Neural Integration of Affective Prosodic and Semantic Cues in Non-literal Forms of Speech Understanding

Adrien Wittmann[1], Leonardo Ceravolo[2], Audrey Mayr[1], & Didier Grandjean[2]

[1] Neuroscience of Emotions and Affective Dynamics lab, Swiss center for Affective Sciences, Department of Psychology and Educational Sciences, University of Geneva, Biotech campus, Chemin des Mines 9, 1202, Geneva, Switzerland

[2] Neuroscience of Emotions and Affective Dynamics lab, Swiss Center for Affective Sciences, Department of Psychology and Educational Sciences, University of Geneva, Unimail, Chemin des Mines 9, 1202, Geneva, Switzerland

Author Note

The authors made the following contributions. Adrien Wittmann: Conceptualization, Methodology, Data curation, Formal analysis, Software, Investigation, Validation, Visualisation, Writing - Original Draft Preparation, Writing - Review & Editing; Leonardo Ceravolo: Methodology, Software, Writing - Review & Editing; Audrey Mayr: Conceptualization, Software, Methodology, Investigation, Validation; Didier Grandjean: Funding acquistion, Supervision, Conceptualization, Writing - Review & Editing.

Correspondence concerning this article should be addressed to Adrien Wittmann, Neuroscience of Emotions and Affective Dynamics lab, Swiss Center for Affective Sciences, Department of Psychology and Educational Sciences, University of Geneva, Biotech campus, Chemin des Mines 9, 1202, Geneva, Switzerland. E-mail: adrienwitt@gmail.com

Abstract

Emotions in speech are conveyed through both semantics and prosody—voice 'melody', yet how listeners integrate these cues in the brain remains unclear. We investigated non-literal forms of speech, such as irony and sarcasm, where understanding beyond literal meaning relies on the dynamic interplay between affective prosodic and semantic cues, alongside theory of mind (ToM). We used functional magnetic resonance imaging while participants listened to short dialogues between two characters, which varied in prosody and semantics to convey either literal or non-literal meanings. Behaviorally, semantics and prosody interacted in shaping participants' evaluations, although the data suggest a prosody dominance effect. Non-literal speech engaged the bilateral inferior frontal gyrus, temporal speech regions, and ToM areas, while integration-demanding tasks (e.g., irony, sarcasm, ToM) evoked stronger ToM-network activation than prosody or semantics tasks. Taken together, our data clarify the behavioral and neural underpinnings of the integration of prosody and semantics in non-literal speech and open new venues in this field.

Keywords: integration, irony, prosody, sarcasm, semantics, ToM

Word count: 7555

Neural Integration of Affective Prosodic and Semantic Cues in Non-literal Forms of Speech Understanding

Human communication employs various modalities, yet language stands out for its capacity to convey precise, abstract, and elaborated expressions. Language may serve many purposes, among them the communication of emotions or attitudes. These can be conveyed in different ways in speech—either through prosody, the intonation and rhythm of the voice, or through semantics, using words and their meanings. These two types of emotional information are differentially communicated (Castelluccio et al., 2016), interactive (Astésano et al., 2004), and activate common brain regions (Buchanan et al., 2000; Meyer et al., 2003; R. L. C. Mitchell, 2007). However, the way these different cues are integrated at the behavioral and brain levels is still little understood. To better understand this semantic and prosodic integration, researchers have often used an "incongruity" paradigm (Kotz et al., 2015; Lin et al., 2020; R. L. C. Mitchell, 2006; Pell et al., 2011; Wittfoth et al., 2009). When emotional prosodic and semantic cues are incongruent, as observed in non-literal forms of speech such as irony or sarcasm, listeners must integrate these cues to effectively decipher the intended emotional message (Matsui et al., 2016). Beyond integration, listeners need to make inferences about others' feelings, attitudes, and mental states—concepts referred to as ToM (Zhu & Wang, 2020). The primary goal of our research was thus to shed light on the brain regions involved in the integration between semantic and prosodic affective cues during the processing of non-literal forms of speech using functional Magnetic Resonance Imaging (fMRI), taking into account the different processes at play, including prosody and semantics decoding, irony and sarcasm inferences, and the related ToM processes.

To date, many studies have examined the neural networks involved in perceiving emotions in speech (Castelluccio et al., 2016; Frühholz & Belin, 2018; Seydell-Greenwald et al., 2020). One key channel is prosody, expressed through suprasegmental features such as

tone modulation (Lin et al., 2020), which involves acoustic changes in pitch (F0), loudness, rhythm, articulation rate, and timbre (Brück et al., 2011; Scherer, 2003; Wildgruber et al., 2006). Processing emotional prosody activates the primary and secondary auditory cortices in the temporal lobe, particularly the superior temporal cortex (STC) or temporal voice-sensitive areas (TVAs; Belin et al., 2000; Grandjean, 2021). The middle STC is consistently activated across fMRI studies comparing vocal emotions to neutral voices, independent of attention or task demands (Ethofer et al., 2006; Grandjean et al., 2005; Witteman et al., 2011). While lesion studies have long suggested right-hemisphere dominance (Borod et al., 2002; Ross & Monnot, 2008; Witteman et al., 2011), this strict lateralization has been challenged (Kotz et al., 2006), with evidence of bilateral STC involvement (Brück et al., 2011). Emotional prosody processing is dynamic and multi-stage, encompassing sensory analysis, emotional inference, and explicit evaluation (Brück et al., 2011; Frühholz & Grandjean, 2012; Kotz & Paulmann, 2011; Schirmer & Kotz, 2006; Wildgruber et al., 2009). First, basic acoustic features are extracted in the bilateral auditory cortices. Second, these features are integrated into an emotional percept, with models differing on the key brain areas involved—either the right posterior superior temporal cortex (Wildgruber et al., 2009) or an anteroventral drift from the bilateral superior temporal gyrus (STG) to the anterior superior temporal sulcus (STS; Schirmer et al.,2006). Finally, emotional prosody is explicitly evaluated, engaging the STG, orbitofrontal cortex (OFC), and inferior frontal gyrus (IFG). While Brück et al. (2011) propose bilateral IFG involvement, Kotz and Paulmann (2011) suggest that the right IFG and OFC primarily handle explicit evaluation. In contrast, the left IFG appears to integrate emotional prosody with other aspects of speech, such as semantics. This role is further supported by greater left IFG activation when integration demands are higher—for example, when prosodic and semantic cues are incongruent (Kotz et al., 2015; R. L. C. Mitchell, 2006; Schirmer et al., 2004; Wiethoff et al., 2008). However, the precise neural mechanisms supporting this integration—particularly in irony and sarcasm, which feature

incongruence between contextual cues, prosody, and semantics—remain poorly understood. The present study addresses this gap using fMRI to examine how prosodic and semantic cues are jointly processed during these forms of non-literal speech.

Emotion in speech is also conveyed through words and utterances whose meanings, associations, and contexts may evoke emotional significance. Neuroimaging and lesion studies have identified key regions for semantic processing, primarily in the left hemisphere. The anterior temporal lobe (ATL), including the anterior superior temporal sulcus and gyrus, is crucial for sentence comprehension (Friederici, 2011; Matchin & Hickok, 2020), although its exact role remains debated. Initially linked to syntactic processing (Friederici et al., 2000; Friederici, 2003; Mazoyer et al., 1993), the ATL has also been shown to support compositional semantic processing (Humphries et al., 2006; Vandenberghe et al., 2002; Westerlund & Pylkkänen, 2014), or both of these processes (Rogalsky & Hickok, 2009). The middle temporal gyrus (MTG) integrates lexical information in context (Binder et al., 2009; Hoffman et al., 2015). The IFG is active during semantic processing and responds more strongly to abstract and emotional words (Hoffman et al., 2015; Pauligk et al., 2019). Processing emotional sentences, with or without prosody, increases activation in the left IFG, the posterior left superior temporal sulcus, and the medial prefrontal cortex—part of the mentalizing network (Beaucousin et al., 2006). Finally, the angular gyrus (AG) acts as a supramodal integration hub for high-level semantic tasks such as concept retrieval and combination (Binder et al., 2009).

Irony and sarcasm are non-literal forms of communication that express ideas opposite to their literal meaning (Bosco et al., 2017; Matsui et al., 2016; Searle, 1979). Irony involves incongruity between explicit and intended meaning, often creating humor or satire. It can express praise through seemingly negative language or, in sarcastic irony, convey derision through superficially positive language. Understanding these non-literal forms of speech hinges on the incongruence between the lexico-semantic content of the utterance and the context or expectations about the speaker's attitude, while paralinguistic

cues—such as tone of voice—play an important role in facilitating the recognition of irony (Mauchand et al., 2020; Nakamura et al., 2022). Several authors have investigated whether a specific ironic or sarcastic tone of voice exists and have found that mean F0 is the most significant acoustic cue for ironic interpretation, with sarcastic utterances generally exhibiting a lower pitch at the perceptual level—though in some languages, such as French, they exhibit a higher pitch (Bryant, 2010; Cheang & Pell, 2008; Lœvenbruck et al., 2013; Rockwell, 2000; Scharrer & Christmann, 2011).

Research on the neural processing of irony and sarcasm has mainly focused on reading, identifying key regions such as the inferior frontal gyrus (IFG) and the medial prefrontal cortex (mPFC; Bosco et al., 2017; Filik et al., 2019; Shibata et al., 2010; Uchiyama et al., 2006). However, only a few studies have examined the auditory processing of irony. Matsui et al. (2016) investigated how contextual and prosodic cues interact during the processing of sarcastic utterances, finding significant activations in the rostroventral area of the left IFG, which they proposed plays a role in integrating discourse context, utterance meaning, and affective prosody. Nakamura et al. (2022) examined how context and prosody interact during sarcasm comprehension, reporting critical activations for context–content incongruity in the anterior rostral medial prefrontal cortex, temporal pole, and bilateral cerebellum, and for content–prosody incongruity in the bilateral amygdala. They further observed interaction effects in the salience network, including the anterior insula and anterior cingulate cortex, extending into the dorsolateral prefrontal cortex and left IFG, which they proposed integrate these incongruity signals to support sarcasm comprehension. Finally, Obert et al. (2016) compared the processing of ironic versus literal utterances and found that irony processing activated several brain regions, including the bilateral IFG, the posterior portion of the left superior temporal gyrus (STG), the mPFC, and the caudate.

Another point of interest in our study concerns ToM—the ability to infer others' intentions, beliefs, perspectives, and emotions, and to recognize that others have mental

states distinct from one's own (R. L. C. Mitchell & Phillips, 2015). ToM is central to irony and sarcasm comprehension, as understanding a speaker's intended meaning requires inferring their mental state. ToM and pragmatic abilities are developmentally linked in children (Bosco & Gabbatore, 2017), and ToM deficits are associated with impaired irony comprehension in individuals with autism spectrum disorder (Happé, 1993; Saban-Bezalel et al., 2019; Song et al., 2024). A broad consensus has emerged from numerous studies supporting the existence of a dedicated neural network underlying ToM. This network encompasses several brain regions, including the bilateral temporoparietal junction (TPJ), associated with perspective-taking and representing others' goals and intentions; the medial prefrontal cortex (mPFC), which engages in reflective reasoning and predicts others' actions (dorsally) and emotions (ventrally) to form enduring trait impressions; and the precuneus (PCun), involved in mental imagery (Frith & Frith, 2006; J. P. Mitchell, 2009; Saxe & Baron-Cohen, 2006; Schurz et al., 2014; Tamir et al., 2016; Van Overwalle, 2009; Van Overwalle & Baetens, 2009). Although lesion studies support the link between ToM and irony processing (Bosco et al., 2018; Channon et al., 2005; Cordonier et al., 2020; Shamay-Tsoory et al., 2005), neuroimaging evidence remains scarce. Notably, (Spotorno et al., 2012) showed that processing irony in healthy adults, compared to literal speech, increases activity in the ToM network—including the mPFC, bilateral TPJ, and, at a more lenient threshold, the PCun—as well as in the bilateral inferior frontal gyrus (IFG), left insula (INS), right dorsolateral prefrontal cortex (dlPFC), and right middle temporal gyrus (MTG). Moreover, connectivity analyses revealed enhanced coupling between the mPFC and left IFG—and, to a lesser degree, the right IFG—during irony comprehension. Taken together, findings from both reading and auditory modalities indicate that irony comprehension engages brain regions supporting semantic processing and ToM, processes that are critical for inferring ironic intent. Nevertheless, how these networks interact to integrate semantic, prosodic, and contextual information—and how such integration relates to higher-order mentalizing functions—remains theoretically unclear and warrants further

neuroimaging investigation.

According to the predictive processing framework (Friston, 2005, 2010), irony comprehension has been proposed to rely on precision-modulated prediction error minimization across pragmatic, social, and emotional components that jointly help infer the speaker's intended, non-literal meaning (Fabry, 2021). In our view, irony comprehension may hinge on identifying and resolving prediction errors that stem from incongruencies between the context and the semantic content, as well as between the semantic content and prosody within an utterance. To minimize these prediction errors, a dynamic integration between contextual, semantic, and prosodic cues is thus needed. To process this integration, the brain engages in evidence accumulation across these different cues in a precision-modulated and bottom-up fashion. These integrated representations are then relayed to higher-order regions that also engage in prediction error minimization through refinement of predictions regarding others' intentions, emotions, and the use of irony given sensory inputs and social contexts. These prediction errors are then back-propagated to lower-order regions to gain precision by assigning greater weight to relevant cues, enhancing future interpretative frameworks for irony comprehension. The left IFG may be a key integration hub that combines prosodic and semantic cues in a bottom-up manner and transmits this information to higher-order and social cognition regions (Tettamanti et al., 2017). The resulting cognitive evaluations are subsequently fed back to the IFG, allowing it to refine perceptual integration and minimize prediction errors through top-down modulation of lower-order processing in temporal regions (Cope et al., 2023; He et al., 2024). This view is supported by evidence of its functional connectivity with superior and middle temporal regions involved in speech comprehension and emotional prosody processing (Ethofer et al., 2012; Frühholz & Grandjean, 2012; Huang et al., 2024; Leitman et al., 2010; Turken et al., 2011; Yue et al., 2013), as well as with the mentalizing network (Mori & Haruno, 2022; Spotorno et al., 2012; Tettamanti et al., 2017). The main goal of our research was thus to gain insight into the brain network underlying the integration of

these different aspects—from lower-order perceptual processing in speech-related regions, through integrative mechanisms, to higher-order cognitive functions such as ToM.

To achieve this, we designed scenarios featuring short conversations between two characters: the first character's sentence provides the context, while the second character's response serves as the target statement. By manipulating the semantics of the context, as well as the semantics and prosody of the target statement, we created distinct conditions that differed in their potential to be perceived as ironic or sarcastic. In the present fMRI study, we used two literal conditions—positive sincerity (positive context, positive statement, positive prosody) and negative sincerity (negative context, negative statement, negative prosody)—and two non-literal conditions—sarcastic irony (negative context, positive statement, negative prosody) and praise irony (positive context, negative statement, positive prosody). We then developed five tasks in which participants were required to attend to and evaluate specific information, with each task focusing on different components involved in understanding non-literal speech, including emotional prosody and semantic decoding, irony and sarcasm inferences, as well as ToM processes. In the prosody and semantic evaluation tasks, participants were specifically required to focus solely on the modality corresponding to the task, whereas in the irony, sarcasm, and ToM tasks, they were instructed to use and integrate all available cues.

In Hypothesis 1 (behavioral data), we predicted an interaction between prosodic and semantic cues across all tasks, with a reduced effect size in tasks where participants were instructed to focus exclusively on a single modality (i.e., prosody or semantics) compared to tasks requiring integration of both cues (i.e., irony, sarcasm, and ToM tasks). Indeed, we still expected that prosodic information would continue to influence semantic judgments, and conversely, even in tasks with unimodal instructions. For fMRI data, we hypothesized that the integration of prosodic and semantic cues during non-literal speech processing would trigger enhanced activity in key brain networks implicated in language comprehension, multimodal integration, and socio-cognitive inference. Specifically, we

predicted heightened activations in: i) the left IFG as a marker of non-literal speech processing (Matsui et al., 2016; Rapp et al., 2012) and highlighting its potential role as a key hub for integrating prosodic, semantic, and contextual cues (Schirmer & Kotz, 2006); ii) the speech processing network (Friederici, 2011, 2012), including the bilateral MTG, STG, STS, and ATL, driven by greater demands for prosodic and semantic decoding and integration; iii) the ToM network (Schurz et al., 2014), including the mPFC, the TPJ, and the PCun supporting increased demands for perspective taking and inference on others' intents. These effects were expected to manifest across three key contrasts. In Hypothesis 2, for the main effect of utterance type (Non-literal > Literal speech), we predicted greater recruitment of the left IFG, speech processing network, and ToM network, driven by the increased demands for resolving incongruencies between prosodic and semantic cues and for inferring social intent in ironic statements. In Hypothesis 3, for the main effect of task demands (Integrative tasks [Irony + Sarcasm + ToM] > Non-integrative tasks [Prosody + Semantic]), we predicted similar increases in these networks, as the integrative tasks explicitly require participants to combine multimodal cues for higher-order inferences. Finally, in Hypothesis 4, for the interaction ([Irony + Sarcasm + ToM] × Non-literal > [Prosody + Semantic tasks] × Literal), we hypothesized the strongest activations in the left IFG, speech processing network, and ToM network, reflecting maximal demands for cue integration when non-literal utterances are paired with tasks that necessitate holistic processing of prosody, semantics, and social intent.

Materials and Methods

The study was approved by the Commission cantonale d'éthique de la recherche (CCER; approval number 2021-01219) and conducted in accordance with the Declaration of Helsinki and local regulations. All participants provided written informed consent and received monetary compensation for their participation.

Participants

In our fMRI study, 45 right-handed participants (23 males and 22 females, M_{age} = 23.04, SD_{age} = 4.65) were recruited at the University of Geneva. With a sample size of 45 participants, the study achieved a statistical power of 0.95 to detect an effect size of *Cohen's* $f^2 = 0.3$ at an alpha level of 0.05 with two predictors. Participation criteria required individuals to be aged from 18 to 45 years, have French as their mother tongue, not being pregnant, not suffering from claustrophobia, having no known hearing or speech impairments, and having no psychiatric history. Furthermore, they had to fill out a safety questionnaire ensuring that the MRI session was risk-free and that security-related MRI compatibility was granted, which was the case for all of our participants.

Materials and Tasks

Our stimuli consisted of recorded sentences in French featuring short dialogues between two characters. Specifically, the first character delivered an utterance, referred to as the context ($M_{duration}$ = *2.10*, $SD_{duration}$ = 0.33), and the second character responded with the target statement (i.e., the one that could be ironic or not; $M_{duration}$ = 1.06, $SD_{duration}$ = 0.21). We created 16 scenarios in which the context's semantics, and the statement's semantics and prosody were manipulated. For a given scenario, the context could be stated either positively or negatively (e.g., "my husband won a lot of money in the lottery" vs. "this player lost a lot of money in the lottery") with a monotone prosody, while the target statement could be stated either positively or negatively (e.g., "he is lucky" vs. "he is unlucky") and delivered with a positive or negative prosody. Systematically combining these factors yielded eight distinct experimental conditions, each with the potential to vary in the degree of perceived irony or sarcasm. For the main study, we narrowed our focus to a subset of these conditions to create a clear contrast between literal and non-literal speech. We first selected the fully literal conditions—those in which

semantic valence and prosody were fully congruent—resulting in the *negative sincerity* condition (negative context, negative statement, negative prosody) and the *positive sincerity* condition (positive context, positive statement, positive prosody). To represent non-literal speech, we then selected the most relevant ironic conditions based on both theoretical considerations and the estimated marginal mean results from Pilot Study 2 (see Table S4). Specifically, we included the *sarcastic irony* condition (negative context, positive statement, negative prosody), which was perceived as the most sarcastic in the sarcasm task of Pilot Study 2. We also included its positive-context counterpart, *praise irony* (positive context, negative statement, positive prosody), which was the condition with a positive context rated as the most ironic in the irony task Pilot Study 2. These choices ensured relevance by providing non-literal forms of speech characterized by incongruence between the context and the target utterance and by featuring incongruence between prosody and semantic cues within the target statement, which was critical to capture their integration. Importantly, they ensured symmetry between sarcastic and praise forms of irony, enabling balanced analyses. Participants then had to evaluate these different conditions in 5 different tasks. In the first task, participants were asked to evaluate the prosodic emotional valence of the target statement on a 5-point Likert scale from *very negative* (1) to *very positive* (5). In a second task, they were asked to evaluate the semantic emotional valence of the target statement from *very negative* (1) to *very positive* (5). In a third task, they were required to first evaluate the degree of perceived irony from *not ironic at all* (1) to *very ironic* (5). In a fourth task, they were asked to evaluate the degree of perceived sarcasm from *not sarcastic at all* (1) to *very sarcastic* (5). Finally, we wanted participants to focus on ToM processes. To achieve this, we drew inspiration from the social faux-pas task (Baron-Cohen et al. 1999). More specifically, a social faux-pas is a task used to measure ToM abilities in which participants are presented with short scenarios and are asked to evaluate social mistakes, specifically whether someone did or said something that could have hurt or made another character feel uncomfortable, without the direct

intention to do so. In the context of our experiment, we merely asked participants whether the second character (i.e., the one making the target statement) could have made feel the first character *not uncomfortable at all* (1) to *very uncomfortable* (5). By doing so, we aimed to target "meta-meta emotional representations" since participants needed to first infer the intent of the second character and then the emotional response of the first character regarding this intent. In the prosodic and semantic tasks, they were explicitly instructed to attend only to the channel relevant to the task while ignoring the other channel. For the irony, sarcasm, and the ToM tasks, they were asked to integrate both channels along with contextual and emotional information available.

fMRI Task Procedure

Before the fMRI session, participants were provided with detailed instructions for the different tasks. Particular attention was devoted to ensuring that participants fully understood the task instructions, especially the distinction between irony and sarcasm. For instance, we clarified that sarcasm represents a specific pejorative subtype of irony—meaning that sarcastic statements should always be perceived as ironic—while irony can also take other forms, such as praise irony. Auditory stimuli were presented via MR-compatible headphones, and the task was displayed on an MR-compatible screen that participants could see through a tilted mirror attached to the head coil. Participants performed the task using a response box in their right hand. Before starting the tasks, a training phase consisting of one random example of each task was conducted within the MRI scanner. This allowed us to adjust the headphone volume according to the MRI noise level. Participants subsequently completed the experiment in five sessions, each corresponding to a different task, with each session lasting approximately 10 minutes. In each task, participants were presented with all scenarios of all conditions, resulting in 64 trials per task and 320 trials in total. For each trial, the context and the statement were separated by a fixation cross on a black screen, with the duration varying between 0.5 and

1.5 seconds to minimize habituation and expectation effects. Participants then evaluated the target statement depending on the task they were performing. The evaluation could last up to a maximum of 5 seconds, but participants could proceed to the next stimulus immediately after validating their response. For each participant and trial, the voices of the context and the statement were assigned randomly. In addition to ensuring that the voice of the context differed from that of the statement, we also implemented the constraint that participants could not encounter the same set of voices for the same scenario of a given condition in another task. This minimized habituation and promoted generalizability. Finally, the task order and the stimuli within each task were randomized. A typical session lasted approximately 1.5 hours.

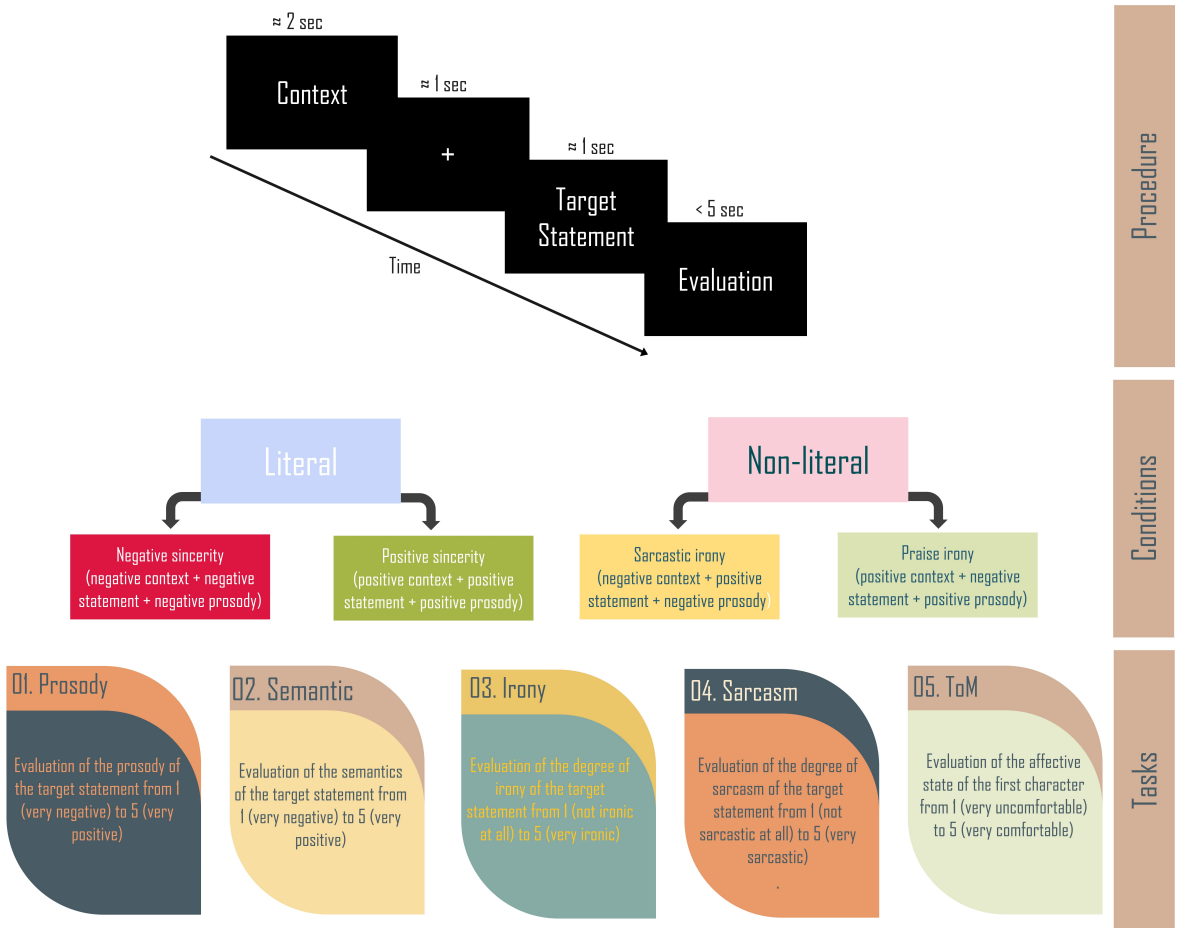

Figure 1. Experimental plan summary

Behavioral analyses

Behavioral data were analyzed using RStudio software (R Core Team, 2021). Mixed-model linear regressions were conducted using the function `lmer` from the `lme4` package (Bates et al., 2015), p-values were derived using the `Anova` function from the `car` package with the Type II sum of squares method (Fox & Weisberg, 2019), the marginal and conditional R^2 values were computed using the `performance` package (Lüdecke et al., 2021), and estimated marginal means were computed using the `emmeans` package (Lenth, 2024). Post-hoc power analysis was conducted using G*Power (Faul et al., 2007). In all our models, we added random intercepts for the participants and for the stimuli of the contexts and the statements. We also calculated the estimated marginal means for the interaction between prosody and semantics to gain a clearer understanding of how each condition was evaluated across tasks. The full set of mixed-model formulas are displayed in Table S6.

Image Acquisition

Imaging was performed at the Brain and Behaviour Laboratory (BBL) of the University of Geneva on a 3T Siemens Trio System (Siemens, Erlangen, Germany) using a T2*-weighted gradient echo planar imaging sequence with 440 volumes per run (EPI; 3.0 × 3.0 × 3.0 mmvoxels, slice thickness = 3 mm, gap = 1 mm, 36 slices, RT = 650 ms, TE = 30 ms, flip angle = 64°, matrix = 64 × 64, field of view = 192 mm). The number of volumes per run varied according to the length of the randomly assigned stimuli and the response time of the participants (Mscans = 832.112, minscans = 738, maxscans = 1024). Additionally, a T1 weighted, magnetization-prepared, rapid-acquisition, gradient echo anatomical scan (0.8 × 0.8 × 0.8 mm voxels, slice thickness = 0.8 mm, 192 slices, RT = 2000 ms, TE = 2.49 ms, flip angle = 9°, matrix = 288 × 288, FOV = 230 mm) was acquired.

Image Analysis

Analysis of functional images was performed using Statistical Parametric Mapping software version 12 (Ashburner et al., 2014). Preprocessing steps included realignment to the first volume of the time series to correct for subject motion, coregistration, normalization into the Montreal Neurological Institute space (Collins et al., 1994), and spatial smoothing with an isotropic Gaussian filter of 6 mm full width at half maximum. Anatomical locations were identified based on a standardized coordinate database using the Automated Anatomical Labeling (AAL) atlas, integrated within the xjView toolbox (http://www.alivelearn.net/xjview). Finally, the brain plots were created using the visualization tool of the CONN toolbox (Whitfield-Gabrieli & Nieto-Castanon, 2012). For the analyses, we used a general linear model to perform the first-level analysis, treating each task as a separate session. Each trial was convolved with the hemodynamic response function, time-locked to the onset of the statement and modeled with a duration extending to the end of the evaluation. Then, separate regressors were created for each condition (positive sincerity, negative sincerity, praise irony, and sarcastic irony) of each task (prosody, semantic, sarcasm, irony, and ToM), resulting in a total of 20 regressors of interest (design matrix: "prosody.positive_sincerity," "prosody.negative_sincerity," ... "ToM.praise_irony," "ToM.sarcastic_irony"). Additionally, for each run, six motion parameters were included as regressors of no interest to account for movement in our data. This resulted in a design matrix of 50 columns plus the 4 constants. To examine our main effects and interaction, we used our regressors of interest to compute 10 simple contrasts for each participant, corresponding to the simple effect of literal and non-literal conditions for each task. We then used these contrasts in a flexible factorial second-level analysis with three factors: the participants' factor (data independence set to "true", variance set to "unequal"), the non-literality factor (independence set to "false", variance set to "unequal"), and the task factor (independence set to "false", variance set to "unequal") to model the two-way interaction between our main factors (non-literality $\times$ task). All results

displayed in the figures were thresholded using a voxel-wise correction for multiple comparisons using a corrected False Discovery Rate (FDR) $p < .05$ and an arbitrary cluster size of $k > 10$ voxels to remove very small clusters.

Results

Before conducting the main study, we carried out two preliminary studies. The first study validated our stimuli by assessing whether the prosody elicited the expected emotional responses in terms of valence and intensity. The second study focused on piloting our experimental tasks, examining how context, semantics, and prosody interacted across the different tasks, and identifying the most relevant conditions to include in the main investigation. The detailed methods and results of these preliminary studies are reported in the Supplementary Material. In the main study presented in this section, participants listened to short dialogues between two characters, where the first character's sentence served as the context and the second character's sentence as the target statement. The semantic valence of the context, the semantic valence of the target statement, and its prosody were manipulated to create literal (positive and negative sincerity) and non-literal (sarcastic and praise irony) conditions. Across five tasks, participants evaluated different aspects of the target statement: prosody, semantics, irony, sarcasm, and ToM. While the prosody and semantic tasks required focus on a single modality, the irony, sarcasm, and ToM tasks required integration of prosodic, semantic, and contextual cues. Building on this, we investigated the interaction between target statement semantics and prosody across our different tasks to better understand their interplay when focusing on different emotional and pragmatics aspects. Also, participants' brain responses were recorded using fMRI and the data were analyzed to examine our fMRI hypotheses. Analyses focused on the time window from the onset of the target statement to the end of the evaluation phase, allowing us to capture both early, lower-level integration mechanisms and higher-order integration involved in social-pragmatic interpretation.

Behavioral Results

To test our first hypothesis, according to which an interaction between prosody and semantics existed across our tasks—with a reduced effect size in tasks where participants were instructed to focus exclusively on a single modality—we computed a linear mixed model including the interaction between prosody and semantics on participants' evaluations. For behavioral results, statistical modeling power is reported for fixed effects (R^2c) and for the combined fixed and random effects (R^2m). We found a significant interaction between semantics and prosody in the semantic task $F(1, 329.51) = 6.54$, $p = .011$, $R^2m = 0.76$, $R^2c = 0.85$; in the irony task $F(1, 403.43) = 4299.90$, $p < .001$, $R^2m = 0.65$, $R^2c = 0.72$; in the sarcasm task $F(1, 428.00) = 2501.43$, $p < .001$, $R^2m = 0.64$, R^2c = not computable due to singularity; in the ToM task $F(1, 419.56) = 600.64$, $p < .001$, $R^2m = 0.26$, $R^2c = 0.50$; but not in the prosody task. These findings are consistent with study piloting results (see Supplementary Material). Please refer to Table S6 for all the mixed model linear regression results, Table S7 for estimated marginal means from the interaction models, Fig. 2 for the plots of the predicted values of our interaction models to allow for clearer interpretation, and Fig. S4 for the plots of the actual values.

fMRI Results

To examine our second hypothesis, we conducted a flexible factorial second-level analysis and examined the main effect contrast Non-literal > Literal. Whole-brain analysis revealed the highest significant clusters of activation in the left IFG (Fig. 3A,E)—with peak activation in the the pars orbitalis (IFGorb), extending into the pars triangularis (IFGtri) and pars opercularis (IFGop). Further activations were found in the bilateral medial superior frontal gyrus (mSFG; Fig. 3B,D,F)—including the mPFC, in the bilateral supplementary motor area, in the right IFG (Fig. 3C,E)—predominantly in the IFGtri and IFGop with lesser involvement of the IFGorb, in the bilateral anterior insula (Fig.

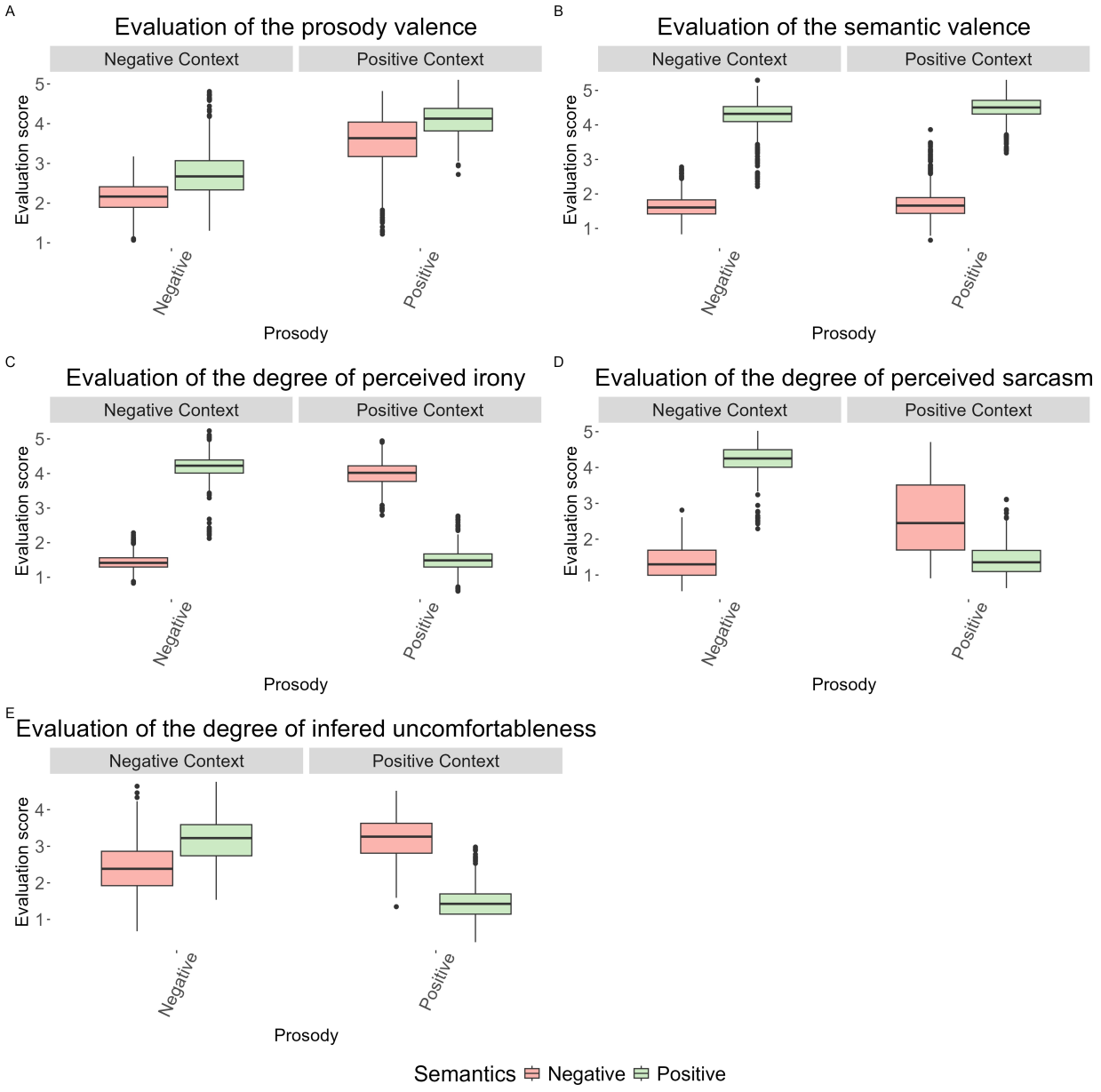

Figure 2. Plots of the predicted values of the interaction model between semantics and prosody in our A) prosody, B) semantic, C) irony, D) sarcasm, and E) ToM tasks. Each boxplot displays the distribution of predicted evaluation scores as a function of prosody (x-axis) and semantic valence (fill color). Green (positive) and red (negative) correspond to the two semantic conditions. Each point represents an individual data value. The boxplots show the median (horizontal line), interquartile range (box), and data spread (whiskers). Panels are faceted by the context (positive or negative).

3A,C)—with left hemisphere dominance, in the bilateral TPJ (Fig. 3A,C)—including the left AG and inferior parietal lobule, as well as the right AG, in the left PCun—extending to the posterior cingulate cortex (PCC; Fig. 3D,F), in the bilateral MTG (Fig. 3A,C)—extending to the STS in the left hemisphere, in the left middle cingulate cortex (MCC; Fig. 3D), and in the bilateral middle frontal gyrus (Fig. 3A,B,C). Please find the full peak MNI coordinates for this contrast in Table S8, corresponding brain activations in Fig. 3, and brain activations for the contrast Literal > Non-literal in Fig. S5.

To examine our third hypothesis, we then examined the flexible factorial model for the main effect contrast Integrative [Irony + Sarcasm + ToM] > Non-Integrative tasks [Prosody + Semantic]. Whole-brain analysis revealed the highest significant clusters of activation in the left TPJ (Fig. 4A)— including the AG and the posterior STS. Further activations were found in the bilateral mSFG (Fig. 4B,D,F)—mainly in the mPFC and extending to the anterior cingulate cortex (ACC), in the bilateral PCun (Fig. 4D,F)—extending to the MCC and PCC, in the right TPJ (Fig. 4C)—including the AG and the posterior STS, in the bilateral MTG (Fig. 4A,C)—extending into the STS and the left STG and temporal pole, in the bilateral precentral gyrus (Fig. 4B), in the bilateral amygdala (Fig. 4D,F), in the left cuneus; in the left IFGorb (Fig. 4A,E), in the right insula (Fig. 4C), in the left superior frontal gyrus (Fig. 4A,B), in the bilateral middle frontal gyrus (Fig. 4A,B,C), and in the right superior parietal lobule (Fig. 4B). Please find the full peak MNI coordinates for this contrast in Table S9, corresponding brain activations in Fig. 4, and brain activations for the contrast [Prosody + Semantic] > [Irony + Sarcasm + ToM] in Fig. S6.

To examine our fourth hypothesis, we then analyzed the flexible factorial model for the interaction contrast [Irony + Sarcasm + ToM] × Non-literal > [Prosody + Semantic] × Literal. Whole-brain analysis revealed no significant activation within our predicted regions, with a significant cluster emerging only in the calcarine sulcus (CS). Full peak MNI coordinates for this contrast are provided in Table S10.

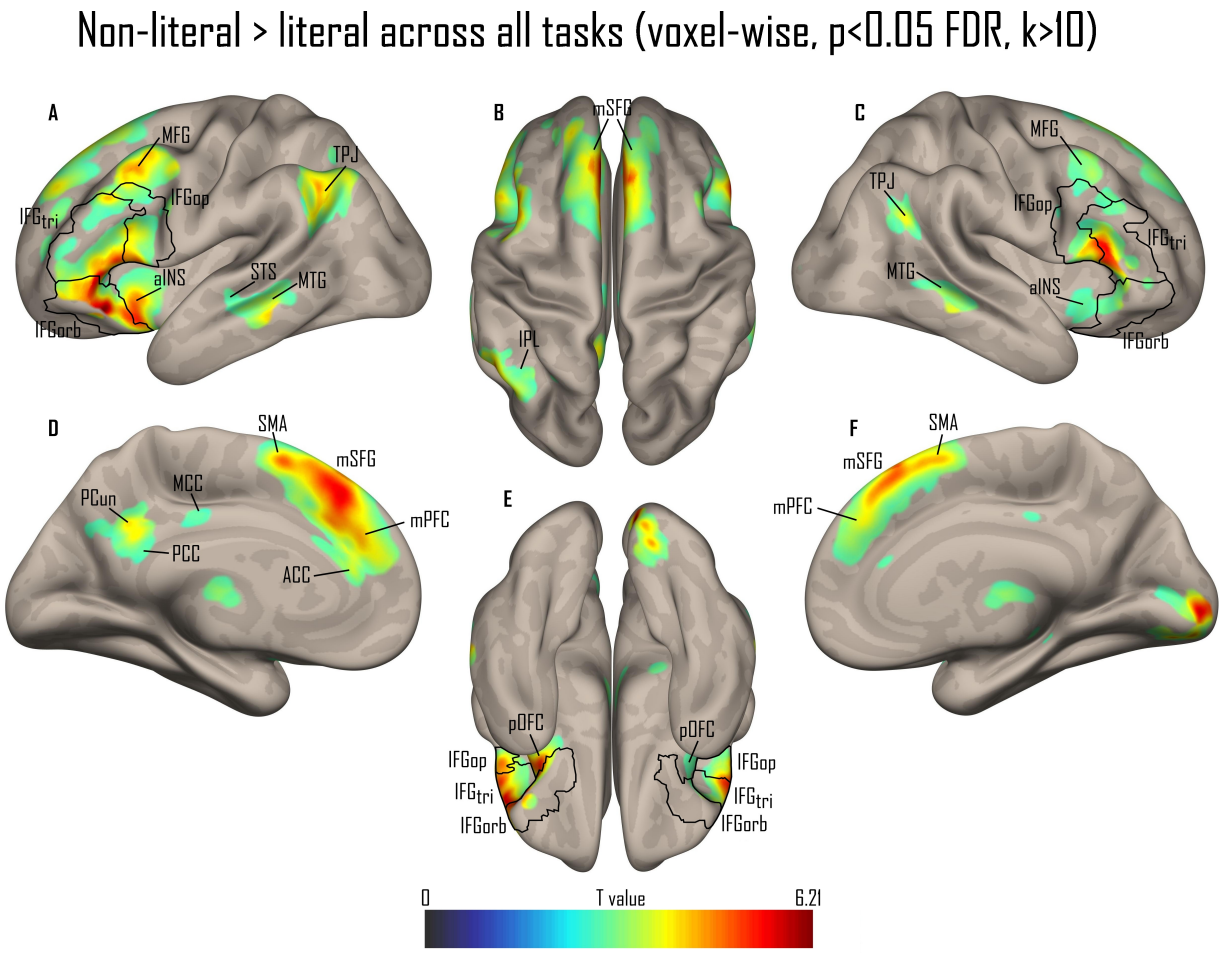

Figure 3. Whole-brain results of the contrast Non-literal > Literal across our different tasks in a sagittal A and C), medial D and F), superior B), and inferior E) view. All activations are thresholded at a voxelwise $p < 0.05$ FDR. The colorbar represents T statistics. IFG: inferior frontal gyrus; tri: pars triangularis; op: pars opercularis; orb: pars orbitalis; aINS: anterior insula; STS: superior temporal sulcus; MTG: mid temporal gyrus; TP: TPJ: temporo-parietal junction; IPL: inferior parietal lobule; PCun: precuneus; PCC: posterior cingulate cortex; MCC: middle cingulate cortex; ACC: anterior cingulate cortex; SMA: supplementary motor area; SFG: superior frontal gyrus; mPFC: medial prefrontal cortex; pOFC: posterior orbitofrontal cortex.

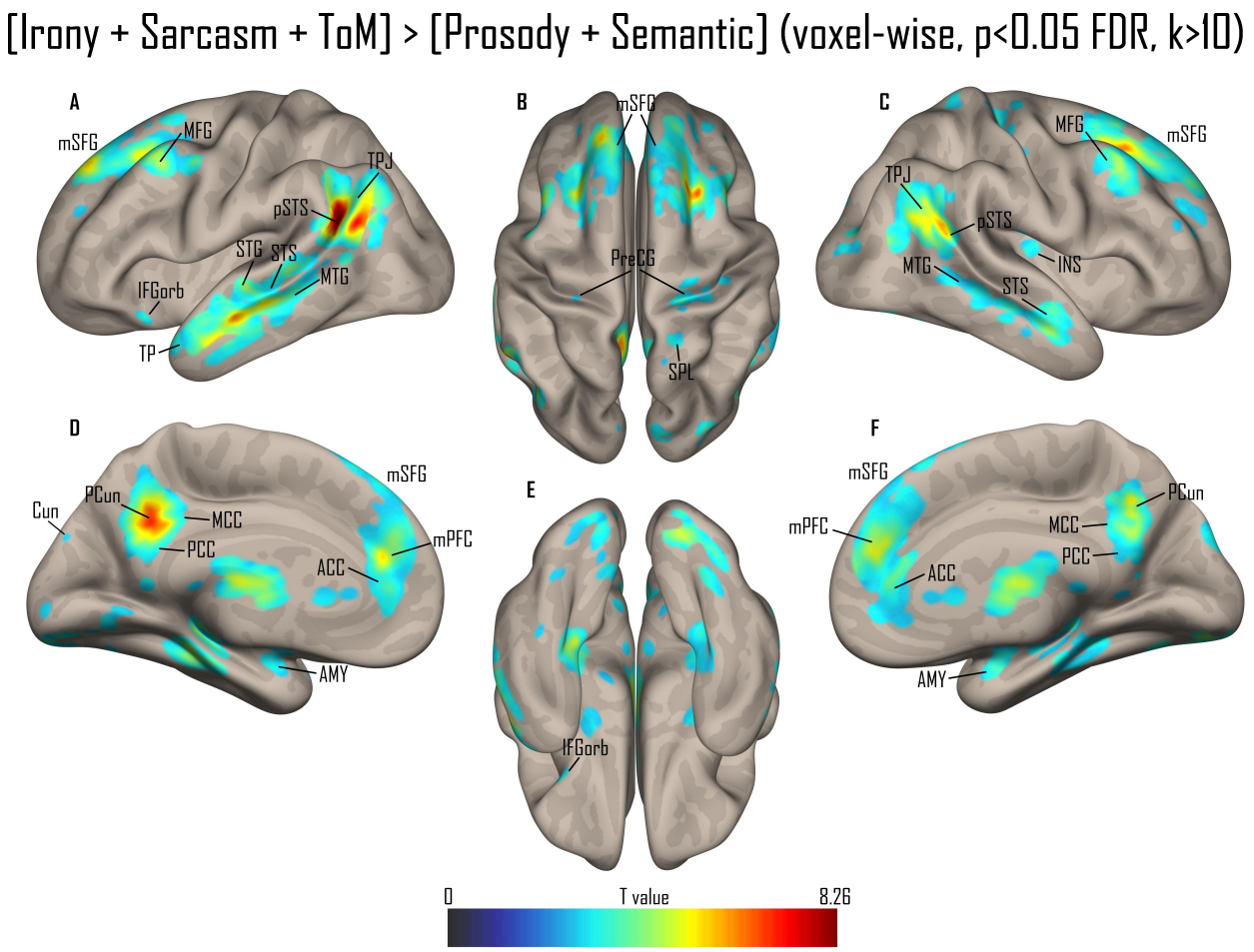

Figure 4. . Whole-brain results of the contrast [Irony + Sarcasm + ToM] > [Prosody + Semantic] in a sagittal A and C), medial D and F), superior B), and inferior E) view. All activations are thresholded at a voxelwise $p < 0.05$ FDR. IFG: inferior frontal gyrus; orb: pars orbitalis; INS: insula; STS: superior temporal sulcus; p : posterior; MTG: mid temporal gyrus; TP: temporal pole; TPJ: temporo-parietal junction; PCun: precuneus; Cun : cuneus; PCC: posterior cingulate cortex; MCC: middle cingulate cortex; ACC: anterior cingulate cortex; mSFG: medial superior frontal gyrus; MFG: middle frontal gyrus; mPFC: medial prefrontal cortex; AMY: amygdala.

Finally, to further investigate brain regions engaged during more demanding integrative processes, we compared Integrative versus Non-integrative tasks within the Non-literal condition. To this end, we conducted a follow-up analysis using the partial interaction contrast: ([Irony + Sarcasm + ToM] > [Prosody + Semantic]) × Non-Literal. Whole-brain analysis revealed the highest significant clusters of activation in the left TPJ (Fig. 5A)—most notably in the AG. Further activations were found in the left PCun—extending to the MCC and PCC (Fig. 5D), in the bilateral mPFC—extending to the ACC, in the bilateral anterior MTG (Fig. 5A,C)—showing large left-hemisphere dominance and extending into the STS, and in bilateral superior frontal gyrus (Fig. 5A,B,C). Please find the full peak MNI coordinates for this contrast in Table S11, corresponding brain activations in Fig. 5, brain activations for the contrast [Prosody + Semantic] > [Irony + Sarcasm + ToM] × Non-literal in Fig. S7, and brain activations for the contrast [Irony + Sarcasm + ToM] > [Prosody + Semantic] × Literal in Fig. S8.

General Discussion

The primary objective of our research was to investigate how prosody and semantics interact in non-literal forms of speech and to elucidate the brain mechanisms supporting their integration—both at the level of combining linguistic channels and in relation to higher-level cognitive processes, such as theory of mind. At the behavioral level, we observed significant interactions between prosodic and semantic factors across all tasks except the prosody task. As predicted, interactions were attenuated in the semantic task compared to integrative tasks (irony, sarcasm, ToM). This asymmetry suggests that while prosody can influence semantic judgments, semantics may not equivalently penetrate prosodic judgments, highlighting different weights of these cues in shaping emotional interpretation, pragmatic understanding, and social cognition. We then conducted whole-brain imaging analyses focusing on three contrasts designed to capture integration processes: (i) Non-literal > Literal speech, (ii) Integrative > Non-integrative tasks, and

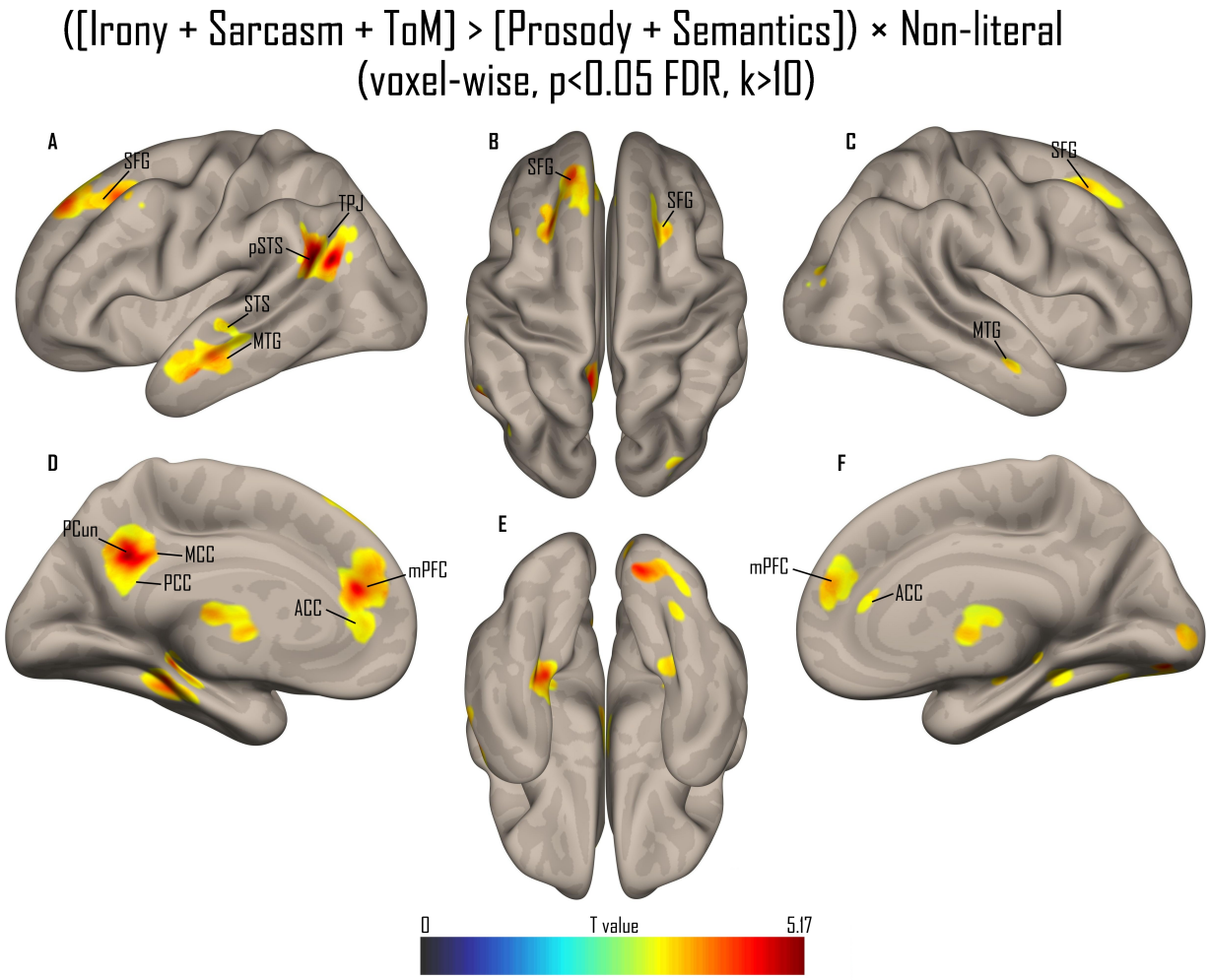

Figure 5. Whole-brain results of the contrast ([Irony + Sarcasm + ToM] > [Prosody + Semantic]) × Non-literal in a sagittal A and C), medial D and F), superior B), and inferior E) view. All activations are thresholded at a voxelwise $p < 0.05$ FDR. STS: superior temporal sulcus; p : posterior; MTG: mid temporal gyrus; TPJ: temporo-parietal junction; PCun: prenuceus; PCC: posterior cingulate cortex; MCC: middle cingulate cortex; ACC: anterior cingulate cortex; SFG: superior frontal gyrus; MFG: middle frontal gyrus; mPFC: medial prefrontal cortex.

(iii) their interaction. We hypothesized that these contrasts would index increasing demands for multimodal integration, with the strongest effects expected for the interaction, and that they would preferentially recruit the left IFG, the speech processing network, and the ToM network. Our hypothesis was partially confirmed for the Non-literal > Literal speech contrast, with significant activity in the bilateral IFG—showing a clear left dominance, the left MTG and STS, the right MTG, as well as the bilateral TPJ, mPFC, and left PCun. For the Integrative > Non-integrative tasks contrast, we again observed partial confirmation, with a small cluster in the left IFGorb, activity in the bilateral MTG, STS, and ATL, and further recruitment of the bilateral TPJ, PCun, and mPFC. By contrast, no predicted regions emerged for the interaction (Integrative × Non-literal > Integrative × Literal). To further investigate this effect, we therefore examined the partial interaction (Integrative > Non-integrative) × Non-literal, which revealed activity in the left MTG, STS, and ATL, smaller clusters in the right MTG, and additional activation in the left TPJ, left PCun, and bilateral mPFC. Overall, these findings suggest that while our predictions were only partially met, non-literal speech primarily engaged the bilateral IFG—with a left dominance, speech network, and ToM regions, whereas integrative tasks engaged more strongly the speech and the ToM network. We discuss and interpret these results in detail below.

In the current study, we investigated the interaction between prosody and semantics on participants' evaluations across multiple tasks. Our findings revealed significant interactions in tasks involving irony, sarcasm, and ToM, with a weaker interaction observed in the semantic task. No interaction was found in the prosody tasks. This result further replicates findings from Pilot Study 2 (see Supplementary Material), which demonstrated a similar pattern. The observation that participants instructed to attend semantic content only while ignoring prosodic cues still had their semantic evaluation modulated by perceived prosody valence, without the reverse effect, supports the prosody dominance hypothesis. For instance, studies using a cross-channel Stroop task with participants asked

to attend only one channel (i.e., prosody or semantics) and evaluate word emotional valence showed that prosody led to faster and more accurate responses than semantics (Filippi et al., 2017; Lin et al., 2020; Schirmer & Kotz, 2003). This effect persisted even in incongruent conditions but also showed less interference of semantics on prosody, highlighting the salience and relevance of prosodic cues (Schirmer & Kotz, 2003). Similarly, Ben-David et al. (2016) developed a new assessment tool that requires participants to attend only one channel to categorize discrete emotions, providing measures of both the dominance of the attended channel and the ability to selectively ignore the other one. Overall, their results demonstrate prosody dominance, also with prosodic cues proving more difficult to selectively ignore, which is in line with the present findings.

Our results show that left IFG activity is enhanced during non-literal versus literal speech, with peak activity in the left IFGorb. A small cluster in the left IFGorb, overlapping with this non-literal peak, was also activated in the contrast Integrative > Non-integrative tasks. Together, these findings suggest that the IFGorb may serve as a convergence hub where semantic and prosodic information are jointly integrated during non-literal language comprehension. This is in line with a large body of evidence showing the convergence of semantics and emotional cues processing—including prosody—within the IFGorb (Belyk et al., 2017). Moreover, the IFGorb has been specifically implicated in combining discourse context and utterance with prosody during sarcasm comprehension (Matsui et al., 2016). In contrast, the left IFG did not exhibit significant activation in either the interaction (Integrative × Non-literal > Integrative × Literal) or the partial interaction (Integrative > Non-integrative) × Non-literal. Unexpectedly, greater activation was observed in the posterior subregions of the IFG—specifically, the IFGtri and IFGop—for the contrast Non-integrative > Integrative tasks (see Supplementary Material). This pattern may be explained by a functional gradient within the left IFG, whereby the IFGop primarily supports lower-order phonological and syntactic processes, the IFGtri engages in syntactic and lower-order semantic processing, and the IFGorb underpins

higher-level semantic integration (Hagoort, 2005; Hagoort & Indefrey, 2014; Van Der Burght et al., 2023). In non-integrative tasks that isolate prosodic and semantic channels, participants must selectively attend to a single linguistic stream while suppressing others, a demand that can amplify activation in IFGop and IFGtri as the brain compensates for the absence of holistic integration by intensifying basic decoding and maintenance processes.

Furthermore, we observed activations within the speech processing network, including the bilateral STG, MTG, STS, and ATL (Friederici, 2011, 2012), with consistent activity across our contrasts in the MTG. This activation pattern fits with the broader functional architecture of the language network, in which temporal regions interface with frontal areas to support speech comprehension and the integration of multiple linguistic cues. Connectivity analyses have shown that the IFG is linked to superior and middle temporal regions involved in both speech comprehension and emotional prosody processing (Ethofer et al., 2012; Frühholz & Grandjean, 2012; Huang et al., 2024; Leitman et al., 2010; Turken et al., 2011; Yue et al., 2013). Such connectivity suggests that the IFG may coordinate information across modalities to integrate prosodic and semantic signals during sentence interpretation. In line with this view, the left IFGorb is recruited when successful performance requires monitoring multiple sources of information and resolving competing interpretations of sentence constituents (Turken et al., 2011). These functions likely depend on reciprocal interactions between the IFGorb and the MTG, supporting the selection and maintenance of appropriate lexical–semantic representations in short-term memory and their integration within the broader contextual framework. Extending this account, analogous IFGorb–MTG mechanisms may facilitate the integration of prosodic information with semantic context, thereby contributing to the comprehension of semantically ambiguous or non-literal speech. Further research is needed to probe the functional and directional nature of the IFGorb–MTG interaction and its role in integrating prosodic and semantic information during speech comprehension.

Consistent activation within the mentalizing network was observed across contrasts,

aligning with previous evidence of co-activation between speech and ToM networks during non-literal speech processing (Hauptman et al., 2023), suggesting that non-literal comprehension engages both core language mechanisms and social inference processes. More specifically, we observed important activation in the bilateral TPJ, primarily located in the angular gyrus and associated with the mentalizing network (Frith & Frith, 2006; J. P. Mitchell, 2009; Saxe & Baron-Cohen, 2006; Schurz et al., 2014; Tamir et al., 2016; Van Overwalle, 2009; Van Overwalle & Baetens, 2009). The enhanced activation of these regions can be interpreted as an effort to represent the others' intentions and goals based on integrated emotionally significant features. Furthermore, these activations may also indicate a broader integration process, particularly involving contextual cues, as evidenced by the role of the left angular gyrus in contextual integration (Branzi et al., 2021; Branzi & Lambon Ralph, 2023). Also, part of the mentalizing network, we observed activation in the mPFC thought to be involved in predictions about social mental states of others, and integrating information about the others' beliefs, intentions, and emotions to facilitate perspective-taking and social understanding (Frith & Frith, 2006; Van Overwalle, 2009). Importantly, the mPFC has been consistently reported in non-literal speech processing (Hauptman et al., 2023; Rapp et al., 2012). Finally, we noted activity in the precuneus, which has previously been linked to irony perception and ToM (Spotorno et al., 2012). This region is strongly associated with ToM and plays a key role in visuospatial mental imagery, motor imagery, mental rotation, route simulation, and perspective-taking (Schurz et al., 2014), suggesting its primary function is mental imagery for understanding the others' perspectives (Cavanna & Trimble, 2006). Further, the precuneus is also engaged during the processing of emotional prosody and affective semantics, particularly for angry speech, indicating its broader involvement in integrating emotional and self-referential aspects of language (Castelluccio et al., 2016).

Even though we did our best to control for potentially confounding factors and variables, several limitations should be acknowledged. In our behavioral and fMRI

analyses, the context and prosody factors were partially nested, as their levels co-varied across the four experimental conditions, thereby limiting the ability to disentangle their unique and independent contributions to the interaction between prosody and semantics. Furthermore, we designed five distinct tasks to target the different mechanisms involved in non-literal speech perception (i.e., prosodic and semantic decoding, irony and sarcasm inference, and theory-of-mind processes). This approach necessarily limited the number of trials per task and, consequently, the statistical power, as we sought to balance the number of tasks with experimental time constraints to minimize habituation and fatigue. We subsequently grouped these tasks across contrasts aimed at investigating different levels of integration between prosody and semantics. However, these design choices may have reduced the power to detect task-specific effects and constrained our ability to examine each task individually. Additionally, we included all voices across our experiment, which were randomly assigned to each trial to ensure the generalizability of our results. Although we included random intercepts for context and statement stimuli in our behavioral analyses, variability in actor performance may still have introduced variability across trials, thereby affecting the reliability and consistency of participants' evaluations and brain response. Lastly, our stimuli consisted of short sentences, which may pose challenges to the ecological validity of the study, as they do not accurately reflect the natural use of irony in everyday conversational contexts. This may have further limited the recruitment of ToM processes and associated brain networks (Spotorno et al., 2012).

Conclusion

In this study, we investigated the neural mechanisms underlying the integration of prosodic and semantic cues in non-literal forms of speech, such as irony and sarcasm. To this end, we designed five tasks targeting distinct cognitive processes, including prosody and semantics evaluation, irony and sarcasm inference, and theory of mind. At the behavioral level, participants' responses revealed systematic interactions between semantic

and prosodic factors across all tasks except the prosody-only condition, suggesting a prosody dominance effect. Our fMRI analyses examined several contrasts designed to probe different levels of integration: (1) Non-literal versus Literal speech, (2) Integrative versus Non-integrative tasks, and (3) their interaction. The non-literal contrast revealed activation in bilateral inferior frontal gyri, as well as in the broader speech and ToM networks. In contrast, the integrative task contrast showed more localized activation in the left IFGorb, alongside robust activity in the speech and ToM networks. The full interaction contrast did not yield significant effects in the predicted regions. To further explore higher-level integration, we analyzed the partial interaction between task type and speech type. This revealed reduced engagement of the speech network but sustained activation within the ToM network. Although our hypotheses were only partially supported across contrasts, main effect contrasts suggest that the left IFGorb may act as a critical hub for integrating prosodic and semantic cues, facilitating the relay of this integrated information to higher-order cognitive systems involved in ToM and decision-making. Future research should further clarify the dynamic interplay between the IFG, speech, and ToM networks—particularly through functional connectivity or electrophysiological approaches—to better understand how these systems cooperate during the comprehension of non-literal language.

Data and Code Availability

Our behavioral and fMRI data, along with the behavioral analysis script will be made available upon acceptance of the manuscript through a unique DOI on the open Yareta repository, which meets the FAIR guidelines for online data.

Author Contributions

W.A. conceived and conducted the second and third experiments, analyzed the data of all experiments, and wrote the manuscript. C.L. contributed to the design of the

experiments and data analysis, and reviewed the manuscript. M.A. created the stimuli and conceived and conducted the first pilot experiment. D.G. acquired the funding, contributed to the design of the experiments, and reviewed the manuscript.

Funding

This work was supported by funds from the NCCR Evolving Language, Swiss National Science Foundation Agreement #51NF40_180888.

Declaration of Competing Interests

The authors declare no conflicts of interest.

References

Ashburner, J., Barnes, G., Chen, C., Daunizeau, J., Flandin, G., Friston, K., Kiebel, S., Kilner, J., Litvak, V., Moran, R., et al. (2014). SPM12 manual. *Wellcome Trust Centre for Neuroimaging, London, UK*, *2464*(4). https://doi.org/10.1109/wincom55661.2022.9966463

Astésano, C., Besson, M., & Alter, K. (2004). Brain potentials during semantic and prosodic processing in french [Journal Article]. *Cognitive Brain Research*, *18*(2), 172–184. https://doi.org/10.1016/j.cogbrainres.2003.10.002

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models Usinglme4 [Journal Article]. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Beaucousin, V., Lacheret, A., Turbelin, M., Morel, M., Mazoyer, B., & Tzourio-Mazoyer, N. (2006). FMRI study of emotional speech comprehension [Journal Article]. *Cerebral Cortex*, *17*(2), 339–352. https://doi.org/10.1093/cercor/bhj151

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex [Journal Article]. *Nature*, *403*(6767), 309–312. https://doi.org/10.1038/35002078

Belyk, M., Brown, S., Lim, J., & Kotz, S. A. (2017). Convergence of semantics and emotional expression within the IFG pars orbitalis [Journal Article]. *NeuroImage*, *156*, 240–248. https://doi.org/10.1016/j.neuroimage.2017.04.020

Ben-David, B. M., Multani, N., Shakuf, V., Rudzicz, F., & Van Lieshout, P. H. H. M. (2016). Prosody and semantics are separate but not separable channels in the perception of emotional speech: Test for rating of emotions in speech [Journal Article]. *Journal of Speech, Language, and Hearing Research*, *59*(1), 72–89. https://doi.org/10.1044/2015_jslhr-h-14-0323

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies

[Journal Article]. *Cerebral Cortex*, *19*(12), 2767–2796. https://doi.org/10.1093/cercor/bhp055

Borod, J. C., Bloom, R. L., Brickman, A. M., Nakhutina, L., & Curko, E. A. (2002). Emotional processing deficits in individuals with unilateral brain damage [Journal Article]. *Applied Neuropsychology*, *9*(1), 23–36. https://doi.org/10.1207/s15324826an0901_4

Bosco, F. M., & Gabbatore, I. (2017). Sincere, deceitful, and ironic communicative acts and the role of the theory of mind in childhood [Journal Article]. *Frontiers in Psychology*, *08*. https://doi.org/10.3389/fpsyg.2017.00021

Bosco, F. M., Gabbatore, I., Angeleri, R., Zettin, M., & Parola, A. (2018). Do executive function and theory of mind predict pragmatic abilities following traumatic brain injury? An analysis of sincere, deceitful and ironic communicative acts [Journal Article]. *Journal of Communication Disorders*, *75*, 102–117. https://doi.org/10.1016/j.jcomdis.2018.05.002

Bosco, F. M., Parola, A., Valentini, M. C., & Morese, R. (2017). Neural correlates underlying the comprehension of deceitful and ironic communicative intentions [Journal Article]. *Cortex*, *94*, 73–86. https://doi.org/10.1016/j.cortex.2017.06.010

Branzi, F. M., & Lambon Ralph, M. A. (2023). Semantic-specific and domain-general mechanisms for integration and update of contextual information [Journal Article]. *Human Brain Mapping*, *44*(17), 5547–5566. https://doi.org/10.1002/hbm.26454

Branzi, F. M., Pobric, G., Jung, J., & Lambon Ralph, M. A. (2021). The left angular gyrus is causally involved in context-dependent integration and associative encoding during narrative reading [Journal Article]. *Journal of Cognitive Neuroscience*, *33*(6), 1082–1095. https://doi.org/10.1162/jocn_a_01698

Brück, C., Kreifelts, B., & Wildgruber, D. (2011). Emotional voices in context: A neurobiological model of multimodal affective information processing [Journal Article]. *Physics of Life Reviews*, *8*(4), 383–403. https://doi.org/10.1016/j.plrev.2011.10.002

Bryant, G. A. (2010). Prosodic contrasts in ironic speech [Journal Article]. *Discourse Processes*, *47*(7), 545–566. https://doi.org/10.1080/01638530903531972

Buchanan, T. W., Lutz, K., Mirzazade, S., Specht, K., Shah, N. J., Zilles, K., & Jäncke, L. (2000). Recognition of emotional prosody and verbal components of spoken language: An fMRI study [Journal Article]. *Cognitive Brain Research*, *9*(3), 227–238. https://doi.org/10.1016/s0926-6410(99)00060-9

Castelluccio, B. C., Myers, E. B., Schuh, J. M., & Eigsti, I. (2016). Neural substrates of processing anger in language: Contributions of prosody and semantics [Journal Article]. *Journal of Psycholinguistic Research*, *45*(6), 1359–1367. https://doi.org/10.1007/s10936-015-9405-z

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: A review of its functional anatomy and behavioural correlates [Journal Article]. *Brain*, *129*(3), 564–583. https://doi.org/10.1093/brain/awl004

Channon, S., Pellijeff, A., & Rule, A. (2005). Social cognition after head injury: Sarcasm and theory of mind [Journal Article]. *Brain and Language*, *93*(2), 123–134. https://doi.org/10.1007/978-1-4419-9422-6_15

Cheang, H. S., & Pell, M. D. (2008). The sound of sarcasm [Journal Article]. *Speech Communication*, *50*(5), 366–381. https://doi.org/10.1016/j.specom.2007.11.003

Collins, D. L., Neelin, P., Peters, T. M., & Evans, A. C. (1994). Automatic 3D intersubject registration of MR volumetric data in standardized talairach space [Journal Article]. *Journal of Computer Assisted Tomography*, *18*(2), 192–205. https://doi.org/10.1097/00004728-199403000-00005

Cope, T. E., Sohoglu, E., Peterson, K. A., Jones, P. S., Rua, C., Passamonti, L., Sedley, W., Post, B., Coebergh, J., Butler, C. R., Garrard, P., Abdel-Aziz, K., Husain, M., Griffiths, T. D., Patterson, K., Davis, M. H., & Rowe, J. B. (2023). Temporal lobe perceptual predictions for speech are instantiated in motor cortex and reconciled by inferior frontal cortex [Journal Article]. *Cell Reports*, *42*(5), 112422.

https://doi.org/10.1016/j.celrep.2023.112422

Cordonier, N., Fossard, M., & Champagne-Lavau, M. (2020). Differential impairments in irony comprehension in brain-damaged individuals: Insight from contextual processing, theory of mind, and executive functions [Journal Article]. *Neuropsychology*, *34*(7), 750–763. https://doi.org/10.1037/neu0000682

Ethofer, T., Anders, S., Erb, M., Herbert, C., Wiethoff, S., Kissler, J., Grodd, W., & Wildgruber, D. (2006). Cerebral pathways in processing of affective prosody: A dynamic causal modeling study [Journal Article]. *NeuroImage*, *30*(2), 580–587. https://doi.org/10.1016/j.neuroimage.2005.09.059

Ethofer, T., Bretscher, J., Gschwind, M., Kreifelts, B., Wildgruber, D., & Vuilleumier, P. (2012). Emotional voice areas: Anatomic location, functional properties, and structural connections revealed by combined fMRI/DTI [Journal Article]. *Cerebral Cortex*, *22*(1), 191–200. https://doi.org/10.1093/cercor/bhr113

Fabry, R. E. (2021). Getting it: A predictive processing approach to irony comprehension [Journal Article]. *Synthese*, *198*(7), 6455–6489. https://doi.org/10.1007/s11229-019-02470-9

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences [Journal Article]. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Filik, R., Ţurcan, A., Ralph-Nearman, C., & Pitiot, A. (2019). What is the difference between irony and sarcasm? An fMRI study [Journal Article]. *Cortex*, *115*, 112–122. https://doi.org/10.1016/j.cortex.2019.01.025

Filippi, P., Ocklenburg, S., Bowling, D. L., Heege, L., Güntürkün, O., Newen, A., & De Boer, B. (2017). More than words (and faces): Evidence for a stroop effect of prosody in emotion word processing [Journal Article]. *Cognition and Emotion*, *31*(5), 879–891. https://doi.org/10.1080/02699931.2016.1177489

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (Third). Sage. https://doi.org/10.32614/cran.package.car

Friederici, A. D. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: Localizing syntactic and semantic processes [Journal Article]. *Cerebral Cortex*, *13*(2), 170–177. https://doi.org/10.1093/cercor/13.2.170

Friederici, A. D. (2011). The brain basis of language processing: From structure to function [Journal Article]. *Physiological Reviews*, *91*(4), 1357–1392. https://doi.org/10.1152/physrev.00006.2011

Friederici, A. D. (2012). The cortical language circuit: From auditory perception to sentence comprehension [Journal Article]. *Trends in Cognitive Sciences*, *16*(5), 262–268. https://doi.org/10.1016/j.tics.2012.04.001

Friederici, A. D., Meyer, M., & Von Cramon, D. Y. (2000). Auditory language comprehension: An event-related fMRI study on the processing of syntactic and lexical information [Journal Article]. *Brain and Language*, *74*(2), 289–300. https://doi.org/10.1006/brln.2000.2313

Friston, K. (2005). A theory of cortical responses [Journal Article]. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Friston, K. (2010). The free-energy principle: A unified brain theory? [Journal Article]. *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing [Journal Article]. *Neuron*, *50*(4), 531–534. https://doi.org/10.1016/j.neuron.2006.05.001

Frühholz, S., & Belin, P. (2018). *The oxford handbook of voice perception* [Book]. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199982295.001.0001

Frühholz, S., & Grandjean, D. (2012). Towards a fronto-temporal neural network for the decoding of angry vocal expressions [Journal Article]. *NeuroImage*, *62*(3), 1658–1666. https://doi.org/10.1016/j.neuroimage.2012.06.015

Grandjean, D. (2021). Brain networks of emotional prosody processing [Journal Article]. *Emotion Review*, *13*(1), 34–43. https://doi.org/10.1177/1754073919898522

Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). The voices of wrath: Brain responses to angry prosody in meaningless speech [Journal Article]. *Nature Neuroscience*, *8*(2), 145–146. https://doi.org/10.1038/nn1392

Hagoort, P. (2005). On broca, brain, and binding: A new framework [Journal Article]. *Trends in Cognitive Sciences*, *9*(9), 416–423. https://doi.org/10.1016/j.tics.2005.07.004

Hagoort, P., & Indefrey, P. (2014). The neurobiology of language beyond single words [Journal Article]. *Annual Review of Neuroscience*, *37*(1), 347–362. https://doi.org/10.1146/annurev-neuro-071013-013847

Happé, F. G. E. (1993). Communicative competence and theory of mind in autism: A test of relevance theory [Journal Article]. *Cognition*, *48*(2), 101–119. https://doi.org/10.1016/0010-0277(93)90026-r

Hauptman, M., Blank, I., & Fedorenko, E. (2023). Non-literal language processing is jointly supported by the language and theory of mind networks: Evidence from a novel meta-analytic fMRI approach [Journal Article]. *Cortex*, *162*, 96–114. https://doi.org/10.1016/j.cortex.2023.01.013

He, Y., Shao, X., Liu, C., Fan, C., Jefferies, E., Zhang, M., & Li, X. (2024). Diverse frontoparietal connectivity supports semantic prediction and integration in sentence comprehension [Journal Article]. *The Journal of Neuroscience*, e1404242024. https://doi.org/10.1523/jneurosci.1404-24.2024

Hoffman, P., Binney, R. J., & Lambon Ralph, M. A. (2015). Differing contributions of inferior prefrontal and anterior temporal cortex to concrete and abstract conceptual knowledge [Journal Article]. *Cortex*, *63*, 250–266. https://doi.org/10.1016/j.cortex.2014.09.001

Huang, C., Li, A., Pang, Y., Yang, J., Zhang, J., Wu, X., & Mei, L. (2024). How the

intrinsic functional connectivity patterns of the semantic network support semantic processing [Journal Article]. *Brain Imaging and Behavior*, *18*(3), 539–554. https://doi.org/10.1007/s11682-024-00849-y

Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension [Journal Article]. *Journal of Cognitive Neuroscience*, *18*(4), 665–679. https://doi.org/10.1162/jocn.2006.18.4.665

Kotz, S. A., Dengler, R., & Wittfoth, M. (2015). Valence-specific conflict moderation in the dorso-medial PFC and the caudate head in emotional speech [Journal Article]. *Social Cognitive and Affective Neuroscience*, *10*(2), 165–171. https://doi.org/10.1093/scan/nsu021

Kotz, S. A., Meyer, M., & Paulmann, S. (2006). *Lateralization of emotional prosody in the brain: An overview and synopsis on the impact of study design* (pp. 285–294) [Book Section]. Elsevier. https://doi.org/10.1016/s0079-6123(06)56015-7

Kotz, S. A., & Paulmann, S. (2011). Emotion, language, and the brain [Journal Article]. *Language and Linguistics Compass*, *5*(3), 108–125. https://doi.org/10.1111/j.1749-818x.2010.00267.x

Leitman, D. I., Wolf, D. H., Ragland, J. D., Laukka, P., Loughead, J., Valdez, J. N., Javitt, D. C., Turetsky, B., & Gur, R. (2010). " it's not what you say, but how you say it": A reciprocal temporo-frontal network for affective prosody [Journal Article]. *Frontiers in Human Neuroscience*, 19. https://doi.org/10.4324/9781315230498-15

Lenth, R. V. (2024). *Emmeans: Estimated marginal means, aka least-squares means.* https://doi.org/10.1080/00031305.1980.10483031

Lin, Y., Ding, H., & Zhang, Y. (2020). Prosody dominates over semantics in emotion word processing: Evidence from cross-channel and cross-modal stroop effects [Journal Article]. *Journal of Speech, Language, and Hearing Research*, *63*(3), 896–912. https://doi.org/10.1044/2020_jslhr-19-00258

Lœvenbruck, H., Jannet, M. B., d'Imperio, M., Spini, M., & Champagne-Lavau, M. (2013). Prosodic cues of sarcastic speech in french: Slower, higher, wider. *Interspeech 2013 – 14th Annual Conference of the International Speech Communication Association*, 3537–3541. https://doi.org/10.21437/interspeech.2013-761

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139

Matchin, W., & Hickok, G. (2020). The cortical organization of syntax [Journal Article]. *Cerebral Cortex*, *30*(3), 1481–1498. https://doi.org/10.1093/cercor/bhz180

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 [Journal Article]. *Neuropsychologia*, *87*, 74–84. https://doi.org/10.1016/j.neuropsychologia.2016.04.031

Mauchand, M., Vergis, N., & Pell, M. D. (2020). Irony, prosody, and social impressions of affective stance. *Discourse Processes*, *57*(2), 141–157. https://doi.org/10.1080/0163853X.2019.1581588

Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., Salamon, G., Dehaene, S., Cohen, L., & Mehler, J. (1993). The cortical representation of speech [Journal Article]. *Journal of Cognitive Neuroscience*, *5*(4), 467–479. https://doi.org/10.1162/jocn.1993.5.4.467

Meyer, M., Alter, K., & Friederici, A. (2003). Functional MR imaging exposes differential brain responses to syntax and prosody during auditory sentence comprehension [Journal Article]. *Journal of Neurolinguistics*, *16*(4), 277–300. https://doi.org/10.1016/S0911-6044(03)00026-5

Mitchell, J. P. (2009). Inferences about mental states [Journal Article]. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1309–1316.

https://doi.org/10.1098/rstb.2008.0318

Mitchell, R. L. C. (2006). How does the brain mediate interpretation of incongruent auditory emotions? The neural response to prosody in the presence of conflicting lexico-semantic cues [Journal Article]. *European Journal of Neuroscience*, *24*(12), 3611–3618. https://doi.org/10.1111/j.1460-9568.2006.05231.x

Mitchell, R. L. C. (2007). FMRI delineation of working memory for emotional prosody in the brain: Commonalities with the lexico-semantic emotion network [Journal Article]. *NeuroImage*, *36*(3), 1015–1025. https://doi.org/10.1016/j.neuroimage.2007.03.016

Mitchell, R. L. C., & Phillips, L. H. (2015). The overlapping relationship between emotion perception and theory of mind [Journal Article]. *Neuropsychologia*, *70*, 1–10. https://doi.org/10.1016/j.neuropsychologia.2015.02.018

Mori, K., & Haruno, M. (2022). Resting functional connectivity of the left inferior frontal gyrus with the dorsomedial prefrontal cortex and temporoparietal junction reflects the social network size for active interactions [Journal Article]. *Human Brain Mapping*, *43*(9), 2869–2879. https://doi.org/10.1002/hbm.25822

Nakamura, T., Matsui, T., Utsumi, A., Sumiya, M., Nakagawa, E., & Sadato, N. (2022). Context-prosody interaction in sarcasm comprehension: A functional magnetic resonance imaging study [Journal Article]. *Neuropsychologia*, *170*, 108213. https://doi.org/10.1016/j.neuropsychologia.2022.108213

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 [Journal Article]. *PLOS ONE*, *11*(11), e0166704. https://doi.org/10.1371/journal.pone.0166704

Pauligk, S., Kotz, S. A., & Kanske, P. (2019). Differential impact of emotion on semantic processing of abstract and concrete words: ERP and fMRI evidence [Journal Article]. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-50755-3

Pell, M. D., Jaywant, A., Monetta, L., & Kotz, S. A. (2011). Emotional speech processing:

Disentangling the effects of prosody and semantic cues [Journal Article]. *Cognition & Emotion*, *25*(5), 834–853. https://doi.org/10.1080/02699931.2010.516915

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

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 [Journal Article]. *NeuroImage*, *63*(1), 600–610. https://doi.org/10.1016/j.neuroimage.2012.06.022

Rockwell, P. (2000). [Journal Article]. *Journal of Psycholinguistic Research*, *29*(5), 483–495. https://doi.org/10.1023/a:1005120109296

Rogalsky, C., & Hickok, G. (2009). Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex [Journal Article]. *Cerebral Cortex*, *19*(4), 786–796. https://doi.org/10.1093/cercor/bhn126

Ross, E. D., & Monnot, M. (2008). Neurology of affective prosody and its functional–anatomic organization in right hemisphere [Journal Article]. *Brain and Language*, *104*(1), 51–74. https://doi.org/10.1016/j.bandl.2007.04.007

Saban-Bezalel, R., Dolfin, D., Laor, N., & Mashal, N. (2019). Irony comprehension and mentalizing ability in children with and without autism spectrum disorder [Journal Article]. *Research in Autism Spectrum Disorders*, *58*, 30–38. https://doi.org/10.1016/j.rasd.2018.11.006

Saxe, R., & Baron-Cohen, S. (2006). Editorial: The neuroscience of theory of mind [Journal Article]. *Social Neuroscience*, *1*(3-4), 1–9. https://doi.org/10.1080/17470910601117463

Scharrer, L., & Christmann, U. (2011). Voice modulations in german ironic speech [Journal Article]. *Language and Speech*, *54*(4), 435–465. https://doi.org/10.1177/0023830911402608

Scherer, K. (2003). Vocal communication of emotion: A review of research paradigms [Journal Article]. *Speech Communication*, *40*(1-2), 227–256.

https://doi.org/10.1016/s0167-6393(02)00084-5

Schirmer, A., & Kotz, S. A. (2003). ERP evidence for a sex-specific stroop effect in emotional speech [Journal Article]. *Journal of Cognitive Neuroscience*, *15*(8), 1135–1148. https://doi.org/10.1162/089892903322598102

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing [Journal Article]. *Trends in Cognitive Sciences*, *10*(1), 24–30. https://doi.org/10.1016/j.tics.2005.11.009

Schirmer, A., Zysset, S., Kotz, S. A., & Yves Von Cramon, D. (2004). Gender differences in the activation of inferior frontal cortex during emotional speech perception [Journal Article]. *NeuroImage*, *21*(3), 1114–1123. https://doi.org/10.1016/j.neuroimage.2003.10.048

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies [Journal Article]. *Neuroscience & Biobehavioral Reviews*, *42*, 9–34. https://doi.org/10.1016/j.neubiorev.2014.01.009

Searle, J. R. (1979). *Expression and meaning* [Journal Article]. https://doi.org/10.1017/cbo9780511609213

Seydell-Greenwald, A., Chambers, C. E., Ferrara, K., & Newport, E. L. (2020). What you say versus how you say it: Comparing sentence comprehension and emotional prosody processing using fMRI [Journal Article]. *NeuroImage*, *209*, 116509. https://doi.org/10.1016/j.neuroimage.2019.116509

Shamay-Tsoory, S. G., Tomer, R., Berger, B. D., Goldsher, D., & Aharon-Peretz, J. (2005). Impaired "affective theory of mind" is associated with right ventromedial prefrontal damage [Journal Article]. *Cognitive and Behavioral Neurology*, *18*(1). https://doi.org/10.1080/17470910600985589

Shibata, M., Toyomura, A., Itoh, H., & Abe, J. (2010). Neural substrates of irony comprehension: A functional MRI study [Journal Article]. *Brain Research*, *1308*,

114–123. https://doi.org/10.1016/j.brainres.2009.10.030

Song, Y., Nie, Z., & Shan, J. (2024). Comprehension of irony in autistic children: <Scp>the</scp> role of theory of mind and executive function [Journal Article]. *Autism Research*, *17*(1), 109–124. https://doi.org/10.1002/aur.3051

Spotorno, N., Koun, E., Prado, J., Van Der Henst, J., & Noveck, I. A. (2012). Neural evidence that utterance-processing entails mentalizing: The case of irony [Journal Article]. *NeuroImage*, *63*(1), 25–39. https://doi.org/10.1016/j.neuroimage.2012.06.046

Tamir, D. I., Thornton, M. A., Contreras, J. M., & Mitchell, J. P. (2016). Neural evidence that three dimensions organize mental state representation: Rationality, social impact, and valence [Journal Article]. *Proceedings of the National Academy of Sciences*, *113*(1), 194–199. https://doi.org/10.1073/pnas.1511905112

Tettamanti, M., Vaghi, M. M., Bara, B. G., Cappa, S. F., Enrici, I., & Adenzato, M. (2017). Effective connectivity gateways to the theory of mind network in processing communicative intention [Journal Article]. *NeuroImage*, *155*, 169–176. https://doi.org/10.1016/j.neuroimage.2017.04.050

Turken, U, & Dronkers, N. F. (2011). The neural architecture of the language comprehension network: Converging evidence from lesion and connectivity analyses [Journal Article]. *Frontiers in System Neuroscience*, *5*. https://doi.org/10.3389/fnsys.2011.00001

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 [Journal Article]. *Brain Research*, *1124*(1), 100–110. https://doi.org/10.1016/j.brainres.2006.09.088

Van Der Burght, C. L., Numssen, O., Schlaak, B., Goucha, T., & Hartwigsen, G. (2023). Differential contributions of inferior frontal gyrus subregions to sentence processing guided by intonation [Journal Article]. *Human Brain Mapping*, *44*(2), 585–598. https://doi.org/10.1002/hbm.26086

Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis [Journal Article]. *Human Brain Mapping*, *30*(3), 829–858. https://doi.org/10.1002/hbm.20547

Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis [Journal Article]. *NeuroImage*, *48*(3), 564–584. https://doi.org/10.1016/j.neuroimage.2009.06.009

Vandenberghe, R., Nobre, A. C., & Price, C. J. (2002). The response of left temporal cortex to sentences [Journal Article]. *Journal of Cognitive Neuroscience*, *14*(4), 550–560. https://doi.org/10.1162/08989290260045800

Westerlund, M., & Pylkkänen, L. (2014). The role of the left anterior temporal lobe in semantic composition vs. Semantic memory [Journal Article]. *Neuropsychologia*, *57*, 59–70. https://doi.org/10.1016/j.neuropsychologia.2014.03.001

Whitfield-Gabrieli, S., & Nieto-Castanon, A. (2012). Conn: A functional connectivity toolbox for correlated and anticorrelated brain networks [Journal Article]. *Brain Connectivity*, *2*(3), 125–141. https://doi.org/10.1089/brain.2012.0073

Wiethoff, S., Wildgruber, D., Kreifelts, B., Becker, H., Herbert, C., Grodd, W., & Ethofer, T. (2008). Cerebral processing of emotional prosody—influence of acoustic parameters and arousal [Journal Article]. *NeuroImage*, *39*(2), 885–893. https://doi.org/10.1016/j.neuroimage.2007.09.028

Wildgruber, D., Ackermann, H., Kreifelts, B., & Ethofer, T. (2006). Cerebral processing of linguistic and emotional prosody: fMRI studies [Book Section]. In S. Anders, G. Ende, M. Junghofer, J. Kissler, & D. Wildgruber (Eds.), *Progress in brain research* (Vol. 156, pp. 249–268). Elsevier. https://doi.org/10.1016/S0079-6123(06)56013-3

Wildgruber, D., Ethofer, T., Grandjean, D., & Kreifelts, B. (2009). A cerebral network model of speech prosody comprehension [Journal Article]. *International Journal of Speech-Language Pathology*, *11*(4), 277–281. https://doi.org/10.1080/17549500902943043

Witteman, J., Ijzendoorn, M. H. van, Velde, D. van de, Heuven, V. J. van, & Schiller, N. O.

(2011). The nature of hemispheric specialization for linguistic and emotional prosodic perception: A meta-analysis of the lesion literature [Journal Article]. *Neuropsychologia*, *49*(13), 3722–3738. https://doi.org/10.1016/j.neuropsychologia.2011.09.028

Wittfoth, M., Schröder, C., Schardt, D. M., Dengler, R., Heinze, H., & Kotz, S. A. (2009). On emotional conflict: Interference resolution of happy and angry prosody reveals valence-specific effects [Journal Article]. *Cerebral Cortex*, *20*(2), 383–392. https://doi.org/10.1093/cercor/bhp106

Yue, Q., Zhang, L., Xu, G., Shu, H., & Li, P. (2013). Task-modulated activation and functional connectivity of the temporal and frontal areas during speech comprehension [Journal Article]. *Neuroscience*, *237*, 87–95. https://doi.org/10.1016/j.neuroscience.2012.12.067

Zhu, N., & Wang, Z. (2020). The paradox of sarcasm: Theory of mind and sarcasm use in adults [Journal Article]. *Personality and Individual Differences*, *163*, 110035. https://doi.org/10.1016/j.bandl.2004.09.002