Brain & Language 141 (2015) 135–141

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Regular Article

First-person and third-person verbs in visual motion-perception regions

Liuba Papeo [a,b,*], Angelika Lingnau [b,c]

[a] Department of Psychology, Harvard University, Williams James Hall, 33 Kirkland Street, Cambridge, 02138 MA, United States
[b] Center for Mind/Brain Sciences (CIMeC), University of Trento, Corso Bettini, 31, 38068 Rovereto, Italy
[c] Department of Psychology and Cognitive Science, University of Trento, Corso Bettini, 84, 38068 Rovereto, Italy

ARTICLE INFO

Article history:
Accepted 22 November 2014
Available online 13 January 2015

Keywords:
Verbs and nouns
Visual motion perception
First and third person
Biological motion
Animacy

ABSTRACT

Verb-related activity is consistently found in the left posterior lateral cortex (PLTC), encompassing also regions that respond to visual-motion perception. Besides motion, those regions appear sensitive to distinctions among the entities *beyond* motion, including that between first- vs. third-person ("third-person bias"). In two experiments, using functional magnetic resonance imaging (fMRI), we studied whether the implied subject (first/third-person) and/or the semantic content (motor/non-motor) of verbs modulate the neural activity in the left PLTC-regions responsive during basic- and biological-motion perception. In those sites, we found higher activity for verbs than for nouns. This activity was modulated by the person (but not the semantic content) of the verbs, with stronger response to third- than first-person verbs. The third-person bias elicited by verbs supports a role of motion-processing regions in encoding information about the entity *beyond* (and independently from) motion, and sets in a new light the role of these regions in verb processing.

1. Introduction

The left posterior lateral temporal cortex (PLTC) is a brain territory particularly suitable to addressing relations between perception and conceptual representation. Nestled in this part of the brain are structures relevant for perception of visual motion (Beauchamp, Lee, Haxby, & Martin, 2002; Beckers & Homberg, 1992; Grossman et al., 2000; Watson et al., 1993) and for conceptual and language tasks (Damasio et al., 2001; Kable, Lease-Spellmeyer, & Chatterjee, 2002; Kemmerer, Castillo, Talavage, Patterson, & Wiley, 2008; Martin & Chao, 2001; Watson, Cardillo, Ianni, & Chatterjee, 2013). The functional boundaries among these regions remain unclear.

The study of verbs and nouns, two fundamental components of the human communication system, has revealed a general *preference* for verbs over nouns in the left PLTC. This effect has a "hot spot" in the posterior middle/superior aspect of the temporal gyrus (e.g. Bedny, Caramazza, Grossman, Pascual-Leone, & Saxe, 2008; Kable et al., 2002; Papeo et al., 2014; Peelen, Romagno, & Caramazza, 2012; Wallentin et al., 2011); in addition, verb-related activity has been shown to encompass the medial temporal area (MT; Saygin, McCullough, Alac, & Emmorey, 2010) and the posterior superior temporal sulcus (pSTS; Bedny et al., 2008; Deen & McCarthy, 2010), regions that are associated with visual perception of basic and biological motion, respectively.

What, in a verb-stimulus, can drive the activity of motion-processing regions?

One possibility is the "amount" of motion implied by the semantics of a verb, which would resolve in higher activity for motor (i.e. motor actions) than non-motor meanings (Deen & McCarthy, 2010; McCullough, Saygin, Korpics, & Emmorey, 2012; Saygin et al., 2010). However, the motor vs. non-motor categorical distinction has not been found consistently across studies (Bedny et al., 2008; Humphreys, Newling, Jennings, & Gennari, 2013). Notably, verb-related activity has been reported for verbs presented in sentential context (i.e. inflected to agree with the sentential subject), but not for verbs presented in isolation, in their infinitive form, i.e. with no subject (see Dravida, Saxe, & Bedny, 2013). Considering these observations, we asked whether the subject – in addition to, or as opposed to the content – of the verb could contribute to drive verb-related effects in motion-processing regions.

While motion is the input that drives the strongest activity in MT and pSTS, both regions also respond to static stimuli; in particular, images of static objects that own the potential to move purposefully and *do things* (e.g. faces, bodies and body parts) trigger greater activity than inanimate objects (Beauchamp et al., 2002; Chao, Haxby, & Martin, 1999; Downing, Jiang, Shuman, & Kanwisher, 2001; Kanwisher, McDermott, & Chun, 1997; Kourtzi & Kanwisher, 2000; Pyles, Garcia, Hoffman, & Grossman, 2007; Saxe, Jamal, & Powell, 2006). These effects suggest sensitivity to a

* Corresponding author at: LP, CIMeC, Corso Bettini, 31, 38068 Rovereto (TN), Italy.
E-mail address: liuba.papeo@gmail.com (L. Papeo).

http://dx.doi.org/10.1016/j.bandl.2014.11.011
0093-934X/

distinction among entities, based on animacy (i.e. animates vs. inanimates).

Besides the animate–inanimate one, self vs. others (or first- vs. third-person) is perhaps the most general distinction among entities. Asymmetric response to first and third-person stimuli was first observed in the monkeys' biological-motion STS region (Hietanen & Perrett, 1996), and then replicated in humans (Allison, Puce, & McCarthy, 2000; Chan, Peelen, & Downing, 2004; Leube et al., 2003; Saxe, Jamal, et al., 2006). For instance, with an elegant perceptual illusion, Leube et al. (2003) showed that visually identical motion elicited greater pSTS activity when perceived as generated by another, than as self-generated. The third-person bias has been extended to nearby regions specialized in visual analysis of bodies (Chan et al., 2004; Saxe, Jamal, et al., 2006).

We used the third-person bias effect as a test to assess the sensitivity of left PLTC regions to information about the entity (i.e. the person *I* vs. *s/he*) implied by verbs.

To be able to relate word-related effects to brain regions responsive to visual motion perception, we used functional magnetic resonance imaging (fMRI) and two functional localizer tasks, to define independently, and individually for each participant (see Saxe, Brett, & Kanwisher, 2006), PLTC activity driven during visual perception of basic and biological visual motion. We then assessed the anatomical relationship of the so-defined visual-perception regions of interest (ROIs) with PLTC regions responsive to words, independently identified with a third functional localizer task. In Experiment 1, we sought to replicate the preference for verbs over nouns across the PLTC-ROIs. In Experiment 2, we tested whether the response to verbs in each ROI was modulated by the semantic content (motor vs. non-motor), and/or by the information about the subject implied by the verb (first- vs. third-person).

2. Materials and methods

2.1. Participants

Participants were 18 healthy native-Italian speakers (eight females; 27-years old ± 10 *SD*), all right-handed, clear of contraindications to fMRI. The study was approved by the ethics committee of University of Trento. Participants gave informed consent and were paid for participation. Three were excluded due to excessive head motion during the scanning session (>4 mm in *x*, *y*, or *z* direction).

2.2. Procedures

For each participant, the whole study was conducted in a single scanning session. In Experiment 1, three localizers were used to define individually ROIs within the left PLTC responsive to: (1) words ([nouns + verbs] vs. baseline), (2) biological motion (biological vs. scrambled motion) and (3) basic motion (moving vs. static dots). We used functional data from the first localizer, including both verbs and nouns, and tested the verb-noun distinction in each identified ROI. Note that, although we used the same data set for defining word-related ROIs ([verbs + nouns] vs. baseline) and for testing the verb-noun effect in those same ROIs (verbs vs. nouns), the contrasts were orthogonal, thus preventing circularity (Kriegeskorte, Simmons, Bellgowan, & Baker, 2009).

In Experiment 2, we studied the effect of semantic content (motor action vs. non-motor) and person (first vs. third) in each ROI.

2.3. Experiment 1

Word localizer. Participants were instructed to read, in a single run, 12 blocks of verbs and 12 blocks of nouns (7 items per block for a total of 84 Italian verbs and 84 Italian nouns, each presented for 1 s, with 1 s inter-trial interval, ITI), randomly intermingled, and separated by 14 s of fixation. Verbs were in their infinitive form (*meditare*, *to wonder*), and nouns were in their singular form, preceded by the appropriate article (*la nuvola*, *the cloud*). Verbs and nouns had concrete (50%) or abstract meaning and were matched for length, $t(1,166) = 1.38$, $P = 0.16$, and written frequency (Bertinetto et al., 2005), $t(1,166) = 1.76$, $P = 0.08$. The list of stimuli is available as Supplementary Information. The experiment began with an instruction screen (5 s) and 15 s of fixation, and ended with 15 s of fixation, for a total duration of 11.43 min. Words appeared in black font on a light-grey background.

Biological Motion Localizer. Participants watched blocks of white point-light displays on a black background, depicting human actions (6 blocks) or scrambled animations (6 blocks; see Lingnau & Petris, 2013). Experimental blocks (14 s each) were separated by 14 s of fixation. Each point-light display lasted 1.5 s and was followed by 0.5 s ITI. The experiment began and ended with a 14-s blank-period, and lasted 5.83 min.

Basic motion localizer. The design was identical to the Biological Motion Localizer, except for the stimuli consisting of blocks of moving and static dots. Moving dots moved outwards along the radial axis at a speed of 4 deg/s.

2.4. Experiment 2

Participants were instructed to read attentively motor-action (hereafter, motor) and non-motor verbs presented in first- and third-person of the present tense. There were four runs of 5.25 min each. Each run began with a 15 s fixation screen followed by 5 s instruction screen (''read attentively''), and included two blocks of 7 items (each item presented for 1 s + 1 s ITI) for each of the four experimental conditions (for a total of 56 items per condition), and two blocks of meaningless letter strings. The very same motor and non-motor verbs were presented in first- (*scrivo*, *I write*) and third-person (*scrive*, s/*he writes*); however, in each run, an item could appear in one form only (either in first- or the third-person). Blocks were separated by 14 s of fixation.

Each run terminated with a testing phase (1.4 min) in which participants had to decide by key-press (yes-or-no response), whether each of eight probe-verbs appeared during the immediately preceding run. For each participant, four probe-items were randomly selected among the 56 presented in the prior run; the remaining four were randomly selected among the items assigned to the other three runs. Participants were instructed to respond ''yes'' when the probe-item matched the form (first- or third-person) in which it had appeared in the immediately preceding run. This demand was meant to encourage participants to attend to the whole stimulus, that is, to both the verb's suffix, defining the person of the verb, and the root-morpheme carrying information about the verb's semantic content.

Verbs for Experiment 2 were selected from a larger set in which each item had been classified by an independent panel ($N = 10$), as action-related or stative and, for those verbs judged as action-related, as related to ''upper limbs'', ''lower limbs'', ''head'' or ''whole body'' (see Papeo, Vallesi, Isaja, & Rumiati, 2009). For the final set of motor verbs, we selected items that were judged as specifically related to manual actions, by at least the 80% of the panel. Verbs across conditions were matched for written frequency (Bertinetto et al., 2005) and length ($Ps > 0.05$). The list of stimuli is provided as Supplementary Information.

Stimuli were back-projected onto a screen by a liquid crystal projector (frame rate: 60 Hz; screen resolution: 1024 × 768 pixels). Participants viewed the stimuli binocularly through a mirror above the head coil. The screen was visible as a rectangular aperture of 17.8 × 13.4°. Stimulus presentation, response collection and synchronization with the scanner relied on ASF software (Schwarzbach, 2011) based on MATLAB Psychtoolbox.