Neuroscience Letters 687 (2018) 71–76

Contents lists available at ScienceDirect

Neuroscience Letters

journal homepage: www.elsevier.com/locate/neulet

Research article

Dorsal premotor cortex is related to recognition of verbal and visual descriptions of actions in the first-person perspective

Hiroshi Shibata[a,*], Kenji Ogawa[b]

[a] Faculty of Medical Science and Welfare, Tohoku Bunka Gakuen University, Japan
[b] Department of Psychology, Hokkaido University, Japan

ARTICLE INFO

Keywords:
Embodied cognition
First-person perspective
Second-person perspective
Dorsal premotor cortex
fMRI

ABSTRACT

This fMRI study examined whether the perspective difference of a verbally and visually descripted action stimulus (i.e., sentence and picture) modulates activity in the motor-related area. The participants were presented with a sentence (e.g., "I grasp an apple" or "You grasp an apple") or a picture (e.g., a picture of grasping an apple in which a right hand appears from the bottom or from the top) as the experimental task. A full factorial analysis of variance model with stimulus modality (verbal vs. visual description) and perspective (first- vs. second-person perspective) was used. The fMRI results showed greater activity in the left dorsal premotor cortex in the first-person perspective than in the second-person perspective for both the verbal and visual descriptions. The results suggest that motor representation is more strongly recruited with the recognition of an action-related stimulus descripted in the first-person perspective than in the second-person perspective, independent of stimulus modality.

1. Introduction

The theory of embodied cognition indicates that cognition is formed through bodily interactions with the environment, and embodied cognition is assumed to involve language comprehension (see [1,2] for a review). For example, some functional magnetic resonance imaging (fMRI) studies have shown effector-specific activations in the motor-related regions during the comprehension of mouth-, arm-, or leg-related words [3–5]. These studies have indicated that motor representation is recruited with the recognition of verbal descriptions of actions.

The recognition of action involves the comprehension of perspective (e.g., who is the agent of an action). A previous fMRI study [6] asked participants to read a short phrase (e.g., "I hammer" and "He admires" in German) silently and reported no significant modulation according to the difference in perspective. By contrast, Papeo et al. [7] showed recruitment of motor simulation with the recognition of action words described in first person but not in the third person. In the study by Papeo et al., the participants' left primary motor cortex was subjected to transcranial magnetic stimulation (TMS) and TMS-induced motor-evoked potentials were recorded from their hand muscles. In their task, the participants were asked to read Italian action and nonaction verbs silently and to judge whether the syntactic subject of the verbs was first or third person. This task required the participants to focus on the syntactic subject; this was one difference from the task utilized by Tomasino et al. [6]. The motor-evoked potentials for the action verbs were significantly greater for first person than for third person subjects.

The difference in perspective also affects brain activity during the recognition of visual descriptions of actions (e.g., videos and static pictures). Previous fMRI studies have found perspective effects (the interaction of laterality of hand and hemisphere) largely in the parietal cortex [8–10]; however, activity modulation in the motor cortex has also been reported. For example, Jackson et al. [11] used video clips of right-hand and right-leg action performed from the first person perspective (1PP) and third person perspective (3PP) and showed a greater increase in activity in the left precentral gyrus with 1PP than with 3PP. Hesse et al. [8] also used video clips of right-hand action performed in 1PP and 3PP and reported that the left precentral gyrus was more activated with 1PP, whereas the right precentral gyrus was more activated with 3PP.

The results of these previous studies predict a modulation of the activity in the motor area by the difference in perspective for both verbal and visual descriptions of actions. This relationship of the recognition of verbal and visual perspective has been investigated in psycholinguistic studies, and some have shown that visual perspectives that are mentally imaged from linguistic stimuli are affected both by the

* Corresponding author at: Department of Rehabilitation, Faculty of Medical Science and Welfare, Tohoku Bunka Gakuen University, 6-45-1 Kunimi, Aoba-ku, Sendai, Miyagi 981-8551, Japan.
E-mail address: hshibata@rehab.tbgu.ac.jp (H. Shibata).

https://doi.org/10.1016/j.neulet.2018.09.025
Received 5 June 2018; Received in revised form 29 August 2018; Accepted 7 September 2018
Available online 15 September 2018

difference in the personal pronoun per se (e.g., “I”, “you”, or “he”) and by the contexts in which the personal pronoun is used [12–14]. For example, Brunyé et al. [12] showed that the use of the first-person pronoun (i.e., “I”) facilitates the image of 1PP (internal or egocentric perspective) in reading a one-sentence action description (e.g., “I am slicing the tomato”), but it facilitates the image of 3PP (external or allocentric perspective) when reading the same action description presented in a discourse context in which sentences that include information about a character (e.g., “I am a 30-year-old deli employee”) are inserted before the action description. Shibata [15] examined the influence of pronouns in a context in which the origin of the sentence was fixed to the participants' own words, i.e., the participants were asked to interpret the presented sentence as if they had uttered that sentence. The results showed that sentences with the first person pronoun (e.g., “I point the finger”) were related to the picture of 1PP (or internal perspective: the action hand was appearing from the bottom of the picture), whereas the second-person pronoun (e.g., “You point the finger”) was related to the image of a second person perspective (2PP) (or external perspective: action hand was appearing from the top of the picture, which was the same as a typical image of 3PP). The current study used first- and second-person pronouns as the sentence stimuli, based on the study by Shibata [15] and a situation in which “I” denoted the participants themselves and “you” denoted some other person.

We predicted that, if participants focus on the perspective of the presented stimuli, activity modulation that is dependent on the difference in perspective would be found for both verbally and visually descripted action stimuli. Previous neuroimaging studies have mainly examined brain activity associated with perspective difference for either verbally or visually descripted action stimuli. Furthermore, there are few neuroimaging studies that have examined a perspective effect for verbally descripted action stimuli, although the TMS study by Papeo et al. [7] showed that motor representations were particularly recruited with action verbs related to the first person. It is therefore unclear which brain areas are related to the perspective difference of verbally descripted action stimuli nor whether the difference in stimulus modality affects the perspective-dependent activity. The current fMRI study manipulated the stimulus modality (verbal vs. visual) and perspective (1PP vs. 2PP) factors simultaneously and examined whether a perspective effect (in particular, strong recruitment in 1PP) was observed in the motor area irrespective of stimulus modality.

2. Material and methods

Fourteen healthy volunteers participated in the experiment. All participants were right-handed, as measured by the Japanese version of the FLANDERS handedness questionnaire [16,17]. Of these, 2 participants with excessive head movements were excluded (> 2 mm). The data from the remaining 12 participants (4 females and 8 males, mean ± SD, age 25.4 ± 5.3) were analyzed. All participants provided written informed consent. The experiment was approved by the ethics committee of Tohoku Bunka Gakuen University and the ethics committee of Hokkaido University.

Eight Japanese sentences were formulated by manipulating the subject (“I” or “you”), predicate (“grasp” or “touch”), and object (“apple” or “orange”). “I” and “you” were used as the sentence stimulus for 1PP and 2PP, respectively. Eight pictures were prepared by manipulating the agent (a right hand appearing from the bottom or from the top for the picture stimulus of 1PP and 2PP, respectively), action (grasping or touching), and object (apple or orange). The pictures for 2PP were the same as typical images used for 3PP, but we labeled them as 2PP following the classification of the sentence stimuli. Each picture could correspond to any one of the eight sentences (see Fig. 1a).

Stimuli were presented on a 32-inch MRI-compatible liquid crystal display (NordicNeuroLab, Bergen, Norway) with a screen resolution of 1920 × 1080, using the Psychopy software (http://www.psychopy.org/). While inside the fMRI scanner, participants observed the stimuli through a mirror mounted above the head coil. The distance between the mirror and the monitor was 170 cm. Each sentence stimulus subtended a horizontal visual angle of 9.7° and a vertical visual angle of 0.8°. Each picture stimulus subtended both horizontal and vertical angles of 5.4°. In the experimental trials, each stimulus was presented at the center of the screen for 2.5 s and was repeated three times, with a blank screen interposed for 0.5 s between stimuli (see Fig. 1b for schematic). In the sentence condition, the participants were asked to read the sentence silently after each sentence was presented and to interpret the sentence as if it was their own words (e.g., as if the participants themselves had said “I grasp an apple.”). This instruction was to clarify the origin of the sentences: “I” denoted participants themselves and “you” denoted another person. In the picture condition, the participants were asked to observe the picture with attention after each picture was presented.

We inserted catch trials into the experimental trials to require the participants to pay attention to the three parts of the stimulus: they needed to attend to the agent information (i.e., 1PP or 2PP) in addition to the information of action (grasping or touching) and the object (apple or orange). In the catch trials of the sentence (or picture) stimulus, any one of the subject (agent), predicate (action), or object (object) was changed at the second or third presentation, e.g., “I grasp an apple” was presented at the first and second presentation, but “You grasp an apple” was presented at the third presentation. The participants were asked to respond by pressing a button with the right index finger immediately after the stimulus was changed. The button press was required only in the catch trials.

The experiment consisted of 3 runs, with each run having 38 trials (32 experimental trials and 6 catch trials, randomly presented). In the experiment trials, sixteen different stimulus trials (8 sentence trials and 8 picture trials) were used twice in each run. In the catch trials, 3 different sentence stimulus trials (subject, predicate, or object was changed) and 3 different picture stimulus trials (agent, action, or object was changed) were used once in each run. Each run began with the presentation of a fixation cross for 9 s. The stimuli were presented in a blocked design with the 8.5 s stimulus periods (3 stimuli and 2 blanks) alternating with 12.5 s of rest.

Before the experimental trials, the participants were given a description of the experimental procedure and were verbally instructed on the task. They then performed up to 16 practice trials outside the fMRI scanner until the task was properly understood. The same stimuli as in the experiment were used in the practice trials. The first four practice trials comprised a sentence experimental, a picture experimental, a sentence catch, and a picture catch trial.

After the participants completed the fMRI trials, they were asked to answer a questionnaire about the relation of perspective between the sentence and picture: they answered whether a 1PP or a 2PP picture was congruent with each sentence.

All scans were performed on a Siemens (Erlangen, Germany) 3-Tesla Prisma scanner with a 64-channel head coil. T2*-weighted echo planar imaging (EPI) was used to acquire a total of 269 scans for each run, with a gradient echo EPI sequence. The first three scans of each run were discarded to allow for T1 equilibration. The scanning parameters were: 3000 ms repetition time (TR); 30 ms echo time (TE); 90° flip angle (FA); 192 × 192 mm field of view (FOV); 94 × 94 matrix; 36 axial slices; and 3.0 mm slice thickness with 0.75 mm gap. T1-weighted anatomical imaging with an MP-RAGE sequence was performed with the following parameters: 2300 ms TR; 2.32 ms TE; 8 °FA; FOV; 256 × 256 mm matrix; 192 axial slices; and 1 mm slice thickness without a gap.

Image preprocessing and statistical analyses were performed using SPM12 (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/). All functional images were realigned to correct for movement artifacts, spatially normalized to the Montreal Neurological Institute (MNI) template with a voxel size of 3 mm^3, and smoothed using a 6 × 6 × 6 mm Gaussian kernel. Condition effects at each voxel were estimated according to the general linear model. The periods of