Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

Research Report

Neural correlates of inferring speaker sincerity from white lies: An event-related potential source localization study

Simon Rigoulot[a,b,*], Karyn Fish[a,b], Marc D. Pell[a,b]

[a]McGill University, Faculty of Medicine, School of Communication Sciences and Disorders, 1266 Avenue des Pins Ouest, Montreal QC, Canada H3G 1A8

[b]McGill Centre for Research on Brain, Language and Music (CRBLM), Canada

ARTICLE INFO

Article history:
Accepted 12 April 2014
Available online 18 April 2014

Keywords:
Speech processing
Deception
Prosody
ERP
P600
Insula

ABSTRACT

During social interactions, listeners weigh the importance of linguistic and extra-linguistic speech cues (prosody) to infer the true intentions of the speaker in reference to what is actually said. In this study, we investigated what brain processes allow listeners to detect when a spoken compliment is meant to be sincere (true compliment) or not ("white lie"). Electroencephalograms of 29 participants were recorded while they listened to Question–Response pairs, where the response was expressed in either a sincere or insincere tone (e.g., "So, what did you think of my presentation?"/"I found it really interesting."). Participants judged whether the response was sincere or not. Behavioral results showed that prosody could be effectively used to discern the intended sincerity of compliments. Analysis of temporal and spatial characteristics of event-related potentials (P200, N400, P600) uncovered significant effects of prosody on P600 amplitudes, which were greater in response to sincere versus insincere compliments. Using low resolution brain electromagnetic tomography (LORETA), we determined that the anatomical sources of this activity were likely located in the (left) insula, consistent with previous reports of insular activity in the perception of lies and concealments. These data extend knowledge of the neurocognitive mechanisms that permit context-appropriate inferences about speaker feelings and intentions during interpersonal communication.

1. Introduction

Even the most routine conversation between friends can be socially complex; often, what a speaker says is not meant to be understood literally by the listener (e.g., in the case of irony). At other times, the intended literal meaning of an utterance does not accurately reflect the speaker's true feelings or beliefs (e.g., in the case of lies). To uncover the underlying speaker meaning, whether intended or hidden, listeners must attend to contextual cues that accompany the

*Corresponding author at: McGill University, Faculty of Medicine, School of Communication Sciences and Disorders, 1266 Avenue des Pins Ouest, Montreal (QC), Canada H3G 1A8.
E-mail address: simon.rigoulot@mail.mcgill.ca (S. Rigoulot).
URL: http://www.mcgill.ca/pell_lab (M.D. Pell).

http://dx.doi.org/10.1016/j.brainres.2014.04.022
0006-8993/

linguistic message to infer the relevance of the utterance in relation to other available input (Sperber and Wilson, 1995). For example, during a face-to-face conversation, meanings assigned to the linguistic code are simultaneously integrated with other cues provided (consciously or accidently) by the speaker, such as facial expressions, body movements, and tone of voice (speech prosody); weighing these different sources of evidence allows the listener to develop a valid hypothesis about the *intentions* of the speaker, and in some situations, to detect when the speaker's true opinion actually differs from what is said: see Sperber and Wilson (2002) for an overview. In cognitive terms, the process of bridging literal utterance meanings and implied speaker meanings appears to be relatively effortful (Noveck and Reboul, 2008). A clear link between pragmatic language interpretation and mentalizing, or 'theory of mind', has also been made (e.g., Spotorno et al., 2012), underlining the complexities of determining 'speaker meaning' and social intentions in speech.

Among the different sources of evidence that reveal a speaker's meaning, prosody often plays a critical role in utterance interpretation. Prosody refers to perceived differences in the temporal/durational properties of speech, global and local fluctuations in pitch and intensity, changes in energy distribution and the formant structure of vowels, and other parameters (Banse and Scherer, 1996; Juslin and Laukka, 2003). Depending on how these cues are used, prosody serves a range of linguistic, emotional, and 'pragmatic' functions that are directly relevant to the listener (Wilson and Wharton, 2006). Of special interest, data show that prosody is important for listeners to process ambiguous words or expressions (Snedeker and Trueswell, 2003) and to detect when a speaker intends to communicate non-literal and figurative meanings of language (e.g., metaphors, idioms, and expressions of irony/sarcasm). For example, idioms are culturally-defined expressions that convey a figurative meaning beyond the literal meaning of the words (e.g., in English: *I bought a lemon*); there is evidence that the prosodic form of idioms is distinct when speakers mean to convey the literal meaning of these expressions (the fruit) versus the idiomatic meaning (an object that is defective), and that these subtle cues can be effectively harnessed by native listeners to differentiate which meaning is intended (Ashby, 2006; Bélanger et al., 2009; Van Lancker-Sidtis, 2003).

Another form of non-literal communication where speakers intend a different, often opposite, meaning to what is literally stated is the case of ironic or sarcastic comments: see Gibbs (1994). If someone listening to a scientific talk being delivered in a hesitant, disorganized manner says to their neighbor, "*Well, this is going smoothly*", the listener will use prosody (among other cues) to infer that a negative, critical attitude is meant by the speaker, despite the fact that the literal message is overtly positive. Acoustic studies report systematic differences in the acoustic profile of otherwise identical utterances produced in an ironic versus literal manner (Bryant, 2010; Bryant and Fox Tree, 2005; Cheang and Pell, 2008, 2009). Ironic utterances are produced slower and with characteristic changes in the speaker's pitch register when compared to their literal variants (Bryant, 2010; Cheang and Pell, 2008), although pitch conventions for signaling irony seem to vary across languages (Cheang and Pell, 2009). Nonetheless, these findings emphasize the importance of prosodic information in the process of inferring sarcastic attitudes, and probably other emotive and interpersonal meanings held by the speaker toward the linguistic message and/or the listener.

Recently, the processes that underlie irony comprehension have been explored in the neurocognitive literature to illuminate how the brain processes speaker intentions. Functional magnetic resonance imaging (fMRI) studies have identified several different brain regions that may be involved during the comprehension of irony, including regions in the temporal lobe, medial prefrontal cortex, and subcortical areas such as the insula or amygdala (Eviatar and Just, 2006; Shibata, et al., 2010; Spotorno, et al., 2012; see also the reviews of Rapp et al. (2012), and Bohrn, et al. (2012) with irony and other forms of figurative language). Moreover, recent studies focusing on event-related potentials (ERPs) have shed light on the cognitive processing structure and time course for interpreting speaker meanings in the context of irony/sarcasm (Regel et al., 2010, 2011). These authors presented sentences such as, "*The game was fantastic*" to German participants, embedded in two types of discourse contexts that biased either an ironic or a literal utterance meaning. Biasing contexts were always presented visually, whereas target sentences were presented either visually (Regel et al., 2010, 2011, Expt 2) or auditorily (Regel et al., 2011, Expt. 1). At the end of each trial, participants were presented a statement and had to judge whether it was true or not by pressing buttons labeled "yes" or "no". Data were analyzed in reference to several ERP components, including those that are widely believed to index on-line difficulties in semantic and syntactic integration processes (N400, P600; Cornejo et al., 2007; Coulson and Van Petten, 2007; Friederici, et al., 2002; Van Berkum et al., 2003).

Results showed that sentences with an ironic meaning elicited early ERP effects (LAN and P200) and significantly larger P600 amplitudes when compared to literal sentences; effects on the P600 component were taken as evidence that additional processing is necessary to interpret ironic meanings in language (Regel et al., 2010, 2011). In contrast, there was no effect of irony on the N400 ERP component, suggesting that a semantic mismatch is not produced when integrating literal sentence meanings with contextual information that signals the opposite, intended speaker meaning: see also Noveck and Posada (2003). Together, these results argue that additional time and cognitive resources are needed to process utterances whose underlying meaning is non-literal or inferred, perhaps due to increased cognitive efforts at the stage of integrating semantic and extra-semantic or contextual information (Regel et al., 2010, 2011). This preliminary work encourages further research that looks at how contextual cues, especially prosody, direct listeners to interpret non-literal or intended speaker meanings that are not literally present in an utterance.

A ubiquitous situation where a speaker's true feelings about the utterance meaning may not be immediately apparent to the listener is when speakers tell a 'white lie'. While many different researchers have studied how "high-stakes" lies are communicated and detected socially (e.g., Vrij and Mann, 2001), only a small proportion of lies told in daily life are high-stakes or fact-related; much more frequently, speakers make "small" socially-motivated lies about feelings, preferences, or opinions to reap psychological rewards such as respect and affection, and to avoid hurt feelings or embarrassing a social partner (DePaulo et al., 1996, 2003). White lies are perceived as so innocuous that they are often grouped with compliments and courtesies (Ekman and O'Sullivan, 2006) or politeness behavior (Talwar et al., 2007). Despite the prevalence of white lies in human social interactions, there is little, if any, research investigating how the brain responds to utterances that represent “white lies”, where listeners infer that the speaker's true feelings or opinions diverge from what is literally said.

In contrast to irony, where the speaker's aim is to indirectly guide the listener to the speaker's intended meaning, when speakers use deception their intended goal is to *conceal* speaker meanings/attitudes about the literal message. When the utterance content matches the speaker's true beliefs, emotions, and/or attitudes, it may be described as ‘sincere’; alternately, when the speaker conveys ideas that differ from his or her true beliefs, it may be described as insincere (and sometimes, a white lie). Researchers have characterized a number of nonverbal cues that tend to be associated with deception, such as subtle changes of expression, pitch of the voice, and body posture (Frank and Ekman 1997; Vrij, 1994; Zuckerman et al., 1979). Although much of the deception research focuses on high-stakes lying and associated forms of physiological arousal—which likely account for many of the verbal and non-verbal indicators of deception—the small lies which are part of normal daily interactions are not typically associated with these effects (DePaulo et al., 2003).

This leaves unanswered whether listeners are able to detect the insincerity of white lies during speech comprehension, and if they can, how they distinguish sincere from insincere utterances based on nonverbal cues such as prosody. This question is meaningful because nonverbal cues are often difficult to control and speakers frequently allow meanings to “leak out” (Ekman and Friesen, 1969), unwittingly conveying their true and opposing attitudes. Recent data indicate that when speakers utter a compliment that is sincere (i.e., represents their true opinions) versus insincere (telling white lies to spare listeners' feelings), utterances perceived as sincere by listeners demonstrated significant acoustic differences in both pitch and intensity from insincere utterances (Fish and Pell, 2011; Villar et al., 2013). Also, there is some evidence that statements about true versus false opinions may be spoken more quickly (Fish and Pell, 2011; Spence et al., 2012). These data confirm that prosody plays a key role in disambiguating sincere from insincere utterances and could help listeners generate inferences about speaker sincerity, in an analogous manner to using prosody to interpret irony and other non-literal meanings.

Building on previous studies of how prosody is used to infer speaker intentions (Noveck and Posada, 2003; Regel et al., 2010, 2011), the goal of this study was to understand how listeners use speech cues to infer speaker sincerity, and to apprehend the underlying cortical activity by gathering data on the temporal processing sequence and source localization of ERPs elicited by sincere versus insincere compliments. We presented Question–Response utterance pairs resembling a casual discourse between friends that always ended in a compliment spoken with a sincere versus insincere prosody (Fish and Pell, 2011). Experimental participants were required to judge the sincerity of the target utterance while the electroencephalogram (EEG) was recorded. Behaviorally, we hypothesized that sincere utterances would be easier to identify than insincere (deceptive) utterances owing to a ‘truth bias’ in sentence processing: see Bond and DePaulo (2006). For the electrophysiological data, we focused on three ERP components identified by previous work (P200, N400, P600). Temporally, we expected that late components (P600), but not the N400, would be sensitive to the sincerity status of target utterances due to the added cognitive demands of inferring speaker attitudes in relation to the utterance, similar to what Regel et al. (2010) observed for irony. Although the P200 is known to be sensitive to the affective salience of acoustic cues, among other features (Paulmann and Kotz, 2008), no clear prediction could be made about the influence of sincerity status on the P200 based on the related literature. In terms of spatial localization, on the basis of fMRI findings that revealed that the frontal lobe and subcortical structures, in particular amygdala and insula, are involved in the production and the detection of lies (e.g., Ganis et al., 2009), we expected that the effects observed in late temporal windows would originate from a differential activity within these structures in response to sincere and insincere compliments.

2. Results

2.1. Behavioral data

As shown in Fig. 1, the participants' ability to judge the sincerity of speakers who uttered the compliment was high overall ($M=82.8\%$ correct responses). There was a significant influence of sincerity condition on accuracy, with more accurate detection of sincere rather than insincere responses (87.0% versus 78.5%; $F(1,24)=11.148$; $p=0.003$). The performance of participants did not differ as a function of the duration of the inter-stimulus interval (82.5% versus 83.0%; $F(1,24)=0.139$; $p=0.713$) and there was no significant interaction between the two factors ($F(1,24)=0.352$; $p=0.558$). As there was no evidence that changes in the inter-stimulus

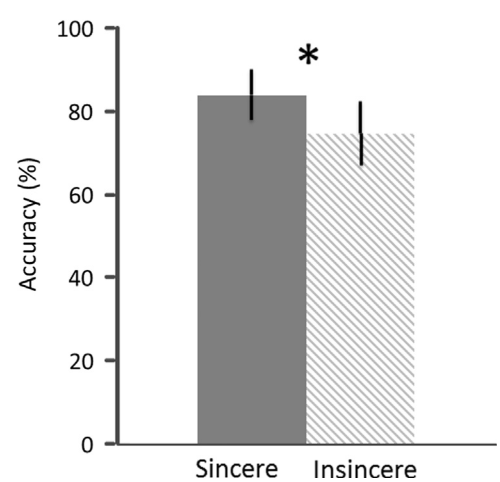

Fig. 1 – Behavioral performance as a function of sentence sincerity (error bars refer to SEM; *: $p<0.05$).

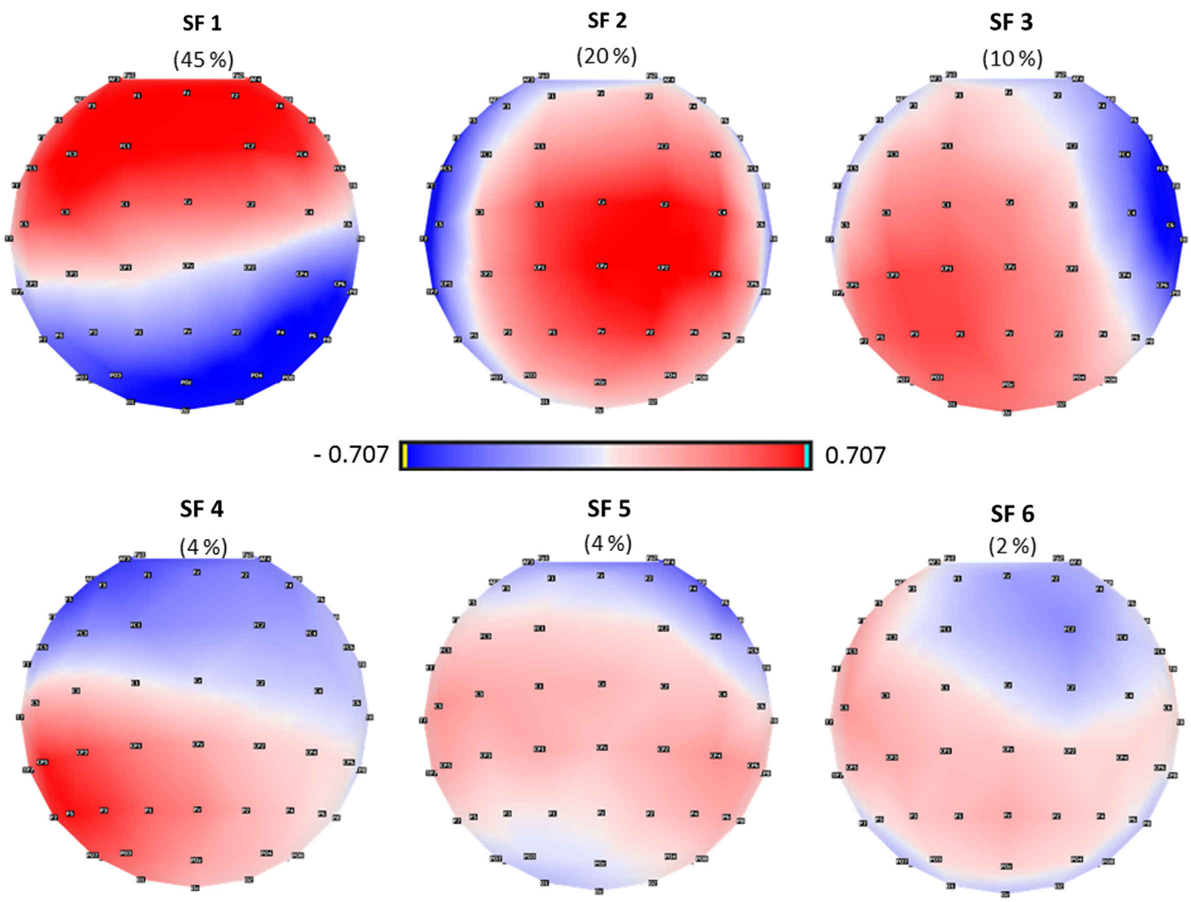

Fig. 2 – Topographic maps of the spatial factor loadings (virtual electrodes). Percentage of variance accounted for by each factor is shown in parentheses.

interval influenced judgements about speaker meaning, this factor was removed from the analysis of the EEG data.

2.2. *EEG data*

After collapsing the data from the two ISI presentations and the exclusion of artefactual trials, 35 insincere and 37 sincere utterances on average for each participant were used for EEG analyses. Peak amplitudes (maximum of amplitude for electrodes within each region of interest, ROI) were analyzed through MANOVA with repeated measures of sincerity (sincere, insincere) and ROI. First, the spatial PCA determined six spatial factors explaining 86% of data variance: see Fig. 2. We then determined groups of electrodes of interest by clustering the electrodes that were accounting for more than 50% of the variance of the data pertaining to each factor. Five groups of electrodes were found, each defining a region of interest: the first ROI was located in the fronto-central region (associated with electrodes F1, F2, F3, F4, Fz, FC1, FC2, FC3); the second ROI was distributed in the right parieto-occipital lobe (electrodes P4, P6, P8, PO4, PO8, PO10, Oz, O2); the third ROI was located in the center of the head (Cz, C2, CPz, CP2); the fourth ROI was in the left temporal areas (FT7, FT9, T7, TP7, TP9); and the fifth ROI was located in the right frontal area (F8, FC6, FT8, FT10, C6, T8). In the second step of the analysis, examination of the grand average of the data served to define three temporal windows encompassing the three components of interest. The first temporal window was between 150 and 300 ms after onset of the response (P200), the second between 300 and 500 ms (N400), and the third between 600 and 900 ms (P600). These time windows are equivalent to those used in previous studies of comparative interest (e.g., Regel et al., 2011).

(a) P200 – For the P200 component, the analysis revealed a significant main effect of ROI ($F(4,21)=19.612$; $p<0.001$). This effect was explained by significantly higher amplitudes in the first and fourth ROIs ($ps<0.015$), thus distributed in the fronto-central regions (there was no difference between the first and fourth ROIs, $p=0.605$). As shown in Fig. 3, the P200 analysis yielded no main effect of sincerity ($F(1,24)=2.660$; $p=0.116$, *ns*) and the interaction of ROI and sincerity was also non-significant ($F(4,21)=0.455$; $p=0.768$, *ns*).[1]

[1]Given that early components are sensitive to acoustic characteristics of complex sentences (Pinheiro et al., 2013), we also conducted classical analyses on N100, focusing on peak amplitudes and latency in response to sincere and insincere utterances at Fz electrode in 50–150 ms temporal window. There was no effect of sincerity neither on peak amplitude ($F(1,24)=0.541$; $p=0.469$) nor peak latency ($F(1,24)=0.419$; $p=0.524$). We ran a similar analysis for P200, without taking into account the results of spatial PCA, since P200 distribution can be more local. Consequently, we analyzed peak amplitude and latency at Cz electrode

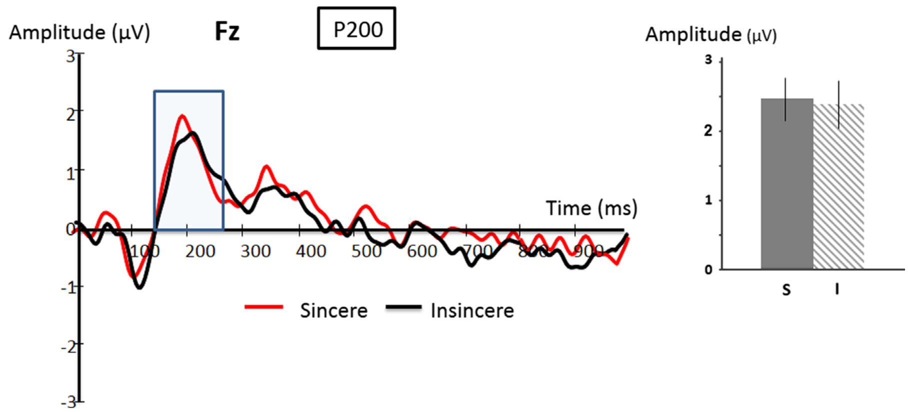

Fig. 3 – P200. Grand-average ERPs ($n=25$ participants) and mean amplitude of P200 at Fz elicited by sincere (blue line) and insincere (red line) sentences (S: sincere; I: insincere).

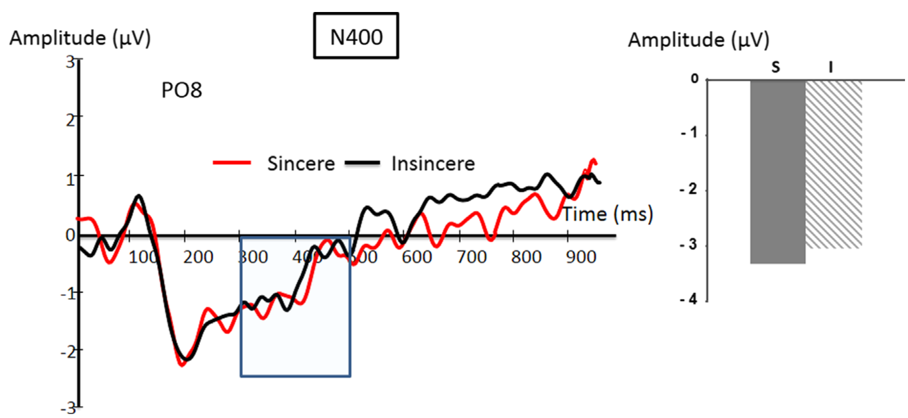

Fig. 4 – N400. Grand-average ERPs ($n=25$ participants) and mean amplitude of N400 at PO8 elicited by sincere (blue line) and insincere (red line) sentences (S: sincere; I: insincere).

(b) N400 – Analysis of the N400 component revealed a similar pattern of effects, characterized by a significant main effect of ROI ($F(4,21)=26.481$; $p<0.001$). Overall, N400 amplitudes were lowest (the most negative) in right parieto-occipital areas ($ps<0.001$). The other ROIs did not differ from each other ($ps>0.111$), except for the left fronto-temporal ROI that was significantly less negative than all others ($ps<0.003$). There was no significant effect of utterance sincerity ($F(1,24)=1.523$; $p=0.229$, *ns*; Fig. 4) or interaction of ROI and sincerity ($F(4,21)=0.827$; $p=0.523$, *ns*) on N400 amplitudes.

(c) P600 – Analysis of the P600 component produced a significant main effect of ROI ($F(4,21)=7.462$; $p=0.001$). While there was no significant main effect of utterance sincerity ($F(1,24)=1.310$; $p=0.264$, *ns*), the interaction of ROI and utterance sincerity approached significance ($F(4,21)=2.711$; $p=0.058$). The interaction allowed us to explore the effect of the sincerity of utterance prosody for each ROI; further analyses revealed that there was a significant effect of sincerity in the right frontal ROI ($F(1,24)=7.865$; $p=0.010$) with higher amplitudes in response to sincere than insincere utterances (illustrated in Fig. 5). No other effects of sincerity were found in the remaining ROIs (frontal ($F(1,24)=0.596$; $p=0.448$)), right parieto-occipital ($F(1,24)=0.061$; $p=0.448$), central ($F(1,24)=0.797$; $p=0.381$) and left fronto-temporal ($F(1,24)=0.183$; $p=0.673$). Interestingly, we found that the P600 activity in the right frontal area was significantly (negatively) correlated to the mean behavioral accuracy of participants ($r(23)=-0.552$; $p<0.01$), with lower peak P600 amplitudes for those participants who were more accurate at inferring speaker sincerity.

To identify cortical regions that may be responsible for the sincerity effect in the right frontal area, we compared the brain activity in response to sincere and insincere sentences using nonparametric randomization tests ($p<0.05$). The highest peak of activity was identified by running tests in all time frames of the 600–900 ms temporal window; as this peak was found to occur at 880 ms, we ran a single test in the [860–900 ms] temporal window (Mulert et al., 2004; Pascual-Marqui, 2002). The data reveal higher activity for sincere than for insincere utterances in the left insula ($X=-30$, $Y=-25$,

---

(*footnote continued*)
in 150–300 ms temporal window. Similarly to what we found with spatial PCA cluster, no effect of sincerity was found for P200 component (amplitude: ($F(1,24)=0.035$; $p=0.852$); latency: ($F(1,24)=0.677$; $p=0.419$)).

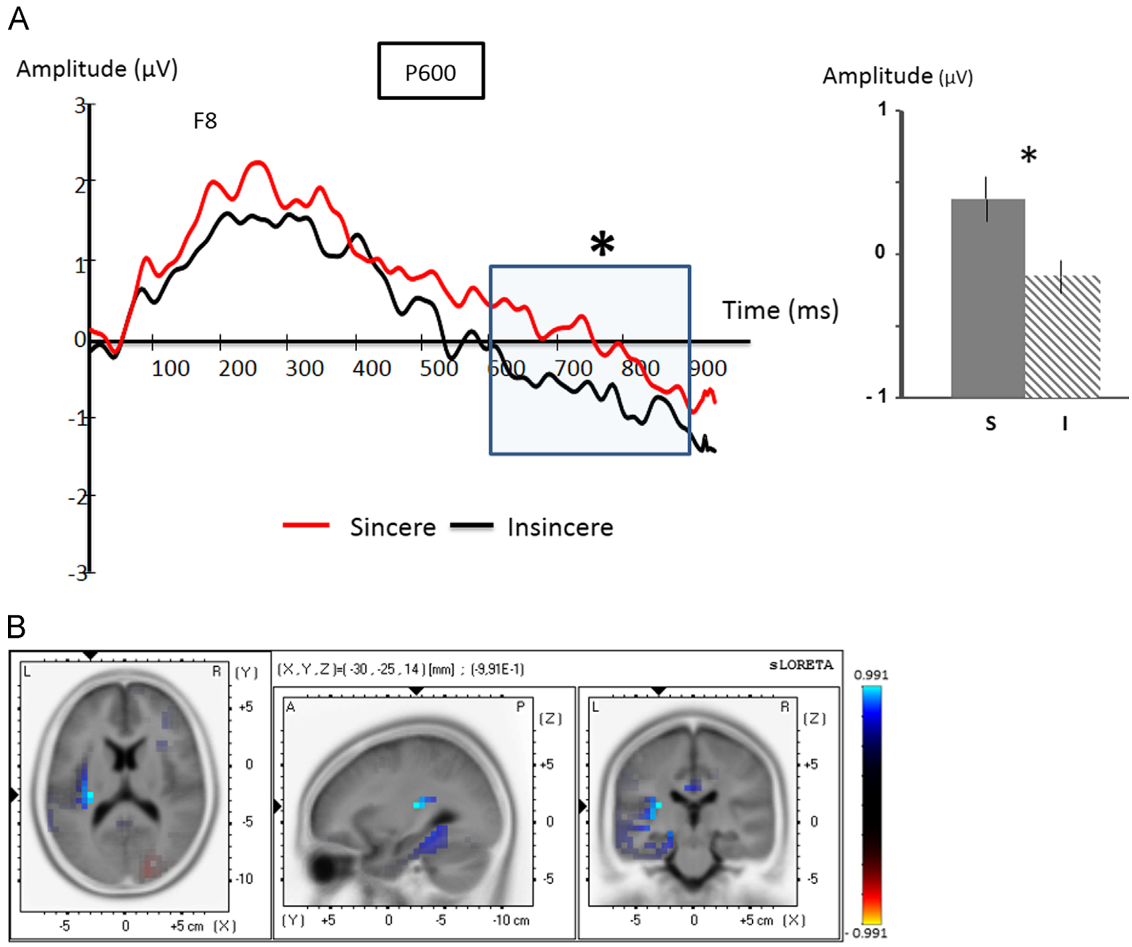

Fig. 5 – P600 Effect. (A) Grand-average ERPs ($n=25$ participants) and mean amplitude of P600 at F8 elicited by sincere (blue line) and insincere (red line) sentences (S: sincere; I: insincere). (B) Results of the sLORETA *t*-statistics comparing neuronal generators underlying the detection of the sincerity status of sentences. The images show LORETA slices in MNI space for the estimated source distributions of activation differences.

Z=15, Brodmann Area 13), the left superior temporal gyrus (X=−35, Y=−36, Z=15, Brodmann Area 41), and the left parahippocampal gyrus (X=−20, Y=−25, Z=−11, Brodmann Area 28). The voxel showing the greatest enhancement of activity for sincere utterances was located in the insula ($p=0.006$; see Fig. 5).

3. Discussion

The meaning of a sentence can go far beyond its strict literal interpretation; this occurs when a speaker intends to convey non-literal or figurative meanings, and also, when contextual cues inform the listener that intended literal meanings do not reflect the speaker's true beliefs. Previous studies show that the decoding of figurative language, like metaphors or irony, involves specific structures in the brain and that the temporal dynamics of the processing of literal and non-literal meanings can differ (Balconi and Amenta, 2010; Cornejo et al., 2007; Coulson, 2007; Regel et al., 2011). Here, we investigated a related but distinct context for interpreting underlying speaker meanings: how listeners interpret the sincerity of compliments produce by a social partner and the role of speech prosodic cues in this process. Our results clearly support the hypothesis that listeners can use prosodic information to accurately judge the sincerity of utterances; moreover, the data imply that these interpretative processes are indexed by a late ERP component, that is modulated by the perceived sincerity of the speaker, which could rely on neural mechanisms involving the insula (among other brain regions). The significance of these patterns is discussed in more detail below.

3.1. *The ability of listeners to judge sincerity*

Our data show that listeners can reliably infer the true opinion of speakers making sincere or insincere compliments from accompanying vocal cues. The overall accuracy of our participants (well over 80%) was considerably higher than that reported in studies of lie–truth discrimination where participants typically perform only slightly better than chance: see Bond and DePaulo (2006) for a review and analysis of this literature. Since we pre-selected items to ensure that prosodic cues were reliably perceived as sincere or insincere (Fish and Pell, 2011) prior to gathering on-line responses to the stimuli, it is not surprising that behavioral judgements of sincerity were more accurate here than in most previous studies of communicating deception: see Levine (2010) for a discussion. Also, Bond and DePaulo (2006) demonstrated that people are more accurate in judging *audible* than visible lies, a factor that may have contributed to the high accuracy of participants in our study focusing on the role of prosodic cues in speech. Another factor that could influence behavioral judgements of sincerity is the ability of individual participants to *produce* lies, which has been recently shown to correlate with the ability to detect truth from lies (Wright et al., 2012). In the absence of broader background measures on lie behavior for our sample, it is possible that some participants were particularly expert at telling (and thus detecting) lies, although other studies have shown that university students do not usually perform at high levels on such tasks (Aamodt and Custer, 2006).

We noted that the detection of sincere compliments was more reliable than the detection of white lies, in line with previous findings of a "truth bias" in communication (Bond and DePaulo, 2006; Levine et al., 1999; Millar and Millar, 1997). Two main factors may have produced evidence of a truth bias in our data. First, this bias is known to be modulated by the familiarity of communication partners, with closer relationships leading to a stronger truth bias in social perception (Van Swol, et al., 2011). For example, McCornack and Parks (1986) suggested that people in close relationships begin to operate under a presumption of honesty that influences all aspects of their interpersonal communication; since our stimuli were constructed to simulate a conversation between friends, participants may have therefore been more likely to consider the responses of the friends (responders) as sincere rather than insincere.

Further data on the nature of response biases in deception detection tasks have been reported by Hurst and Oswald (2012); they manipulated the weighting of errors by presenting videos of individuals interrogated by police where participants had to judge whether the person was trustworthy or lying. Participants received points for accurate responses. They could watch the video sequences up to four times but the more they watched the videos, the fewer points they earned. The number of watched videos was then used as a measure of error weighting. If someone watched more video segments before making a final truth judgement than before making a final lie judgement, then this meant that that person was more concerned about 'Misses' than about 'False Alarms' (or vice-versa). The authors found that people who made more judgements of truth (i.e., who displayed a truth bias) were also more concerned about false alarms. Conversely, participants who were more concerned about misses produced more lie judgements (i.e., a lie bias). If true here, Hurst and Oswald's (2012) results suggest that participants in our study tended to assume an honest behavior from the friend, yielding a truth bias when performing our task. This tendency may reflect the broader reality that in everyday life, people are more often confronted with truthful than with deceptive statements during social interactions with friends, and continually questioning the veracity of utterances would arguably be both resource-demanding and socially maladaptive in this context. To test this further, new studies will be needed to study the role of contextual factors (e.g., speaker gender, social distance) that have consequences for interpreting sincere and insincere utterances and detecting lies, and to determine how judgement biases influence the interpretative process in specific contexts.

3.2. *On-line neural responses to speaker sincerity*

With regard to the electrophysiological activity recorded during the task, we found a late positivity occurring 600 ms after the onset of the response (the compliment), sensitive to the sincerity status of the compliment. This positivity, located in right frontal areas of the scalp, exhibited increased amplitudes in response to sincere versus insincere utterances. Previous studies have reported late positive components associated with similar processes for discriminating literal and non-literal sentence meanings. Using ironic and non-ironic sentences, Regel et al. (2010) found a significant effect on a late positive component, the P600, and suggested that this component could reflect the integration of all information (semantic and extra-semantic) needed to establish a coherent pragmatic interpretation of a sentence: see Brouwer et al. (2012). Given our observation that cues to speaker sincerity modulated the P600 component, combined with our finding that the ability of participants to judge sincerity correlated significantly with the peak of the P600, our results supply further evidence that the integration of semantic information, cued here via speech prosody, is indexed by a late positive component sensitive to the alternative meanings or social intentions of a speaker, including inferences about their sincerity.

In several studies investigating the time course of the production of deceptive responses, it has been found that the LPC—a late positive component usually observed between 300 and 500 ms, located in central-parietal areas—decreased in amplitude when a response conflicts with the truth (Dong and Wu, 2010; Johnson et al., 2003; Proverbio et al., 2013). Interestingly, Proverbio et al. asked their participants to lie or answer truthfully in response to questions on various topics; when a deceptive response was made, they found a smaller late positivity, occurring between 550 and 750 ms after stimulus onset, over frontal areas (especially in the right hemisphere). This pattern closely mirrors the one observed here, where we witnessed higher amplitudes for *sincere* than insincere utterances in the right frontal areas of the scalp. While subject to further study, these data suggest that the late positive component we found in the right frontal areas indexes the stage in which participants assess the speaker's true (or false) beliefs in the context of the utterance (Regel et al., 2010), allowing them to infer when the speaker is being sincere.

Similar to Regel et al. (2010) in their study of vocal irony, we did not find any significant effects of the sincerity status of utterances when the P200 and N400 components were analyzed (experiment 1; to note, these authors did find an effect with *visual* material in experiment 2). As acoustic properties have long been known to influence the P200 component (Crowley and Colrain, 2004), it may have been expected that the different acoustic profiles of sincere and insincere sentences would influence amplitudes of the P200. However, this is likely not the case for more complex acoustic stimuli like spoken utterances that require integration of complex cue configurations. For example, while Paulmann and Kotz (2008) reported significant emotional effects on the P200, the modulations observed could not be explained *only* by the different acoustic profiles of emotion stimuli presented in their experiment. As for the N400, this component has long been associated with the contextual processing of semantic cues and detection of semantic anomaly, where increased N400 amplitudes are usually observed following presentation of a contextually-incongruous versus congruous word (e.g., "I like my coffee with cream and sugar/socks",: see Lau et al. (2008) for a review. The absence of an N400 effect here is not surprising, as it has been previously found that pragmatic processes, such as producing implicatures that revolve around truth judgements, do not impact the stage of semantic processing and integration of a sentence (e.g., Noveck and Posada, 2003; but see Proverbio et al. (2013) who found an effect on N400 when a deceptive act was performed). Rather, our data exemplify that these effects are observed somewhat later, through positive deflections in the 600–900 ms time window, when listeners judge speaker sincerity in reference to the semantic content of an utterance.

This absence of N400 effect, and our interpretation that P600 component could reflect pragmatic processes, fit well with a recent model of language comprehension proposed by Brouwer et al. (2012). These authors argue that current models do not fully explain the findings observed in studies using ERPs and propose a model revisiting the functional role traditionally attributed to the N400 and P600 components. In particular, on the basis of several studies which found P600 effects specifically when listeners had to infer elements within structurally complex sentences (Burkhardt, 2006, 2007), they suggested that P600 component reflects the integration of lexical information into the existing current mental representation of an unfolding sentence. The interpretation of our results, that P600 component reflects integration of semantic information which here is informed by prosodic cues of a speaker, is consistent with this proposition and appears to extend it to a broader range of communication contexts.

3.3. Source localization analysis

Providing further clues about the neural specification of processing at this stage, a source localization analysis revealed that a likely generator of the P600 effect in our data was associated with the insula (as well as limbic regions such as the parahippocampal gyrus), left in particular (which suggests that the source dipole is oriented transversally since the effects on the scalp were observed in the right part of the scalp).[2] The insula has been ascribed a wide range of functions, including those related to emotion and social behavior (Craig, 2009; Lamm and Singer, 2010; Nieuwenhuys, 2012). In this domain, insular activations have been recorded during aversive emotional experiences associated with strong visceral and somatic sensations, such as the experience of unfairness (Corradi-Dell'Acqua et al., 2013; Guo et al., 2013; Güroğlu et al., 2010; Sanfey et al., 2003; Tabibnia et al., 2008), and during the anticipation of negative and unknown emotional events (Herwig et al., 2007).

Of key interest here, several studies have described specific links between the insula and deception behavior, in particular when people *tell* lies (Baumgartner et al., 2013; Ding et al., 2012; Ito et al., 2012; Kireev et al., 2012; Lee et al., 2010; Spence et al., 2008). Ito et al. employed a modified recognition-memory paradigm to identify with fMRI neural networks involved in the preparation and execution of deception, reporting an increase of insula activity when participants were lying (Ito et al., 2012); see also Kireev et al. (2012). In another study, Baumgartner et al. (2013) built a deception paradigm in which subjects play the role of a trustee and interact with investors. The trustee had to make a promise decision at the beginning of three subsequent trust game trials, indicating that he/she always, mostly, sometimes, or never planned to be trustworthy and return the money. The authors recorded resting EEG to measure individual differences in the neural baseline activation and observed an important (negative) relationship between neural baseline activity in the insula and the participants' propensity to deceive. There are also data suggesting involvement of the insula when people process or *perceive* lies; Ding et al. found that concealing, but not faking, identity was associated with increased activity in the insula (among other regions; Ding et al., 2012). As well, in an fMRI study where participants were presented true and false written statements about themselves or another person, the insula was activated in both conditions, suggesting that this structure is recruited during the perception of a lie involving a third party (Ganis et al., 2009). Altogether, these studies argue that the insula serves an important role in the act of producing and perceiving lies and other deceptive social behavior, consistent with our finding that insular activity was associated with on-line neural responses for differentiating sincere versus insincere utterances.

There is also evidence that the insula could play a broader role in certain inferential processes. In a meta-analysis on the functional role of the inferior frontal gyrus that included the insular cortex, Liakakis et al. (2011) found that insular regions were engaged when participants made *predictive* inferences during various tasks. Evidence that the insula provides a substrate for social inferencing is also suggested by studies involving high-Machiavellian individuals—people who are highly adept at inferring others' intentions, beliefs and knowledge and using these predictions to their own advantage—who were shown to exhibit greater insula activation

[2]It should be noted that since statistical analyses were run around the peak of EEG activity within the time window of the P600 and not on the full time window (600–900 ms), it could be argued that the results of this analysis do not identify the generator of the P600 effect. However, source analyses using sLORETA use typically very short time windows: see Hasko et al. (2012); Jung et al. (2013); Wang et al. (2014) and when we analyzed the full time window, we still found that the source of ERP activity was localized in the insula, reinforcing the idea that our results are indeed representative of our data.

than low-Machiavellian individuals when playing a Trust Game (Bereczkei et al., 2013). Heightened activity in the insula was also witnessed in conditions where listeners were able to correctly detect subtle errors in sentences with "semantic illusions" (e.g., *It was two animals of each kind that Moses took on the ark*); this latter finding was interpreted in light of the insula's role in inhibiting competing concepts for message-level integration and resolution of sentence meaning (Raposo and Marques, 2013). While our data provide no direct indications of how the insula may have been engaged by our task, the prospect that our participants needed to suppress the pragmatic assumption that a compliment was true ("truth bias"), while integrating linguistic and prosodic cues to infer speaker sincerity, is not incompatible with ideas raised by Raposo and Marques (2013) as well as Regel et al. (2010).

3.4. *Conclusions*

By demonstrating that listeners can reliably detect the sincerity of speakers and determine when speakers are telling a "white lie", our study emphasizes the role of extra-semantic cues, in particular speech prosody, in the ability to infer speaker intentions and mental states, and provides further data on the neural basis of social communication. These data have a logical connection to accumulating evidence of how humans mentalize or adopt a "Theory of Mind" about another person's state of mind (Abu-Akel, 2003; Frith and Frith, 2001; Premack and Woodruff, 1978). Sperber and Wilson (2002) predicted that the brain areas necessary to infer speaker meaning and to derive figurative meanings in language would include regions associated with Theory of Mind processing; consistent with this idea, recently Spotorno et al. (2012) reported that the brain network involved in Theory of Mind processing (e.g., Frith and Frith, 2006) is activated when listeners process verbal irony. Given the extensive links between the insula and temporal and prefrontal regions involved in Theory of Mind, more research will be needed to understand how the brain registers the feelings, opinions, and beliefs of speakers from their social context, and how these neurocognitive mechanisms are linked to basic processes for mentalizing as well as executive functions (working memory, inhibitory control) which appear to draw upon overlapping neural systems (Christ et al., 2009).

The prospect that inferring a speaker's true opinions and beliefs can in some cases trigger an emotional or empathic process—for example, when listeners infer that a compliment given to a friend is really insincere—is an intriguing idea for further study in this area. Lamm and Singer (2010) proposed that the insula is activated by many aspects of social interactions (empathy, compassion, fairness among others) and in the representation of bodily emotional states, which lead several authors to suggest that the insula plays a role in the inference of others' emotional states (Craig, 2009; Lamm and Singer, 2010; Singer et al., 2009). It should be noted that many tasks of pragmatic language processing have been found to activate typical emotional processing networks involving the insula, the anterior cingulate cortex, and the amygdala (e.g., Mashal et al., 2005; Schmidt and Seger, 2009). For example, Shibata et al. (2010) found that ironic and literal sentences were associated with differential activation of the amygdala and the left insula, a pattern that partially resembles the one observed here when processing sincere versus insincere utterances. Future studies may be in the position to show how the neurocognitive apparatus, especially the insula, plays a role not only in the integration of socially-relevant cues that inform the nature of intended "speaker meanings", but also how these inferential processes are linked to concomitant emotional effects that are an inherent part of social interactions.

4. Experimental procedures

4.1. *Participants*

The participants were 29 native English speakers (16 men/13 women aged 18–34 years; mean age 22.5 ± 3.8 years old) who were recruited through campus advertisements. Based on self-report, 27 of the participants were right-handed and all had normal hearing and normal/corrected-to-normal vision. Before the experiment, each participant completed a questionnaire to establish basic demographic information (age, handedness, and language abilities). The study was ethically approved by the Faculty of Medicine Institutional Review Board at McGill University (Montréal, Canada) and informed written consent was obtained from each participant prior to entering the study.

4.2. *Apparatus*

Stimuli were presented via Presentation software (Neurobehavioural Systems) on a DELL monitor with an AMD Athlon™ computer (Processor 3700, Windows XP). Electroencephalogram (EEG) was recorded with a 72-channel amplifier (QuickAmp, 24 bit AD-converter, Brain Products system) connected to an Intel AMD Athlon™ computer (Processor 3700, Windows XP).

4.3. *Materials*

All experimental stimuli were presented in the auditory modality. Trials consisted of Question–Response utterance pairs that depicted a short conversation between two female friends; in each case, one friend asked the other for her opinion about something, and the second friend replied with a compliment (e.g., '*So, what do you think of my new haircut?*'/'*I think you look really amazing.*'). For each discourse context, the response was spoken in two distinct ways to convey the fact that the speaker giving the compliment truly believed what she said (sincere condition) or was trying to hide her true, negative feelings to spare the questioner's feelings (insincere condition). These meanings could only be determined by how the target response was uttered, i.e. from its prosody. All compliments were 9–11 syllables in length and centered on topics that might be discussed between female friends, falling under four themes: *appearance* (i.e., hairstyle, makeup, bikini, glasses); *creative acts* (painting, acting, writing poetry); *interpersonal acts* (hosting a party, making a presentation, giving a gift); and *third parties* (someone in a close relationship, such as a *boyfriend*, *mother*, or *brother*). Examples of the discourse stimuli with target utterances (compliments) are provided in Table 1.

Table 1 – List of stimuli (question–response pairs) presented in the study. Each question was paired with the corresponding response produced in a sincere and an insincere manner.

Question	Response (compliment)
So, what do you think of how this bikini looks on me?	*It looks really beautiful on you*
So, what do you think of my new glasses?	*I think they look really wonderful on you*
So, what did you think of the painting I made?	*I think you're a very talented artist*
So, what do you think of my new haircut?	*I think you look really amazing*
So, what do you think of the poems I wrote?	*I think you're a really great writer*
So, what did you think of the gift I got you?	*I think you made an excellent choice*
So, what did you think of the movie I chose?	*I found it incredibly funny*
So, what did you think of my party?	*I think it was a really fun night*
So, what do you think of my muscles?	*I think they look very impressive*
So, what do you think of my new boyfriend?	*I think he's a genuinely nice guy*
So, what did you think of your date with my brother?	*I had an unbelievable time*
So, what do you think of my mom as a teacher?	*I think she's phenomenally good*

(a) Stimulus recording and perceptual validation – as reported by Fish and Pell (2011), experimental stimuli were recorded by five female speakers (24±5 years old) who were native speakers of Canadian English. A single speaker produced all of the questions, whereas the other four speakers produced the responses (compliments). Stimuli were restricted to female speakers because they are more likely than men to be the deliverers and recipients of other-oriented, protective lies (i.e., insincere compliments), with the highest occurrence of these other-centered lies involving female–female dyads (DePaulo et al., 1996). In addition, intimate relationships tend to be associated with increased other-oriented dishonesty or "white lies" (DePaulo and Kashy, 1998), justifying why our trials were constructed to simulate a conversation between close female friends.

During the recording session, each of the four speakers who produced the compliments was shown the written target utterance (*I think you look amazing*) and provided contextual information about their true feelings before producing the utterance; each then heard the target question presented free-field in a sound-attenuated booth (*So, what do you think of my haircut?*) and responded in a way that was meant to be sincere or to spare the other person's feelings, according to the biasing context provided. Each target utterance was produced in both a sincere and insincere condition. All utterances were digitally recorded using a head-mounted microphone, saved as individual audio files, normalized to a peak amplitude of 70 dB HL (to correct for slight differences in recording levels across recording sessions), and then validated by a listener group to ensure that the underlying sincerity of the speakers could be accurately perceived. These recordings were presented in a perceptual validation study to 30 participants (15 male and 15 female; $M=22\pm3$ years old) who did not take part in the current study (Fish and Pell, 2011). Participants were asked to rate 256 Question–Response pairs, separated by a 750 ms inter-stimulus interval, that ended in a sincere or insincere compliment; their task was to rate the response on a scale from -2 to $+2$ in reference to "how much the second friend really meant the compliment". Negative ratings indicated that they believed the compliment to be insincere, not reflecting the friend's true opinion, whereas positive scores indicated that the compliment was judged to be sincere.

Based on Fish and Pell's (2011) data, we selected 52 Question–Response pairs for the current experiment: 26 pairs ended with a sincere compliment and 26 semantically-identical pairs ended with an insincere compliment. When selecting the stimuli, we ensured that the prosody of the responses for each item was reliably interpreted as sincere or insincere; utterances were considered 'sincere' if they were rated by at least 60% of participants in the validation study in positive regions of the scale (+1, +2) with less than 30% of participants providing negative ratings of the item (-1, -2). An inverse strategy was used to define 'insincere' utterances. The mean target recognition of selected items in the sincere and insincere condition based on these ratings was equivalent ($M=72\%$ correct judgements in each condition for the 30 listeners). To control for these factors, compliments produced by only three of the responders were selected for each condition (in equal numbers), in addition to the one speaker who produced all of the questions in both the validation study and in the current study.

(b) Acoustic analyses – to further characterize the stimuli, target responses were acoustically analyzed using Praat speech analysis software (Boersma and Weenink, 2008) to determine which vocal patterns were most associated with listeners' perception of sincerity and insincerity. Acoustic measures of the 52 compliments (26 sincere, 26 insincere) were analyzed to characterize five major global (i.e., utterance-level) prosodic features that may have cued listeners: fundamental frequency ($f0$) mean and standard deviation, to index perceived differences in speaker pitch and pitch variability; amplitude mean and standard deviation, to index corresponding differences in perceived loudness; and speech rate (measured in syllables per second) to explore potential differences in timing of sincere versus insincere utterances. One-way ANOVAs on each acoustic measure with response type (sincere versus insincere) as the within-subject factor showed that sincere and insincere utterances differed significantly with respect to mean $f0$ ($F(1,51)=11.07$, $p=0.002$) and $f0$ variation/ standard deviation ($F(1,51)=6.94$, $p=0.011$). Utterances perceived as insincere demonstrated a higher pitch and more pitch variation overall than utterances rated as sincere. There were no differences in the speech rate or amplitude of sincere versus insincere compliments for our dataset. Acoustic features of the experimental stimuli are provided in Table 2.

Table 2 – Acoustic features of the experimental stimuli for sincere and insincere compliments.

	Sincere compliment	White Lie (insincere)
Mean *f*0 (Hz)	254.76 (16.7)	273.66 (30.2)
*f*0 variation (sd in Hz)	50.75 (13.5)	41.35 (12.2)
Mean amplitude (dB)	67.54 (0.8)	67.75 (1.3)
Amplitude variation (sd in dB)	6.43 (0.8)	6.16 (1.1)
Speech rate (syllables/s)	6.10 (0.8)	5.87 (0.8)

*f*0 = fundamental frequency

4.4. *Experimental design/procedure*

Participants were invited to take part in a study of "communication and emotion"; they were seated in a sound-attenuated booth at a distance of 75 cm from the computer screen. After the setup of the cap and electrodes, the quality of the EEG signal was checked, participants were informed about the problem of artifacts and how to reduce them, and the experimental procedures were explained. Presentation software (Neurobehavioural Systems) was used for the auditory presentation of Question–Response pairs via headphones (ER-1 Tubephone, Etymotic Research). Given suggestions that the duration of a pause before a response provides listeners important information for inferring responder attitudes (e.g., willingness to comply with requests and agreement with assessments), the 26 items in each sincerity condition were presented twice in two separate presentation blocks; in one block, the inter-stimulus interval between the question and response was 800 ms, and in the other block it was 1200 ms: see Roberts et al. (2006) for a discussion. Also, deception has been shown to increase feelings of guilt, anxiety, and cognitive load (Caso et al., 2005), resulting in longer response latencies when lying than when telling the truth (Walczyk et al., 2003), motivating an exploration of these two conditions. The order of the blocks was counterbalanced across participants.

Each trial started with a central fixation cross to reduce eye movements (500 ms duration) followed by the auditory Question stimulus; after a delay of either 800 or 1200 ms, the auditory Response was triggered: see Fig. 6. Responses were followed by a visual screen displaying the question, "*Did she mean what she said?*" and the participant was instructed to click on one of two answer boxes (yes/no) at the bottom of the screen to record their decision. Each block contained the same number of sincere versus insincere responses, and thus yielded similar numbers of "yes" and "no" judgements. The next trial started after a 2500 ms delay. Participants completed five practice trials before each recording session to familiarize themselves with the procedures and materials. Participants were paid at the end of the experiment ($30 CAD) which lasted approximately two hours.

4.5. *EEG recording and analysis*

EEG was recorded from 64 cap-mounted active electrodes (10/10 System, actiCap) with AFz electrode as the ground and FCz electrode used as reference. Four additional electrodes were placed for vertical and horizontal electro-oculogram recording: two at the outer canthi of eyes and one above and below each eye. The impedance for all the electrode sites was kept below 5 kΩ. The EEG was digitized at 1024 Hz in continuous and was down-sampled to 250 Hz for analyses EEGLab software, version 9, with MATLAB (R2010b, 7.11). After the recording, a bandpass was set offline between 0.016 and 40 Hz. The rejection of artifacts, in particular eye movements and blinks, was performed using EEGLab. Only trials answered correctly were included in the analyses of the EEG signal: first, we manually removed all noisy parts of the signal for each participant; second, we decomposed the signal by an Independent Component Analysis (ICA, runica algorithm). Components identified by the ICA were studied by the first author to remove eye movement artifacts, blinks, or muscular movements from the signal. In a last step, we excluded all trials with a voltage exceeding $\pm 40\,\mu V$ from the analysis. Across conditions, 16% of trials (only correctly answered trials were included in the ERP analysis) were rejected by this procedure. Average referenced EEG epochs (−100–1000 ms) were time-locked to the onset of the response sentence, baseline corrected (−100–0 ms), and averaged offline according to the sincerity status of the stimuli (sincere/insincere).

4.6. *Statistical analyses*

Data of 25 participants were considered in all statistical analyses. Four participants (two males/two females) were excluded because of their poor EEG signal (more than 40% of trials were excluded for these participants). We analyzed the behavioral accuracy of the participants by running a multivariate ANOVA (MANOVA: see Vasey and Thayer (1987) with repeated measures of sincerity (sincere, insincere) and inter-stimulus interval (800, 1200 ms). For the statistical analysis of the EEG data, we used the properties of principal component analysis (PCA) to define spatial regions of interest: see Spencer et al. (1999), (2001). We performed a spatial PCA with 64 electrode sites as dependent variables and time points (249), participants (25), and conditions (sincere/insincere) as observations (Varimax rotation, SPSS V.20 software; Pourtois, et al., 2008). Each Spatial Factor represents a specific spatial configuration of brain activation and the factor loading corresponds to the Spatial Factor's contribution to the original variables (i.e. how much the spatial factor accounts for the voltage recorded at each electrode). These spatial configurations can be visualized by topographic maps of factor loadings (Cartool software v.3.52, D. Brunet, https://sites.google.com/site/fbmlab/cartool) and are usually defined by considering electrodes with the highest factor loadings

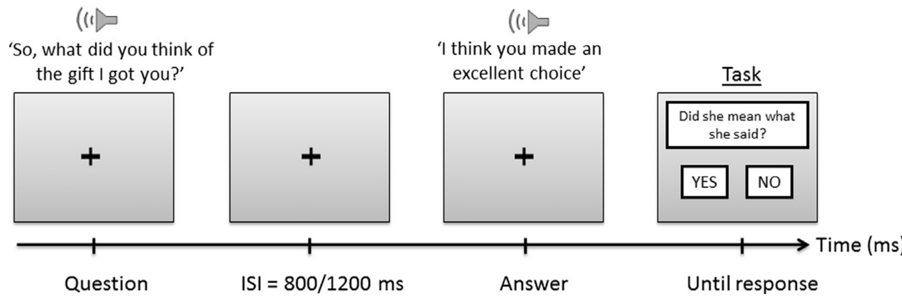

Fig. 6 – Illustration of a trial sequence used in the study.

(Delplanque et al., 2006; D'Hondt et al., 2010; Rigoulot et al., 2008, 2011, 2012). Here, a group of electrodes was identified as a region of interest (ROI) when the loadings of these electro des were superior to 0.707, signifying that these electrodes explained more than 50% of data variance.

Following the definition of ROIs, three temporal windows that included the components of interest (P200, N400, and P600) were defined after examination of the grand average of the data. We then performed a MANOVA on the peak amplitudes of P200, N400 and P600, with repeated measures of sincerity (sincere, insincere), ISI (800,1200) and ROI as within-subject factors. In a last step, in order to identify the neural sources underlying the detection of sincerity, source localization was conducted for ERP components that differed between sincere and insincere sentences using standardized low-resolution brain electromagnetic tomography (sLORETA; Pascual-Marqui, 2002; www.unizh.ch/keyinst/NewLORETA/sLORETA/sLORETA.htm). Basically, sLORETA gives a single linear solution to the inverse problem of localization of brain function based on extra-cranial measurements (Marco-Pallares et al., 2005) and produces images of standardized current density with no localization bias (Pascual-Marqui, 2002).

Although solutions provided by EEG-based source–location algorithms should generally be interpreted with caution due to their potential error margins, the localization accuracy of sLORETA has been validated in simultaneous EEG/fMRI studies (Mulert et al., 2004; Olbrich et al., 2009; Vitacco et al., 2002). For sLORETA, the intracerebral volume is partitioned in 6239 voxels at 5 mm spatial resolution and the standardized current density at each voxel is then calculated in a realistic head model (Fuchs et al., 2002) using the Montreal Neurological Institute (MNI) template (Mazziotta et al., 2001). In the present study, non-parametric statistical analyses (Statistical non-Parametric Mapping, SnPM) were used, employing a log of F ratio for average statistic for paired groups (sincere and insincere target utterances). The results of SnPM are described as t-values for each voxel with Bonferroni correction for multiple comparisons. Corrected $p<0.05$ values were accepted as statistically significant.

Acknowledgments

Thanks to Hope Valeriote and Sameer Sethi for running the experiment. This research was funded by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (RGPIN/203708 to MDP).

REFERENCES

Aamodt, M.G., Custer, H., 2006. Who can best catch a liar? A meta-analysis of individual differences in detecting deception. Forensic Exam. 15, 6–11.
Abu-Akel, A., 2003. A neurobiological mapping of theory of mind. Brain Res. Rev. 43, 29–40.
Ashby, M., 2006. Prosody and idioms in english. J. Pragmat. 38, 1580–1597.
Balconi, M., Amenta, S., 2010. "A fighter is a lion" Neuropsychological indexes in comprehending frozen metaphors. J. Pragmat. 42 (12), 3246–3257.
Banse, R., Scherer, K.R., 1996. Acoustic profiles in vocal emotion expression. J. Personal. Soc. Psychol. 70 (3), 614–636.
Baumgartner, T., Gianotti, L.R.R., Knoch, D., 2013. Who is honest and why: baseline activation in anterior insula predicts inter-individual differences in deceptive behavior. Biol. Psychol. 94 (1), 192–197.
Bélanger, N., Baum, S.R., Titone, D., 2009. Use of prosodic cues in the production of idiomatic and literal sentences by individuals with right- and left-hemisphere damage. Brain Lang. 110, 38–42.
Bereczkei, T., Deak, A., Papp, P., Perlaki, G., Orsi, G., 2013. Neural correlates of Machiavellian strategies in a social dilemma task. Brain Cogn. 82 (1), 108–116.
Boersma, P., & Weenink, D. (2008). Praat: doing phonetics by computer (Version 5.0.22) [Computer program]. [retrieved 30.04.08]. Available from: ⟨http://www.praat.org/⟩.
Bohrn, I., 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, 2669–2683.
Bond, C.F., DePaulo, B.M., 2006. Accuracy of deception detection. Personal. Soc. Psychol. Rev. 10, 214–234.
Brouwer, H., Fitz, H., Hoeks, J., 2012. Getting real about semantic illusions: rethinking the functional role of the P600 in language comprehension. Brain Res. 1446, 127–143.
Bryant, G.A., 2010. Prosodic contrasts in ironic speech. Discourse Process. 47, 545–566.
Bryant, G.A., Fox Tree, J.E., 2005. Is there an ironic tone of voice?. Lang. Speech 48, 257–277.
Burkhardt, P., 2006. Inferential bridging relations reveal distinct neural mechanisms: evidence from event-related brain potentials. Brain Lang. 98, 159–168.
Burkhardt, P., 2007. The P600 reflects cost of new information in discourse memory. Neuroreport 18, 1851–1854.
Caso, L., Gnisci, A., Vrij, A., Mann, S., 2005. Processes underlying deception: an empirical analysis of truth and lies when manipulating the stakes. J. Investig. Psychol. Offender Profiling 2, 195–202.
Cheang, H.S., Pell, M.D., 2008. The sound of sarcasm. Speech Commun. 50, 366–381.

Cheang, H.S., Pell, M.D., 2009. Acoustic markers of sarcasm in cantonese and english. J. Acoust. Soc. Am. 126, 1394–1405.
Christ, S.E., Van Esse, D.C., Watson, J.M., Brubaker, L.E., McDermott, K.B., 2009. The contributions of prefrontal cortex and executive control to deception: evidence from activation likelihood estimate meta-analyses. Cereb. Cortex 19, 1557–1566.
Cornejo, C., Simonetti, F., Aldunate, N., Ibanez, A., Lopez, V., Melloni, L., 2007. Electrophysiological evidence of different interpretative strategies in irony comprehension. J. Psycholinguist. Res. 26, 411–430.
Corradi-Dell'Acqua, C., Civai, C., Rumiati, R.I., Fink, G.R., 2013. Disentangling self- and fairness-related neural mechanisms involved in the ultimatum game: an fMRI study. Soc. Cogn. Affect. Neurosci. 8 (4), 424–431.
Coulson, S., 2007. A special role for the right hemisphere in metaphor comprehension? ERP evidence from hemifield presentation. Brain Res. 1146 (1), 128–145.
Coulson, S., Van Petten, C., 2007. A special role for the right hemisphere in metaphor comprehension?: ERP evidence from hemifield presentation. Brain Res. 1146, 128–145.
Craig, A.D., 2009. How do you feel—now? The anterior insula and human awareness. Nat. Rev. Neurosci. 10, 59–70.
Crowley, K.E., Colrain, I.M., 2004. A review of the evidence for P2 being an independent component process: age, sleep and modality. Clin. Neurophysiol. 115, 732–744.
D'Hondt, F., Lassonde, M., Collingon, O., Dubarry, A.-S., Robert, M., Rigoutlot, Sequiera, H., 2010. Early brain–body impact of emotional arousal. Hum. Neurosci. 4, 1–10.
Delplanque, S., Silvert, L., Hot, P., Rigoulot, S., Sequeira, H., 2006. Arousal and valence effects on event-related P3a and P3b during emotional categorization. Int. J. Psychophysiol. 60, 315–322.
DePaulo, B.M., Kashy, D.A., Kirkendol, S.E., Wyer, M.M., Epstein, J.A., 1996. Lying in everyday life. J. Personal. Soc. Psychol. 70, 979–995.
DePaulo, B.M., Kashy, D.A., 1998. Everyday lies in close and casual relationships. J. Personal. Soc. Psychol. 74, 63–79.
DePaulo, B.M., Lindsay, J.J., Malone, B.E., Muhlenbruck, L., Charlton, K., Cooper, H., 2003. Cues to deception. Psychol. Bull. 129, 74–118.
Ding, X.P., Du, X., Lei, D., Hu, C.S., Fu, G., Chen, G., 2012. The neural correlates of identity faking and concealment: an FMRI study. PloS One 7 (11), e48639.
Dong, G., Wu, H., 2010. Attempting to hide our real thoughts: electrophysiological evidence from truthful and deceptive responses during evaluation. Neurosci. Lett. 1, 1–5.
Ekman, P., Friesen, W.V., 1969. Nonverbal leakage and clues to deception. Psychiatry 32, 88–105.
Ekman, P., O'Sullivan, M., 2006. From flawed self-assessment to blatant whoppers: the utility of voluntary and involuntary behavior in detecting deception. Behav. Sci. Law 24, 673–686.
Eviatar, Z., Just, M.A., 2006. Brain correlates of discourse processing: an fMRI investigation of irony and conventional metaphor comprehension. Neuropsychologia 44, 2348–2359.
Fish, K. & Pell, M.D. (2011). Prosodic markers of sincere compliments and white lies. In: Experimental and Theoretical Advances in Prosody 2, September 2011, Montréal, pp. 23–25.
Frank, M.G., Ekman, P., 1997. The ability to detect deceit generalizes across different types of high-stake lies. J. Personal. Soc. Psychol. 72, 1429–1439.
Friederici, A.D., Hahne, A., Saddy, D., 2002. Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. J. Psycholinguist. Res. 31, 45–63.
Frith, C.D., Frith, U., 2006. The neural basis of mentalizing. Neuron 50, 531–534.
Frith, U., Frith, C., 2001. The biological basis of social interaction. Curr. Dir. Psychol. Sci. 10, 151–155.
Fuchs, M., Kastner, J., Wagner, M., Hawes, S., Ebersole, J.S., 2002. A standardized boundary element method volume conductor model. Clin. Neurophysiol. 113, 702–712.
Ganis, G., Morris, R.R., Kosslyn, S.M., 2009. Neural processes underlying self- and other-related lies: an individual difference approach using fMRI. Soc. Neurosci. 4 (6), 539–553.
Gibbs, R.W., 1994. Figurative thought and figurative language. In: Gernsbacher, M.A. (Ed.), Handbook of Psycholinguistics. Academic Press, San Diego, CA, pp. 411–446.
Guo, X., Zheng, L., Zhu, L., Li, J., Wang, Q., Dienes, Z., Yang, Z., 2013. Increased neural responses to unfairness in a loss context. NeuroImage 77, 246–253.
Güroğlu, B., van den Bos, W., Rombouts, S.A.R.B., Crone, E.A., 2010. Unfair? It depends: neural correlates of fairness in social context. Soc. Cogn. Affect. Neurosci. 5 (4), 414–423.
Hasko, S., Bruder, J., Bartling, J., Schulte-Korne, G., 2012. N300 indexes deficient integration of orthographic and phonological representations in children with dyslexia. Neuropsychologia 50, 640–654.
Herwig, U., Kaffenberger, T., Baumgartner, T., Jancke, L., 2007. Neural correlates of a 'pessimistic' attitude when anticipating events of unknown emotional valence. NeuroImage 34, 848–858.
Hurst, M., Oswald, M., 2012. Mechanisms underlying response bias in deception detection. Psychol. Crime Law 18, 759–778.
Ito, A., Abe, N., Fujii, T., Hayashi, A., Ueno, A., Mugikura, S., Mori, E., 2012. The contribution of the dorsolateral prefrontal cortex to the preparation for deception and truth-telling. Brain Res. 1464, 43–52.
Johnson, R., Barnhardt, J., Zhu, J., 2003. The deceptive response: effects of response conflict and strategic monitoring on the late positive component and episodic memory-related brain activity. Biol. Psychol. 64, 217–253.
Jung, E.K., Kang, K.Y., Kim, Y.Y., 2013. Frontoparietal activity during deceptive responses in the P300-based guilty knowledge test: an sLORETA study. Neuroimage 78, 305–315.
Juslin, P.N., Laukka, P., 2003. Communication of emotions in vocal expression and music performance: different channels same code?. Psychol. Bull. 129 (5), 770–814.
Kireev, M.V., Korotkov, A.D., Medvedev, S.V., 2012. Functional magnetic resonance study of deliberate deception. Hum. Physiol. 38 (1), 32–39.
Lamm, C., Singer, T., 2010. The role of anterior insular cortex in social emotions. Brain Struct. Funct. 214, 579–591.
Lau, E.F., Phillips, C., Poeppel, D., 2008. A cortical network for semantics:(De)constructing the N400. Nat. Rev. Neurosci. 9, 920–933.
Lee, T.M.C., Lee, T.M.Y., Raine, A., Chan, C.C.H., 2010. Lying about the valence of affective pictures: an fMRI study. PloS One 5 (8), e12291.
Levine, T.,R., 2010. A few transparent liars: explaining 54% accuracy in deception detection experiments. Commun. Yearb. 34, 41–59.
Levine, T.R., Park, H.S., McCornack, S.A., 1999. Accuracy in detecting truths and lies: documenting the 'Veracity Effect'. Commun. Monogr. 66, 125–144.
Liakakis, G., Nickel, J., Seitz, R.J., 2011. Diversity of the inferior frontal gyrus—a meta-analysis of neuroimaging studies. Behav. Brain Res. 225 (1), 341–347.
Marco-Pallares, J., Grau, C., Ruffini, G., 2005. Combined ICA-LORETA analysis of mismatch negativity. NeuroImage 25, 471–477.

Mashal, N., Faust, M., Hendler, T., 2005. The role of the right hemisphere in processing nonsalient metaphorical meanings: application of principal components analysis to fMRI data. Neuropsychologia 43 (14), 2084–2100.
Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Mazoyer, B., 2001. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). Philos. Trans. R. Soc. Lond. B Biol. Sci. 1412 (1293-1322).
McCornack, S.A., Parks, M.R., 1986. Deception detection and relationship development: the other side of trust. In: McLaughlin, M.L. (Ed.), Communication Yearbook, 9. Sage, Beverly Hills, CA, pp. 377–389.
Millar, M.G., Millar, K.G., 1997. The effects of cognitive capacity and suspicion on truth bias. Commun. Res. 24, 556–570.
Mulert, C., Jagger, L., Schmitt, R., Bussfeld, P., Pogarell, O., Moller, H.J., Juckel, G., Hegerl, U., 2004. Integration of fMRI and simultaneous EEG: towards a comprehensive understanding of localization and time-course of brain activity in target detection. NeuroImage 22, 83–94.
Nieuwenhuys, R., 2012. The insular cortex: a review. Prog. Brain Res. 195, 123–163.
Noveck, I., Posada, A., 2003. Characterizing the time course of an implicature: an evoked potentials study. Brain Lang. 85, 203–210.
Noveck, I., Reboul, A., 2008. Experimental pragmatics: a Gricean turn in the study of language. Trends Cogn. Sci. 12, 425–431.
Olbrich, S., Mulert, C., Karch, S., Trenner, M., Leicht, G., Pogarell, O., Hegerl, U., 2009. EEG-vigilance and BOLD effect during simultaneous EEG/fMRI measurements. NeuroImage 45, 319–332.
Pascual-Marqui, R.D., 2002. Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. Methods Find. Exp. Clin. Pharmacol. 24, 5–12.
Paulmann, S., Kotz, S., 2008. Early emotional prosody perception based on different speaker voices. Neuroreport 19, 209–213.
Pinheiro, A.P., Del Re, E., Mezin, J., Nestor, P.G., Rauber, A., McCarley, R.W., Goncalves, O.F., Niznikiewicz, M.A., 2013. Sensory-based and higher-order operations contribute to abnormal emotional prosody processing in schizophrenia: an electrophysiological investigation. Psychol. Med. 43 (3), 603–618.
Pourtois, G., Delplanque, S., Michel, C., Vuilleumier, P., 2008. Beyond conventional event-related brain potential (ERP): Exploring the time-course of visual emotion. Brain Topogr. 20, 266–277.
Premack, D., Woodruff, G., 1978. Does the chimpanzee have a theory of mind?. Behav. Brain Sci. 1, 515–526.
Proverbio, A.M., Vanutelli, M.E., Adorni, R., 2013. Can you catch a liar? How negative emotions affect brain responses when lying or telling the truth. PloS One 8 (3), e59383.
Raposo, A., Marques, J.F., 2013. The contribution of fronto-parietal regions to sentence comprehension: insights from the Moses illusion. NeuroImage 83C, 431–437.
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, 600–610.
Regel, S., Gunter, T.C., Friederici, A.D., 2011. Isn't it ironic? An electrophysiological exploration of figurative language processing. J. Cogn. Neurosci. 23, 277–293.
Regel, S., Coulson, S., Gunter, T.C., 2010. The communicative style of a speaker can affect language comprehension? ERP evidence from the comprehension of irony. Brain Res. 1311, 121–135.
Rigoulot, S., Delplanque, S., Despretz, P., Defoort-Dhellemmes, S., Honoré, J., Sequeira, H., 2008. Peripherally presented emotional scenes: a spatiotemporal analysis of early ERP responses. Brain Topogr. 20, 216–223.
Rigoulot, S., D'Hondt, F., Defoort-Dhellemmes, S., Despretz, P., Honoré, J., Sequeira, H., 2011. Perceiving fearful faces in peripheral vision: behavioral and ERP evidence. Neuropsychologia 49, 2013–2021.
Rigoulot, S., D'Hondt, F., Honoré, J., Sequeira, H., 2012. Implicit emotional processing in peripheral vision: behavioral and neural evidence. Neuropsychologia 12, 2887–2896.
Roberts, F., Francis, A.L., Morgan, M., 2006. The interaction of inter-turn silence with prosodic cues in listener perceptions of 'Trouble' in conversation. Speech Commun. 48, 1079–1093.
Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., Cohen, J.D., 2003. The neural basis of economic decision-making in the Ultimatum Game. Science 300, 1755–1758.
Schmidt, G.L., Seger, C.A., 2009. Neural correlates of metaphor processing: the roles of figurativeness, familiarity and difficulty. Brain Cogn. 71 (3), 375–386.
Shibata, M., Toyomura, A., Itoh, H., Abe, J., 2010. Neural substrates of irony comprehension: a functional MRI study. Brain Res. 1308, 114–123.
Singer, T., Critchley, H.D., Preuschoff, K., 2009. A common role of insula in feelings, empathy and uncertainty. Trends Cogn. Sci. 13, 334–340.
Snedeker, J., Trueswell, J., 2003. Using prosody to avoid ambiguity: effects of speaker awareness and referential context. J. Mem. Lang. 48, 103–130.
Spence, K., Villar, G., Arciuli, J., 2012. Markers of deception in italian speech. Front. Psychol. 3, 453.
Spence, S., Kaylor-Hughes, C., Farrow, T.F.D., Wilkinson, I.D., 2008. Speaking of secrets and lies: the contribution of ventrolateral prefrontal cortex to vocal deception. NeuroImage 40 (3), 1411–1418.
Spencer, K.M., Dien, J., Donchin, E., 1999. A componential analysis of the ERP elicited by novel events using a dense electrode array. Psychophysiology 36, 409–414.
Spencer, K.M., Dien, J., Donchin, E., 2001. Spatiotemporal analysis of the late ERP responses to deviant stimuli. Psychophysiology 38, 343–358.
Sperber, D., Wilson, D., 1995. Relevance: Communication and Cognition. Basil Blackwell, Oxford.
Sperber, D., Wilson, D., 2002. Pragmatics, modularity and mindreading. Mind Lang. 17, 3–23.
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, 25–39.
Tabibnia, G., Satpute, A.B., Lieberman, M.D., 2008. The sunny side of fairness. Psychol. Sci. 19, 339–347.
Talwar, V., Murphy, S.M., Lee, K., 2007. White lie-telling in children for politeness purposes. Int. J. Behav. Dev. 1, 1–11.
Vasey, M.W., Thayer, J.F., 1987. The continuing problem of false positives in repeated measures ANOVA in psychophysiology: a multivariate solution. Psychophysiology 4, 479–486.
Van Berkum, J.J.A., Brown, C.M., Hagoort, P., Zwitserlood, P., 2003. Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. Psychophysiology 40, 235–248.
Van Lancker-Sidtis, D., 2003. Auditory recognition of idioms by native and nonnative speakers of english: it takes one to know one. Appl. Psycholinguist. 24 (1), 45–47.
Van Swol, L.M., Braun, M.T., Malhotra, D., 2011. Evidence for the Pinocchio effect: linguistic differences between lies, deception by omission, and truth. Discourse Process. 49, 79–106.
Villar, G., Arciuli, J., Paterson, H., 2013. Vocal Pitch production during lying: beliefs about deception matter. Psychiatry Psychol. Law 20 (1), 123–132.

Vitacco, D., Brandeis, D., Pascual-Marqui, R., Martin, E., 2002. Correspondence of event-related potential tomography and functional magnetic resonance imaging during language processing. Hum. Brain Mapp. 17, 4–12.

Vrij, A., 1994. The impact of information and setting on detection of deception by police detectives. J. Nonverbal Behav. 18, 117–136.

Vrij, A., Mann, S., 2001. Telling and detecting lies in a high-stake situation: the case of a convicted murderer. Appl. Cogn. Psychol. 15, 187–203.

Walczyk, J.J., Roper, K.S., Seemann, E., Humphrey, A.M., 2003. Cognitive mechanisms underlying lying to questions: response time as a cue to deception. Appl. Cogn. Psychol. 17, 755–774.

Wang, Y., Deng, Y., Sui, D., Tang, Y.Y., 2014. Neural correlates of cultural differences in moral decision making: a combined ERP and sLORETA study. Neuroreport 25 (2), 110–116.

Wilson, D., Wharton, T., 2006. Relevance and prosody. J. Pragmat. 38, 1559–1579.

Wright, G.R., Berry, C.J., Bird, G., 2012. You can't kid a kidder: association between production and detection of deception in an interactive deception task. Front. Hum. Neurosci. 6, 1–7.

Zuckerman, M., DeFrank, R.S., Hall, J.A., Larrance, D.T., Rosenthal, R., 1979. Facial and vocal cues of deception and honesty. J. Exp. Soc. Psychol. 15, 378–396.