Brain & Language 147 (2015) 30–40

Contents lists available at ScienceDirect

Brain & Language

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

Semantic brain areas are involved in gesture comprehension: An electrical neuroimaging study

Alice Mado Proverbio [a,*], Veronica Gabaro [a], Andrea Orlandi [a,b], Alberto Zani [b]

[a] *NeuroMI-Milan Center for Neuroscience, Dept. of Psychology, University of Milano-Bicocca, Piazza dell'Ateneo Nuovo 1, 20126 Milan, Italy*
[b] *Institute of Bioimaging and Molecular Physiology, IBFM-CNR, Milan, Italy*

ARTICLE INFO

Article history:
Received 4 July 2014
Accepted 2 May 2015

Keywords:
ERPs
Action processing
Language
Semantic violation
N400
Body language
Mirror neurons

ABSTRACT

While the mechanism of sign language comprehension in deaf people has been widely investigated, little is known about the neural underpinnings of spontaneous gesture comprehension in healthy speakers. Bioelectrical responses to 800 pictures of actors showing common Italian gestures (e.g., emblems, deictic or iconic gestures) were recorded in 14 persons. Stimuli were selected from a wider corpus of 1122 gestures. Half of the pictures were preceded by an incongruent description. ERPs were recorded from 128 sites while participants decided whether the stimulus was congruent. Congruent pictures elicited a posterior P300 followed by late positivity, while incongruent gestures elicited an anterior N400 response. N400 generators were investigated with swLORETA reconstruction. Processing of congruent gestures activated face- and body-related visual areas (e.g., BA19, BA37, BA22), the left angular gyrus, mirror fronto/parietal areas. The incongruent–congruent contrast particularly stimulated linguistic and semantic brain areas, such as the left medial and the superior temporal lobe.

1. Introduction

In this study, the neural mechanisms underlying normal speakers' ability to understand spontaneous gestures was investigated. Gesture language comprises a set of actions, mostly involving facial mimicry and hand movements (but also of other body parts) that are used automatically either while talking with others to emphasize the message meaning (Beattie & Shovelton, 1999, 2002; Dick, Goldin-Meadow, Hasson, Skipper, & Small, 2009), or in replacement of oral speech (e.g., in noisy environments and/or with distant interlocutors). For example, nodding, very frequently used by infants to signify negation, and later on followed by hand shaking ("baby signs"), are both emblematic gestures (Kirk, Howlett, Pine, & Fletcher, 2013). Interestingly, it has been shown that people are better at understanding ambiguous utterances if allowed to see the accompanying gestures performed by speakers in videos (Guellai, Langus, & Nespor, 2014). For example, Holle and Gunter (2007) recorded EEG as participants watched videos of a person gesturing and speaking simultaneously. They found that N400 to target words was smaller after a congruent gesture and larger after an incongruent gesture, suggesting that listeners can use gestural information to disambiguate speech. In contrast to standard sign language (e.g., American Sign Language, ASL, British Sign Language, BSL, which is a formal language employing a system of hand gestures for communication, as by the deaf), spontaneous sign language is not formally taught and, although it does have a sort of grammatik and a relatively fixed set of rules, it is used in a rather flexible way. It presents considerable individual, regional, and cultural differences and it's implicitly learned and understood by speakers. Spontaneous sign language is defined as any means of communication through bodily movements, especially of the hands and arms, rather than through speech, which is spontaneously used by humans, without formal training. For this reason it has been conceptualized as halfway between a formal sign language and an emotional body language (EBL) system, for example by Andric et al. (2013).

If, on one hand, the neural systems underlying standard sign language comprehension have been widely investigated by neuroimaging and electromagnetic studies (Braun, Guillemin, Hosey, & Varga, 2001; Hänel-Faulhaber et al., 2014; Husain, Patkin, Kim, Braun, & Horwitz, 2012; Levänen, Uutela, Salenius, & Hari, 2001), the same does not hold for spontaneous gesture comprehension. As for the BSL, MacSweeney et al. (2002) identified a set of regions commonly activated in deaf and hearing sign language users, which included inferior prefrontal regions bilaterally (including Broca's area) and superior temporal regions bilaterally (including

---

* Corresponding author at: NeuroMI-Milan Center for Neuroscience, Department of Psychology, University of Milano-Bicocca, Piazza dell'Ateneo Nuovo 1, U6, 20126 Milan, Italy. Fax: +39 026448 3788.

E-mail address: mado.proverbio@unimib.it (A.M. Proverbio).

http://dx.doi.org/10.1016/j.bandl.2015.05.002
0093-934X/

Wernicke's area). Sign language (vs. audiovisual speech) generated enhanced activation in the posterior occipito-temporal regions (V5), most likely because of its dynamic nature. Interestingly, deaf native signers demonstrated greater activation in the left superior temporal gyrus in response to BSL than hearing native signers. The role of Broca's area for sign language production (Braun et al., 2001) and of Wernicke's area for sign language comprehension (Petitto et al., 2000) have been demonstrated by a series of neuroimaging and clinical studies (Hickok, Bellugi, & Klima, 1996; Poizner, Klima, & Bellugi, 1987).

To address the neural mechanism supporting the comprehension of spontaneous body language, some ERP studies have used the N400 linguistic paradigm as a contrast to ERPs for correct vs. incongruent gestures (e.g., Bach, Gunter, Knoblich, Prinz, & Friederici, 2009; Gunter & Bach, 2004), which represents a valuable tool for determining the neural processing time course for conceptual, linguistic and semantic information (Kutas & Federmeier, 2000; Kutas & Hillyard, 1980). Interestingly, this component has been shown to reflect the detection of semantic violations in sign language among deaf individuals (e.g., in American sign language: Neville et al., 1997, or in German sign language: Hänel-Faulhaber et al., 2014), thus suggesting the recruitment of auditory language-devoted structures in sign language processing.

To investigate the electrophysiological correlates of gesture comprehension in normal hearing speakers Bach et al. (2009) presented two hand actions as consecutive frames, one showing an instrument and the other one a potential target object of the action. Two mismatches were possible: a tool orientation mismatch or a misuse case (e.g., a screwdriver followed by a keyhole). Both types of violation were associated with the occurrence of similar N400 responses. Although these studies demonstrated the temporal component of brain processing, they only concerned hand actions, which are highly relevant to body language but cannot convey the complex pattern of affective and gestural indices provided by the entire body. One recent neuroimaging study (Andric et al., 2013) compared the neural processing of object-directed actions (such as grasping) with the processing of emblems, but the study only considered a limited set of hand actions (while the full body of the agent was not visible).

To pursue a more ecological approach, in a very recent ERP study on the mechanism of emotional body language (EBL) comprehension (Proverbio, Calbi, Manfredi, & Zani, 2014), we recorded ERPs in 30 Italian University students while they evaluated 280 full-body pictures displaying typical EBL patterns acted out by 8 male and female Italian actors. Half of the stimuli were incongruent with a previous verbal description (for example, "Come here, let me hug you!" followed by the portrait of visibly hostile man. ERP responses showed an anterior N400 response indicating the detection of incongruent body language, starting as early as 300 ms post-stimulus. SwLORETA was performed on the N400 difference by subtracting ERPs to congruent actions from ERPs to incongruent ones to identify the strongest generators of this effect in the right rectal gyrus (BA11) of the ventromedial orbitofrontal cortex, the bilateral uncus (limbic system) and the cingulate cortex, and in cortical areas involved in face and body processing (superior temporal sulcus (STS), fusiform face area (FFA) and extra-striate body area (EBA) and the premotor cortex (BA6), which is involved in action understanding. These findings are consistent with fMRI literature on brain response to emotional (not symbolic) body language (De Gelder et al., 2009). Recently Kana and Travers (2012) found significant activation in brain areas associated with visual representation (EBA and FFA), with action processing (inferior frontal gyrus (IFG) and STS, see also Wurm, Hrkać, Morikawa, & Schubotz, 2014) and inferior parietal lobule (IPL), and with emotion processing (anterior insula, medial prefrontal cortex (MPFC), striatum, superior colliculus and pulvinar) when interpreting the actions and emotions of stick figures. In this case as well, stimuli were not very ecologically valid because stick figures were used instead of real pictures.

The purpose of this study was multifold: (1) We wished to investigate the neural mechanisms underlying normal speakers' ability to understand spontaneous gestures. (2) We intended to determine whether the comprehension of gestures in healthy users was closer to the mechanism of sign language processing in non-hearing speakers, involving linguistic brain areas, or more similar to affective body language comprehension, as investigated in a previous ERP study (Proverbio et al., 2014). (3) By applying the N400 paradigm we aimed at studying for the first time the processing of a large corpus of spontaneous gestures (including deictic, transformative, metaphoric, emblematic, iconic and motor gestures (see Goodwin, 2003) engaging both body and face processing. (4) In order to provide the richest stimulation as possible individual and dialectal variations of bodily expressions were considered by recruiting 6 different male and female actors. To accomplish these goals, more than one thousand pictures were taken of actors displaying a large variety of distinctive and highly consistent gestures, which normally accompany social communication in the Italian culture.

Emblematic gestures represent a meaning without relying on spoken context (Ekman & Friesenm, 1969; McNeill, 2005) and, unlike formal sign language, are not combined into complex gesture strings to make longer sentences (Goldin-Meadow, 1999). On the other hand, iconic gestures (Holler & Beattie, 2003) are hand gestures that represent meaning that is closely related to the semantic content of the segments of speech that they accompany (McNeill, 2005). For example, the utterance: "she's eating the food", might be accompanied by the iconic gesture "left hand moves toward the mouth" (Beattie & Shovelton, 1999). Iconic gestures are typically large complex movements that are performed relatively slowly and carefully in the central gesture space, but they can also be small and fast; speech and gesture refer to the same event, but each presents a somewhat different aspect of it (McNeill, 1992). The meaning of words and gestures is processed in an integrated manner, as suggested by an ERP study showing larger N400 responses to target words preceded by incongruent gestures (Holle & Gunter, 2007).

In the present study, gestures used as stimuli did not express an emotional state, but delivered precise semantic information about speakers' plans, desires, motivations, attitudes, believes and thoughts (such as: "I am getting out of here", "I am not letting you do this", "it's cold", "it was so tall", "he is lying", "it happened a long time ago", "there were so many").

Each gesture was performed by 6 different actors (3 males and 3 females) and specifically validated for its clarity and comprehensibility by an independent group of 18 judges before ERP testing. Each selected gesture was associated with a short descriptive label that could be either congruent or incongruent with the gesture itself. We expected to observe an N400 modulation of ERP responses as a function of sentence-image congruence. Electromagnetic tomography (LORETA) was applied to the bioelectrical activity recorded during the processing of congruent and incongruent gestures and to the differential activity. Because the gestures and related body language were semantically, rather than emotionally, pregnant in nature, we expected to find a scarce involvement of emotional areas compared to the previous study in which actors displayed emotional states (e.g., "I am rather disappointed", "I am in love with you", "I am deeply disgusted" "I am so ashamed", etc."; Proverbio et al., 2014). Furthermore, we expected to find a greater involvement of linguistic brain areas, with some similarity to neural circuits involved in sign-language processing (Levänen et al., 2001). Therefore, although the two set of stimuli (people gestures involving face and body expressions) and