



## Context-prosody interaction in sarcasm comprehension: A functional magnetic resonance imaging study

Tagiru Nakamura<sup>a,b</sup>, Tomoko Matsui<sup>c</sup>, Akira Utsumi<sup>d</sup>, Motofumi Sumiya<sup>e</sup>, Eri Nakagawa<sup>e</sup>, Norihiro Sadato<sup>e,\*</sup>

<sup>a</sup> Faculty of Data Science, Musashino University, 3-3-3 Ariake, Kotoku, Tokyo, 135-8181, Japan

<sup>b</sup> Research Center for Liberal Education, Musashino University, 1-1-20 Shimmachi, Nishitokyo, Tokyo, 202-8585, Japan

<sup>c</sup> Center for Research in International Education, Tokyo Gakugei University, 4-1-1 Nukuikitamachi, Koganei, Tokyo, 184-8501, Japan

<sup>d</sup> Department of Informatics, Graduate School of Informatics and Engineering, The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo, 182-8585, Japan

<sup>e</sup> Division of Cerebral Integration, Department of System Neuroscience, National Institute for Physiological Sciences (NIPS), 38 Nishigonaka, Myodajji, Okazaki, Aichi, 444-8585, Japan

### ARTICLE INFO

#### Keywords:

Incongruity  
Functional magnetic resonance imaging  
Mentalizing network  
Relevance network  
Salience network

### ABSTRACT

During conversation, sarcasm is perceived as an incongruity between the context, content, and prosody of the utterance. We hypothesized that prosody modifies the context–content incongruity effect. Thus, we conducted a functional magnetic resonance imaging study with an auditory sarcasm detection task in 22 healthy adult participants. The participants listened to a short conversation according to which they had done either a good or bad deed, about which their conversational partner made a positive comment. When the context was positive (congruent with the content of utterance), positive prosody lessened the sarcasm rating, whereas negative prosody enhanced this rating. When the context was negative, the positive prosody effect disappeared, while negative prosody increased the sarcasm rating. Thus, context–content incongruity is the primary determinant of sarcasm comprehension; and is modified by prosody in a context-dependent manner. Neuroimaging results showed that the context–content incongruity effect was notable in the cerebellum and the mentalizing network, representing what was uttered in a particular context. The content–prosody incongruity effect was observed in the bilateral amygdala, representing the manner of utterance. The interaction between these incongruity effects was found in the bilateral dorsolateral prefrontal cortex, extending to the inferior frontal gyrus and the salience network, including the anterior insular cortex and the caudal part of the dorso-medial prefrontal cortex. These findings indicate that two distinct incongruity detection systems for sarcasm comprehension are integrated in the prefrontal cortices through the salience network.

### 1. Introduction

Sarcasm mediates implicit criticism of the listener by provoking negative emotions with disapproval, contempt, and scorn. The term sarcasm is a subcategory of the term irony, which aims to transfer the intention that cannot be conveyed literally, either for positive goals (e.g., humor, emphasis) or for opposing goals (e.g., sarcasm, criticism) (Utsumi, 2000). This study focuses on sarcasm. Speakers express negative emotion (e.g., disappointment, anger, reproach, envy) regarding the difference between their expectations and failure. They implicitly display these emotions by various means: an allusion to the expectation, a pragmatic insincerity by intentionally violating one of the pragmatic

principles, and an indirect expression of the negative emotion toward failure to meet their expectation (Utsumi, 2000). During the conversation, sarcasm is perceived as a multi-layered incongruity among the utterance's context, content, and prosody. The incongruity between the context and the content of utterance enables the hearer to understand the ironic intent of the speaker (Ackerman, 1983; Colston, 2002; Ivanko and Pexman, 2003; Katz and Lee, 1993; Katz and Pexman, 1997; Kreuz and Glucksberg, 1989). The larger the disparity between the content and the discourse context is, the more disapproval is perceived in the irony by the hearer (Colston and O'Brien, 2000; Gerrig and Goldvarg, 2000). Previous studies pointed out the importance of prosody in ironic understanding (Bryant and Fox Tree, 2005; Caillies et al., 2019; Capelli

\* Corresponding author.

E-mail address: [sadato@nips.ac.jp](mailto:sadato@nips.ac.jp) (N. Sadato).

<https://doi.org/10.1016/j.neuropsychologia.2022.108213>

Received 1 September 2021; Received in revised form 9 March 2022; Accepted 11 March 2022

Available online 12 March 2022

0028-3932/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

et al., 1990; Laval and Bert-Erboul, 2005; Le Gall and Iakimova, 2018; Tobe et al., 2016; Wickens and Perry, 2015). The ironic prosody is often enough for a hearer to identify an utterance as an instance of irony (Bryant and Fox Tree, 2005; Capelli et al., 1990). As children first use ironic prosody as an effective clue for irony comprehension at around the age of 5 years, before they can also make use of the discourse contexts as a clue (Laval and Bert-Erboul, 2005), ironic prosody is a distinctive cue for irony comprehension. Furthermore, the ironic prosody interacts with the discourse context in sarcasm comprehension (Rivière et al., 2018; Voyer et al., 2016; Woodland and Voyer, 2011).

Consider the following example from Utsumi (2000):

A mother asked her son to clean up his messy room, but he was lost in a comic book. After a while, she discovered that his room was still messy, and said to her son: "This room is totally clean!"

The son initially assumed that his mother spoke directly and honestly. However, when he recognized the incongruity between what he had expected to hear from his mother (i.e., the discourse context) and what he had heard (i.e., the utterance content), he was geared toward resolving context–content incongruity. The hearer also recognizes incongruity by detecting adverse effects displayed implicitly by paralinguistic cues such as affective prosody, facial expression, and verbal cues such as hyperbole (Utsumi, 2000).

The "implicit display theory" (Utsumi, 2000) holds that two factors are critical for an utterance to be interpreted ironically. The first factor is the ironic environment in which the discourse situation is located. The second factor is the negative attitude that is implicitly displayed by the speaker. The theory suggests that comprehension of sarcasm requires at least one of these factors, while another factor enhances this understanding (Utsumi, 2000). Within this theoretical framework, context–content incongruity corresponds to the implicit display of an ironic environment, and content–prosody incongruity corresponds to the indirect expression of negative attitudes. In other words, this theory predicts that the degrees of sarcasm due to context–content incongruity and content–prosody incongruity additively determine the overall degree of sarcasm (Utsumi, 2000). For example, the negative context of a messy room in the example above anticipates the mother's utterance with some negative content. However, in reality, when a mother's utterance with positive content is received, the utterance acts in a directional effect that increases the degree of sarcasm due to pragmatic insincerity, i.e., the context–content incongruity. When the utterance has negative prosody, the utterance acts in a directional effect that increases the degree of sarcasm due to the indirect expression of negative attitude, i.e., the content–prosody incongruity. As a result, the two effects additively provide the strongest sarcasm (Utsumi, 2000). When the utterance has positive prosody, the utterance acts in a directional effect that decreases the degree of sarcasm. The effects of the two different directions may cancel each other out, and the degree of sarcasm is weaker than in negative prosody. However, in the absence of context–content incongruity, the degree of sarcasm will be determined solely by the content–prosody incongruity. Importantly, sarcasm is understood as a clue to some incongruity, and the degree of sarcasm is related to the additive result of each incongruity (Utsumi, 2000).

Previous neuroimaging studies of the comprehension of sarcastic and non-sarcastic irony with context–content incongruity have revealed the involvement of mentalizing networks (Eviatar and Just, 2006; Rapp et al., 2010; Shibata et al., 2010; Spotorno et al., 2012; Wakusawa et al., 2007; Wang et al., 2006a), in addition to the classical language areas, including the left inferior frontal gyrus (IFG) (Spotorno et al., 2012; Uchiyama et al., 2006; Varga et al., 2013). These findings have been confirmed by meta-analyses (Bohn et al., 2012; Rapp et al., 2012; Reyes-Aguilar et al., 2018) and subsequent studies (Akimoto et al., 2014; Bosco et al., 2017; Filik et al., 2019; Herold et al., 2018; Obert et al., 2016). These findings are consistent with the notion that the successful comprehension of irony depends on the perceiver's ability to infer other

people's mental states, thoughts, and feelings (Channon et al., 2005; Filik et al., 2019; Frith and Frith, 2003). They have also suggested that the left IFG, including Brodmann's area (BA) 47, might be the brain region where mentalizing and language processes interact during sarcasm detection (Uchiyama et al., 2006).

Regarding negative attitudes that is the characteristics of sarcasm (Utsumi, 2000), previous studies have reported activation of the amygdala, which suggests emotional processing related to social behavior (Akimoto et al., 2014; Uchiyama et al., 2012). The amygdala forms part of the relevance network (Nakamura et al., 2018; Ousdal et al., 2008; Sander et al., 2003). The affective prosody network consists of the medial prefrontal cortex (MPFC), IFG, and the anterior insula (AI) (Frühholz et al., 2016). Thus, during a conversation, the affective prosody network is expected to interact with some areas implicated in mentalizing and semantics. In the auditory modality, Matsui et al. (2016) found a context–prosody incongruity effect in the left IFG, with an additional negative prosody effect in the anterior cingulate cortex (ACC) and bilateral AI. The ACC and AI comprise a saliency network that marks salient events (Seeley et al., 2007) and initiates appropriate control signals for additional processing, such as the maintenance and implementation of task sets (Dosenbach et al., 2006; Nelson et al., 2010) and the coordination of behavioral responses (Medford and Critchley, 2010), recruiting the executive network (Menon and Uddin, 2010; Sridharan et al., 2008). In this context, Matsui et al. (2016) speculated that the saliency network is involved in the detection of content–prosody incongruity, which in turn initiates further processing involving context incorporation, which is mediated by the rostral IFG (BA 47). It remains unclear, however, how these two factors, context–content incongruity and content–prosody incongruity, are integrated to improve sarcasm understanding; moreover, its neural underpinning remains unknown. As the majority of the neuroimaging studies of sarcasm comprehension adopted reading materials, the context–prosody interaction has rarely been investigated (Matsui et al., 2016; Uchiyama et al., 2006).

Based on the abovementioned findings, we hypothesized that content–prosody incongruity modifies the context–content incongruity effect. We considered that the areas implicated in semantic processes would integrate these effects through the saliency network. To test this hypothesis, we conducted a functional magnetic resonance imaging (fMRI) study with an auditory sarcasm detection task. We extended the auditory modality-based task utilized in a previous study (Matsui et al., 2016, in which the participants observed daily conversational interactions wherein a child did either good or bad, about which a parent made a positive comment), in two ways. First, Matsui et al. (2016) previously applied experimental stimuli that can be used by both children and adults in an experiment involving adults; thus, the provided context might have been too easy for adults. In this study, we made the context situation more complex, to enhance involvement of the mentalizing network. Second, we converted the materials to the second person, that is, critical comments were directed at the participants themselves, rather than at a character in a scenario, thus making the situation more conducive to sarcasm comprehension.

## 2. Methods

### 2.1. Participants

Twenty-three healthy individuals participated in the study as paid volunteers for the fMRI experiment. One participant was excluded due to high rates of response errors in the judgment phase, leaving 22 participants for the final analysis (11 women and 11 men; mean age, 21.7 years; range, 19–36 years). All participants had normal or corrected-to-normal visual acuity, and were right-handed (including one who was ambidextrous) according to the Edinburgh handedness inventory (mean score: 80.5; range, 11.1–100) (Oldfield, 1971), and had no history of neurological or psychiatric illness.

All participants provided written informed consent for participation in the study. This study was approved by the Ethics Committee of the National Institute for Physiological Sciences, Japan. The experiments were undertaken in compliance with national legislation and the Code of Ethical Principles for Medical Research Involving Human Subjects of the World Medical Association (Declaration of Helsinki).

## 2.2. Preparation of task materials

### 2.2.1. Auditory discourse task

The auditory discourse task consisted of two parts: the first (S1) was a three-sentence story that explained the situation of the participant who was in the scanner (you), whereas the second (S2) commented on you. S2 represented the target sentence. The target sentence could be interpreted differently, depending on the context provided in S1 (Mano et al., 2009; Matsui et al., 2016; Uchiyama et al., 2012) and the prosody of S2. We manipulated the context via S1 to make sentence S2 either congruent or incongruent with the context. Furthermore, S2 presented positive, negative, or monotone prosody, while the prosody of S1 was always monotonous. Thus, during fMRI scanning, participants listened to short narratives (S1), followed by a target sentence (S2). The task was to judge how sarcastic the target sentence S2 sounds. An example of an everyday conversation is given below.

S1. (1) You and your friend sing together in the same opera. (2) During your performance, you often sing off-key. (3) After the show, your friend says to you: S2. (4) “Tonight, you gave a superb performance.”

In this example, S1 indicates that you (fMRI participant) were in error. In this context, the friend’s comment on S2 should be interpreted as an example of sarcasm. In contrast, in the following example, S1 provides a positive situation, and the same S2 should be interpreted literally.

S1. (1) You and your friend sing in the same opera. (2) The show was excellent, and you received a long ovation. (3) After the show, your friend says to you: S2. (4) “Tonight, you gave a superb performance.”

We fixed the valence of the S2, the uttered contents, as positive. Thus, there were six patterns of context–content–prosody combination for S2 (abbreviations: P, positive; N, negative; and m, monotone): PPP, PPN, PPM, NPP, NPN, NPM, creating a 2 (Context) × 3 (Prosody) design matrix. All auditory stimuli were in Japanese.

### 2.2.2. Recording of auditory stimuli

The stimuli were recorded by a female actor in a silent room, using a microphone (SM58; Shure, Evanston, IL), an audio interface (0202; E-MU Systems, Scotts Valley, CA), and a personal computer (ThinkPad X201; Lenovo, Morrisville, NC).

### 2.2.3. Norming study

To confirm that our experimental stimuli with these auditory utterances were indeed perceived as sarcasm in the negative context (NPP, NPN, and NPM) conditions, 45 volunteers (28 females and 17 males; mean age, 36.5 y; range, 22–64 years) participated in a norming study. We presented the experimental stimuli in a pseudo-random order and asked the participants to rate how sarcastic the partner’s comment sounded in S2, using a five-point scale (5 = sarcasm; 1 = literal, i.e., not sarcastic). The mean scores of these sarcasm judgments were (Fig. 1(a)): 1.06 for the PPP condition, 2.88 for the PPN condition, 1.73 for the PPM condition, 3.87 for the NPP condition, 4.94 for the NPN condition, and 4.18 for the NPM condition. A two-way analysis of variance (ANOVA) of the discourse context (positive and negative) and the affective prosody (positive, negative, and monotone) with a Bonferroni’s correction for multiple comparisons conducted on the sarcasm judgments revealed that the main effects of discourse context ( $F(1, 44) = 391.247$ , mean squared error ( $MSE$ ) = 1.029,  $p < 0.001$ ), affective prosody ( $F(2, 88) = 78.672$ ,  $MSE = 0.619$ ,  $p < 0.001$ ) and the interaction between discourse context and affective prosody ( $F(2, 88) = 9.907$ ,  $MSE = 0.316$ ,  $p < 0.001$ ) were significant. The nature of this interaction was such that when a discourse context involved a positive event, an utterance with negative prosody was judged as significantly more sarcastic than one with monotone prosody (Bonferroni’s,  $p < 0.001$ ), which was more sarcastic than the one with positive prosody (Bonferroni’s,  $p < 0.001$ ) ( $F(2, 176) = 81.507$ ,  $MSE = 0.467$ ,  $p < 0.001$ ). However, for a discourse context involving a negative event, an utterance accompanied by negative prosody was judged as significantly more sarcastic than an utterance with monotone prosody (Bonferroni’s correction for multiple comparisons,  $p < 0.001$ ) which was more sarcastic than one with positive prosody (Bonferroni’s,  $p < 0.05$ ) ( $F(2, 176) = 29.386$ ,  $MSE = 0.467$ ,  $p < 0.001$ ). These results demonstrated that our experimental stimuli were well-controlled in that the S2 with positive content was perceived as sarcastic in the negative context.

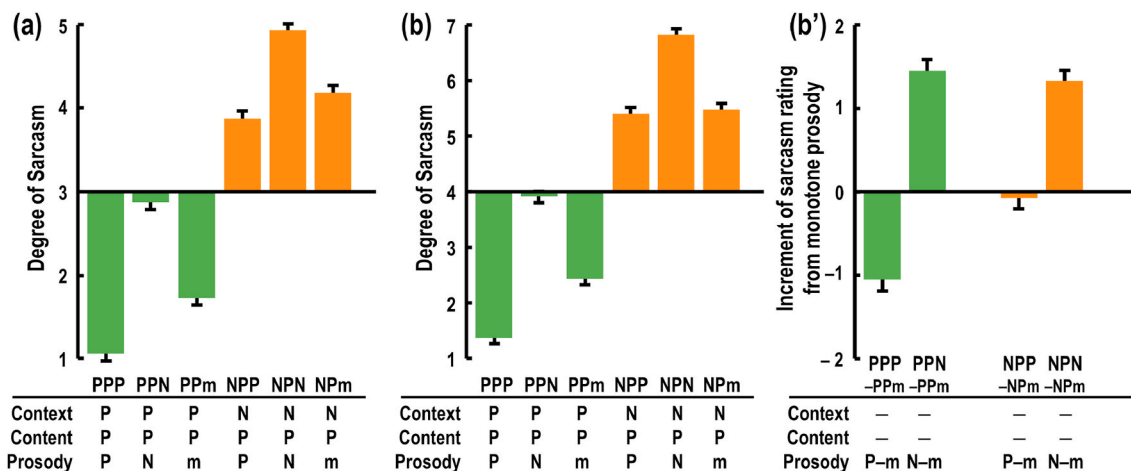


Fig. 1. (a) In a norming study, the mean scores of sarcasm judgments (5 = sarcasm; 1 = literal, i.e., not sarcasm). In an fMRI study, (b) the mean scores of sarcastic responses (7 = sarcasm; 1 = literal, i.e., not sarcasm) and (b') the increment of sarcasm ratings from monotone prosody. Error bars indicate the standard error of the mean. Conditions (see Table 1 for details) with a positive context are indicated by green, and those with negative context by orange. The main effect of context, prosody, and their interaction were significant. Abbreviations: P, positive; N, negative; m, monotone. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

### 2.3. fMRI procedures

Before the fMRI session, the participants were given detailed instructions on the task procedure. To familiarize participants with the task, they were provided with examples of stimuli that did not appear during the fMRI session. All task stimuli were prepared and presented using Presentation® 19.0, software (Neurobehavioral Systems, Albany, CA) running on a personal computer (dc7900; Hewlett-Packard Inc., Palo Alto, CA). Using a liquid crystal display projector (CP-SX12000J; Hitachi Ltd., Tokyo, Japan), the visual stimuli were projected onto a half-transparent viewing screen located behind the head coil of the MRI scanner. The participants viewed the stimuli using a mirror attached to the head coil. The spatial resolution of the projector was  $1024 \times 768$  pixels, and used a 60-Hz refresh rate. The distance between the screen and the participant's eyes was approximately 175 cm, and the visual angle was  $13.8^\circ$  (horizontal)  $\times$   $10.4^\circ$  (vertical). Auditory stimuli were presented via MR-compatible headphones (Kiyohara-Kougaku, Tokyo, Japan).

Participants were instructed to play the role of the protagonist in a scenario, in which the protagonist had a short conversation with another person (colleagues, etc.). Within the scenario, the participant was the person who was evaluated (positively or negatively) by the utterance. In other words, there was no process of empathy for a third person who was evaluated by the sarcastic utterance; as a result, the experimental situation was more suitable for depicting the neural substrates of sarcasm. Each of the four phases of each trial was set to correspond to the length of the presented sentences (the first phase, 4.5 s [range: 3.5–5.3 s]; the second phase, 4.6 s [range: 3.0–6.6 s]; the third phase, 2.4 s [range: 2.0–3.7 s]; and the fourth phase (S2), 2.7 s [range: 1.7–3.8 s]), followed by presentation of a fixation cross on a black screen (visual angle,  $0.6^\circ \times 0.6^\circ$ ) for 0.9 s. Thereafter, when determining the degree of sarcasm, the instruction was to integrate the contextual effect and the prosodic effect. That is, the participant was required to judge how sarcastic the S2 utterance sounded, and respond to a seven-point scale (7 = sarcasm; 1 = literal, i.e., not sarcastic) that appeared on the screen for 4 s (judgment phase), as quickly as possible, by pressing a button with their right index or middle finger. After the participant responded, a fixation cross was again shown on the screen for 4 s (range: 1.5–6.5 s).

An event-related design was used in this study. Each run consisted of the first 25 s of a fixation cross, followed by 24 task trials, and the final 15 s of a fixation cross. The 72 task trials (12 scenarios  $\times$  two discourse contexts  $\times$  three affective prosodies [positive, negative, and monotone]) were presented in a pseudo-random order; that is, all context–prosody combinations were presented. The experiment consisted of three runs, each consisting of 24 experimental trials. The presentation order of the three runs was counterbalanced across participants.

All images were acquired using a 3-T MR scanner (Verio®; Siemens, Erlangen, Germany). An ascending T2\*-weighted gradient-echo echo-planar imaging (EPI) procedure was used to produce 66 continuous trans-axial slices that covered the entire cerebrum and the cerebellum (repetition time [TR], 1000 ms; echo time [TE], 30.8 ms; flip angle,  $55^\circ$ ; field-of-view [FoV], 216 mm;  $216 \times 216$  matrix; voxel dimensions,  $2.4 \times 2.4$  mm in plane; 2.0-mm slice thickness with a 0.4-mm gap). Six slices were acquired simultaneously using a multi-band sequence (Moeller et al., 2010) to maximize the acquisition speed. Oblique scanning was used to exclude the eyeballs from the images. Each run consisted of a continuous series of 640 vol acquisitions, resulting in a total acquisition duration of 10 min 40 s. For anatomical imaging, T1-weighted magnetization-prepared rapid-acquisition gradient echo anatomical images were also obtained (TR, 2400 ms; TE, 2.24 ms; flip angle,  $8^\circ$ ; FoV, 256 mm; 1 slab; number of slices per slab, 224; voxel dimensions,  $0.8 \times 0.8 \times 0.8$  mm) for each participant.

### 2.4. Data analysis

#### 2.4.1. Behavioral performance

A two-way ANOVA with two within-subject factors, namely discourse context and affective prosody, with a Bonferroni's correction for multiple comparisons, was conducted on the seven-point scale of how sarcastic the utterance in S2 sounded. The analysis was performed using SPSS® version 26.0 software (IBM, Armonk, NY).

#### 2.4.2. Imaging data

Imaging data were preprocessed as follows. The 640 EPI volumes per run (a total of 1920 EPI volumes per participant) were analyzed using Statistical Parametric Mapping 12 (SPM12; Wellcome Department of Imaging Neuroscience, London, UK; Friston et al., 2007) implemented in MATLAB® (MathWorks, Natick, MA). The EPI volumes were spatially realigned to correct for head motion. Next, a mean EPI image was coregistered with the T1-weighted anatomical image, and the parameter was then applied to all EPI images. The anatomical image was spatially normalized to the Montréal Neurological Institute (MNI) T1 template using a segmentation-normalization method. The normalization parameters of the T1-weighted anatomical image were applied to all EPI volumes, and then spatially smoothed in three dimensions using an 8 mm full-width half-maximum Gaussian kernel.

After preprocessing, the EPI data obtained for each participant were analyzed using a general linear model. The S2 of the six auditory conditions (PPP, PPN, PPM, NPP, NPN, and NPM) were separately modeled by convolution with a hemodynamic response function. S1 was also modeled as a regressor of no interest convolved with a hemodynamic response function. Additionally, button responses were modeled as independent regressors using a convolved delta function. We applied high-pass filters (128 s) to the time-series data. An autoregressive model was used to estimate the temporal autocorrelation. Six regressors for the head movement parameters obtained in the realignment process were entered into the model.

The contrast images, which consisted of the weighted sum of parameter estimates and which represented the normalized task-related increment of the MR signal obtained in the individual analyses, were then subjected to group analysis. In total, data from 22 participants and six contrasts (PPP, PPN, PPM, NPP, NPN, and NPM) were incorporated into the 2 (discourse context)  $\times$  3 (affective prosody) within-subject factorial design (Friston et al., 2007). Using the flexible factorial design model (Friston et al., 2007), the subject factor was set as independent to take different individuals into account. The error variance was set to be equal across participants because they were sampled from the same underlying population. In contrast, the two condition factors were set as dependent because the different factor levels were correlated within the subject, with equal error variances because they were taken from the same subject. Given our hypotheses, we evaluated the following predefined contrasts of *t*-tests (Table 1):

The resulting set of voxel values for each contrast constituted statistical parametric mapping (SPM) of the *t* statistic, which was transformed into normal distribution units with a height threshold set at

**Table 1**  
Predefined contrasts for the second level analysis.

	PPP	PPN	PPm	NPP	NPN	NPM
(a) Main effect of context–content incongruity	–1	–1	–1	1	1	1
(b) Main effect of content–prosody incongruity	–1	1	0	–1	1	0
(c) Interaction (i.e., context–prosody incongruity)	–1	1	0	1	–1	0

Note: Three letters of each condition name (e.g., PPP, NPP) indicate the set of valences of your (the participant's) situation (context; left), partner's comment (content; middle), and the prosody of the utterance (right). Abbreviations: P, positive; N, negative; m, monotone.

uncorrected  $p < 0.001$  and  $p < 0.05$  with a family-wise error (FWE) correction at the cluster level for the entire brain (Friston et al., 1996). Since the amygdala shows prominent activation of a small region with sharp boundaries, however, to observe the amygdala alone, we also exploratorily used a more conservative FWE correction at the voxel level of  $p < 0.05$  (Friston et al., 1996). The activated area was determined by MarsBaR version 0.44 software (Brett et al., 2002) along with MarsBaR AAL ROIs version 0.2 software (Tzourio-Mazoyer et al., 2002) (<http://marsbar.sourceforge.net>); and the locations were checked using the Atlas of the Human Brain (Mai et al., 2015).

### 3. Results

#### 3.1. Behavioral performance

As expected, the partner's comment for a negative event (NPP, NPN, and NPM conditions) was interpreted as sarcasm, whereas the partner's comment for a positive event (PPP, PPN, and PPM conditions) was interpreted as literal. In other words, the mean scores of sarcastic responses (7 = sarcasm; 1 = literal, i.e., not sarcasm) were 1.37 for PPP condition, 3.91 for PPN condition, 2.44 for PPM condition, 5.40 for NPP condition, 6.82 for NPN condition, and 5.48 for NPM condition (Fig. 1(b)). A two-way ANOVA of discourse context (positive and negative) and affective prosody (positive, negative, and monotone) conducted on the sarcasm judgments revealed a significant main effect of discourse context (i.e., the effect of context-content incongruity) ( $F(1, 21) = 404.153$ ,  $MSE = 0.903$ ,  $p < 0.001$ ), a significant main effect of affective prosody (i.e., the effect of content-prosody incongruity) ( $F(2, 42) = 132.089$ ,  $MSE = 0.345$ ,  $p < 0.001$ ), and a significant interaction between these two factors (i.e., the effect of context-prosody incongruity) ( $F(2, 42) = 17.938$ ,  $MSE = 0.231$ ,  $p < 0.001$ ). The nature of this interaction was such that when a discourse context involved a positive event, an utterance with negative prosody was judged as significantly more sarcastic than one with monotone prosody (Bonferroni's,  $p < 0.001$ ), which was more sarcastic than one with positive prosody (Bonferroni's,  $p < 0.001$ ) ( $F(2, 84) = 124.093$ ,  $MSE = 0.288$ ,  $p < 0.001$ ). However, for a discourse context involving a negative event, an utterance accompanied by negative prosody was judged as significantly more sarcastic than an utterance with monotone or positive prosody (Bonferroni's correction for multiple comparisons,  $p < 0.001$ ; there was no significant difference between monotone and positive prosody) ( $F(2, 84) = 48.542$ ,  $MSE = 0.288$ ,  $p < 0.001$ ).

To reveal how prosody modifies the context-content incongruity effect, we calculated the increment of sarcasm ratings from monotone prosody (Fig. 1(b')). The increment was  $-1.07$  for [PPP - PPM],  $1.47$  for [PPN - PPM],  $-0.08$  for [NPP - NPM], and  $1.34$  for [NPN - NPM]. A two-way ANOVA of discourse context and affective prosody showed that the main effect of discourse context ( $F(1, 21) = 15.386$ ,  $MSE = 0.267$ ,  $p < 0.001$ ), affective prosody ( $F(1, 21) = 141.947$ ,  $MSE = 0.606$ ,  $p < 0.001$ ), and their interaction ( $F(1, 21) = 18.546$ ,  $MSE = 0.373$ ,  $p < 0.001$ ) were significant. The nature of this interaction was such that an utterance with negative prosody had a significantly larger increment of sarcasm ratings than an utterance with positive prosody regardless of the discourse context ( $F(1, 42) = 144.775$ ,  $MSE = 0.489$ ,  $p < 0.001$  for a positive context; and  $F(1, 42) = 45.112$ ,  $MSE = 0.489$ ,  $p < 0.001$  for a negative context). In contrast, when an utterance involved positive prosody, the degree of increment of sarcasm ratings was negative for both discourse contexts, and a discourse context involving a negative event showed a significantly smaller decrease than a discourse context involving a positive event ( $F(1, 42) = 33.886$ ,  $MSE = 0.320$ ,  $p < 0.001$ ). However, when an utterance was accompanied by negative prosody, the degree of increment was positive for both discourse contexts and there was no significant difference between a discourse context involving a negative event and one involving a positive event ( $F(1, 42) = 0.571$ ,  $MSE = 0.320$ , not significant). In other words, when the context was positive (congruent with the content of utterance), positive prosody

lessened the sarcasm rating, whereas negative prosody enhanced it. When the context was negative, the positive prosody effect disappeared, while negative prosody increased the sarcasm rating. Thus, context-content incongruity is the primary determinant of sarcasm comprehension, and is modified by prosody in a context-dependent manner.

#### 3.2. Group analysis of fMRI data

We observed a corresponding main effect of context-content incongruity in an auditory modality task in the utterance-related neural activation of the anterior rostral zone of the MPFC (arMPFC; BA 9; MNI coordinates  $[-4\ 56\ 40]$ ), the right temporal pole (TP; BA 38;  $[48\ 4\ -34]$ ), and the bilateral cerebellum ( $[-24\ -76\ -36]$ ;  $[26\ -76\ -32]$ ) (Fig. 2 and Table 2(a)). In the activated cluster (382 voxels), including the left cerebellum, 219 voxels (57.3%) were in the left Crus I and 97 voxels (25.4%) were in the left Crus II. In the activated cluster (524 voxels), including the right cerebellum, 362 voxels (69.1%) were in the right Crus I and 26.7% (140 voxels) were in the right Crus II.

Furthermore, we found a main effect of content-prosody incongruity in the bilateral amygdala ( $[-14\ -4\ -16]$ ;  $[20\ -4\ -10]$ ) (Fig. 3 and Table 2(b)).

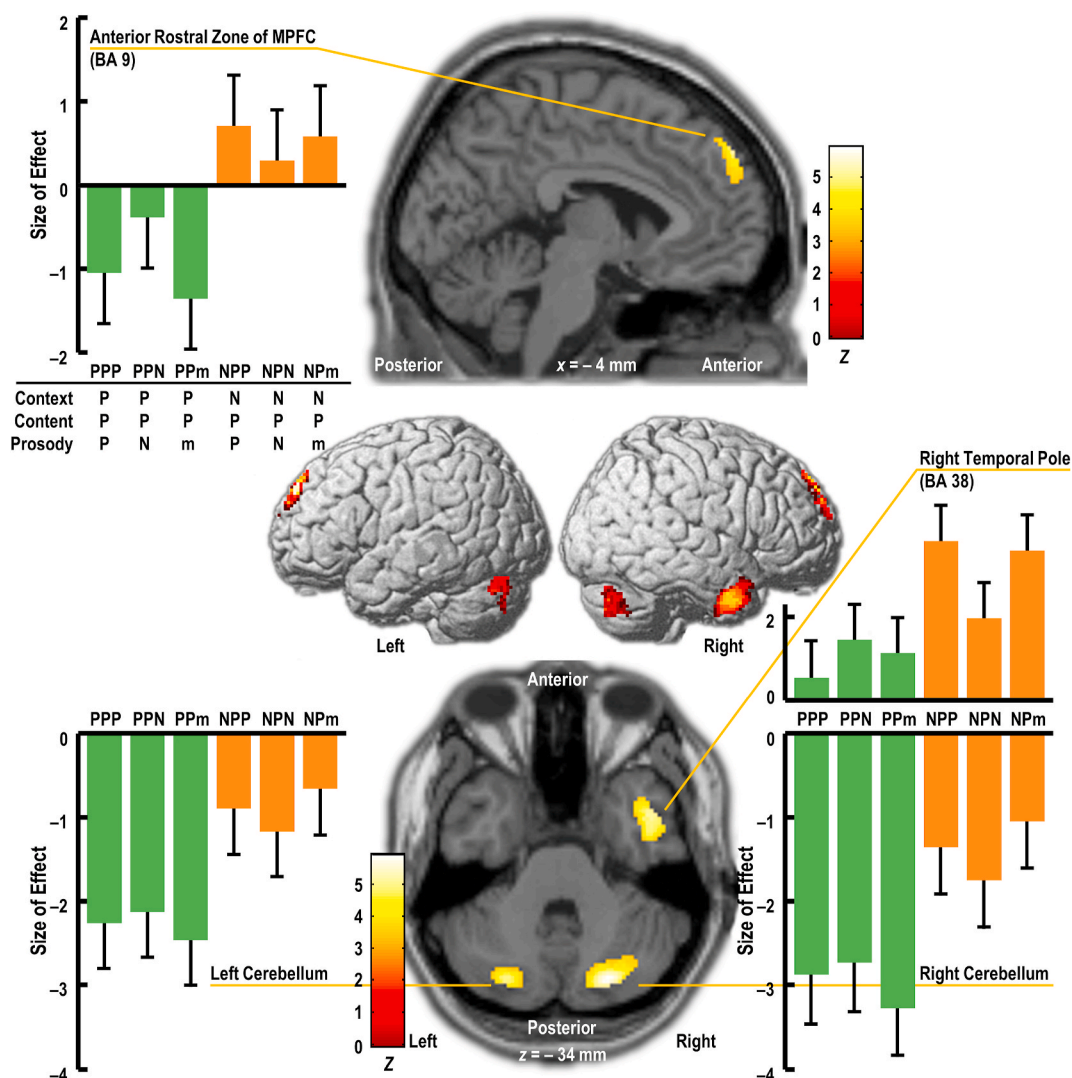
The interaction effect between context-content incongruity and content-prosody incongruity (i.e., context-prosody incongruity) was found in the posterior rostral zone of the MPFC (prMPFC; BA 8/32;  $[-8\ 18\ 52]$ ), which included the ACC, the bilateral AI (BA 13;  $[-38\ 20\ -2]$ ;  $[48\ 18\ -10]$ ) extending to the dorsolateral prefrontal cortex (DLPFC; BA 46), the left rostral part of the premotor cortex (BA 6;  $[-34\ 4\ 62]$ ), and the left superior parietal lobule extending to the inferior parietal lobule (BA 7/39;  $[-28\ -64\ 52]$ ) (Fig. 4 and Table 2(c)). In the activated cluster (2813 voxels), including the left AI, 336 voxels (11.9%) were in the left BA 44, 1120 voxels (39.8%) were in the left BA 45, 288 voxels (10.2%) were in the left BA 46, and 437 voxels (15.5%) were in the left BA 47. In the activated cluster (1489 voxels) including the right AI, 303 voxels (20.3%) were in the right BA 44, 221 voxels (14.8%) were in the right BA 45, 145 voxels (9.7%) were in the right BA 46, and 390 voxels (26.2%) were in the right BA 47.

### 4. Discussion

The present study investigated the neural mechanisms underlying sarcasm comprehension during conversation, and found that it required two distinct incongruity detection systems: the mentalizing network and the relevance network, which are integrated in the region implicated in the semantic processes, including the prefrontal cortices, through the salience network.

#### 4.1. Behavioral results

When the context was positive (congruent with the content of the utterance), positive prosody lessened the sarcasm rating, whereas negative prosody enhanced this rating. When the context was negative, the positive prosody effect disappeared, while negative prosody increased the sarcasm rating. Thus, context-content incongruity is the primary determinant of sarcasm comprehension and is modified by prosody in a context-dependent manner. This finding reflects the fact that the context precedes the utterance. That is, according to the predictive coding schema (Friston, 2005), the context provokes the generative model that predicts the upcoming event, and if there is incongruity between these factors, it results in prediction error. Thus, the context-induced generative model is the major determinant of sarcasm rating that is understood as the prediction error. The participants judged the prediction error, the deviation from the literal, as sarcastic degree. For example, PPN items provided positive context that makes listeners predict the positive attitude of the speaker. Thus, negative prosody generates prediction error reflected in the increased sarcasm rating compared with PPM or PPP, both of which are concordant with the



**Fig. 2.** The neural substrates of context–content incongruity (the contrast of Table 1(a)) in sarcasm comprehension, superimposed on the sagittal ( $x = -4$  mm; top) and transaxial ( $z = -34$  mm; bottom) sections and on the rendered surface (middle) of the standard MNI template. Bar graphs of the task-related activation in the anterior rostral zone of the medial prefrontal cortex (MPFC; BA 9; MNI coordinates  $[-4\ 56\ 40]$ ; top left), the left cerebellum ( $[-24\ -76\ -36]$ ; bottom left), the right temporal pole (BA 38;  $[48\ 4\ -34]$ ; middle right), and right cerebellum ( $[26\ -76\ -32]$ ; bottom right) are plotted. Conditions with positive context are indicated by green, and those with negative context by orange. Abbreviations: MNI, Montréal Neurological Institute; BA, Brodmann’s area. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

prediction of positive attitude of the speaker. Sarcasm mediates implicit criticism of the listener by provoking negative emotions (Sperber and Wilson, 1995). As negative prosody is the most direct means to display a critical cue of a negative attitude of the speaker, the increased sarcasm rating should be interpreted as discrepancy between the expected (positive) and actual (negative) attitudes of the speaker, that is, sarcasm.

#### 4.2. Neural representation

##### 4.2.1. Context–content incongruity

We confirmed the context–content incongruity effect, representing what is uttered in a particular context, in the mentalizing network (Frith and Frith, 2003) including the arMPFC, TP, and cerebellum. This finding is consistent with those of previous studies showing the importance of context in ironic understanding (Ackerman, 1983; Colston, 2002; Ivanko and Pexman, 2003; Katz and Lee, 1993; Katz and Pexman, 1997; Kreuz and Glucksberg, 1989). Activation of the MPFC and TP is consistent with the findings of previous neuroimaging studies of sarcasm comprehension during a reading task (Uchiyama et al., 2006, 2012) and when using the auditory modality (Herold et al., 2018; Varga et al., 2013; Wang

et al., 2006a). Activation of the mentalizing network during sarcasm detection represents the prediction error, that is, the incongruity between the prediction of the meaning of the utterance of the speaker that was pragmatically relevant to the context, and the actual utterance. Cerebellar activation is consistent with the previous review study showing that the Crus I and Crus II in the cerebellum are involved in social cognition, particularly in mentalizing processing related to emotional cognition in the Crus II (Van Overwalle et al., 2020).

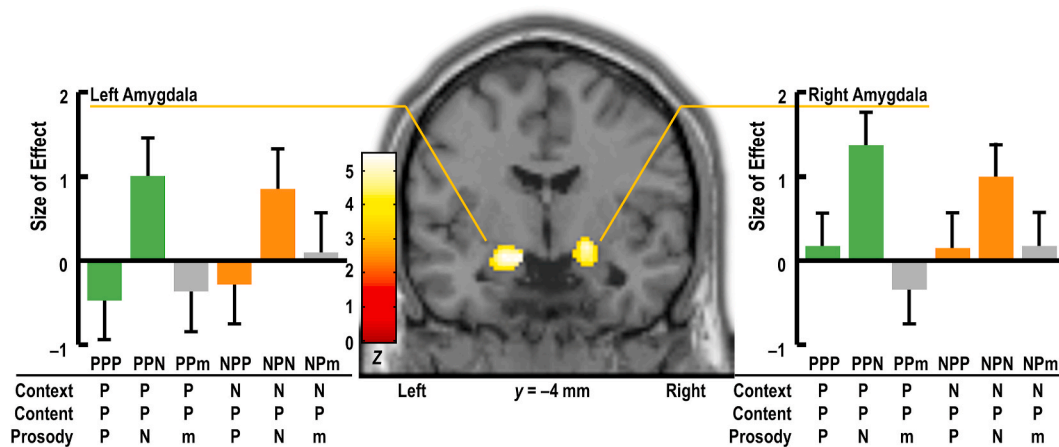
##### 4.2.2. Content–prosody incongruity

We also confirmed that content–prosody incongruity in sarcasm comprehension is related to activation in the bilateral amygdala. This finding is consistent with those of previous studies showing the importance of prosody in ironic understanding (Bryant and Fox Tree, 2005; Caillies et al., 2019; Capelli et al., 1990; Laval and Bert-Erboul, 2005; Le Gall and Iakimova, 2018; Tobe et al., 2016; Wickens and Perry, 2015). It is also consistent with those of previous neuroimaging studies reporting the involvement of the amygdala in sarcasm-specific process (Akimoto et al., 2014; Uchiyama et al., 2012) and affective sound processing in the detection of emotional and social valence (Frühholz et al., 2016).

**Table 2**  
Activation related to sarcasm comprehension.

Cluster <i>p</i> -FWE	Peak <i>p</i> -FWE	Cluster size (number of voxels)	Z-value	MNI coordinates			Side	Location	BA
				x	y	z			
<b>(a) Main effect of context–content incongruity</b>									
0.043	0.050	347	4.49	−4	56	40	L	arMPFC	9
0.029	<u>0.052</u>	393	4.48	48	4	−34	R	TP	38
0.032	0.023	382	4.68	−24	−76	−36	L	Cerebellum (Crus I/II)	
0.010	0.003	524	5.15	26	−76	−32	R	Cerebellum (Crus I/II)	
<b>(b) Main effect of content–prosody incongruity</b>									
<u>0.104</u>	0.012	250	4.84	−14	−4	−16	L	Amygdala	
<u>0.172</u>	0.040	198	4.54	20	−4	−10	R	Amygdala	
<b>(c) Interaction (context–prosody incongruity)</b>									
0.000	0.000	2268	5.66	−8	18	52	L	prMPFC	8/32
0.000	0.013	2813	4.81	−38	20	−2	L	AI	13
0.000	0.013	1489	4.82	48	18	−10	R	AI	13
0.004	0.011	629	4.86	−34	4	62	L	PMdr	6
0.016	<u>0.320</u>	466	3.94	−28	−64	52	L	SPL/IPL	7/39

Abbreviations: MPFC, medial prefrontal cortex; arMPFC, anterior rostral zone of MPFC; prMPFC, posterior rostral zone of MPFC; TP, temporal pole; AI, anterior insula; PMdr, rostral portion of the dorsal premotor cortex; SPL, superior parietal lobule; IPL, inferior parietal lobule; MNI, Montréal Neurological Institute; BA, Brodmann’s area; L, left; R, right; FWE, family wise error correction for the entire brain; underlined, not significant.



**Fig. 3.** The neural substrates of content–prosody incongruity (the contrast of Table 1(b)) in sarcasm comprehension, superimposed on the coronal ( $y = -4$  mm; middle) section of the standard MNI template. Bar graphs of the task related activation in the left (MNI coordinates  $[-14 -4 -16]$ ; left) and right ( $[20 -4 -10]$ ; right) amygdala are plotted in the same format as in Fig. 2, except for monotone prosody indicated by gray.

Furthermore, Nakamura et al. (2018) found that humor comprehension activated the left amygdala. Sarcasm comprehension was similar to humor comprehension in that it consists of both cognitive and emotional components of incongruity resolution. Given that the amygdala is involved in relevance detection (Ousdal et al., 2008; Sander et al., 2003), it is conceivable that the amygdala is involved in the content–prosody incongruity process.

#### 4.2.3. Incongruity interaction

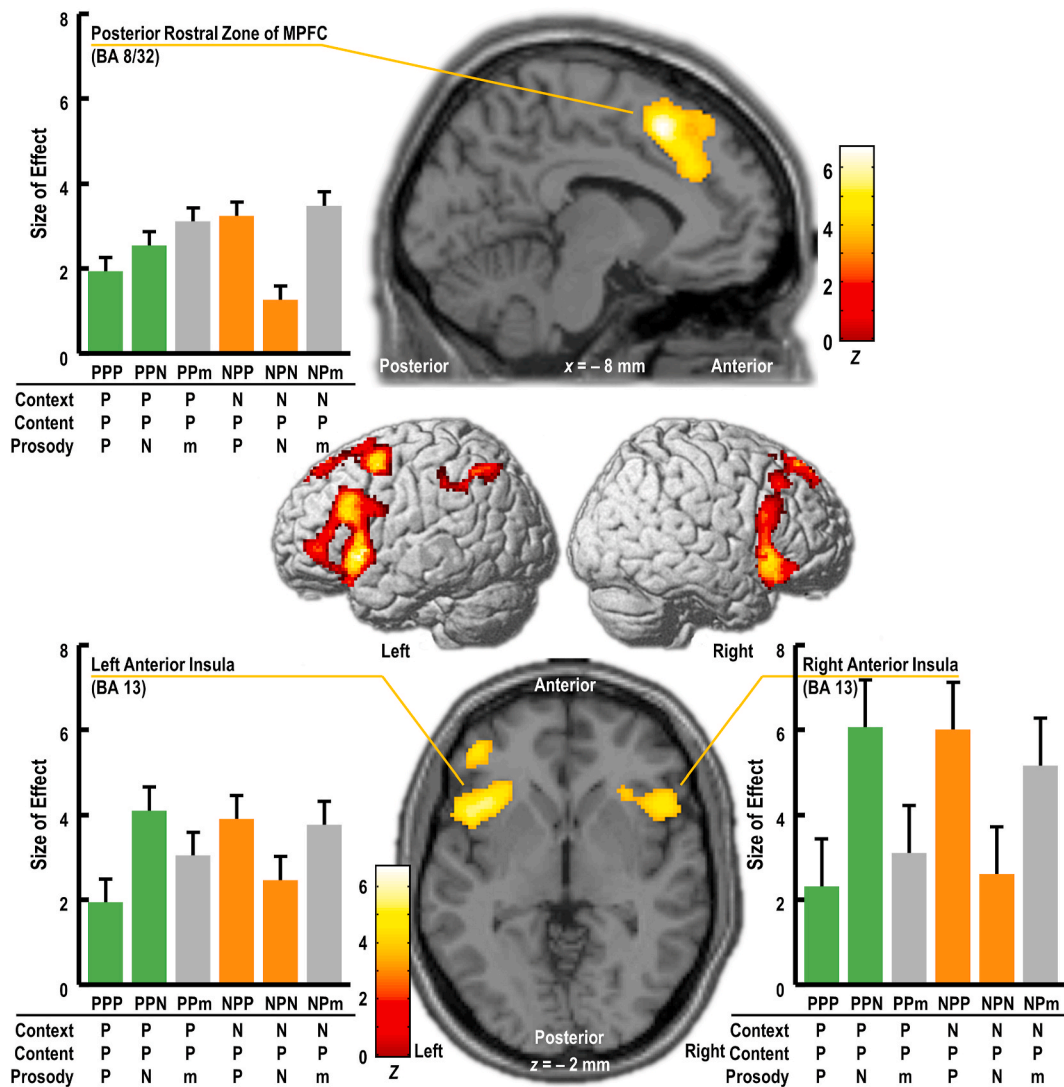
We found an interaction between context–content incongruity and content–prosody incongruity in the salience network, which extended to the IFG and DLPFC, consistent with our pre-experiment hypothesis that the regions implicated in semantic processes would be involved in integrating the main effects of incongruities through the salience network. This finding is consistent with those of previous studies showing the importance of prosody in ironic understanding (Rivière et al., 2018; Voyer et al., 2016; Woodland and Voyer, 2011).

Activation of the prefrontal regions was also consistent with the results of previous studies. The IFG is the main neural basis for affective prosody networks (Frühholz et al., 2016) along with a unification space (Hagoort, 2005). Uchiyama et al. (2006) concluded that the IFG, particularly BA 47, is the site where integration of semantic and mentalizing processes occur during sarcasm detection. We also found activation in the DLPFC, which has also been implicated in sarcasm

comprehension (Bosco et al., 2017; Spotorno et al., 2012; Varga et al., 2013) and has been suggested to be involved in non-literal language processing (Rapp et al., 2012) and executive functioning to integrate information related to sarcasm comprehension (Bosco et al., 2017).

The AI is critical for empathy (Carr et al., 2013). A meta-analysis of emotion activation studies (Phan et al., 2002) showed that the ACC and AI, the core nodes of the salience network, are involved in emotional recall or imagery of personally relevant events and emotional tasks that exert a cognitive demand. The amygdala and AI are co-activated for interpreting emotional facial expressions (Phan et al., 2002). For empathy to arise through processing of emotional facial expression, the AI is critical in the communication between the action representation network and the limbic areas, including the amygdala (Carr et al., 2013). Thus, the salience network is involved in emotional comprehension by linking with the relevance detection system.

Compared with a previous study by Matsui et al. (2016) that showed the content–prosody incongruity effect, we used a complex context, which might enhance the effect of interaction between context–content incongruity and content–prosody incongruity in the salience network. Another factor was that we used the second-person perspective in the experimental setup: sarcastic comments were directed at the participants themselves rather than at a character in a scenario, thus making the situation more self-related for sarcasm comprehension. The salience network, including the AI and ACC, is known to be involved in



**Fig. 4.** The neural substrates of interaction of context–content incongruity and content–prosody incongruity (i.e., context–prosody incongruity; the contrast of Table 1(c)) in sarcasm comprehension, superimposed on the sagittal ( $x = -8$  mm; top) and transaxial ( $z = -2$  mm; bottom) sections and on the rendered surface (middle) of the standard MNI template. Bar graphs of the task-related activation in the posterior rostral zone of medial prefrontal cortex (MPFC; BA 8/32;  $[-8\ 18\ 52]$  which included the anterior cingulate cortex; top left), the left anterior insula (BA 13;  $[-38\ 20\ -2]$ ; bottom left), and the right anterior insula ( $[48\ 18\ -10]$ ; bottom right) are plotted in the same format as in Fig. 3.

self-related processing, such as self-face processing. Individuals can experience embarrassment when exposed to self-feedback images, depending on the extent of divergence from the internal representation of the standard self. By utilizing the fact that embarrassment is enhanced by observation by others, Morita et al. (2014) showed that the AI and ACC are involved in emotional processing for self-face recognition. More specifically, they found enhanced functional connectivity of the ACC with the dorsal and ventral parts of the MPFC, and the left lateral prefrontal cortex, including the middle frontal gyrus, IFG, and AI, when viewing self-face images while being observed than when doing so without observation. They argued that this enhanced connectivity of the ACC with the MPFC and the left prefrontal cortex represented processing of the reflective self or narrative self (Gallagher, 2000). In contrast, the right AI appears to be involved in creating the subjective experience of embarrassment, probably through the comparison with the “standard self” (Morita et al., 2014). This notion is consistent with a “predictive coding” framework (Seth, 2013; Seth and Friston, 2016) that postulates the prediction error caused by comparison of the prior model with the perception of the visually presented self-face evoked emotion. In a similar vein, during sarcasm comprehension in the present experiment,

the salience network was involved in self-relevant, self-evaluative emotional processing triggered by the sarcastic utterance directed toward the participant.

#### 4.2.4. Two distinct incongruity detection systems integrated by salience network

These results suggest that sarcasm is perceived as a multi-layered incongruity among the context, content, and prosody of the utterance (Utsumi, 2000). Our findings indicated that there are two distinct incongruity detection systems and an integration system corresponding to the implicit display theory for sarcasm comprehension (Utsumi, 2000). One is the mentalizing system (Frith and Frith, 2003; Spotorno et al., 2012; Uchiyama et al., 2006, 2012) involving the cerebellum for context–content incongruity (i.e., what is uttered in a particular context), which corresponds to a pragmatic insincerity by intentionally violating one of the pragmatic principles (Utsumi, 2000). The other is the relevance detector system (Nakamura et al., 2018; Ousdal et al., 2008; Sander et al., 2003) for content–prosody incongruity (i.e., how it is spoken), which corresponds to an indirect expression of the negative effect toward the failure to meet their expectation (Utsumi, 2000). An



interaction effect of these systems with the salience network was found in the prefrontal regions (Menon and Uddin, 2010; Sridharan et al., 2008), including the AI and ACC. This finding indicates that, for sarcasm comprehension, different incongruity sources are integrated in the prefrontal cortices through the salience network.

#### 4.2.5. Clinical implication

It is well known that individuals with Autism Spectrum Disorder (ASD) have difficulty in appreciating irony (Happé, 1993, 1994; Kaland et al., 2002, 2005; Leekam and Prior, 1994; Martin and McDonald, 2004; Tantam, 1991). This impairment could be related to deficits in using both prosodic and contextual information to make inferences about a speaker's communicative intent (Wang et al., 2006b). Wang et al. (2006b) showed using fMRI that children with ASD showed significantly greater activity than control group in the right IFG and bilateral temporal region, suggesting "more effortful processing needed to interpret the intended meaning of an utterance." In contrast, a meta-analysis of the functional MRI with individuals with ASD showed hypoactivation of the AI by the social tasks (Di Martino et al., 2009). Given that the AI is positioned as "a hub mediating interactions between large-scale networks involved in externally and internally oriented cognitive processing," Uddin and Menon (2009) concluded that dysfunctional AI connectivity plays an important role in ASD. Consistent with these previous studies, the present study revealed that the AI, along with the lateral prefrontal cortices, is related to integrating both prosodic and context information. To directly elucidate the pathogenesis of ASD individuals, an fMRI study of individuals with ASD applying the present verbal sarcasm comprehension task is warranted for future study.

#### 4.3. Study limitations

This study had some limitations. For example, to maintain the total duration of the fMRI experiment within a reasonable interval, we restricted our experimental stimuli to those with moderate difficulty. Future research should investigate the neural mechanisms underlying different types of neural bases for incongruity detection in sarcasm comprehension based on an auditory modality using stimuli of multi-level difficulties. For the same reason, we also restricted our target utterances to those with a positive semantic valence, and it may have been possible to predict whether the utterance was sarcastic at the contextual phase. Thus, future research should include those with negative semantic valence. We did not conduct connectivity analysis among several networks such as the relevance network, affective prosodic network, mentalization network, the regions implicated in semantic processes, salience network, executive network, and default mode network. The function of the salience network in the connection among the executive network and default mode network has been proposed (Menon and Uddin, 2010). The detailed interaction among these network warrants future study. As shown in Table 2, two thresholds were used together to show the results of the imaging data. Since the amygdala is a relatively small neural substrate, it may have been difficult for the cluster size to grow. This may be related to the number of participants in this study (22). Some other potentially interesting questions regarding incongruity detection in sarcasm comprehension were excluded from the current study. For instance, a theoretical study suggested that a hearer detects incongruity not only by affective prosody (paralinguistic cues), but also by verbal cues, such as hyperbole and interjection, and speech that expresses counterfactual pleased affect, as well as by nonverbal cues, such as facial expression and behavioral cues (Utsumi, 2000). Future research should investigate these neural mechanisms using visual and auditory stimuli.

## 5. Conclusions

This study revealed that two distinct incongruity detection systems for sarcasm comprehension are integrated in the prefrontal cortices

through the salience network.

## Author contributions

**Tagiru Nakamura:** Conceptualization, Investigation, Formal analysis, Writing – original draft, Visualization, Project administration, and Funding acquisition. **Tomoko Matsui:** Conceptualization, Investigation, Formal analysis, Writing – reviewing & editing, and Funding acquisition. **Akira Utsumi:** Conceptualization, Investigation, Formal analysis, Writing – reviewing & editing, and Funding acquisition. **Eri Nakagawa:** Investigation and Writing – reviewing & editing. **Motofumi Sumiya:** Investigation and Writing – reviewing & editing. **Norihiro Sadato:** Conceptualization, Investigation, Writing – reviewing & editing, Supervision, Funding acquisition.

## Data availability

The datasets generated in this study are available from the corresponding author on reasonable request.

## Declaration of competing interest

No potential conflict of interest was reported by the authors.

## Acknowledgments

This work was supported by the National Institute for Physiological Science ["the Cooperative Study Program" (T.N., T.M., and A.U.)] and the Japan Society for the Promotion of Science (Grants-in-Aid for Early-Career Scientists #JP20K13034 (T.N.)). Parts of this study represent the results of the Japan Agency for Medical Research and Development under Grant Numbers JP18dm0107152 and JP18dm0307005 (N.S.).

## References

- Ackerman, B.P., 1983. Form and function in children's understanding of ironic utterances. *J. Exp. Child Psychol.* 35, 487–508. [https://doi.org/10.1016/0022-0965\(83\)90023-1](https://doi.org/10.1016/0022-0965(83)90023-1).
- Akimoto, Y., Sugiura, M., Yomogida, Y., Miyauchi, C.M., Miyazawa, S., Kawashima, R., 2014. Irony comprehension: social conceptual knowledge and emotional response. *Hum. Brain Mapp.* 35 (4), 1167–1178. <https://doi.org/10.1002/hbm.22242>.
- Bohrn, I.C., Altmann, U., Jacobs, A.M., 2012. Looking at the brains behind figurative language: a quantitative meta-analysis of neuroimaging studies on metaphor, idiom, and irony processing. *Neuropsychologia* 50 (11), 2669–2683. <https://doi.org/10.1016/j.neuropsychologia.2012.07.021>.
- Bosco, F.M., Parola, A., Valentini, M.C., Morese, R., 2017. Neural correlates underlying the comprehension of deceitful and ironic communicative intentions. *Cortex* 94, 73–86. <https://doi.org/10.1016/j.cortex.2017.06.010>.
- Brett, M., Anton, J.-L., Valabregue, R., Poline, J.-B., 2002. Region of interest analysis using an SPM toolbox. *Neuroimage* 16 (2), 1140–1141. [https://doi.org/10.1016/S1053-8119\(02\)90013-3](https://doi.org/10.1016/S1053-8119(02)90013-3).
- Bryant, G.A., Fax Tree, J.E., 2005. Is there an ironic tone of voice? *Lang. Speech* 48 (3), 257–277. [10.1177/002238309050480030101](https://doi.org/10.1177/002238309050480030101).
- Caillies, S., Gobin, P., Obert, A., Terrien, S., Coutté, A., Iakimova, G., Besche-Richard, C., 2019. Asymmetry of affect in verbal irony understanding: what about the N400 and P600 components? *J. Neurolinguistics* 51, 268–277. <https://doi.org/10.1016/j.jneuroling.2019.04.004>.
- Capelli, C.A., Nakagawa, N., Madden, C.M., 1990. How children understand sarcasm: the role of context and intonation. *Child Dev* 61 (6), 1824–1841. <https://doi.org/10.2307/1130840>.
- Carr, L., Iacoboni, M., Dubeau, M.C., Mazziotta, J.C., Lenzi, G.L., 2003. Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proc. Natl. Acad. Sci. U.S.A.* 100 (9), 5497–5502. <https://doi.org/10.1073/pnas.0935845100>.
- Channon, S., Pellieff, A., Rule, A., 2005. Social cognition after head injury: sarcasm and theory of mind. *Brain Lang* 93 (2), 123–134. <https://doi.org/10.1016/j.bandl.2004.09.002>.
- Colston, H., O'Brien, J., 2000. Contrast and pragmatics in figurative language: anything understatement can do, irony can do better. *J. Pragmat.* 32 (11), 1557–1583. [https://doi.org/10.1016/S0378-2166\(99\)00110-1](https://doi.org/10.1016/S0378-2166(99)00110-1).
- Colston, H., 2002. Pragmatic justifications for nonliteral gratitude acknowledgments: oh sure, anytime. *Metaphor Symb.* 17 (3), 205–226. [https://doi.org/10.1207/S15327868MS1703\\_3](https://doi.org/10.1207/S15327868MS1703_3).
- Dosenbach, N.U., Visscher, K.M., Palmer, E.D., Miezin, F.M., Wenger, K.K., Kang, H.C., Burgund, E.D., Grimes, A.L., Schlaggar, B.L., Petersen, S.E., 2006. A core system for

- the implementation of task sets. *Neuron* 50 (5), 799–812. <https://doi.org/10.1016/j.neuron.2006.04.031>.
- Di Martino, A., Ross, K., Uddin, L.Q., Sklar, A.B., Castellanos, F.X., Milham, M.P., 2009. Functional brain correlates of social and nonsocial processes in autism spectrum disorders: an activation likelihood estimation meta-analysis. *Biol. Psychiatr.* 65 (1), 63–74. <https://doi.org/10.1016/j.biopsych.2008.09.022>.
- Eviatar, Z., Just, M.A., 2006. Brain correlates of discourse processing: an fMRI investigation of irony and conventional metaphor comprehension. *Neuropsychologia* 44 (12), 2348–2359. <https://doi.org/10.1016/j.neuropsychologia.2006.05.007>.
- Filcik, R., Turcan, A., Ralph-Nearman, C., Pitiot, A., 2019. What is the difference between irony and sarcasm? An fMRI study. *Cortex* 115, 112–122. <https://doi.org/10.1016/j.cortex.2019.01.025>.
- Friston, K., 2005. A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360 (1456), 815–836. <https://doi.org/10.1098/rstb.2005.1622>.
- Friston, K.J., Holmes, A., Poline, J.-B., Price, C.J., Frith, C.D., 1996. Detecting activations in PET and fMRI: levels of inference and power. *Neuroimage* 4 (3 Pt 1), 223–235. <https://doi.org/10.1006/nimg.1996.0074>.
- Friston, K.J., Ashburner, J., Kiebel, S.J., Nichols, T.E., Penny, W.D., 2007. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Elsevier Academic Press, London. <https://www.elsevier.com/books/statistical-parametric-mapping-the-analysis-of-functional-brain-images/penny/978-0-12-372560-8>.
- Frith, U., Frith, C.D., 2003. Development and neurophysiology of mentalizing. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358 (1431), 459–473. <https://doi.org/10.1098/rstb.2002.1218>.
- Frühholz, S., Trost, W., Kotz, S.A., 2016. The sound of emotions: towards a unifying neural network perspective of affective sound processing. *Neurosci. Biobehav. Rev.* 68, 96–110. <https://doi.org/10.1016/j.neubiorev.2016.05.002>.
- Gallagher, S., 2000. Philosophical conceptions of the self: implications for cognitive science. *Trends Cognit. Sci.* 4 (1), 14–21. [https://doi.org/10.1016/s1364-6613\(99\)01417-5](https://doi.org/10.1016/s1364-6613(99)01417-5).
- Gerrig, R., Goldvarg, Y., 2000. Additive effects in the perception of sarcasm: situational disparity and echoic mention. *Metaphor Symb.* 15 (4), 197–208. [https://doi.org/10.1207/S15327868MS1504\\_1](https://doi.org/10.1207/S15327868MS1504_1).
- Hagoort, P., 2005. On Broca, brain, and binding: a new framework. *Trends Cognit. Sci.* 9 (9), 416–423. <https://doi.org/10.1016/j.tics.2005.07.004>.
- Happé, F.G., 1993. Communicative competence and theory of mind in autism: a test of relevance theory. *Cognition* 48 (2), 101–119. [https://doi.org/10.1016/0010-0277\(93\)90026-r](https://doi.org/10.1016/0010-0277(93)90026-r).
- Happé, F.G., 1994. An advanced test of theory of mind: understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *J. Autism Dev. Disord.* 24 (2), 129–154. <https://doi.org/10.1007/BF02172093>.
- Herold, R., Varga, E., Hajnal, A., Hamvas, E., Berecz, H., Tóth, B., Tényi, T., 2018. Altered neural activity during irony comprehension in unaffected first-degree relatives of schizophrenia patients: an fMRI study. *Front. Psychol.* 8, 2309. <https://doi.org/10.3389/fpsyg.2017.02309>.
- Ivanko, S.L., Pexman, P.M., 2003. Context incongruity and irony processing. *Discourse Process* 35 (3), 241–279. [https://doi.org/10.1207/S15326950DP3503\\_2](https://doi.org/10.1207/S15326950DP3503_2).
- Kaland, N., Møller-Nielsen, A., Callesen, K., Mortensen, E.L., Gottlieb, D., Smith, L., 2002. A new advanced test of theory of mind: evidence from children and adolescents with Asperger syndrome. *JCPP (J. Child Psychol. Psychiatry)* 43 (4), 517–528. <https://doi.org/10.1111/1469-7610.00042>.
- Kaland, N., Møller-Nielsen, A., Smith, L., Mortensen, E.L., Callesen, K., Gottlieb, D., 2005. The Strange Stories test: a replication study of children and adolescents with Asperger syndrome. *Eur. Child Adolesc. Psychiatr.* 14 (2), 73–82. <https://doi.org/10.1007/s00787-005-0434-2>.
- Katz, A.N., Lee, C.J., 1993. The role of authorial intent in determining verbal irony and metaphor. *Metaphor Symb.* 8 (4), 257–279. [https://doi.org/10.1207/s15327868ms0804\\_1](https://doi.org/10.1207/s15327868ms0804_1).
- Katz, A.N., Pexman, P.M., 1997. Interpreting figurative statements: speaker occupation can change metaphor to irony. *Metaphor Symb.* 12 (1), 19–41. [https://doi.org/10.1207/s15327868ms1201\\_3](https://doi.org/10.1207/s15327868ms1201_3).
- Kreuz, R.J., Glucksberg, S., 1989. How to be sarcastic: the echoic reminder theory of verbal irony. *J. Exp. Psychol. Gen.* 118 (4), 374–386. <https://doi.org/10.1037/0096-3445.118.4.374>.
- Laval, V., Bert-Erboul, A., 2005. French-speaking children's understanding of sarcasm: the role of intonation and context. *J. Speech Lang. Hear. Res.* 48 (3), 610–620. [https://doi.org/10.1044/1092-4388\(2005\)042](https://doi.org/10.1044/1092-4388(2005)042).
- Leekam, S.R., Prior, M., 1994. Can autistic children distinguish lies from jokes? A second look at second-order belief attribution. *JCPP (J. Child Psychol. Psychiatry)* 35 (5), 901–915. <https://doi.org/10.1111/j.1469-7610.1994.tb02301.x>.
- Le Gall, E., Iakimova, G., 2018. Cognition sociale dans la schizophrénie et les troubles du spectre de l'autisme: points de convergences et différences fonctionnelles [Social cognition in schizophrenia and autism spectrum disorder: points of convergence and functional differences]. *L'Encephale* 44 (6), 523–537. <https://doi.org/10.1016/j.encep.2018.03.004>.
- Mai, J.K., Majtanik, M., Paxinos, G., 2015. *Atlas of the Human Brain*, fourth ed. Elsevier Academic Press, London <https://www.elsevier.com/books/atlas-of-the-human-brain/mai/978-0-12-802800-1>.
- Mano, Y., Harada, T., Sugiura, M., Saito, D.N., Sadato, N., 2009. Perspective-taking as part of narrative comprehension: a functional MRI study. *Neuropsychologia* 47 (3), 813–824. <https://doi.org/10.1016/j.neuropsychologia.2008.12.011>.
- Martin, I., McDonald, S., 2004. An exploration of causes of non-literal language problems in individuals with Asperger Syndrome. *J. Autism Dev. Disord.* 34 (3), 311–328. <https://doi.org/10.1023/b:jadd.0000029553.52889.15>.
- Matsui, T., Nakamura, T., Utsumi, A., Sasaki, A.T., Koike, T., Yoshida, Y., Harada, T., Tanabe, H.C., Sadato, N., 2016. The role of prosody and context in sarcasm comprehension: behavioral and fMRI evidence. *Neuropsychologia* 87, 74–84. <https://doi.org/10.1016/j.neuropsychologia.2016.04.031>.
- Medford, N., Critchley, H.D., 2010. Conjoint activity of anterior insular and anterior cingulate cortex: awareness and response. *Brain Struct. Funct.* 214 (5–6), 535–549. <https://doi.org/10.1007/s00429-010-0265-x>.
- Menon, V., Uddin, L.Q., 2010. Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* 214 (5–6), 655–667. <https://doi.org/10.1007/s00429-010-0262-0>.
- Moeller, S., Yacoub, E., Olman, C.A., Auerbach, E., Strupp, J., Harel, N., Ugurbil, K., 2010. Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magn. Reson. Med.* 63 (5), 1144–1153. <https://doi.org/10.1002/mrm.22361>.
- Morita, T., Tanabe, H.C., Sasaki, A.T., Shimada, K., Kakigi, R., Sadato, N., 2014. The anterior insular and anterior cingulate cortices in emotional processing for self-face recognition. *Soc. Cognit. Affect Neurosci.* 9 (5), 570–579. <https://doi.org/10.1093/scan/nst011>.
- Nakamura, T., Matsui, T., Utsumi, A., Yamazaki, M., Makita, K., Harada, T., Tanabe, H.C., Sadato, N., 2018. The role of the amygdala in incongruity resolution: the case of humor comprehension. *Soc. Neurosci.* 13 (5), 553–565. <https://doi.org/10.1080/17470919.2017.1365760>.
- Nelson, S.M., Dosenbach, N.U., Cohen, A.L., Wheeler, M.E., Schlaggar, B.L., Petersen, S.E., 2010. Role of the anterior insula in task-level control and focal attention. *Brain Struct. Funct.* 214 (5–6), 669–680. <https://doi.org/10.1007/s00429-010-0260-2>.
- Obert, A., Gierski, F., Calmus, A., Flucher, A., Portefaix, C., Pierot, L., Kaladjian, A., Caillies, S., 2016. Neural correlates of contrast and humor: processing common features of verbal irony. *PLoS One* 11 (11), e0166704. <https://doi.org/10.1371/journal.pone.0166704>.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the edinburgh inventory. *Neuropsychologia* 9 (1), 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4).
- Ousdal, O.T., Jensen, J., Server, A., Hariri, A.R., Nakstad, P.H., Andreassen, O.A., 2008. The human amygdala is involved in general behavioral relevance detection: evidence from an event-related functional magnetic resonance imaging go-nogo task. *Neuroscience* 156 (3), 450–455. <https://doi.org/10.1016/j.neuroscience.2008.07.066>.
- Phan, K.L., Wager, T., Taylor, S.F., Liberzon, I., 2002. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage* 16 (2), 331–348. <https://doi.org/10.1006/nimg.2002.1087>.
- Rapp, A.M., Mutschler, D.E., Wild, B., Erb, M., Lengsfeld, I., Saur, R., Grodd, W., 2010. Neural correlates of irony comprehension: the role of schizotypal personality traits. *Brain Lang* 113 (1), 1–12. <https://doi.org/10.1016/j.bandl.2009.11.007>.
- Rapp, A.M., Mutschler, D.E., Erb, M., 2012. Where in the brain is nonliteral language? A coordinate-based meta-analysis of functional magnetic resonance imaging studies. *Neuroimage* 63 (1), 600–610. <https://doi.org/10.1016/j.neuroimage.2012.06.022>.
- Reyes-Aguilar, A., Valles-Capetillo, E., Giordano, M., 2018. A quantitative meta-analysis of neuroimaging studies of pragmatic language comprehension: in search of a universal neural substrate. *Neuroscience* 395, 60–88. <https://doi.org/10.1016/j.neuroscience.2018.10.043>.
- Rivière, E., Klein, M., Champagne-Lavau, M., 2018. Using context and prosody in irony understanding: variability amongst individuals. *J. Pragmat.* 138, 165–172. <https://doi.org/10.1016/j.pragma.2018.10.006>.
- Sander, D., Grafman, J., Zalla, T., 2003. The human amygdala: an evolved system for relevance detection. *Rev. Neurosci.* 14 (4), 303–316. <https://doi.org/10.1515/revneuro.2003.14.4.303>.
- Seeley, W.W., Menon, V., Schatzberg, A.F., Keller, J., Glover, G.H., Kenna, H., Reiss, A.L., Greicius, M.D., 2007. Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* 27 (9), 2349–2356. <https://doi.org/10.1523/JNEUROSCI.5587-06.2007>.
- Seth, A.K., 2013. Interoceptive inference, emotion, and the embodied self. *Trends Cognit. Sci.* 17 (11), 565–573. <https://doi.org/10.1016/j.tics.2013.09.007>.
- Seth, A.K., Friston, K.J., 2016. Active interoceptive inference and the emotional brain. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371, 20160007. <https://doi.org/10.1098/rstb.2016.0007>.
- Shibata, H., Toyomura, A., Itoh, H., Abe, J., 2010. Neural substrates of irony comprehension: a functional MRI study. *Brain Res* 1308, 114–123. <https://doi.org/10.1016/j.brainres.2009.10.030>.
- Sperber, D., Wilson, D., 1995. *Relevance: Communication and Cognition*, second ed. Blackwell, Oxford.
- Spotorno, N., Koun, E., Prado, J., Van Der Henst, J.B., Noveck, I.A., 2012. Neural evidence that utterance-processing entails mentalizing: the case of irony. *Neuroimage* 63 (1), 25–39. <https://doi.org/10.1016/j.neuroimage.2012.06.046>.
- Sridharan, D., Levitin, D.J., Menon, V., 2008. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proc. Natl. Acad. Sci. U.S.A.* 105 (34), 12569–12574. <https://doi.org/10.1073/pnas.0800005105>.
- Tantam, D., 1991. Asperger syndrome in adulthood. In: Frith, U. (Ed.), *Autism and Asperger Syndrome*. Cambridge University Press, Cambridge, pp. 147–183. <https://doi.org/10.1017/CBO9780511526770.005>.
- Tobe, R.H., Corcoran, C.M., Breland, M., MacKay-Brandt, A., Klim, C., Colcombe, S.J., Leventhal, B.L., Javitt, D.C., 2016. Differential profiles in auditory social cognition deficits between adults with autism and schizophrenia spectrum disorders: a preliminary analysis. *J. Psychiatr. Res.* 79, 21–27. <https://doi.org/10.1016/j.jpsychires.2016.04.005>.

- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15 (1), 273–289. <https://doi.org/10.1006/nimg.2001.0978>.
- Uchiyama, H., Seki, A., Kageyama, H., Saito, D.N., Koeda, T., Ohno, K., Sadato, N., 2006. Neural substrates of sarcasm: a functional magnetic-resonance imaging study. *Brain Res* 1124 (1), 100–110. <https://doi.org/10.1016/j.brainres.2006.09.088>.
- Uchiyama, H.T., Saito, D.N., Tanabe, H.C., Harada, T., Seki, A., Ohno, K., Koeda, T., Sadato, N., 2012. Distinction between the literal and intended meanings of sentences: a functional magnetic resonance imaging study of metaphor and sarcasm. *Cortex* 48 (5), 563–583. <https://doi.org/10.1016/j.cortex.2011.01.004>.
- Uddin, L.Q., Supekar, K., Menon, V., 2013. Reconceptualizing functional brain connectivity in autism from a developmental perspective. *Front. Hum. Neurosci.* 7 (458), 1–11. <https://doi.org/10.3389/fnhum.2013.00458>.
- Utsumi, A., 2000. Verbal irony as implicit display of ironic environment: distinguishing ironic utterances from nonirony. *J. Pragmat.* 32 (12), 1777–1806. [https://doi.org/10.1016/S0378-2166\(99\)00116-2](https://doi.org/10.1016/S0378-2166(99)00116-2).
- Van Overwalle, F., Manto, M., Cattaneo, Z., Clausi, S., Ferrari, C., Gabrieli, J., Guell, X., Heleven, E., Lupo, M., Ma, Q., Michelutti, M., Olivito, G., Pu, M., Rice, L.C., Schmahmann, J.D., Siciliano, L., Sokolov, A.A., Stoodley, C.J., van Dun, K., Vandervert, L., Leggio, M., 2020. Consensus paper: cerebellum and social cognition. *Cerebellum* 19 (6), 833–868. <https://doi.org/10.1007/s12311-020-01155-1>.
- Varga, E., Simon, M., Tényi, T., Schnell, Z., Hajnal, A., Orsi, G., Dóczy, T., Komoly, S., Janszky, J., Füredi, R., Hamvas, E., Fekete, S., Herold, R., 2013. Irony comprehension and context processing in schizophrenia during remission: a functional MRI study. *Brain Lang* 126 (3), 231–242. <https://doi.org/10.1016/j.bandl.2013.05.017>.
- Voyer, D., Thibodeau, S.H., Delong, B.J., 2016. Context, contrast, and tone of voice in auditory sarcasm perception. *J. Psycholinguist. Res.* 45, 29–53. <https://doi.org/10.1007/s10936-014-9323-5>.
- Wakusawa, K., Sugiura, M., Sassa, Y., Jeong, H., Horie, K., Sato, S., Yokoyama, H., Tsuchiya, S., Inuma, K., Kawashima, R., 2007. Comprehension of implicit meanings in social situations involving irony: a functional MRI study. *Neuroimage* 37 (4), 1417–1426. <https://doi.org/10.1016/j.neuroimage.2007.06.013>.
- Wang, A.T., Lee, S.S., Sigman, M., Dapretto, M., 2006a. Developmental changes in the neural basis of interpreting communicative intent. *Soc. Cognit. Affect Neurosci.* 1 (2), 107–121. <https://doi.org/10.1093/scan/nsl018>.
- Wang, A.T., Lee, S.S., Sigman, M., Dapretto, M., 2006b. Neural basis of irony comprehension in children with autism: the role of prosody and context. *Brain* 129 (4), 932–943. <https://doi.org/10.1093/brain/awl032>.
- Wickens, S., Perry, C., 2015. What do you mean by that?! An electrophysiological study of emotional and attitudinal prosody. *PLoS One* 10 (7), e0132947. <https://doi.org/10.1371/journal.pone.0132947>.
- Woodland, J., Voyer, D., 2011. Context and intonation in the perception of sarcasm. *Metaphor Symb.* 26 (3), 227–239. <https://doi.org/10.1080/10926488.2011.583197>.