organization of the semantics of action in the mirror neuron system.

**MATERIALS AND METHODS**

**Pilot Study**

A pilot study was completed to establish which gestures being presented as primes were actually symbolic as opposed to meaningless.

**Gesture and Landscape video selection.** Fifteen participants were involved in the pilot. They all spoke English as their first language. Participants were placed in front of a 20” CRT monitor at a viewing distance of 56cm. They viewed 145 symbolic gesture videos, 78 meaningless gesture videos, and 88 landscape videos through the Eprime Studio 2.0 program. The gesture videos consisted of a person wearing a skin-colored mask performing a hand gesture for 1500ms. The mask was necessary because it ensured that the actor’s facial expressions were not conferring any meaning (see figure 1.1). The videos of landscapes were downloaded from the internet and

---

**Figure 1.1.** Example of a symbolic gesture prime.

**Figure 1.2.** Example of a landscape prime.
shown in a random order with the gestures (see figure 1.2).

The participants’ task consisted of rating each video on a Likert scale from 1 to 5, 1 meaning that the gesture had no recognizable meaning, and 5 meaning that it had a highly recognizable meaning. If the participant rated the video between 3 and 5 they were asked to type out its perceived meaning in three or less words.

If 75% or more of the pilot study participants agreed on a gesture as semantic (Likert rating of 4-5) then it was considered a “symbolic gesture”. Likewise, 75% or more of the participants had to rate a gesture with a 1 or 2 to make the gesture a “meaningless gesture”. For landscape videos, participants did not rate the video and only replied with what they believed the prominent object in the video was.

At the conclusion of the pilot study we selected 87 of the videos as symbolic gestures and 62 as meaningless gestures.

**Target words.** With our list of agreed upon gestures, we next had to find congruent words, unrelated words, and pseudo-words to associate with each video. We tallied each word for lexical agreement first—that is, the same root words were being answered by participants. We then inferred the semantic meaning that the subjects were stating for the gesture (e.g. halt and stop sharing a similar semantic meaning, though being different lexically), and then tallied those for each video between subjects. Here is an example of one of our stimuli and the results we received: a symbolic gesture video of a person holding out both outstretched hands towards the viewer led 8 subjects to respond to it as ‘stop’, 5 responded ‘slow down’, and 1 said ten. These were the lexical agreement tallies, whereas the semantic agreement tally would be 13 for ‘stop/halt/slow down’. Another example is a landscape video of lava bursting from a volcano being called ‘volcano’ by 11, ‘explosion’ by 2, and ‘sparks’ by 1 subject. In this instance, both the lexical and semantic agreement tallies were simply 11 for ‘volcano’.

At the end of the pilot we were able to select the stimuli for the future experiments. For congruent words, we used the most frequently tallied semantic meaning from the pilot. To select unrelated words, we used three criteria to choose our words: same part of speech as the congruent word, same frequency of appearance in general English literature as the congruent word, and within +/- one character length as the congruent word. Finally, the pseudo-words were created to follow the same character length requirement as the unrelated words, to use the same first letter of either the congruent or unrelated word, and to follow the basic English word structure rules. An example for the target words associated with the symbolic gesture used above: congruent word = stop, unrelated word = beat, pseudo-word = refi.

Within each category (gesture or landscape) the congruent and unrelated words were controlled for the frequency in which they occurred in English to control for the possibility that people respond faster to congruent words simply because they are seen more in regular life than the unrelated words. We used the Celex database and SubtlexUS website to find frequency data for each of our words. Both of these databases presented “frequency” as the number of times the word occurred per 10 million words in English literature.

**Experiment 1**

The first experiment after the pilot study sought to establish that videos of symbolic gestures indeed had a semantic priming effect on subjects. Eighteen subjects completed the study, eleven were female and seven male, all right-handed, and they received college course credit as compensation. The subjects sat at a desk with a desktop computer and keyboard, with a CRT monitor (50cm diagonal length). They were placed at a viewing distance of 56cm from the screen. Eprime Studio 2.0 was the software used to present the stimuli and record reaction times. Using a lexical decision task as the explicit instructions, we were able to measure participants’ reaction times across a number of conditions. The setup of each trial was as follows: Fixation (500ms) --> Video (1500ms) --> Target (150ms) --> Response (3000ms maximum time allotted to respond).

The various combinations of primes and targets that occurred in random order throughout each experimental block included: video primes of symbolic gestures, meaningless gestures, or landscapes; and target words that were either congruent or unrelated to the prime, and pseudo-words. The meaningless gesture-congruent word condition was obviously omitted because there are no semantically congruent words to gestures without meaning.

These trials were presented in 3 blocks: two blocks of 78 gestures each and one block of 80 landscape trials. The order of the blocks was counterbalanced
between subjects, as were the keyboard buttons used to make the lexical decision. Summing across the three blocks, there were 42 symbolic gesture-congruent trials, 28 symbolic gesture-unrelated, 24 meaningless gesture-unrelated, 46 symbolic gesture-pseudo-word, 16 meaningless gesture-pseudo-word, 24 landscape-congruent, 24 landscape-unrelated, and 32 landscape-pseudo-word. The experiment instructed participants to press the ‘green button’ if they thought the target word was a real word and the ‘red button’ if they thought it was not a word. These two buttons alternated every other subject between the ‘b’ and ‘m’ keys on a standard keyboard.

The meaningless gestures were included as a control to make sure that random arm movements weren’t priming anything despite the fact that they would have presumably activated mirror neurons through observation.

**Experiment 2**

In order to test the somatotopic strength of the semantic representation we ran a second experiment that changed the effector that subjects made the lexical decision task with. Seventeen subjects completed the study, ten were female and seven male, all right-handed, and they received college course credit as compensation. Experiment 2 repeated the precise procedure of experiment 1 except that subjects made their lexical decisions on a foot pedal that was colored green and red. A foot pedal was programmed through Eprime Studio’s e-basic scripting language to record reaction times after the target word appeared (exactly as the keyboard would do for experiment 1). Subjects were counterbalanced just like in the first experiment so that the right and left foot pedal buttons each were the green and red buttons half of the time.

**RESULTS**

**Experiment 1**

We first filtered the data for only accurate responses (making the correct lexical decision). There were no subjects with a significantly low proportion of correct responses, and there were no conditions that led to particularly low accuracy. The average accuracy across all subjects and conditions was 95%.

Figure 2.1 depicts the priming pattern for the eight conditions in this experiment. The reaction times for the symbolic gestures-congruent condition was significantly faster than the symbolic gestures-unrelated condition (t(17) = -4.282, p < .001). Additionally, landscape-congruent was significantly faster than landscape-unrelated (t(17) = -3.9434, p < .001). Because of the different frequency rate of the congruent words for the symbolic gestures and landscapes we were not able to compare directly across these two conditions, however the data of these two categories are comparable when analyzing the magnitude of the priming effect. Another finding to note is that all real word responses were significantly faster than pseudo-word responses. This could be attributed to a frequency or familiarity effect, as pseudo-words are never seen in normal reading and might be more difficult for people to recognize.

**PRIMING EFFECT.** The priming effect of the symbolic gesture and landscape videos on reaction time for experiment 1 is shown in figure 3.1. The priming effect was calculated by subtracting the average reaction time in the symbolic gestures-congruent word condition from the average reaction time in the symbolic gestures-unrelated word condition. This was done for each subject, resulting in 17 independent priming effects. These priming effects were then averaged to give us an average priming effect. This process was repeated for the landscape-congruent word and landscape-unrelated word conditions.

The average size of the priming effect of symbolic gestures on reaction time in experiment 1 was 38ms, and 57ms for landscapes. Though the priming effect was apparently larger for landscapes than for symbolic gestures, this was not a statistically significant difference (t(17) = -1.208, p < .25).

**Semantic and Lexical Agreement.** The pilot study also provided data on the strength of semantic and lexical agreement for each stimulus used in this experiment. Thus, we could separate the symbolic gesture-congruent word and landscape-congruent word conditions further into high and low agreement both semantically and lexically. We calculated the proportion of pilot subjects that agreed on the meaning (semantic) or specific word (lexical) that the stimulus was depicting. For example, in the pilot study, 93% of subjects agreed that the stimulus shown in figure 1.1 meant “stop, halt, get away”; and 81% of the pilot subjects submitted that the stimulus meant “stop”. These proportions were above the median proportion of agreement across all symbolic gestures and landscapes (85% for semantic agreement, 60% for lexical agreement). Thus this stimulus was designated as a high semantic agreement and high lexical
We filtered our reaction time data in experiment 1 by stimuli with high semantic agreement, low semantic agreement, high lexical agreement, and low lexical agreement; and we did this for symbolic gestures and landscapes. We could then average our subjects’ reaction times in each of these conditions (i.e. high semantic agreement symbolic gestures—congruent word, low semantic agreement landscapes—congruent word, etc.). Figure 4 depicts the average reaction times for each of the conditions.

A dependent means t-test on the reaction times for symbolic gestures with high semantic agreement versus low semantic agreement showed a faster reaction time for high semantic agreement symbolic gestures that was statistically significant ($t(17) = -4.1652$, $p < .001$). A dependent means t-test on symbolic gestures with high versus low lexical agreement revealed a statistically significant result as well ($t(17) = -3.8216$, $p < .005$).

Interestingly, this effect is not present for landscapes.
for both semantic agreement ($t(17) = 0.9127, p < .38$) and lexical agreement ($t(17) = -0.7059, p < .49$).

We repeated the analysis for priming effects on these data. Comparing the priming effect of high semantic agreement symbolic gestures with high semantic agreement landscapes did not lead to a statistically significant difference, nor did high lexical agreement symbolic gestures versus high lexical agreement landscapes ($p < .80$ and $p < .62$, respectively).

**EXPERIMENT 2**

We first filtered the data for only accurate responses (making the correct lexical decision). There were no subjects with a significantly low proportion of correct responses, and there were no conditions that led to particularly low accuracy. The average accuracy across all subjects and conditions was 97%.

Experiment 2 had similar results to experiment 1. Figure 2.2 shows the reaction times for each condition. There was statistically significant priming for both symbolic gestures-congruent word ($t(16) = -7.2689, p < .0001$) and landscape-congruent word ($t(16) = -4.4756, p < .001$).

As in experiment 1, the priming effect was larger for landscapes (71ms) than symbolic gestures (57ms) (figure 3.2), but again there was no significant difference between these two means ($t(16) = -0.8128, p < .4282$).

These results suggest that the same priming occurred from the symbolic gestures and landscapes regardless of the effector being used to respond. While these results are not what we expected, they do not rule out the possibility of somatotopically specific semantic representation in the mirror neuron system. In fact, these data provide a launching pad for future TMS studies that can be directed at specific regions of the premotor cortex to test the specificity of semantic representation in the system. This is discussed in further detail below.

**DISCUSSION**

To summarize, our question was whether there is semantic representation in the mirror neuron system for symbolic gestures. To test this, we used a priming paradigm with a lexical decision task. We found priming when subjects viewed videos of symbolic gestures and then responded to congruent words in a lexical decision task, relative to unrelated words.

We established landscape trials as a baseline measure; however our experiment found that significant priming occurred for the landscape-congruent condition as well. Moreover, the symbolic gesture-congruent word trials were not significantly different from the landscape-congruent word; and so no conclusions could be made specifically about semantic representation in the mirror neuron system. We recognized that these behavioral experiments alone were insufficient to answer our question, and must be paired with a series of TMS experiments to examine the neural substrates of the mirror neuron system and the mechanism of semantic representation.

Further analysis of experiment 1 by separating the trials of the experiment according to pilot study data on semantic and lexical agreement revealed that high agreement in these two categories led to faster reaction times in congruent trials for symbolic gestures, but not for landscapes. The fact that highly agreed upon gestures had an additional priming effect while highly agreed upon landscape primes did not suggests some inherent differences between the two types of primes. While embodiment theory suggests that this boost may be because the symbolic gestures activate mirror neurons which house semantic representation whereas landscapes do not, our set of experiments offer no direct conclusion (Aziz-Zadeh et al., 2006). Our findings for experiment 2 closely matched those of experiment 1, except that the priming effect for both symbolic gestures and landscapes were larger. This was to be expected, however, as the timeframe in making a foot response compared to making a hand response is longer. One proposed answer to receiving the same priming patterns in experiment 2 was that semantic activation is extremely diffuse, particularly if the video being viewed was 1500ms long. This diffuse activation would allow subjects to make quicker decisions on semantically congruent trials regardless of whether the mirror neurons were being activated somatotopically.

An unexpected finding that we must discuss is the priming effect of meaningless gestures. In experiments 1 and 2, the meaningless gesture-unrelated word condition was significantly faster than the meaningless gesture-pseudoword condition. One explanation of this result is that observing a person perform any bodily gesture at all activates the mirror neuron system, which primes people to respond to real words regardless of whether the gesture has any
semantic meaning. A more specific interpretation along these lines is that semantic content in the mirror neuron system is heterogeneous, and diffusely activated—leading even meaningless gestures to prime for real words at large.

There are thus two ideas that we theorized for the semantic representation of symbolic gestures. The first idea is that people say the meaning of the gesture in their head once they see it, which leads to understanding the gesture. This is one explanation for the priming effect found for congruent trials in our experiment, as the word conjured in a person’s head may effectively lead to a word priming effect. Another hypothesis is that symbolic gestures activate mirror neurons in the observer, and that semantic representation within this neural substrate itself leads to comprehension.

The second experiment we did sought to explore this second pathway hypothesis indirectly. Experiment 2 was designed to discover whether there was effector specificity in the priming from the symbolic gesture videos. We performed the same experiment as experiment 1, with the exception that subjects responded to the lexical decision task with their foot rather than their hand. If there was indeed highly specific semantic representation in the mirror neuron system, then there would not be priming when responding with a foot since subjects were viewing hand and arm gestures. Our results provided no answer to this question, and we look to TMS studies to explore the neural substrates underlying semantic representation in the mirror neuron system.

CONCLUSIONS

Symbolic gestures and landscapes were both successful semantic primes in our experiments, irrespective of which effector was used to respond. The lack of effector specificity does not eliminate the possibility that symbolic gestures activate the mirror neuron system somatotopically however, as semantic activation from the video primes may simply have been diffuse enough to cause the priming effect for any effector. Lastly, while videos of symbolic gestures and landscapes can both successfully serve as semantic primes, it is only the highly agreed upon symbolic gestures that have an additional benefit to comprehension. These experiments were insufficient to address the relationship between the mirror neuron system and semantic representation. However, the paradigm used in these behavioral experiments provides a solid framework to test the effects of TMS on various regions of the cortex.

FUTURE DIRECTIONS

Transcranial Magnetic Stimulation (TMS) has been used to create virtual lesions in the brain, and several studies have validated its use in disrupting (or enhancing, depending on technique) verbal responses to gesture production (Gentilucci et al., 2006). A future study employing repetitive TMS (rTMS) will be conducted to discover whether creating virtual lesions in the premotor cortex or Broca’s area can successfully abolish the gesture priming found in these behavioral experiments.

rTMS involves 40 seconds of high frequency bursts directed at a particular part of the cortex, and has been found to impair that region’s functioning for 45 to 60 minutes. For the purposes of this future study, rTMS will be directed at the premotor cortex for one cohort of subjects and Broca’s area for another cohort. Faux rTMS will be used on a control group. After receiving rTMS, subjects will perform the behavioral experiments laid out in this experiment. We will then employ the data analysis techniques described in this paper to look for changes in the priming pattern after rTMS. If the results confirm some impairment or change in semantic priming due to rTMS, we will repeat these experiments on increasingly specific regions of the cortex, guided by the somatotopic organization of the premotor cortex.

REFERENCES

5. Bernardis, Paolo, Salillas, Elena and...


