available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

BRAIN RESEARCH

Research Report

Neural substrates of irony comprehension: A functional MRI study

*Midori Shibata**, *Akira Toyomura, Hiroaki Itoh, Jun-ichi Abe*

Department of Psychology, Hokkaido University, Kita 10, Nishi 7, Kita-ku, Sapporo, 060-0810, Japan

ARTICLE INFO

Article history:
Accepted 14 October 2009
Available online 22 October 2009

Keywords:
Irony comprehension
Language
Pragmatic
Echoic mention
Mentalizing
fMRI

ABSTRACT

In daily communication, we sometimes use ironic expressions to convey the opposite meaning. To understand these contradictory statements, we have to infer contextual implications and the speaker's mental state. However, little is known about how our brains carry out these complex processes. In this study, we investigated the neural substrates involved in irony comprehension using echoic utterance (Sperber and Wilson, 1986, 1995). Participants read a short scenario that consisted of five sentences. The first four sentences explained the situation of the protagonists. The fifth connoted either an ironic, literal, or unconnected meaning. The participants had to press a button to indicate whether or not the final sentence expressed irony. In the ironic sentence condition, the bilateral superior frontal gyrus, middle frontal gyrus, inferior frontal gyrus, medial prefrontal cortex, superior temporal gyrus, inferior parietal lobule, caudate, thalamus, the left insula, and amygdala were activated. In the literal sentence condition, the right superior frontal gyrus, the bilateral middle frontal gyrus, inferior frontal gyrus, medial prefrontal cortex, superior temporal gyrus, inferior parietal lobule, caudate, the left insula, the right thalamus, and the left amygdala were activated. However, in the ironic sentence condition minus the literal sentence condition, we observed higher activation in the right medial prefrontal cortex (BA 10), the right precentral (BA 6), and the left superior temporal sulcus (BA 21). Our results suggest that irony comprehension is strongly related to mentalizing processes and that activation in these regions might be affected by higher-order cognitive operations.

1. Introduction

Irony is a pragmatic statement that conveys a meaning opposite to its literal meaning. An ability to comprehend irony is required to infer the intentions, beliefs, and feelings of a speaker who expresses such an opposite meaning. Many cognitive studies have investigated the processing of irony comprehension over the past 20 years (Colston, 1997, 2000; Colston and Gibbs, 2002; Dews and Winner, 1995, 1999; Gibbs, 1986a,b; Giora and Fein, 1999a,b; Grice, 1975; Jorgensen et al., 1984; Kumon-Nakamura et al., 1995; Sperber and Wilson, 1986, 1995; Winner, 1988). In the traditional model of irony comprehension, ironic utterances flout the first maxim of quality, and the listener tries to determine the connotation of an utterance only after failing to find a literal meaning (Grice, 1975). This model implies that irony comprehension requires more steps than literal comprehension. In contrast, Gibbs and colleagues criticized the Gricean account and proposed that both figurative and literal language comprehension follow similar processing mechanisms and assume an initial

* *Corresponding author.* Fax: +81 11 706 3047.
E-mail address: shibata@let.hokudai.ac.jp (M. Shibata).

0006-8993/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.brainres.2009.10.030

influence of context on lexical access, which allows for direct comprehension of the context-relevant (i.e., ironic) meaning without an incompatibility phase. Sperber and Wilson (1986, 1995) also criticized Gricean analysis of communication and proposed the Relevance Theory, which defined verbal irony as an "echoic mention" of the prior utterance. According to this theory, irony is not the "language use" to express the speaker's thought, but rather the mention of apparent blame or criticism where specific desires or expectation must be involved. Through "echoic mention", the speaker achieves relevance by informing the hearer of the fact that he has in mind the content of prior utterance, and has a certain attitude to it. That is, an ironical utterance involves the implicit expression of an attitude that is mostly of the rejecting or disapproving kind. To understand the ironical utterance, the hearer must infer the implied meaning of the utterance echoed. This suggests that an understanding of irony depends on the ability to read another's mind, such as a speaker's situational awareness and beliefs (mentalizing). Hence, the relationship between an irony comprehension and mentalizing (theory of mind) functions is remarkably close.

Previous developmental and neuropsychological studies with patients have also suggested that irony comprehension is correlated to a theory of mind or mentalizing (Frith and Frith, 2003; Wimmer et al., 1985; Winner, 1988). In developmental studies, Winner (1988) indicated that irony comprehension requires a recognition of the speaker's beliefs and attitudes. She also indicated that non-literal language comprehension (as in the use of metaphor and irony) is established via three steps: (1) detection of non-literal intent, (2) detection of the relationship between the sentence and speaker meanings, and (3) detection of the speaker meaning. Of these three steps, step 1 is more difficult in irony comprehension than in metaphor comprehension. With regard to irony comprehension, young children often tend to confuse such an utterance with deception. If the message is recognized as contrary to the speaker's beliefs, and thus as intentionally false, the listener has to infer the speaker's intention (second-order belief). It is difficult for young children to comprehend the mental states of others (Winner, 1988).

Neuropsychological studies have been performed with patients who cannot understand the figurative use of language (Winner et al., 1998; Frith and Happé, 1994; Happé, 1993; Leekam and Prior, 1994; Shamay-Tsoory et al., 2005; Wang et al., 2006). Happé (1993) investigated specific difficulties of communication in autistic people using relevance theory (Sperber and Wilson, 1986). Relevance theory (Sperber and Wilson, 1986) explicitly defines a role for the comprehension of intentions in human communication and predicts about the levels of communicative competence which required not only first-order theory of mind ability (another person's belief) but also second-order theory of mind ability (the attribution of a belief about another person's belief). Three types of autistic patients and a normal control group were tested for their ability to understand figurative language (metaphor, simile and irony) which requires some understanding of intentions. The results showed that second-order metarepresentation is necessary to comprehend irony as an expression of the speaker's attitude to an attributed thought and supported Sperber and Wilson's relevance theory. Pediatric functional MRI (fMRI) studies in children and adolescents with autism spectrum disorders (ASD) also suggested that similar neural mechanisms were involved in irony comprehension and mentalizing. Wang et al. (2006) investigated the neural circuitry underlying deficits in understanding irony in children and adolescents with autism spectrum disorders (ASD) using functional magnetic resonance imaging (fMRI). Reliable activity was observed in the medial prefrontal cortex (MPFC) of the typically developing (TD) group but not in the ASD group. While both groups showed activity in the inferior frontal gyrus (IFG) in the left hemisphere, only the ASD group showed reliable prefrontal activity in the right hemisphere. The authors concluded that the increased activity in the ASD group was reduced in the network recruited in the TD group and may have reflected a greater effort in the processing needed to interpret the intended meaning of an utterance. These results are consistent with the notion that individuals with ASD have difficulty interpreting the communicative intent of others. Thus, these neuropsychological studies as described above suggest that irony (sarcasm) comprehension is associated with the regions that play a role in the theory of mind or mentalizing.

In recent neuroimaging studies, a few researchers have investigated the neural basis of irony (sarcasm) comprehension with normal participants (e.g., Eviatar and Just, 2006; Uchiyama et al., 2006; Wakusawa et al., 2007). In contrast to metaphoric and literal utterances, Eviatar and Just (2006) investigated neural activity using ironic utterances in brief three-sentence stories. Metaphoric utterances resulted in significantly higher levels of activation in the left IFG and in bilateral inferior temporal cortex than the literal and ironic utterances. On the other hand, ironic utterances resulted in significantly higher levels of activation in the right superior and middle temporal cortices than metaphoric statements. Thus, the results showed a difference in hemispheric sensitivity in relation to these aspects of figurative utterances. Wakusawa et al. (2007) investigated the cortical mechanisms that underlie the processing of implicit meanings (particularly irony) in social situations using photographs of daily situations. Participants were shown pictures depicting daily communicative situations and were asked to judge whether the presented utterance was situationally appropriate or not (the situational task) and whether the presented utterance was literally correct or not (the literal task). Their results showed that the right temporal pole was activated task independently during irony-specific processing and that the medial orbitofrontal cortex was activated task-dependently during irony processing in situational tasks. Their results also showed that the left MPFC showed significantly greater activation during the situational task than the literal task. They concluded that each region had a different function in processing implicit meaning and irony in social situations. Uchiyama et al. (2006) investigated cerebral activity with fMRI in healthy participants using scenario-reading tasks. The participants read sarcastic, non-sarcastic, or contextually unconnected sentences. The detection of sarcasm activated the left temporal pole, the superior temporal sulcus (STS), the MPFC, and the IFG (Brodmann's area 47) in both sarcastic and non-sarcastic sentences. Their findings indicated that sarcasm detection (sarcastic and non-sarcastic sentences contrasted with unconnected sentences) activated the neural