Brain & Language 125 (2013) 77–81

Contents lists available at SciVerse ScienceDirect

Brain & Language

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

Short Communication

Co-speech gesture production in an animation–narration task by bilinguals: A near-infrared spectroscopy study

Misato Oi [a], Hirofumi Saito [a,*], Zongfeng Li [b], Wenjun Zhao [a]

[a] Department of Cognitive Informatics, Graduate School of Information Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan
[b] Hitachi Systems, Ltd., Ohsaki, Shinagawa-ku 141-8672, Japan

ARTICLE INFO

Article history:
Accepted 14 January 2013
Available online 28 February 2013

Keywords:
Co-speech gesture
Gesture production
Speech load
Inferior frontal gyrus (IFG)
Posterior superior temporal sulcus (pSTS)
Near-infrared spectroscopy (NIRS)

ABSTRACT

To examine the neural mechanism of co-speech gesture production, we measured brain activity of bilinguals during an animation–narration task using near-infrared spectroscopy. The task of the participants was to watch two stories via an animated cartoon, and then narrate the contents in their first language (Ll) and second language (L2), respectively. The participants showed significantly more gestures in L2 than in L1. The number of gestures lowered at the ending part of the narration in L1, but not in L2. Analyses of concentration changes of oxygenated hemoglobin revealed that activation of the left inferior frontal gyrus (IFG) significantly increased during gesture production, while activation of the left posterior superior temporal sulcus (pSTS) significantly decreased in line with an increase in the left IFG. These brain activation patterns suggest that the left IFG is involved in the gesture production, and the left pSTS is modulated by the speech load.

1. Introduction

In a conversation, speakers often spontaneously and unwittingly produce "gestures" that are movements of their hands and arms in conjunction with their speech (McNeill, 1992). The function of co-speech gestures has been discussed from a listener-oriented perspective (Graham & Argyle, 1975; Kendon, 1987) versus speaker-oriented perspective (Alibali, Kita, & Young, 2000; Kita, 2000; Krauss, Chen, & Chawla, 1996; Morrel-Samuels & Krauss, 1992; Rauscher, Krauss, & Chen, 1996). The aim of the present study is to explore the neural basis of co-speech gestures from the speaker-oriented perspective.

Two hypotheses were proposed to explain how co-speech gestures aid speech production: the lexical access hypothesis and the information packaging hypothesis (Alibali et al., 2000; Kita, 2000). According to the lexical access hypothesis, gesture facilitates lexical retrieval (Krauss et al., 1996; Morrel-Samuels & Krauss, 1992; Rauscher et al., 1996). Instead of the lexical retrieval, the information packaging hypothesis claims that "[the production of] gesture helps speakers organize rich spatio-motoric information into package suitable for speaking" (Kita, 2000, p. 163). In other words, "gesture is involved in the conceptual planning" (Alibali et al., 2000, p. 593).

Despite essential differences in the two hypotheses, if gesticulation may reduce cognitive load during speech, this would be consistent with the finding that more gestures were observed when bilinguals speak in their second language (L2), which is their less proficient language, than when they speak in their first language (L1) (e.g., Marcos, 1979; Nagpal, Nicoladis, & Marentette, 2011; Nicoladis, Pika, Yin, & Marentette, 2007). However, the results are controversial on types of gestures (see reviews in Nagpal et al., 2011; Nicoladis et al., 2007).

Recently gesture perception studies have set out to investigate the neural correlates of co-speech gestures, and most of them have investigated the neural integration of speech and gesture (e.g., Gentilucci, Bernardis, Crisi, & Dalla Volta, 2006; Holle, Obleser, Rueschemeyer, & Gunter, 2010; Skipper, Goldin-Meadow, Nusbaum, & Small, 2007; Skipper, Goldin-Meadow, Nusbaum, & Small, 2009; Willems, Özyürek, & Hagoort, 2007).

For example, Holle et al. (2010) presented participants three types of movies during functional magnetic resonance imaging (fMRI) scanning. In a gesture and speech movie, an actress verbally described a manual object action (e.g., "And now I grate the cheese") while simultaneously producing a gesture that illustrated the meaning of the sentence. In a gesture only or a speech only movie, the actress produces gestures without speech or speech without gestures. They found higher brain activation in the bilateral posterior superior temporal sulci (pSTS) of the observers in the gesture and speech condition than that in the speech only or the gesture only conditions. The activity of the left pSTS increased when a gesture accompanied an utterance in noise. The results suggest that the left pSTS is involved in the facilitation of speech comprehension by concurrent gestural input.

* Corresponding author. Fax: +81 052 789 5452.
E-mail address: saito@is.nagoya-u.ac.jp (H. Saito).

0093-934X/$ - see front matter
http://dx.doi.org/10.1016/j.bandl.2013.01.004

Using fMRI, Willems et al. (2007) observed that the left inferior frontal gyrus (IFG) is more active when gestural information did not match the preceding sentence context (e.g., the sentence, "He should not forget the items that he *wrote* on the shopping list," was accompanied by the incompatible gesture "hit") than when the gesture matched the context. They argued that the higher activation of the left IFG in the mismatch condition reflects an increase in the semantic integration load.

Taken together, these gesture perception studies suggest that the left IFG (Gentilucci et al., 2006; Skipper et al., 2007; Willems et al., 2007) and the left pSTS (Holle et al., 2010) would be involved in the gesture-speech integration. In addition, Gentilucci, Dalla Volta, and Gianelli (2008) in a review proposed that the system governing both speech and gesture is located in the left IFG.

In contrast to gesture perception research using brain imaging, there have been relatively few studies concerning gesture production research due to the limitation of the measuring technique such as fMRI and positron emission tomography (PET) that are unable to assess cortical function in ambulant participants.

One exception is an fMRI study of Hermsdörfer, Terlinden, Mühlau, Goldenberg, and Wohlschläger (2007), in which the movement of the participants' upper and lower arms, was restricted so that they were only allowed forearm pronation and supination. The participants were "required" to demonstrate the use of tools (e.g., hammer) either as pantomimes or with the tool in hand. In both the actual tool use and the pantomime conditions, they found widespread activation in the parietal, posterior temporal, frontal, and subcortical areas. It is noteworthy that they were concerned with the commonality and difference of the neural mechanism between pantomime and actual tool use, but not the neural mechanism of "spontaneous" co-speech gesture.

Unlike other neuroimaging techniques, near-infrared spectroscopy (NIRS) imposes few physical constraints on the participant and is relatively unaffected by motion artifact, which permits serial assessments of tasks in relaxed and natural environments (Leff et al., 2011; Liu, Saito, & Oi, 2012).

Our purpose in the present study was to investigate the neural bases of spontaneous co-speech gesture production using NIRS. To address this question, we asked participants who were unbalanced Chinese (L1)–Japanese (L2) bilinguals to narrate animated cartoons in their Ll and L2 during NIRS measurement. We focused on the function of conceptualization by gestures. In the animation–narration task, gestures would facilitate conceptualizing (reframing) the cartoon which they watched, that is, the participants recall actions and intentions of the characters, and prepare for the contents and the frame of their narration. We examined whether brain activity involved in the gesture production would differ between the L1 and the L2 conditions.

Previous studies indicate that the left IFG (Gentilucci et al., 2006; Skipper et al., 2007; Willems et al., 2007) and the left pSTS (Holle et al., 2010) are involved in gesture perception. These areas correspond to the "classical language areas" referred to as Broca's area and Wernicke's area, respectively. Despite a classical view that Wernicke's area is involved in speech perception, recent neurocognitive studies reported Wernicke's area is related to speech production as well as Broca's area (see reviews by Indefrey, 2011; Indefrey & Levelt, 2004). Based on the close connection between gesture and speech (McNeill, 1992), we predicted that Broca's area (the left IFG) and Wernicke's area (the left pSTS) would be concerned with co-speech gesture production. More specifically, we supposed that the left IFG is involved in gesture production, as suggested by Gentilucci et al. (2008). If this is the case, then the left IFG would show increased activation when the participants produce a gesture. If we consider that Broca's area and Wernicke's area are connected by the arcuate fasciculus which transmits information both anteriorly and posteriorly (Crosson, 1985), it is plausible that the left pSTS would show decreased activation which reflects reduction of the cognitive load in line with an increase in the left IFG.

To investigate distinctive activation patterns of the left IFG and pSTS for the gesture production, we measured brain activity of bilinguals during an animation–narration task using NIRS. The task of the participants was to watch two stories of an animated cartoon, and then narrate the contents in their Ll and L2, respectively.

2. Results

2.1. Behavioral data

We averaged speech time (second) and number of words across the four sections and participants. Neither the speech time (L1: $M = 76.16$, $SD = 35.03$; L2: $M = 116.50$, $SD = 57.10$) nor the number of words (L1: $M = 177.56$, $SD = 59.66$; L2: $M = 188.81$, $SD = 108.71$) showed significant difference between the L1 and the L2 conditions (speech time: $t(7) = 2.10$, $p = .074$, $r = .62$, two-tailed; number of words: $t(7) = 0.36$, $p = .726$, $r = .14$, two-tailed).

To examine whether the participants produced gestures more frequently in their L2 than L1, we compared the average gesture rate (per 100 words) across the four narration sections in the L1 ($M = 4.91$, $SD = 3.18$) and the L2 conditions ($M = 8.95$, $SD = 4.09$). The participants showed a significantly higher gesture rate in L2 than in L1 ($t(7) = 3.81$, $p = .003$, $r = .82$, one-tailed).

Fig. 1 shows mean number of gestures and words in each session in the L1 and the L2 conditions. A 2 (L1, L2) $\times$ 4 (four sessions) chi-square test revealed significantly different patterns in terms of number of gestures between the L1 and the L2 conditions ($\chi^2(3) = 11.80$, $p = .008$). The fourth session showed significantly lower number of gestures than the second session in the L1 condition ($\chi^2(1) = 28.77$, $p < .001$), but not in the L2 condition. The number of words did not show significant difference between these sessions in the L1 and the L2 conditions.

2.2. NIRS data

To examine whether the brain activities in the left IFG and the pSTS are modulated by the gesture production, we compared the *z*-scores of CoxyHb of these regions before (pre) and after (post) gesture production. Two-way analyses of variances (ANOVA) for *z*-scores of Coxy-Hb with Gesture (gesture, no-gesture) and Period (pre, post) as the within factors were performed in the L2 condition, then in the L1 condition. The analyses were carried out respectively, on channels 15, 18, 19, and 22 presumably located

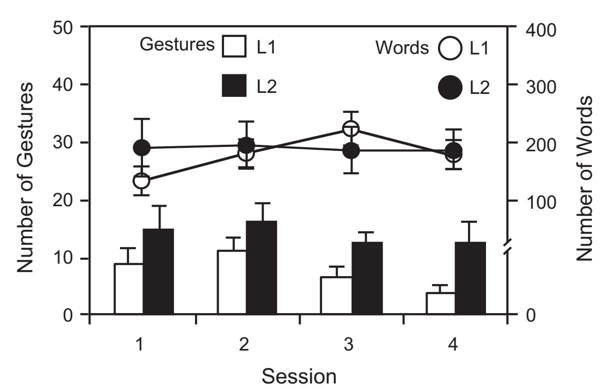

Fig. 1. Average number of gestures and number of words in the first language (L1) and in the second language (L2) conditions. Error bars represent standard error values.