



Immediate online use of prosody reveals the ironic intentions of a speaker: neurophysiological evidence

Maël Mauchand¹ · Jonathan A. Caballero¹ · Xiaoming Jiang² · Marc D. Pell¹

Accepted: 8 November 2020 / Published online: 9 January 2021
© The Psychonomic Society, Inc. 2021

Abstract

In social interactions, speakers often use their tone of voice (“prosody”) to communicate their interpersonal stance to pragmatically mark an ironic intention (e.g., sarcasm). The neurocognitive effects of prosody as listeners process ironic statements in real time are still poorly understood. In this study, 30 participants judged the friendliness of literal and ironic criticisms and compliments in the absence of context while their electrical brain activity was recorded. Event-related potentials reflecting the uptake of prosodic information were tracked at two time points in the utterance. Prosody robustly modulated P200 and late positivity amplitudes from utterance onset. These early neural responses registered both the speaker’s stance (positive/negative) and their intention (literal/ironic). At a later timepoint (You are such a great/horrible cook), P200, N400, and P600 amplitudes were all greater when the critical word valence was congruent with the speaker’s vocal stance, suggesting that irony was contextually facilitated by early effects from prosody. Our results exemplify that rapid uptake of salient prosodic features allows listeners to make online *predictions* about the speaker’s ironic intent. This process can constrain their representation of an utterance to uncover nonliteral meanings without violating contextual expectations held about the speaker, as described by *parallel-constraint satisfaction* models.

Keywords Speech communication · Vocal expression · Nonliteral meanings · Irony · ERPs

Introduction

Often, what a speaker says does not align with their expressed attitude, promoting nonliteral meanings (Grice, 1989; Searle, 1965; Sperber & Wilson, 1981). To ensure that their intention is accurately recognized, speakers use different forms of contextual and pragmatic cues, such as “tone of voice” (speech *prosody*), to help listeners detect their interpersonal stance and to retrieve meanings that go beyond the semantic content of the verbal message (Gibbs & Colston, 2007; Pexman, 2008). Prosodic information furnishes powerful cues about the

affective disposition, mental state, and social (e.g., politeness) intentions of a speaker as listeners process language (Belin, Fecteau, & Bédard, 2004; Jiang & Pell, 2015; Van Lancker Sidtis, Pachana, Cummings, & Sidtis, 2006; Vergis & Pell, 2020).

In this study, we focused selectively on prosody as a major pragmatic cue for inferring a speaker’s stance or “sincerity” (Kumon-Nakamura, Glucksberg, & Brown, 1995) as listeners process familiar ironic statements, such as sarcastic or teasing comments (“You’re such a great driver”). Following recent calls to more rigorously define how prosody influences irony comprehension (Cornejo et al., 2007; Deliens, Antoniou, Clin, Ostashchenko, & Kissine, 2018; Pexman, 2008), our design sheds new light on the neurocognitive mechanisms that contribute to verbal irony processing and how prosody influences the time course of the neural response.

Verbal Irony and Prosody

Two commonly studied forms of verbal irony are *ironic criticisms* (i.e., sarcasm) and *ironic compliments* (i.e., ironic praise, teasing, or banter, Bruntsch & Ruch, 2017; Slugoski & Turnbull, 1988). Ironic criticisms, which have been studied

✉ Maël Mauchand
mael.mauchand@mail.mcgill.ca

Marc D. Pell
marc.pell@mcgill.ca

¹ School of Communication Sciences and Disorders, Faculty of Medicine, McGill University, 2001, McGill College, Montréal, Quebec H3A1G1, Canada

² Institute of Linguistics, Shanghai International Studies University, Shanghai, China

more extensively and are considered more prototypical than ironic compliments, make use of a positive statement to convey criticism; hence, they display a negative stance towards the listener. For example, a speaker who says “You’re such a great driver” to someone who just drove through a red light is likely to convey a critical, sarcastic attitude (highlighting a failed positive expectation that the listener would drive carefully). For this to be understood by their interlocutor, speakers create an *implicature* (Grice, 1989; Wichmann, 2000; Wilson, 2017), for example, by supplying pragmatic cues showing that the positive statement does not reflect their true opinion (i.e., the speaker is “pragmatically insincere”; Kumon-Nakamura et al., 1995).

In fact, a primary reason for using verbal irony, as opposed to making a literal statement, is for the speaker to highlight their positive or negative attitude toward the referent of the ironic utterance (Kumon-Nakamura et al., 1995; Sperber & Wilson, 1981). Research shows that when listeners detect ironic criticism, even in the absence of prosody (i.e., based on written statements), they evaluate the speaker’s stance as being significantly more negative than when the same statement is produced as a literal compliment (Dews, Kaplan, & Winner, 1995; Gibbs & Colston, 2007; Pexman & Olineck, 2002; Pexman & Zvaigzne, 2004). Yet, when compared to *literal* criticisms that phrase the utterance in a negative manner (“You’re such a lousy driver”), ironic criticisms are judged to be less critical, more polite, and more friendly (Dews et al., 1995; Mauchand, Vergis, & Pell, 2020; Pexman & Zvaigzne, 2004). This suggests that the indirect nature of ironic criticisms “tinges” and softens listeners’ perception of the speaker’s critical attitude, in the absence of explicit negative verbal cues (Dews et al., 1995; Pexman & Olineck, 2002). These findings underscore the notion that cues referring to the speaker’s attitude or stance are of central importance as listeners process and interpret ironic language.

Compared with ironic criticism, *ironic compliments* are characterized by the opposite structure: the speaker makes a negative statement to convey a positive stance (to “praise” the listener, for example, when negative expectations are exceeded, Bruntsch & Ruch, 2017). Saying “You’re such a lousy driver” to a friend who is clearly very good implies an intent to playfully compliment, not to criticize. Despite similarities in structure, the intent of ironic compliments does not seem to be recognized as accurately as ironic criticisms or processed in the same manner (Bruntsch & Ruch, 2017; Caillies et al., 2019; Caffarra et al., 2019). This may be due to the “asymmetry of affect” (Matthews, Hancock, & Dunham, 2006; Pexman & Olineck, 2002). While listeners seem to know that ironic compliments are *meant* to tease (i.e., demonstrate the speaker’s positive stance), they still judge these utterances as somewhat critical, impolite, and/or unfriendly (Alberts, Kellar-Guenther, & Corman, 1996; Kreuz & Link, 2002; Matthews et al., 2006; Pexman & Olineck,

2002). Due to the social risks of being misinterpreted, speakers thus tend to avoid producing ironic compliments except in intimate relationships due to their strong reliance on contextual constraints (Matthews et al., 2006; Pexman & Zvaigzne, 2004; Sally, 2003).

What is known about the prosodic form of ironic criticisms and compliments? Prosody refers to dynamic changes in pitch, loudness, voice quality, and duration which together create meaningful contrasts at the suprasegmental level of speech. Perceptual-acoustic studies have investigated whether there is a specific “sarcastic tone of voice” (Bryant & Fox Tree, 2005; Cheang & Pell, 2008; Wichmann, 2000). While this topic remains unresolved, there is strong evidence that ironic criticisms are associated with a particular set of acoustic features. Speakers tend to produce these utterances slower, with lower pitch, a restriction of pitch variation, and harsher voice quality than (otherwise identical) literal compliments (Anolli, Ciceri, & Infantino, 2000; Bryant, 2010; Cheang & Pell, 2008, 2009; Mauchand, Vergis, & Pell, 2018). Attempts to define a “teasing tone of voice” have reported acoustic-perceptual features such as laughter or “smiled speech,” but relative to sarcasm, these cues seem to depend more on the linguistic content and context of the utterance (Alberts et al., 1996; Cheang & Pell, 2008; Keltner, Capps, Kring, Young, & Heerey, 2001). This literature points to ways that prosody could serve as a pragmatic marker of a speaker’s intent to be ironic, while implying that vocal expressions of ironic criticism (sarcasm) may be systematized to a greater extent and easier to detect than those for expressing ironic compliments.

The extent to which prosodic information is used to retrieve ironic meanings during spoken language processing in relation to other forms of context is not clear. Research implies that the salience of prosodic features during irony perception is diminished when background information (e.g., verbal descriptions) for inferring the speaker’s ironic intent is already known (Deliens et al., 2018; Regel, Coulson, & Gunter, 2010; Regenbogen et al., 2012). Recently, Mauchand et al. (2020) investigated the perception of ironic criticisms and compliments from prosody under different conditions of attentional focus when listeners had no background information about the speaker. Participants rated the affective *stance* of the speaker along a friendliness scale when focusing on the speaker’s prosody, their statement, or in a more holistic (i.e., presumably integrative) manner. When focusing only on the speaker’s prosody or their statement, participants were always successful in recognizing the positive or negative characteristics of cues in each channel (for example, they rated ironic criticisms as negative/unfriendly and ironic compliments as positive/friendly when ignoring the statement, but with the opposite valence when ignoring the prosody). However, when processed holistically, prosody led to stronger differentiation of literal vs. ironic utterances when the speaker was being critical/sarcastic, and the statement was positive (You’re such a

great driver!), than when the speaker was teasing and the statement was negative (You're such a lousy driver!). Thus, while it can be said that prosody is sufficient to point listeners to a nonliteral interpretation for both types of irony in the absence of other cues, listeners seem to accord greater perceptual weight to the negative semantic content of ironic compliments and less to the speaker's stance for this type of remark (see also, Carretié, Mercado, Tapia, & Hinojosa, 2001; Kreuz & Link, 2002; Kumon-Nakamura et al., 1995; Leary, 2000). This raises the possibility that distinct neurocognitive mechanisms act on prosody when processing ironic criticisms and ironic compliments (Caillies et al., 2019). Moreover, the *time course* of prosodic effects on the neurocognitive system may be unique as listeners build an ironic interpretation for positive and negative statements, although this question has not been comprehensively addressed.

Neurocognitive Studies of Verbal Irony

To characterize how listeners incrementally construct representations of ironic utterances in daily interactions, it is necessary to establish a time course of irony processing that considers the uptake of prosodic cues that express different types of verbal irony (e.g., criticisms, compliments) and different time intervals during which listeners encounter pragmatic markers of irony and integrate them with linguistic information. Event-related potentials (ERPs) are well suited to this task, because they allow us to capture fine-grained differences in cognitive function as prosodic cues are first registered and then integrated into an utterance representation in real time (Jiang & Pell, 2015; Rigoulot, Vergis, Jiang & Pell, 2020).

Most ERP studies of verbal irony have investigated how background context affects the processing of *written* sentences with possible ironic meanings, usually ironic criticism (positive statements). When target sentences are incongruent with previous descriptions of events and imply irony, an increased P600/late positivity has been reported from the onset of a critical word in ironic versus literal sentences (Cornejo et al., 2007; Regel, Gunter, & Friederici, 2011; Regel, Meyer, & Gunter, 2014; Spotorno, Cheylus, Van Der Henst, & Noveck, 2013; Weissman & Tanner, 2018). Some work also suggests that irony modulates the N400 under certain conditions (Caillies et al., 2019; Cornejo et al., 2007; Filik, Leuthold, Wallington, & Page, 2014; Kowatch, Whalen, & Pexman, 2013) or may even be registered earlier (in the 200–400ms time window, Caffara et al., 2019; Filik et al., 2014; Regel et al. 2010, 2011, 2014; Spotorno et al., 2013). Typically, N400 and P600 effects are elicited by mismatches in meaning, with larger peaks attributed to greater processing effort and/or unexpected information (Brouwer, Crocker, Venhuizen, & Hoeks, 2017; Brouwer, Fitz, & Hoeks, 2012; Kutas & Federmeier, 2010; Thierry, Berkum, Brouwer, & Crocker, 2017). During irony interpretation, it

was proposed that the P600 reflects a late process of pragmatic inference and reanalysis, thus demonstrating a time window in which listeners reintegrate the meaning of the semantic content with extralinguistic information to arrive at (more effortful) nonliteral meanings (Regel et al., 2011). However, for the most part, these claims are based on how contextual knowledge—i.e., information gleaned from a verbal description of events—impacts irony processing, often in the visual modality.

Increasingly, the impact of prosody on ERP responses to orally produced ironic remarks is being studied (Caillies et al., 2019; Filik et al., 2014; Regel et al., 2011; see Matsui et al., 2016 for fMRI data). Regel et al.'s (2011) influential study included a condition in which both the preceding context (a written discourse) and differences in prosody (sarcastic vs. “normal” voice) were manipulated. Participants performed a task that queried their understanding of the background context. They found no influence of prosody, only context, on the P600 elicited by the critical word (ironic > literal). While informative, these methods (which were similar in Filik et al., 2014) were likely insensitive to the effects of prosody due to their participants' attentional focus to the background context and its task-relevance. This design is known to minimize the salience of prosodic information during irony processing (Deliens et al., 2018). These arguments are justified by recent data which presented literal and ironic statements in the complete absence of any background context (Caillies et al., 2019). When participants could use only prosody to arrive at ironic interpretations, and when they focused on the intention of the speaker (*does the speaker think what he says?*), significant prosody-related changes were observed in both the N400 and P600 components to the target word (the authors did not look at any other time windows). These findings call for new studies which track the effects of prosody during irony processing with even greater precision.

Measuring the neural response at the end of an utterance—the approach taken to date—is effective for illuminating how expectations derived from various types of context (including prosody) alter semantic processing of the critical word to create implicatures or pragmatic inferences about ironic meaning (Regel et al., 2011). However, such measures are unlikely to capture effects of prosody that signal the speaker's stance and intentions as they first emerge to the listener, nor how these effects evolve to hypothetically construct an initial representation of ironic intent. Documenting this dynamic process will require analysis of ERPs at multiple time points during irony processing (Kowatch et al., 2013). In an adjacent literature, it has been shown that vocal expressions that encode various facets of a speaker's affective or mental state are registered rapidly and automatically by listeners directly from speech onset (Jiang, Gossack-Keenan, & Pell, 2020; Jiang & Pell, 2015; Paulmann & Kotz, 2008). Notably, motivationally salient prosodic cues increase the P200 amplitude from

utterance onset in a range of communicative contexts, depending on their potential relevance to the listener and/or the task they are engaged in (Hajcak, Weinberg, MacNamara, & Foti, 2012; Liu, Rigoulot, & Pell, 2015; Paulmann & Kotz, 2008; Pell et al., 2015). Ongoing cognitive analysis and elaboration of prosodic meanings over time can promote differences in the late positivity, within the 300-ms to 800-ms time window following speech onset (Hajcak et al., 2012; Pell et al., 2015).

While none of this work has considered ironic speech, it can be predicted that prosodic features that convey a speaker's affective stance when communicating irony also are registered by listeners from a very early time point. These effects would produce ERP differences (P200, late positivity) *well before* the semantically disambiguating target word is encountered at the end of the utterance. Moreover, if prosodic cues are highly salient and associated with expressing specific forms of irony (e.g., sarcasm; Cheang & Pell, 2008), it may be found that these cues constrain an ironic interpretation of the utterance from an early time point, as predicted by *parallel-constraint-satisfaction* models and recent data (Deliens et al., 2018; Katz, Blasko & Kazmerski, 2004; Pexman, 2008). Our study set out to test these ideas.

Objectives

This study was designed to draw a time course of ironic speech processing by illuminating the role of prosody at early and late stages of constructing a representation of the speaker's stance and (non) literal intention. We studied two distinct and potentially "asymmetrical" types of familiar irony: ironic criticisms (sarcasm) and ironic compliments (teasing). Our task required listeners to attend cues that would allow them to *socially* evaluate the speaker (in terms of friendliness), rather than focus on linguistic comprehension. We predicted that listeners would immediately register prosodic attributes referring to the speaker's positive or negative stance in literal and ironic statements, modulating the P200 and/or late positive component from utterance onset. Due to its task relevance, prosodic cues signaling positive stance should increase P200 responses, while the ongoing monitoring of negative, threatening stance would be indexed by a later sustained positivity (Paulmann & Kotz, 2008; Pell et al., 2015; Wang, Bastiaansen, & Yang, 2015). We speculated that prosody might reveal more than just a speaker's stance and inform listeners about their actual ironic intention at an early timepoint, especially for sarcasm, which is highly salient in the vocal channel (Bryant, 2010; Mauchand et al., 2018).

Once registered by the cognitive system, we expected meaningful prosodic contrasts to shape neural responses to the target word (e.g., *great* or *horrible* driver), at which point initial expectations created by prosody are confirmed or disconfirmed by the listener. These responses would thus depend on how prosody is processed from utterance onset, and

whether participants only registered speakers' stance (thus expecting a congruent, literal target) or also identified speakers' intentions (literal or ironic). Greater difficulty integrating the two information sources when speakers are pragmatically insincere would increase the P600 for ironic as opposed to literal target words (Regel et al., 2011, 2014). Ironic meanings could also increase the semantic N400 (Cornejo et al., 2007; Filik et al., 2014) or possibly earlier attentional processing stages (e.g., P200; Filik et al., 2014; Regel et al., 2011, 2014). On the other hand, if prosody supplies salient cues to the possible ironic intention of the speaker from utterance onset, operations for accessing and contextually integrating the critical word in ironic versus literal statements may not lead to differential processing demands in the P200, N400, and/or P600 time windows or may show a reversal in previously reported patterns due to on-line *facilitation* of ironic meanings at the critical word (Caillies et al., 2019; Deliens et al., 2018; Kowatch et al., 2013; Regel et al., 2010).

Methods

Participants

Thirty native English speakers (12 males and 18 females, age: $M = 22.4$ years, $SD = 3.7$) were recruited on the campus of McGill University, following a power analysis to determine the sample size (run with G*power, Faul, Erdfelder, Lang, & Buchner, 2007).¹ All participants were right-handed and reported no history of major psychiatric or neurological illness or speech/hearing problems. Participants voluntarily consented to take part in the study that was ethically approved by the Faculty of Medicine Institutional Review Board (McGill University, Montreal, Canada).

Stimuli

Utterances were taken from a recording database constructed for a previous study (see Mauchand et al., 2020 for full details). Stimuli were based on 48 statements expressing a judgement addressed to the listener in the form, "*You are such an -adjective- -noun-*." Half of the statements were positive (*You are such a great cook*), and the other half were negative (*You are such a terrible cook*). Each were formed by substituting the valence of the adjective in the corresponding root sentence. These sentence pairs were taken from a previous study (Vergis & Terkourafi, 2015), in which they were matched for offensiveness, emotional damage, and emotional

¹ Assuming medium -effect sizes (Jiang & Pell, 2015), 24 participants were required to achieve power over 0.8 with ANOVAs. As because our analyses were actually done using linear mixed effects models, this estimation is assumed to be a lower bound of the actual power achieved by the study (Kuznetsova et al., 2017).

Table 1 Definition of experimental conditions in the study with examples of stimuli (ironic conditions are highlighted in shaded cells)

| | | Prosody (stance) | |
|-----------|---|--|---|
| | | Positive | Negative |
| Statement | Positive <i>You are such a great cook</i> | Literal Compliment | Ironic Criticism <i>(e.g. sarcasm)</i> |
| | Negative <i>You are such a horrible cook</i> | Ironic Compliment <i>(e.g. teasing)</i> | Literal Criticism |

state of the speaker. There were 12 positive adjectives (syllables: $M = 2.83$, standard deviation [sd] = 1.27; frequency: $M = 9.24$, $sd = 1.69$) and 12 negative adjectives (syllables: $M = 2.50$, $sd = 0.90$; frequency: $M = 8.79$, $sd = 0.92$), each used in two sentences. There was no significant difference in the frequency ($t(16.96) = -0.81$, $p = 0.428$) or syllable length ($t(19.90) = 0.74$, $p = 0.467$) of positive versus negative adjectives.²

Each sentence was uttered in a literal and ironic manner by four different speakers (2 males and 2 females) in a sound-attenuated testing booth with a head-mounted microphone. During recording, speakers were instructed to express each statement in a way that was natural so that listeners would understand their intended meaning (literal or ironic). These methods produced 24 token sets, which varied in the valence of the statement (positive, negative) and its prosody (positive, negative) to communicate literal versus ironic compliments and criticisms (Table 1). Given that the stimuli were not created for the purpose of ERPs, and key words were not precisely matched in structure and duration between conditions, any remaining variability in the recorded duration of key words was rectified at the ERP analysis stage by realigning latencies across ERP events through Residue Iteration Decomposition (Ouyang, Sommer, & Zhou, 2016), as described below.

Utterances were perceptually validated by 20 English-speaking Canadian participants in a previous online study (Mauchand et al., 2020) using the recruiting platform Prolific Academic (Peer, Brandimarte, Samat, & Acquisti, 2017) and the LimeSurvey online testing software (Schmitz, 2012). For each utterance, participants answered two questions evaluating the literality of the statement (“Does the speaker mean what they say?”) and the attitude of the speaker (“Is the attitude of the speaker positive?”) on 5-point Likert scales from “Not at all” to “Very much.” This validation showed that on the literality scale, literal compliments and criticisms were predictably rated very high and ironic

criticisms (sarcasm) was rated very low; ironic compliment (teasing) evaluations fell in between. In terms of attitude, literal compliments were rated as very positive and literal criticisms as very negative; ironic criticisms ratings fell around the middle of the scale, but ironic compliments were perceived as conveying a slightly more negative attitude (Mauchand et al., 2020). The token sets that best fitted the expectation of each type of attitude (e.g., low literality and low positivity for ironic criticisms) while maximizing the rating differences between an ironic statement and its literal counterpart were selected. Nine different root sentences (token sets) were selected per speaker, resulting in minimal repetition of particular statements in the final stimulus set (and no direct repetition of any stimulus since tokens were produced by different speakers). In total, 144 acoustically unique stimuli conveying literal vs. ironic intentions were used in the experiment (4 speakers \times 9 sentence roots \times 4 utterances per token set). For the purpose of a broader study, these statements were presented along with 144 filler stimuli (short requests), which varied in expressed politeness.³ When both experimental and filler stimuli are considered, ironic and literal statements each comprised 25% of the total stimuli presented. The acoustic onset/offset of each stimulus were precisely marked using Praat (Boersma & van Heuven, 2001), and each .wav audio file was normalized to a peak intensity of 70 dB to control for slight differences in sound recording levels.

Acoustic information about the stimuli is summarized in Table 2. Analyses of acoustic parameters derived across the full utterance suggest that a speaker's stance influenced their mean fundamental frequency/F0M ($F(1, 140) = 13.22$, $p <$

² Log-transformed HAL frequency norms taken from the English Lexicon Project (Balota et al., 2007)

³ Filler stimuli were interpersonal requests of similar duration to the irony stimuli (e.g., Please/lend me a nickel), recorded by a different group of 2two male and 2two female speakers. Like the irony stimuli, the 144 filler items were factorially manipulated according to the speaker's prosodic stance (positive, negative) and the type of Statement (requests with and without the explicit politeness marker “Please”). This means that when the entire stimulus set is considered, including both irony ($n = 144$) and filler politeness ($n = 144$) stimuli, half of all tokens conveyed positive vs. negative stance at the prosodic level, and half contained linguistic structures that were positive vs. negatively valenced. Construction of the politeness stimuli is described by Vergis & Pell (2020).

Table 2 Acoustic features of the selected stimuli, for each utterances type produced by males and females

| Sex | Utterance type | F0M (Hz) | | F0SD (Hz) | | IntSD (dB) | | HNR | | Duration (s) | |
|---------|--------------------|----------|-------|-----------|-------|------------|------|------|------|--------------|------|
| | | M | SD | M | SD | M | SD | M | SD | M | SD |
| Females | Literal Compliment | 235.45 | 27.93 | 57.81 | 20.31 | 8.31 | 1.03 | 9.02 | 1.31 | 2.34 | 0.29 |
| | Literal Criticism | 179.65 | 29.99 | 32.53 | 20.75 | 7.23 | 1.11 | 7.99 | 2.82 | 1.97 | 0.34 |
| | Ironic Compliment | 210.10 | 42.14 | 55.32 | 31.59 | 7.42 | 0.99 | 9.45 | 1.91 | 1.98 | 0.37 |
| | Ironic Criticism | 171.78 | 20.16 | 24.26 | 10.92 | 7.07 | 1.04 | 7.53 | 2.03 | 2.58 | 0.48 |
| Males | Literal Compliment | 131.63 | 17.47 | 30.58 | 8.76 | 7.51 | 1.21 | 6.65 | 1.83 | 1.96 | 0.35 |
| | Literal Criticism | 137.96 | 14.55 | 35.05 | 12.17 | 8.40 | 1.44 | 6.50 | 2.57 | 2.02 | 0.52 |
| | Ironic Compliment | 128.61 | 21.98 | 29.06 | 13.87 | 8.02 | 1.19 | 7.29 | 1.96 | 2.08 | 0.32 |
| | Ironic Criticism | 128.16 | 27.77 | 26.55 | 13.43 | 8.43 | 1.38 | 6.89 | 2.98 | 2.85 | 0.59 |

F0M mean fundamental frequency, F0SD standard deviation of the fundamental frequency, IntSD standard deviation of Intensity, HNR Harmonics-to-Noise Ratio

0.001), fundamental frequency variability/F0SD ($F(1, 240) = 12.56, p < 0.001$), and Harmonics-to-Noise Ratio/HNR ($F(1, 140) = 7.67, p = 0.006$). Perceptually, unfriendly utterances had lower pitch, restricted pitch variability, and increased noise in the pitch signal compared to friendly utterances. Speaker intention also was represented by reductions in F0M ($F(1, 140) = 7.56, p = 0.007$) and F0SD ($F(1, 140) = 4.40, p = 0.038$) for ironic compared with literal utterances. Speaker stance and intention conjointly affected utterance duration ($F(1, 140) = 50.27, p < 0.001$) such that ironic criticisms were longer than all other utterances. These results roughly correspond to a previous acoustic analyses using predictive methods on a restricted set of the stimuli (Mauchand et al., 2018), where it was demonstrated that only reduced F0SD and increased duration accurately differentiated ironic criticisms from the other utterances. Also, perceptual data for these stimuli revealed that the prosody of each of the four intentions expressed in the current study was accurately discriminated and identified by participants based on their prosody, regardless of the verbal statement produced (Mauchand et al., 2020). Exemplars of the literal/ironic stimuli are available through the Open Science Framework (Foster & Deardorff, 2017).⁴

Task and EEG Recording Procedure

The experiment was conducted in an electrically shielded, sound-attenuating booth. After electrode preparation, participants were seated in a comfortable chair in front of a computer. Stimuli from all conditions were intermixed and presented over headphones in a pseudorandomized order that prevented direct repetition of tokens from a given speaker or statements from the same token set. For each trial, the target stimulus was

preceded by a fixation point of jittered duration (500ms to 1500ms). After hearing the statement, participants were prompted to answer the question “How friendly is the speaker?” using a 5-point scale from *Not at all* to *Very much* that appeared 500 ms after stimulus offset. Friendliness ratings were chosen to focus listeners’ attention to the affective stance of the speaker without drawing explicit attention to the notion of ironic versus literal meanings (see also Mauchand et al., 2020). Responses were recorded with a 5-button response box programmed to correspond to a visual scale shown on the computer screen. The order of the scale was reversed for half of the participants. Trials ended after a response from the participant or 5 seconds, and the next trial started after a 1,500-ms blank screen. The experiment always began with eight practice trials, which did not appear in the main experiment.

While performing the task, the electroencephalograms (EEGs) were recorded continuously from 64 Ag/AgCl electrodes using the ActiCap System (Brain Products, Germany). The vertical electrooculograms (VEOG) were recorded from above and below the right eye and the horizontal electrooculograms (HEOG) were recorded from the outer canthus of both eyes. The recordings were online referenced to FCz and re-referenced offline to the bilateral mastoids. The EEGs were digitized at 500 Hz and filtered with a band-pass from 0.016 Hz to 100 Hz. After the EEG experiment, participants completed short questionnaires to assess their demographic characteristics and level of social anxiety for the purpose of a companion study. The whole session lasted approximately 2.5 hours, including EEG preparation and completion of the post-tests. Participants were compensated a small amount at the end of the study.

EEG Data Processing

All pre-processing procedures were performed using EEGLAB (Delorme & Makeig, 2004) and ERPLAB

⁴ https://osf.io/4geun/?view_only=77cdfba3fed744b6b4eef39976048615. If you wish to access the whole stimulus set for research purposes, please contact the Neuropragmatics and Emotion lab at McGill University (https://www.mcgill.ca/pell_lab/).

(Lopez-Calderon & Luck, 2014). The continuous EEGs were first visually inspected. Signals with excessive movement artifact, alpha activity, or amplifier saturation were manually excluded from the analysis. The subsequent EEGs were filtered using a 40-Hz low-pass and a 0.1-Hz high-pass Butterworth of the fourth order and then decomposed with an ICA algorithm (Makeig, Bell, Jung, & Sejnowski, 1996) to remove ocular artifacts. Given our hypotheses about the time course of prosody-related effects, we defined two separate epochs. The first, time-locked to the acoustic onset of the utterance (200–1,200 ms), examined how prosody influences irony processing before the target word. The second, time-locked to the onset of the positive or negative target adjective (–200 to 800 ms), should reveal how prosody is integrated with semantic information to confirm ironic intentions (virtually all previous studies of irony have only examined this later processing interval). Epochs were baseline corrected based on the mean EEG activity in the respective prestimulus interval. Segments with signal peak-to-peak voltage exceeding 100 mV within a 200-ms sliding window in steps of 100 ms were automatically rejected. According to a predefined criterion, any subject with more than 40% trials rejected in any of the four matching stimulus conditions (literal compliment, ironic criticism, literal criticism, ironic compliment) in either of the two time-locked epochs (utterance onset, target word onset) was excluded from further analysis. Four subjects were excluded for this reason. On average, this left approximately 30 trials per speaker per condition for measures derived at utterance onset (literal compliments: $M = 29.85$, $sd = 5.26$; ironic criticisms: $M = 29.92$, $sd = 5.31$; literal compliments: $M = 30.19$, $sd = 5.59$; ironic compliments: $M = 30.58$, $sd = 4.21$) and approximately 31 trials per speaker per condition for measures taken at the target word onset (literal compliments: $M = 30.85$, $sd = 5.21$; ironic criticisms: $M = 31.35$, $sd = 4.47$; literal compliments: $M = 30.88$, $sd = 5.01$; ironic compliments: $M = 30.96$, $sd = 4.00$).

ERP Analysis

As trial-to-trial latency jitter creates single-trial variability that reduces component discrimination, attenuates component amplitudes, and may yield erroneously significant effects, a RIDE (Residue Iteration Decomposition) procedure was performed on the ERP data before analysis (Ouyang et al., 2016). RIDE uses the latency variability and time markers to separate ERP components into predicted component clusters with a stimulus-locked cluster S (corresponding here to N100 and P200) and one or more central clusters with unknown latency C or C1 and C2, as will be further elaborated below (e.g., corresponding to later components N400, P600, or late positivity). RIDE was performed independently on all EEG epochs, with a different setting for utterance onset-locked epochs and target word-locked epochs. For utterance onset-

locked epochs, the S cluster had a time window of 0–400 ms, and there was one C cluster with a time window of 100–900 ms. For target word-locked epochs, the S cluster had a time window of 0–400 ms, then a C1 cluster corresponding to N400 with a window of 200–600 ms and a C2 cluster corresponding to P600 with a window of 400–800 ms. The latency of S for each trial was set to be locked to the stimulus onset. The latency of C clusters was first estimated by Woody's method within the time windows. Then, ERPs were subjected to RIDE into the component clusters associated with the latency sets; these two steps were iterated until convergence. After resynchronization of the subcomponent clusters to their own latency across single trials, ERPs were reconstructed accounting for variability of latency across trials (Ouyang, Herzmann, Zhou, & Sommer, 2011; Ouyang, Sommer, & Zhou, 2015; Ouyang et al., 2016). From the resulting ERPs, we extracted the mean amplitude calculated per subject per condition in relevant time windows from the two onsets. From utterance onset (prosody-related effects): 230–290 ms for P200 (determined with peak detection); and 600–1,000 ms for late positivity. From the critical word onset (prosody x statement effects): 230–290 ms for the P200; 300–500 ms for N400; and 450–800 ms for P600.

Statistical Analysis

Statistical analysis of measures taken from both utterance onset and target word onset was conducted in two stages. First, we separately analyzed responses to the positive statements (ironic criticisms and literal compliments) and the negative statements (ironic compliments and literal criticisms). Given that most research has looked only at ironic criticism (i.e., positive statements), and the weighting of prosodic cues when processing ironic criticism and ironic compliments may be “asymmetrical” (Matthews et al., 2006; Mauchand et al., 2020; Pexman & Olineck, 2002), analyzing each type of irony separately allowed us to compare how a speaker's positive/negative vocal stance influences neural responses in potentially unique contexts of “sarcasm” and “teasing.” At a second stage, which considered the entire dataset, we recoded our stimuli according to the speaker's intention to be literal or ironic. These analyses were designed to shed light on how the speaker's stance (prosody) and their statement (semantic choice) each affect ERP responses according to the speaker's intention to communicate in an ironic manner.

Linear Mixed-Effects Models (LMEM) were built, considering the effects of Prosody (positive, negative), Statement (positive, negative), and/or Intention (literal, ironic), according to the analysis. We included Region of Interest (ROI), defined by 9 ROIs represented each by 6–8 electrodes (Jiang & Pell 2015). The ROIs were: left anterior (AF3, FP1, F7, F5, F3, FT7, FC5, FC3), left central (T7, C5, C3, TP7, CP5, CP3), left posterior (P7, P5, P3, PO9, PO7, PO3), medial anterior

(F1, FZ, F2, FC1, FCZ, FC2), medial central (C1, CZ, C2, CP1, CPZ, CP200), medial posterior (P1, PZ, P2, O1, POZ, O2), right anterior (AF4, FP2, F4, F6, F8, FC4, FC6, FT8), right central (C4, C6, T8, CP4, CP6, TP8), and right posterior (P4, P6, P8, PO4, PO8, PO10). Participants and channels were included as random factors. Models were examined using F tests for main effects and interactions and t tests for specific contrasts, using the Tukey correction. All analyses were performed in R-studio (R Version 3.4.3, <http://cran.r-project.org>) with the *lme4* (Bates, Machler, Bolker, & Walker, 2015) and *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2017) packages.

Results

Behavioral Results: Impressions of Speaker Friendliness

LMEMs were built to analyze the friendliness ratings given to speakers when they produced literal versus ironic utterances, according to their Prosody (positive, negative) and Statement (positive, negative), including participants as random intercepts. Overall, speakers were perceived as friendlier when they made a positive versus negative statement ($F(1, 3281.1) = 851.23, p < 0.001, \beta = 1.00, SE = 0.03, p < 0.001$), and when they used a positive versus negative sounding prosody ($F(1, 3280) = 1125.94, p < 0.001, \beta = 1.15, SE = 0.03, p < 0.001$). Speaker impressions were significantly influenced by the interaction of the two factors (Statement \times Prosody: $F(1, 3280.5) = 22.94, p < 0.001$). While a positive-sounding prosody increased impressions of friendliness for all statements, this effect was significantly smaller when the statement was negative ($\beta = 0.98, SE = 0.05, t(3279.64) = 20.47, p < 0.001$) versus positive ($\beta = 1.31, SE = 0.05, t(3280.88) = 26.94, p < 0.001$). This implies that when the utterance was phrased negatively, prosodic cues marking the speaker's stance had less influence on decisions about whether the speaker meant to be

friendly when making the remark (Figure 1). It is noteworthy that mean ratings of ironic compliments (negative statement-positive prosody) fell below the midpoint of the scale ($M = 2.93, SE = 0.14$), suggesting that these utterances were generally evaluated as *unfriendly* or negative.

ERP Effects from Utterance Onset

Separate models were first built to examine ERPs to positive statements (literal compliment vs. ironic criticism) versus negative statements (literal criticism vs. ironic compliment). Focusing on measures derived from utterance onset, we examined how speaker stance was encoded from prosodic information in theoretically motivated time windows *prior* to the onset of the critical target word. New models were then built on ERPs derived from the adjective onset to understand the contextual effects of prosody on semantic processing of the critical target word. The effect of prosody at early and late time points during irony processing is summarized in Figures 2 and 3, respectively.

Positive Statement (Literal Compliment/Ironic Criticism)

In the 230-ms to 290-ms time window, the P200 amplitude was influenced by Prosody ($F(1, 2933) = 192.20, p < 0.001$), significant at all anterior ROIs (Prosody \times ROI, $F(8, 2933) = 2.91, p < 0.001$: left: $\beta = -1.79, SE = 0.24, t(2932.98) = -7.32, p < 0.001$; medial: $\beta = -2.10, SE = 0.28, t(2932.98) = -7.43, p < 0.001$; right: $\beta = -1.80, SE = 0.24, t(2932.98) = -7.37, p < 0.001$). The P200 response increased when the speaker's stance was positive (literal compliment) versus negative (ironic criticism, $\beta = -1.27, SE = 0.09, t(2932.98) = -13.86, p < 0.001$). Differences in Prosody also were registered in the late positivity time window (600-1,000 ms, $F(1, 2990) = 10.61, p = 0.001$). As shown in Figure 2a, negative prosody (ironic criticism) was associated with an increased late positivity response ($\beta = 0.29, SE = 0.09, t(2990) = 3.26, p = 0.001$).

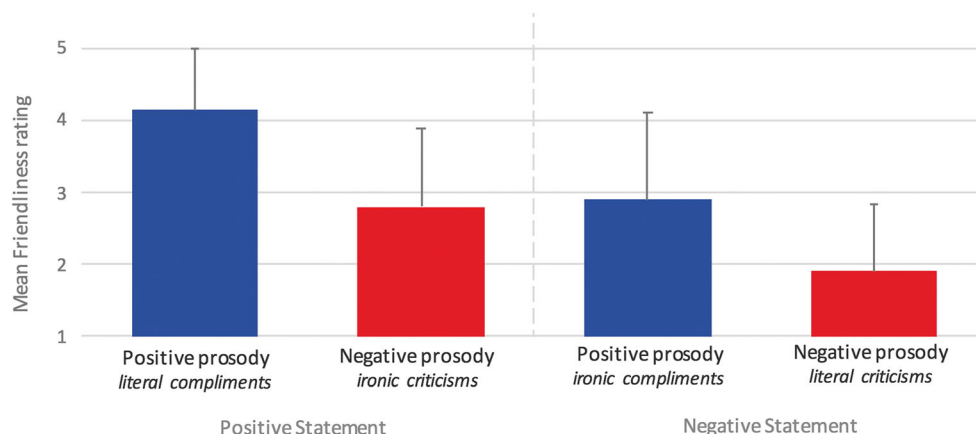


Fig. 1 Average friendliness ratings (and standard deviations) for each type of utterance

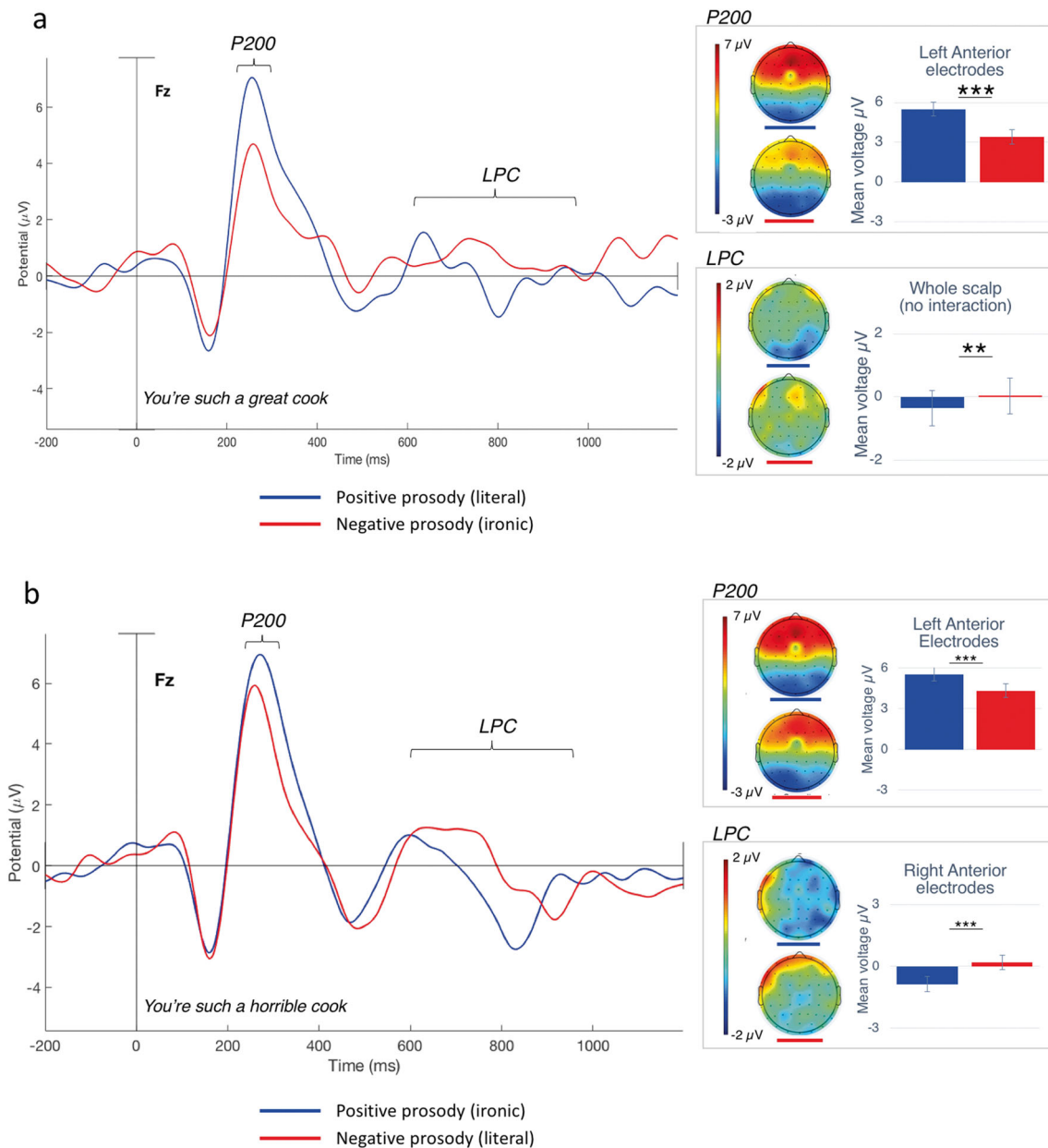


Fig. 2 Effects of prosody on ERPs time-locked to the utterance onset when speakers made: **a** Positive statements (literal compliments vs. ironic criticisms); and **b** Negative statements (literal criticisms vs. ironic compliments). The right panels show the scalp maps and average potentials for each type of utterance at each observed time-window (P200: 230-

290ms, LPC: 600-1000ms). The average potentials shown are extracted from the most significant ROI; if no interaction between ROIs and prosody was found, the average is from the whole scalp. Error bars indicate the standard error of the means

Negative Statement (Literal Criticism/Ironic Compliment)

Prosody again modulated the P200 component ($F(1, 2933) = 21.74, p < 0.001$), with increased amplitudes when the speaker's expressed stance was positive (ironic compliment) than negative (literal criticism, $\beta = 0.39, SE = 0.08, t(2932.98) = 4.66, p < 0.001$). Prosody-related effects on the P200 were significant at left anterior ($\beta = 1.24, SE = 0.22, t(2932.98) = -5.56, p < 0.001$) and surrounding ROIs (Prosody \times ROI, $F(8,$

$2933) = 4.36, p < 0.001$). Analysis of the late positivity revealed effects of Prosody ($F(1, 2990) = 24.34, p < 0.001$) and Prosody \times ROI ($F(8, 2990) = 2.33, p = 0.017$). Prosodic information conveying negative stance, associated with literal criticisms in this context, evoked a more positive-going wave as the utterance unfolded ($\beta = -0.42, SE = 0.09, t(2990) = -4.93, p < 0.001$).

This effect was detectable at right anterior ($\beta = -1.06, SE = 0.23, t(2990) = -4.66, p < 0.001$) and right central ($\beta = 0.71,$

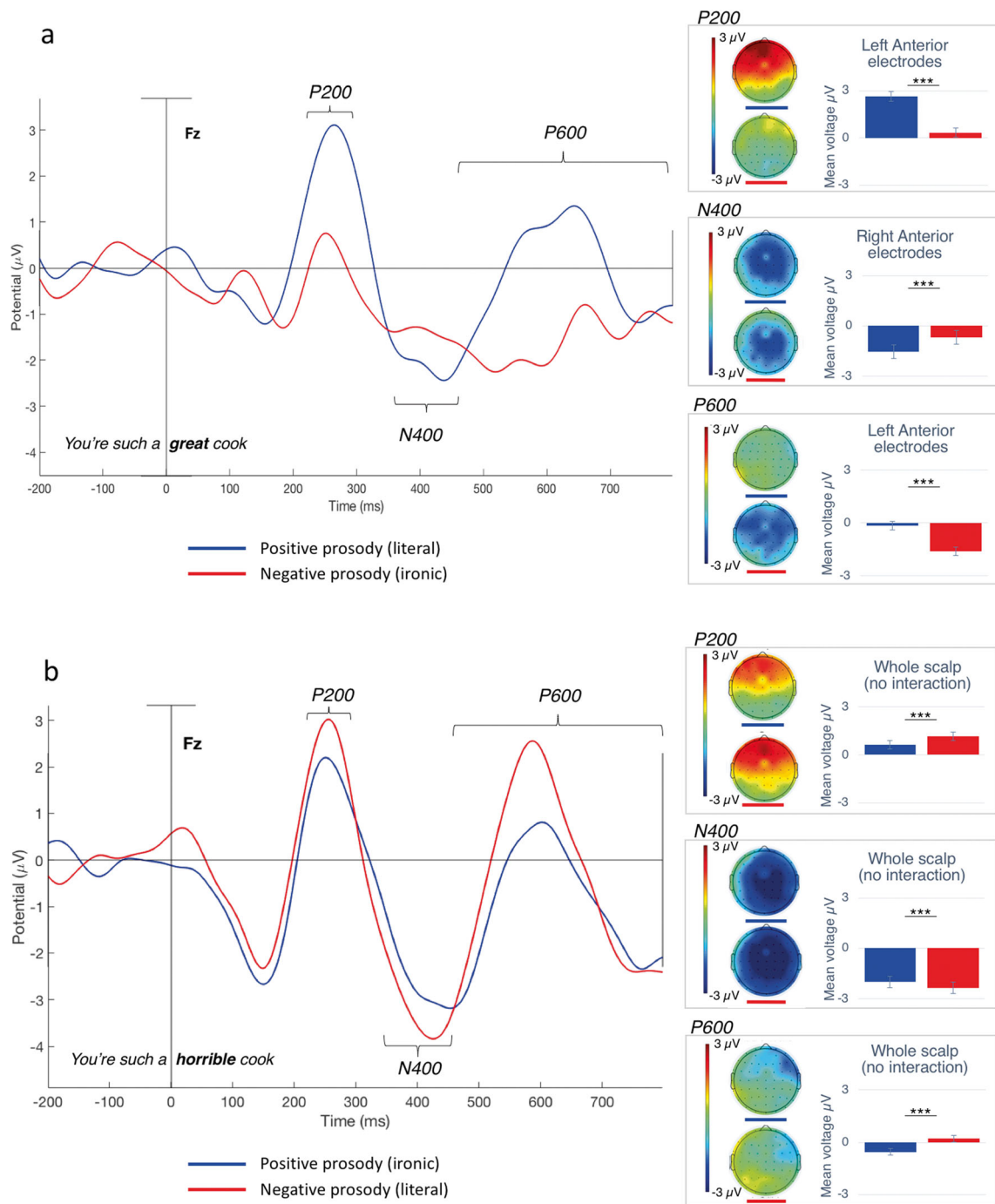


Fig. 3 Effects of prosody on ERPs time-locked to the critical word onset when speakers made: **a** Positive statements (literal compliments vs. ironic criticisms); and **b** Negative statements (literal criticisms vs. ironic compliments). The right panels show the scalp maps and average potentials for each type of utterance at each observed time window (P200: 230-

290ms, N400: 350-450ms, P600:450-800ms). The average potentials shown are extracted from the most significant ROI; if no interaction between ROIs and prosody was found, the average is from the whole scalp. Error bars indicate standard error of the means

SE = 0.26, $t(2990) = 2.72$, $p = 0.007$) ROIs. These results are displayed in Figure 2b.⁵

⁵ Following research on emotional prosody (e.g., Pell et al., 2015), we explored potential effects of prosody in the N100 time window (130-160 ms, determined by peak detection). No significant effect of Prosody on the N100 was found for either analysis (p 's > 0.70).

ERP Effects from Critical Word Onset

Positive Statement (Literal Compliment/Ironic Criticism)

The speaker's prosody significantly modulated P200, N400, and P600 amplitudes evoked by the target adjective (P200:

$F(1,2990) = 417.51, p < 0.001$; N400: $F(1, 2933) = 7.76, p = 0.005$; P600: $F(1, 2990) = 379.70, p < 0.001$). For each component, an increase in the neural response was observed when the speaker's stance also was positive (literal compliment) versus negative (ironic criticism)(P200: $\beta = -1.42, SE = 0.07, t(2990) = -20.4, p < 0.001$); N400: $\beta = -0.22, SE = 0.08, t(2932.98) = -2.78, p < 0.005$; P600: $\beta = -1.18, SE = 0.06, t(2990) = -19.49, p < 0.001$). The P200 effect was significant in all ROIs but largest at left anterior ($\beta = -2.31, SE = 0.19, t(2990) = -12.40, p < 0.001$) and surrounding ROIs (Prosody x ROI, $F(8, 2990) = 7.58, p < 0.001$). The N400 effect was anteriorly distributed (Prosody x ROI, $F(8, 2933) = 2.70, p = 0.006$), largest at right anterior electrodes ($\beta = -0.85, SE = 0.21, t(2932.98) = -3.96, p < 0.001$). Prosodic effects on the P600 were broadly distributed but most pronounced at left anterior ($\beta = -1.45, SE = 0.16, t(2990) = -8.99, p < 0.001$), medial anterior ($\beta = -1.44, SE = 0.19, t(2990) = -7.72, p < 0.001$), and medial central ($\beta = -1.63, SE = 0.19, t(2990) = -8.76, p < 0.001$) sites. These patterns are displayed in Figure 3a.

Negative Statement (Literal Criticism/Ironic Compliment)

Prosody significantly modulated P200, N400, and P600 amplitudes to the target adjective when speakers made a negative statement (P200: $F(1, 2933) = 83.35, p < 0.001$; N400: $F(1, 2932.98) = 22.36, p < 0.001$; P600: $F(1, 2932.95) = 54.39, p < 0.001$). For each component, responses to the critical word increased when the speaker's stance was negative (literal criticism) versus positive (ironic compliment (P200: $\beta = 0.52, SE = 0.06, t(2932.97) = 9.13, p < 0.001$; N400: $\beta = 0.34, SE = 0.07, t(2932.96) = -4.15, p < 0.001$; P600: $\beta = -0.44, SE = 0.06, t(2932.95) = -7.28, p < 0.001$). These results are displayed in Figure 3b.

ERP Effects as a Function of Ironic Intention

Initial findings point to rapid differentiation of the speaker's prosody from utterance onset, with increased P200 amplitudes when speakers expressed a positive stance, followed by an increased late positivity to negative prosody as the utterance unfolded. At the critical target word, larger P200, N400, and P600 waves were observed when the speaker's stance was *congruent* in valence with their statement; that is, when the speaker was being "pragmatically sincere" and literally meant what they said (irrespective of the valence of their attitude). The impact of prosody was similar and robust when speakers made either positive or negative statements.

To better understand how a speaker's intention to be literal or ironic was influenced by our independent variables (Prosody, Statement), irrespective of the "type" of irony used, follow-up analyses were run on the full dataset grouping stimuli according to the intention of the speaker. The first analysis

considered potential early effects of Prosody (positive, negative) on the registration of speaker Intention (literal, ironic) from utterance onset. A second analysis, time-locked to the target word onset, examined effects of Statement type (positive, negative) on the processing of Intention at a late time point in the utterance after information from different cue sources had accumulated.

Utterance Onset (Prosody x Intention)

Focusing on novel effects involving differences in Intention⁶, it was found that literal utterances increased the P200 over ironic utterances overall (Intention: $F(1, 5949) = 44.99, \beta = 0.44, SE = 0.07, p < 0.001$), although this depended on the speaker's stance (Prosody x Intention, $F(1, 5949) = 7.39, p = 0.006$). The prosody-related P200 was larger for literal than ironic utterances only when stance was negative or critical ($\beta = 0.62, SE = 0.07, t(5949.00) = 6.67, p < 0.001$). No differences in intention were registered at this early stage when prosody was positive (i.e., literal vs. ironic compliments).

Interactive effects of Prosody and Intention also influenced the late positivity (Prosody x Intention, $F(1, 6006) = 10.08, p = 0.002$), significant at right and medial anterior electrodes (Prosody x Intention x ROI, $F(8, 6006) = 2.75, p = 0.005$). Here, when speakers displayed a positive stance, the intention to be literal was associated with a more positive-going wave in the 600- to 1,000-ms time window than when speakers meant to be ironic ($\beta = 0.27, SE = 0.09, t(6006.00) = 2.98, p = 0.003$). There was no impact of Intention on the late positivity when speakers displayed a negative stance ($\beta = -0.14, SE = 0.09, t(6006.00) = -1.52, p = 0.13$). These patterns imply that prosody allowed listeners to register the ironic intentions of the speaker in a salient manner directly from utterance onset; however, ironic meanings emerged at *distinct time points* depending on the speaker's stance (i.e., much earlier when prosody was negative, see Figure 4).

Critical Word Onset (Statement x Intention)

The appearance of the critical adjective (e.g., great/horrible) yielded a significant main effect of Intention in the P200 window ($F(1, 5949) = 399.63, \beta = -0.97, SE = 0.05, p < 0.001$), literal target words yielding increased P200 compared with ironic target words. A main effect of Statement was also found ($F(1, 5949) = 23.93, \beta = -0.24, SE = 0.05, p < 0.001$),

⁶ Consistent with data reported above, positive stance increased the P200 over negative stance (Stance effect: $F(1, 5949) = 158.54, \beta = 0.83, SE = 0.07, p < 0.001$) at anterior electrodes (left: $\beta = 1.51, SE = 0.17, t(5949.00) = 8.61, p < 0.001$; medial: $\beta = 1.42, SE = 0.20, t(5949.00) = 6.98, p < 0.001$; right: $\beta = 1.07, SE = 0.18, t(5949.00) = 6.09, p < 0.001$). Stance also affected the LPC, with negative prosody increasing the late positivity over positive prosody overall (Stance effect: $F(1, 6006) = 29.22, \beta = -0.35, SE = 0.07, p < 0.001$).

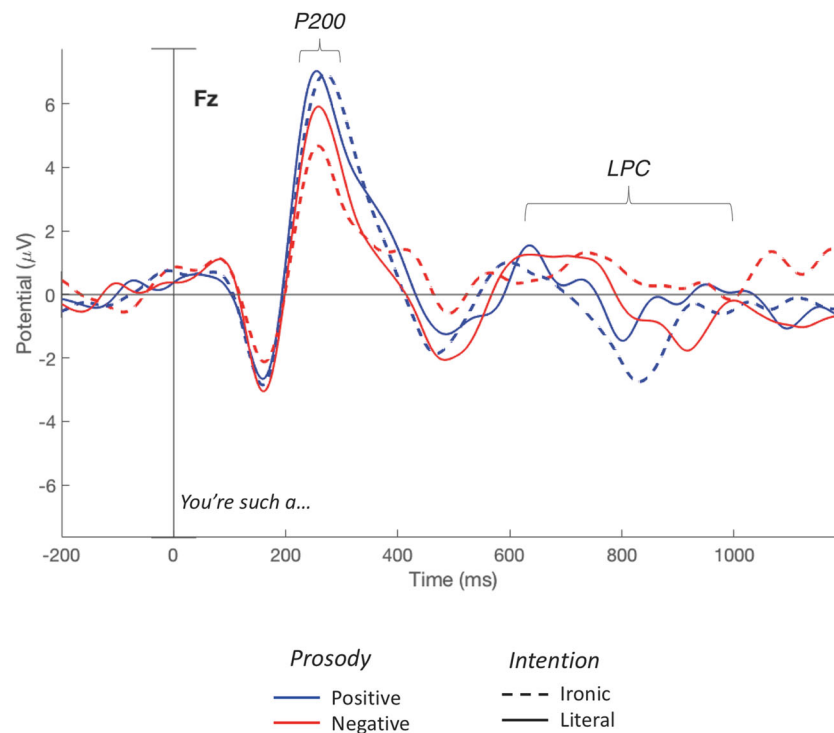


Fig. 4 Effects of the speaker's intention to be literal versus ironic on ERPs according to prosodic differences measured from the utterance onset

suggesting that negative adjectives increased P200 relative to positive adjectives, but a significant interaction ($F(1, 5949) = 86.10, p < 0.001$) revealed that this effect was only in ironic utterances, ($\beta = -0.69, SE = 0.07, t(5948.97) = -10.02, p < 0.001$), while literal utterances showed a smaller, opposite effect ($\beta = 0.21, SE = 0.07, t(5948.97) = 3.10, p = 0.002$).

In the N400 window, processing the critical adjective yielded significant main effects of Statement ($F(1, 5949) = 219.27, \beta = -0.91, SE = 0.06, p < 0.001$) and Intention ($F(1, 5949) = 21.15, \beta = -0.28, SE = 0.27, p < 0.001$) in the absence of an interaction ($p = 0.321$). Negative words markedly increased the N400 over positive words overall. Also, N400 to literal versus ironic target words was greater overall. The effect of Statement was widespread on the whole scalp, largest at medial central electrodes ($\beta = -1.42, SE = 0.19, t(5948.97) = -7.49, p < 0.001$). The effect of Intention was only present at anterior ROIs (left: $\beta = -0.62, SE = 0.16, t(5948.97) = -3.79, p < 0.001$; medial: $\beta = -0.56, SE = 0.19, t(5948.97) = -2.97, p = 0.003$; right: $\beta = -0.61, SE = 0.16, t(5948.97) = -3.73, p < 0.001$).

In general, the P600 displayed qualitatively similar tendencies to the N400 due to Statement ($F(1, 5949) = 75.24, \beta = -0.39, SE = 0.04, p < 0.001$) and Intention ($F(1, 5949) = 329.62, \beta = -0.81, SE = 0.04, p < 0.001$), but this analysis revealed an interaction of the two variables ($F(1, 5949) = 67.40, p < 0.001$). As shown in Figure 5, when the speaker made a negative statement, differences in Intention (literal > ironic) were significantly reduced in the P600 time window ($\beta = 0.44, SE = 0.06, t(5948.97) = 7.03, p < 0.001$) than when

speakers made a positive statement ($\beta = 1.18, SE = 0.06, t(5948.97) = 18.64, p < 0.001$). These results could reflect an 'asymmetry' in how prosody is weighted at late interpretative stages when speakers make a negative versus positive statement (Bruntsch & Ruch, 2017; Mauchand et al., 2020).

Discussion

Our findings provide new insights into how ironic speech is processed, illuminating on-line effects of prosody on the neurocognitive system and their time course as listeners evaluate short compliments and criticisms. By examining neural responses at the acoustic onset of the utterance and the onset of target words that listeners use to decipher whether the speaker is being ironic, our data begin to reveal how mental impressions of interpersonal stance and speaker intentions are incrementally formed, based on the immediate use of prosodic information that serves as a pragmatic marker for the listener.

Registering Stance and Ironic Intentions from Utterance Onset

The role of prosody in expressing a speaker's stance in discourse has long been discussed (Argyle et al., 1971; Argyle, Salter, Nicholson, Williams, & Burgess, 1970; Pell et al., 2018), but is only now being operationalized in neurocognitive studies of spoken language. According to

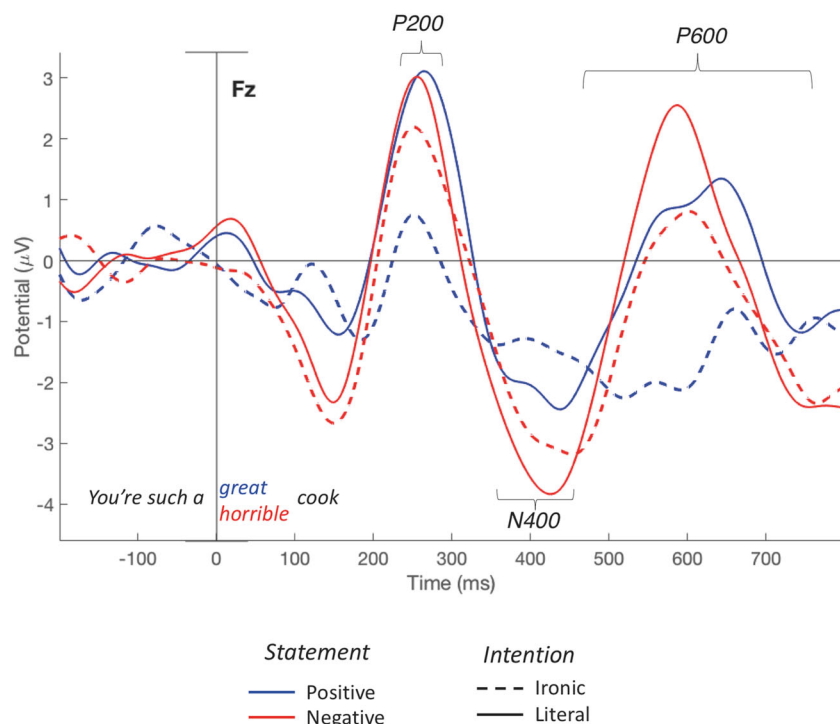


Fig. 5 Effects of the speaker's intention to be literal versus ironic on ERPs according to differences in statement type measured from the critical word onset

three-step models of vocal expression processing (Jiang, Gossack-Keenan, & Pell, 2020; Schirmer & Kotz, 2006), socio-affective information conveyed by a speaker's prosody is assigned significance incrementally and without delay from the acoustic onset of speech. Early processing stages (N100, P200) serve to categorize the auditory event and deploy attention to acoustic properties that are motivationally salient in the processing environment, promoting a coarse semantic analysis of the stimulus. Ongoing monitoring of the prosodic input allows listeners to refine their analysis of vocally-expressed meanings and to make mental associations (late positive component), as stimulus properties and task processing demands continue to evolve (Paulmann & Kotz, 2008; Pell et al., 2015).

As predicted, we found evidence that prosodic information encoding the speaker's stance toward the listener as they produced (potentially) ironic comments was differentiated at two distinct processing stages, beginning approximately 200 ms after speakers initiated a compliment or criticism (Jiang & Pell, 2015; Paulmann & Kotz, 2008; Pell et al., 2015; Vergis, Jiang & Pell, 2020). P200 amplitudes increased in anterior scalp regions when the speaker's stance was positive (compliments) versus negative (criticisms), an effect that could not be linked to linguistic features of the stimuli (which were identical for each irony type). This result implies that cues marking a speaker's positive stance were preferentially encoded at an initial stage, which is congruent with the current task demands of rating speaker friendliness (a positive social attribute).

However, subsequent effects in the late positivity point to heightened analysis of vocal cues encoding the speaker's negative (unfriendly) stance, which exhibited a more positive-going wave 600–1,000 ms post-onset of the utterance. This shift in the late positivity could reflect increased attention and monitoring of emotional negativity in the speech signal (Pell et al., 2015; Wang, Bastiaansen, & Yang, 2015) to assess the relevance of these cues as a mental representation of speaker meaning is being built (Martinelli et al., 2019; Stewart et al., 2010).

The data show that speaker stance was not the only information reliably extracted from the prosodic form of utterances prior to the critical word onset. P200 amplitudes were generally larger for literal compared to ironic utterances, which could not occur based solely on understanding the speaker's stance (see Rigoulot, Fish, & Pell, 2014 for related findings). However, upon closer inspection it was found that early sensitivity to speaker intentions (literal vs. ironic) was limited to one type of verbal irony, ironic criticism (i.e., sarcastic prosody). These utterances showed a marked reduction in the prosody-related P200, unlike ironic compliments. This pattern fits with data exemplifying that sarcastic prosody is acoustically and perceptually much more salient than many other forms of social expression in the voice (demonstrating a slower rate, lower pitch, changes in voice quality, etc., Cheang & Pell, 2008, 2009; Rockwell, 2000; Wilson, 2017). Furthermore, it underscores that sarcastic prosody, which

serves as a pragmatic marker of ironic criticism, can be differentiated at the neurocognitive level at a very early time point, much sooner than previously reported (Kowatch et al., 2013).

In contrast, a core set of features strongly associated with a “teasing” voice has been more difficult to identify (Caillies et al., 2019; Cheang & Pell, 2008; Mauchand et al., 2018), which could explain the less fine-grained differentiation of ironic compliments in the prosody-related P200. However, our data show that when prosody was positive, intention effects were simply delayed; with increasing exposure to vocal attributes of the stimuli (late positivity time window), listeners robustly differentiated the speaker’s intention to make a literal versus ironic compliment in the period 600–1,000 ms post-onset of the utterance (still before the critical target word). Intention-related differences on the late positivity, which were detected at right anterior electrodes, had a similar distribution to a P600-like effect reported by Rigoulot et al. (2014), who observed increased responses to prosodically sincere versus insincere compliments (“*I think you look really amazing*”). When put together, it is evident that listeners use prosody to register not only the attitude of the speaker, but also, to form more fine-grained impressions or predictions about the speaker’s (ironic) intent. Importantly, these operations construct meaning *before* the onset of the critical target word, and in the absence of any background context in our study, cannot be linked to processes for disconfirming expectations about the speaker’s intention as the process ironic statements. Rather, it seems clear that prosody can prompt a rapid attribution of ironic intentions to a speaker at very early stages of spoken language processing (Kowatch et al., 2013), although the *timing* of prosodic effects likely differs according to irony type (with earlier detection of ironic messages meant to be critical).

Integrating Prosodic and Semantic Meaning at the Critical Word

As noted, most ERP studies on verbal irony report the effects of context incongruence on measures derived from a target word at the end of written statements that could have literal or ironic interpretations. In most of these studies, target words that conflict with previous descriptions of events elicit an increased P600/late positivity when the speaker is being ironic versus literal (Spotorno, Cheylus, Van Der Henst, & Noveck, 2013; Weissman & Tanner, 2018). This effect has been attributed to continued processing and costly re-analysis of unexpected information when the speaker is pragmatically insincere (Brouwer et al., 2012, 2017; Kuperberg, Brothers, & Wlotko, 2019; Kutas & Federmeier, 2010; Thierry et al., 2017; Van Petten & Luka, 2012), thus representing an important time window in which listeners engage in pragmatic inferences to uncover irony and other nonliteral meanings (Regel et al., 2011, 2014).

Our study, in which only prosody served as a context for accessing ironic meanings at the critical word, supports previous research demonstrating that ironic intentions modulate the P600 component (Regel et al., 2010, 2011, 2014; Spotorno et al., 2013; Weissman & Tanner, 2018). However, in marked contrast to the literature, we found that the P600 was not greater in conditions of irony processing, but rather, when the speaker was being literal, i.e., when the valence of the critical word matched the stance expressed through their prosody. These patterns emphasize that our ironic stimuli, unlike previous work, did not create a strong pragmatic expectancy violation or increase demands on processes for contextual integration, which are the typical source of P600 differences at the critical word (Brouwer et al., 2012, 2017; Kutas & Federmeier, 2010; Thierry et al., 2017). Rather, given that the neurocognitive system had registered facets of both stance and speaker intentions from utterance onset, it would seem here that prosody created a strong contextual constraint that facilitated access and integration of the target word when speakers intended to be ironic (Katz et al., 2004; Pexman, 2008). This pattern could have been facilitated by the familiar lexical construction of our utterances, notably the word “such,” which when pragmatically emphasized may have acted as an additional marker of irony (Attardo, Eisterhold, Hay, & Poggi, 2003; Burgers, van Mulken, & Schellens, 2012). Still, the early registration of both stance and intention from utterance onset *before* the appearance of the word “such” suggests that pragmatic constraints were originally driven by features of the prosodic signal. Interestingly, in the only other study that we know focused strictly on how prosody influences irony processing, Caillies et al. (2019) reported a similar reversal in the direction of the P600 (literal > ironic) in one of their two conditions (when French speakers conveyed ironic compliments or “praise,” but not ironic criticism). While the specific patterns need to be resolved, and other methods between studies varied quite significantly, they both exemplify that prosody can at times point listeners directly to a speaker’s ironic intentions without the need to violate previously held expectations.

Indeed, it is noteworthy that P200, N400, *and* P600 responses to the critical word were all systematically larger when the speaker was being literal versus ironic, i.e., when the target word matched the affective stance of the speaker. While these responses (especially P600) are traditionally maximal at central and posterior electrodes, language-related components often are found to be more anteriorly distributed for speech compared to written stimuli, as is the case here (Jiang & Pell, 2015; Kotz & Paulmann, 2007). These data underscore that prosody influenced how meaning was derived from the target word at multiple processing stages. According to recent accounts, N400 indexes lexical retrieval and access to a target word, which can be primed by previous contextual cues, whereas P600 reflects integration of that word in the

meaning of the sentence (Delogu, Brouwer, & Crocker, 2019). In this light, our patterns suggest that the characteristic use of prosody, processed from utterance onset, was enough for listeners to correctly infer that the speaker meant to express irony; this reduced demands on lexical access to ironic versus literal meanings of the target word (N400) and subsequent procedures related to pragmatic reanalysis (P600). This means that (somewhat counterintuitively) literal statements may have produced relatively greater uncertainty about speaker intentions in our design, because early prosody effects allowed listeners to constrain and *predict* ironic meanings (see Caillies et al., 2019 for a similar reversal in the N400). The fact that P200 amplitudes to the target word also were greater for literal versus ironic meanings may be viewed as further evidence that ironic prosody facilitated processing of the ironic meaning of target words at an early stage of attentional deployment; the ironic meaning, which had already been activated, did not require as much attentional resources (Jiang & Pell, 2016).

These results mark a clear difference between context-cued irony and prosody-cued irony, as predicted by Allusion Pretense Theory (Kumon-Nakamura et al., 1995). When participants are provided background details about a speaker or event (Regel et al., 2011; Spotorno et al., 2013), context acts as an implicature during irony processing. It is only when it is compared to the semantic content that the incongruence can be made clear and the utterance apprehended as ironic (Grice, 1989; Schwoebel, Dews, Winner, & Srinivas, 2000; Sperber & Wilson, 1981; Wilson, 2017). This situation appears to place special demands on processes of pragmatic reanalysis for ironic messages in the P600 time window (Regel et al., 2010). In contrast, our results and others (Kowatch et al., 2013) show that prosody can already hint at the intention of the speaker, allowing listeners to build an initial representation of the speaker's stance and predictions about speaker meaning. This situation does not always produce a mismatch or highlight failed expectations when the critical word expressing irony is encountered (an "after-the-fact" pragmatic reinterpretation of ironic comments; Regel et al., 2010; Schwoebel et al., 2000).

The fact that our stimuli were presented in the absence of any context, and that listeners focused on social attributes of the speaker, likely amplified the observed effects of prosody on neurocognitive processes during irony processing. As the only reliable cue for communicating nonliteral intentions, speakers would need to provide clear signals marking when their intention was ironic, and listeners would give stronger weight to these cues as utterances are processed (Bryant, 2010), explaining why the impact of prosody is often masked in context-focused designs (Deliens et al., 2018; Filik et al., 2014; Regel et al., 2011). Indeed, while prosodically marked statements which lack context do seem to occur in spontaneous speech (Bryant, 2010; Bryant & Fox Tree, 2002, 2005),

irony is usually produced in specific situations (e.g. a failed expectation) in which recognition of various contextual parameters for communicating irony influence interpretative processes. Certainly, the manner in which context has been operationalized in this and previous studies of irony does not accurately represent the level of complexity and *ambiguity* that often characterizes more spontaneous ironic communications between individuals. By isolating prosody effects, the present results highlight the potential contributions of an ironic tone of voice as one source of "context" which can now be studied in more ecologically valid, ambiguous situations that contain multiple cues for evaluating speaker intentions, such as less formulaic ironic conversations.

The idea that prosody is immediately and incrementally structured by the brain as a predictive cue to ironic intentions, whether the speaker is expressing a positive or negative stance, is consistent with parallel-constraint satisfaction models of verbal irony (Katz et al., 2004; Pexman, 2008). Evidence for this position is accumulating in multiple domains. For example, based on word-by-word reading times, Pexman et al. (2000) reported that participants immediately use available cues to build a representation of the speaker's ironic intent, concluding that ironic meanings are considered as soon as there is sufficient evidence to support them. In keeping with this idea, listeners displayed similar decision times about literal versus ironic criticisms when they had been exposed to prosody in a visual world paradigm (Kowatch et al., 2013), and ironic prosody *decreased* response times to the sentence-final word of ironic remarks in recent work (Deliens et al., 2018). These results bolster the claim that ironic meanings can be accessed *directly* if sufficiently constrained (Gibbs, 2002). Along these lines, it has been found that speaker-specific information, such as the tendency of certain individuals to use irony, promotes very early effects on irony processing in the 150-ms to 300-ms time range (Regel et al., 2010), and the intention to be ironic seems to produce rapid changes in low gamma band energy in the 280-ms to 400-ms time window following a target word (Spotorno et al., 2013).

These findings conflict with the standard pragmatic view (e.g., Grice, 1989), suggesting that listeners can quickly and directly access pragmatically salient meanings such as irony and do not need to first analyze the literal meaning of linguistic expressions (Gibbs, 2002; Pexman, 2008). Other views, such as the graded salience/defaultness hypothesis, suggest that irony can be accessed as a default interpretation provided the context or lexical construction of an utterance is biased towards an ironic meaning (Giora, Givoni, & Fein, 2015). While our stimuli were constructed as judgements that could easily be transformed into ironic comments given the proper tone of voice, they do not follow patterns that are clearly biased towards irony (such as negative constructions). Our data provide instead strong empirical support for a parallel-

constraint satisfaction model of irony by demonstrating that prosody activates and *constrains* ironic meanings at multiple neural processing stages as listeners process spoken language. This model also explains differences between the present data and previous context-based irony research in which prosody effects were limited (Deliens et al., 2018; Regel et al., 2011). In the latter case, a written description of a situation presented before an utterance would constitute a major constraint on the literal interpretation of the utterance. It not only prevails temporally on prosody (earlier and longer exposition), it also is a simpler constraint to satisfy in an experimental setting (the utterance is either congruent or incongruent). Here, we show that prosody does contain salient information that can constrain interpretations towards an ironic meaning. This potential parallel processing of prosody was not investigated at utterance onset or at critical words in other studies; future research on ironic prosody in context should investigate these time windows in which prosody effects seem to occur before the contextual constraint is resolved.

On the Negativity Bias During Irony Processing

Another issue raised here and elsewhere (Kumon-Nakamura et al., 1995; Sperber & Wilson, 1981) refers to the asymmetrical effects of negative information on the integration of speech-related cues during irony processing ("negativity bias"). Previous data on ironic compliments suggest that the explicit negativity of these statements (*You are a horrible driver!*) promotes asymmetric perception and impressions of irony (Bruntsch & Ruch, 2017; Caillies et al., 2019; Kreuz & Link, 2002; Matthews et al., 2006; Pexman & Olineck, 2002), due to attentional biases towards threatening cues (Carretié et al., 2001; Wabnitz, Martens, & Neuner, 2015). In our study, negative statements reduced the modulating effects of prosody on the critical word in the P200, N400 and P600 time-windows; prosody had less impact on the neural response when speakers used a negative statement to tease (vs. literally criticize) than when they used a positive statement to convey sarcasm (vs. literally compliment). These results fit with the idea of an asymmetry of affect when processing irony from positive versus negative statements and that pragmatic cues from detecting ironic attitudes (here, cued by prosody) tend to be less salient when speakers use an explicit negative statement (Kumon-Nakamura et al., 1995). On the other hand, this pattern can still be viewed as evidence that prosody attenuates the force of a negative, face-threatening statement to a certain extent when speakers make ironic compliments, through the strategic use of indirect communication (Slugoski & Turnbull, 1988).

Thus, while the speakers' interpersonal stance when producing an ironic compliment is meant to be positive, from the listener's point of view these remarks are often interpreted as softened criticisms, consistent with our perceptual ratings and previous literature (Dews et al., 1995; Mauchand et al., 2020;

Pexman & Olineck, 2002). As widely noted, the general prevalence of positive norms/expectations renders ironic compliments less prototypical in human communication (because they imply a "better than expected" response to a negative expectation, Sperber & Wilson, 1981), and due to their risk of being misinterpreted, these comments are mostly used in close relationships in which knowledge of the speaker is highly constrained (Matthews et al., 2006; Pexman & Zvaigzne, 2004). As our stimuli were produced by completely unfamiliar speakers, our methods would place special demands on the processing of ironic compliments versus criticisms (see also Caillies et al., 2019). Given ERP evidence that social distance has important impacts on neural processing (Katz, Blasko, & Kazmerski, 2004; Perry, Rubinsten, Peled, & Shamay-Tsoory, 2013; Yu, Hu, & Zhang, 2015), this factor should be operationalized in future studies of how listeners apprehend and process ironic intentions. Moreover, in light of findings showing that a speaker's communicative style (Regel et al., 2010) or linguistic background (Caffarra et al., 2019) influence how the brain processes irony, future studies should continue to address speaker identity issues (e.g., responses towards voices of familiar and unfamiliar speakers) while accounting for prosody effects.

Conclusions

Our study extends the literature on ironic speech processing by highlighting the often-ignored effects of speech prosody. Based on temporally fine-grained evidence of brain activity measured at multiple time points in an utterance, our data reveal unique mechanisms that are engaged by a speaker's tone of voice, which suggest that ironic prosody does not work as an implicature like other pragmatic and contextual cues (e.g., background knowledge about a speaker). Rather, the immediate on-line uptake and structuring of prosodic cues (especially sarcastic prosody) can point directly to nonliteral meanings, allowing listeners to predict and constrain a representation of the speaker's ironic intent as a "best fit" solution (Gibbs, 2002). This process is best captured by the parallel-constraint satisfaction model (Katz et al., 2004; Pexman, 2008), which is receiving increasing support (Caffarra et al., 2019; Deliens et al., 2018; Kowatch et al., 2013; Spotorno et al., 2013). When prosodic cues are salient and continuously processed, initial representations of ironic intent appear to facilitate linguistic processing which does not require pragmatic reinterpretation, although these operations are somewhat altered for less prototypical, negative statements communicating irony. Because the lack of context in our study likely enhanced the importance of prosody in our results, new studies should continue to examine how the neurocognitive system responds to prosody and different forms of context to map the time course of these effects as an utterance unfolds to the listener.

Acknowledgments The authors thank Alexandra Topilova for her help recruiting and testing participants. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) to M.D. Pell under grant # RGPIN-2016-04373 (“Neurocognitive studies of vocal emotions in speech”). Declarations of interest: none. Data and materials are available [here](#), and this experiment was not preregistered.

References

- Alberts, J. K., Kellar-Guenther, Y., & Corman, S. R. (1996). That’s not funny: Understanding recipients’ responses to teasing. *Western Journal of Communication*, 60(4), 337–357. <https://doi.org/10.1080/10570319609374553>
- Anolli, L., Ciceri, R., & Infantino, M. G. (2000). Irony as a game of implicitness: Acoustic profiles of ironic communication. *Journal of Psycholinguistic Research*, 29(3), 275–311. <https://doi.org/10.1023/A:1005100221723>
- Argyle, M., Alkema, F., & Gilmour, R. (1971). The communication of friendly and hostile attitudes by verbal and non-verbal signals. *European Journal of Social Psychology*, 1(3), 385–402. <https://doi.org/10.1002/ejsp.2420010307>
- Argyle, M., Salter, V., Nicholson, H., Williams, M., & Burgess, P. (1970). The Communication of Inferior and Superior Attitudes by Verbal and Non-verbal Signals *. *British Journal of Social and Clinical Psychology*, 9(3), 222–231. <https://doi.org/10.1111/j.2044-8260.1970.tb00668.x>
- Attardo, S., Eisterhold, J., Hay, J., & Poggi, I. (2003). Multimodal markers of irony and sarcasm. *Humor*, 16(2), 243–260.
- Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., ... & Treiman, R. (2007). The English lexicon project. *Behavior research methods*, 39(3), 445–459
- Bates, D., Machler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: Neural correlates of voice perception. *Trends in Cognitive Sciences*, 8(3), 129–135. <https://doi.org/10.1016/j.tics.2004.01.008>
- Boersma, P., & van Heuven, V. (2001). Speak and unSpeak with Praat. *Glott International*, 5(9–10), 341–347.
- Brouwer, H., Crocker, M. W., Venhuizen, N. J., & Hoeks, J. C. J. (2017). A Neurocomputational Model of the N400 and the P600 in Language Processing. *Cognitive Science*, 41, 1318–1352. <https://doi.org/10.1111/cogs.12461>
- Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about Semantic Illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research*, 1446, 127–143. <https://doi.org/10.1016/j.brainres.2012.01.055>
- Bruntsch, R., & Ruch, W. (2017). The role of humor-related traits and broad personality dimensions in irony use. *Personality and Individual Differences*, 112, 139–143. <https://doi.org/10.1016/j.paid.2017.03.004>
- Bryant, G. A. (2010). Prosodic Contrasts in Ironic Speech. *Discourse Processes*, 47(7), 545–566. <https://doi.org/10.1080/01638530903531972>
- Bryant, G. A., & Fox Tree, J. E. (2002). Recognizing Verbal Irony in Spontaneous Speech. *Metaphor and Symbol*, 17(2), 99–119. https://doi.org/10.1207/S15327868MS1702_2
- Bryant, G. A., & Fox Tree, J. E. (2005). Is there an ironic tone of voice? *Language and Speech*, 48(3), 257–277. <https://doi.org/10.1177/00238309050480030101>
- Burgers, C., van Mulken, M., & Schellens, P. J. (2012). Type of evaluation and marking of irony: The role of perceived complexity and comprehension. *Journal of Pragmatics*, 44(3), 231–242. <https://doi.org/10.1016/j.pragma.2011.11.003>
- Caffarra, S., Motamed Haeri, A., Michell, E., & Martin, C. D. (2019). When is irony influenced by communicative constraints? ERP evidence supporting interactive models. *European Journal of Neuroscience*, 50(10), 3566–3577. <https://doi.org/10.1111/ejn.14503>
- Caillies, S., Gobin, P., Obert, A., Terrien, S., Coutté, A., Iakimova, G., & Besche-Richard, C. (2019). Asymmetry of affect in verbal irony understanding: What about the N400 and P600 components? *Journal of Neurolinguistics*, 51, 268–277. <https://doi.org/10.1016/J.JNEUROLING.2019.04.004>
- Carretié, L., Mercado, F., Tapia, M., & Hinojosa, J. A. (2001). Emotion, attention, and the ‘negativity bias’, studied through event-related potentials. *International Journal of Psychophysiology*, 41(1), 75–85. [https://doi.org/10.1016/S0167-8760\(00\)00195-1](https://doi.org/10.1016/S0167-8760(00)00195-1)
- Cheang, H. S., & Pell, M. D. (2008). The sound of sarcasm. *Speech Communication*, 50(5), 366–381. <https://doi.org/10.1016/j.specom.2007.11.003>
- Cheang, H. S., & Pell, M. D. (2009). Acoustic markers of sarcasm in Cantonese and English. *The Journal of the Acoustical Society of America*, 126(3), 1394–1405. <https://doi.org/10.1121/1.3177275>
- Cornejo, C., Simonetti, F., Aldunate, N., Ibáñez, A., López, V., & Melloni, L. (2007). Electrophysiological evidence of different interpretative strategies in irony comprehension. *Journal of Psycholinguistic Research*, 36(6), 411–430. <https://doi.org/10.1007/s10936-007-9052-0>
- Deliens, G., Antoniou, K., Clin, E., Ostashchenko, E., & Kissine, M. (2018). Context, facial expression and prosody in irony processing. *Journal of Memory and Language*, 99, 35–48. <https://doi.org/10.1016/j.jml.2017.10.001>
- Delogu, F., Brouwer, H., & Crocker, M. W. (2019). Event-related potentials index lexical retrieval (N400) and integration (P600) during language comprehension. *Brain and Cognition*, 135, 103569. <https://doi.org/10.1016/J.BANDC.2019.05.007>
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21
- Dews, S., Kaplan, J., & Winner, E. (1995). Why Not Say It Directly? The Social Functions of Irony. *Discourse Processes*, 19(August), 347–367. <https://doi.org/10.1080/01638539509544922>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191.
- Filik, R., Leuthold, H., Wallington, K., & Page, J. (2014). Testing theories of irony processing using eye-tracking and ERPs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(3), 811–828. <https://doi.org/10.1037/a0035658>
- Foster, E. D., & Deardorff, A. (2017). Open Science Framework (OSF). *Journal of the Medical Library Association*, 105(2), 203–206. <https://doi.org/10.5195/JMLA.2017.88>
- Gibbs, R. W. (2002). A new look at literal meaning in understanding what is said and implicated. *Journal of Pragmatics*, 34(4), 457–486. [https://doi.org/10.1016/S0378-2166\(01\)00046-7](https://doi.org/10.1016/S0378-2166(01)00046-7)
- Gibbs Jr, R. W., Gibbs, R. W., & Colston, H. L. (Eds.). (2007). *Irony in language and thought: A cognitive science reader*. Psychology Press.
- Giora, R., Givoni, S., & Fein, O. (2015). Defaultness Reigns: The Case of Sarcasm. *Metaphor and Symbol*, 30(4), 290–313. <https://doi.org/10.1080/10926488.2015.1074804>
- Grice, P. (1989). *Studies in the Way of Words*. Harvard University Press.
- Hajcak, G., Weinberg, A., MacNamara, A., & Foti, D. (2012). ERPs and the study of emotion. In S. J. Luck & E. S. Kappenman (Eds.), *Oxford library of psychology. The Oxford handbook of event-*

- related potential components (p. 441–472). Oxford University Press.
- Jiang, X., Gossack-Keenan, K., & Pell, M. D. (2020). To believe or not to believe? How voice and accent information in speech alter listener impressions of trust. *Quarterly Journal of Experimental Psychology*, 73(1), 55–79. <https://doi.org/10.1177/1747021819865833>
- Jiang, X., & Pell, M. D. (2015). On how the brain decodes vocal cues about speaker confidence. *Cortex*, 66, 9–34. <https://doi.org/10.1016/J.CORTEX.2015.02.002>
- Jiang, X., Pell, M. D. (2016). Neural responses towards a speaker's feeling of (un)knowing. *Neuropsychologia*, 81:79–93
- Katz, A. N., Blasko, D. G., & Kazmerski, V. A. (2004). Saying What You Don't Mean: Social Influences on Sarcastic Language Processing. *Current Directions in Psychological Science*, 13(5), 186–189. <https://doi.org/10.1111/j.0963-7214.2004.00304.x>
- Keltner, D., Capps, L., Kring, A. M., Young, R. C., & Heerey, E. A. (2001). Just Teasing: A Conceptual Analysis and Empirical Review. *Psychological Bulletin*, 127(2), 229–248. <https://doi.org/10.1037/0033-2909.127.2.229>
- Kotz, S. A., & Paulmann, S. (2007). When emotional prosody and semantics dance cheek to cheek: ERP evidence. *Brain Research*, 1151, 107–118. <https://doi.org/10.1016/j.brainres.2007.03.015>
- Kowatch, K., Whalen, J. M., & Pexman, P. M. (2013). Irony Comprehension in Action: A New Test of Processing for Verbal Irony. *Discourse Processes*, 50(5), 301–315. <https://doi.org/10.1080/0163853X.2013.799934>
- Kreuz, R. J., & Link, K. E. (2002). Asymmetries in the use of verbal ironies. *Journal of Language and Social Psychology*, 21(2), 127–143
- Kumon-Nakamura, S., Glucksberg, S., & Brown, M. (1995). How About Another Piece of Pie: The Allusional Pretense Theory of Discourse Irony. *Journal of Experimental Psychology: General*, 124(1), 3–21. <https://doi.org/10.1037/0096-3445.124.1.3>
- Kuperberg, G. R., Brothers, T., & Wlotko, E. W. (2019). A tale of two positivities and the N400: Distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. *Journal of Cognitive Neuroscience*, 32(1), 12–35. https://doi.org/10.1162/jocn_a_01465
- Kutas, M., & Federmeier, K. D. (2010). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP) ms: millisecond. *Annual Review of Psychology*, 62, 621–647. <https://doi.org/10.1146/annurev.psych.093008.131123>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Leary, M. R. (2000). *Affect, cognition, and the social emotions*. In J. P. Forgas (Ed.), *Studies in emotion and social interaction, second series. Feeling and thinking: The role of affect in social cognition* (p. 331–356). Cambridge University Press.
- Liu, P., Rigoulot, S., & Pell, M. D. (2015). Culture modulates the brain response to human expressions of emotion: Electrophysiological evidence. *Neuropsychologia*, 67, 1–13. <https://doi.org/10.1016/j.neuropsychologia.2014.11.034>
- Lopez-Calderon, J., Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, 8
- Makeig, S., Bell, A. J., Jung, T.-P., & Sejnowski, T. J. (1996). Independent Component Analysis of Electroencephalographic Data. *Advances in Neural Information Processing Systems*, 145–151. <https://doi.org/10.1109/ICOSP.2002.1180091>
- Martinelli, A., Kreifelts, B., Wildgruber, D., Ackermann, K., Bernhard, A., Freitag, C. M., & Schwenck, C. (2019). Aggression modulates neural correlates of hostile intention attribution to laughter in children. *NeuroImage*, 184, 621–631. <https://doi.org/10.1016/J.NEUROIMAGE.2018.09.066>
- Matsui, T., Nakamura, T., Utsumi, A., Sasaki, A. T., Koike, T., Yoshida, Y., ... Sadato, N. (2016). The role of prosody and context in sarcasm comprehension: Behavioral and fMRI evidence. *Neuropsychologia*, 87, 74–84. <https://doi.org/10.1016/j.neuropsychologia.2016.04.031>
- Matthews, J. K., Hancock, J. T., & Dunham, P. J. (2006). The Roles of Politeness and Humor in the Asymmetry of Affect in Verbal Irony. *Discourse Processes*, 41(1), 3–24. https://doi.org/10.1207/s15326950dp4101_2
- Mauchand, M., Vergis, N., & Pell, M. D. (2018). Ironic tones of voices. In *Proc. 9th International Conference on Speech Prosody 2018* (pp. 443–447). <https://doi.org/10.21437/SpeechProsody.2018-90>
- 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>
- Ouyang, G., Herzmann, G., Zhou, C., & Sommer, W. (2011). Residue iteration decomposition (RIDE): A new method to separate ERP components on the basis of latency variability in single trials. *Psychophysiology*, 48(12), 1631–1647. <https://doi.org/10.1111/j.1469-8986.2011.01269.x>
- Ouyang, G., Sommer, W., & Zhou, C. (2015). A toolbox for residue iteration decomposition (RIDE)-A method for the decomposition, reconstruction, and single trial analysis of event related potentials. *Journal of Neuroscience Methods*, 250, 7–21. <https://doi.org/10.1016/j.jneumeth.2014.10.009>
- Ouyang, G., Sommer, W., & Zhou, C. (2016). Reconstructing ERP amplitude effects after compensating for trial-to-trial latency jitter: A solution based on a novel application of residue iteration decomposition. *International Journal of Psychophysiology*, 109, 9–20. <https://doi.org/10.1016/j.ijpsycho.2016.09.015>
- Paulmann, S., & Kotz, S. A. (2008). An ERP investigation on the temporal dynamics of emotional prosody and emotional semantics in pseudo- and lexical-sentence context. *Brain and Language*, 105(1), 59–69. <https://doi.org/10.1016/j.bandl.2007.11.005>
- Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70, 153–163. <https://doi.org/10.1016/j.jesp.2017.01.006>
- Pell, M. D., Rothermich, K., Liu, P., Paulmann, S., Sethi, S., & Rigoulot, S. (2015). Preferential decoding of emotion from human non-linguistic vocalizations versus speech prosody. *Biological Psychology*, 111, 14–25. <https://doi.org/10.1016/j.biopsycho.2015.08.008>
- Pell, M. D., Vergis, N., Caballero, J., Mauchand, M., & Jiang, X. (2018). Prosody as a window into speaker attitudes and interpersonal stance. *The Journal of the Acoustical Society of America*, 144(3), 1840–1840. <https://doi.org/10.1121/1.5068107>
- Perry, A., Rubinsten, O., Peled, L., & Shamay-Tsoory, S. G. (2013). Don't stand so close to me: A behavioral and ERP study of preferred interpersonal distance. *NeuroImage*, 83, 761–769. <https://doi.org/10.1016/J.NEUROIMAGE.2013.07.042>
- Pexman, P. M. (2008). It's Fascinating Research. *Current Directions in Psychological Science*, 17(4), 286–290. <https://doi.org/10.1111/j.1467-8721.2008.00591.x>
- Pexman, P. M., & Olineck, K. M. (2002). Does Sarcasm Always Sting? Investigating the Impact of Ironic Insults and Ironic Compliments Penny. *Discourse Processes*, 33(3), 199–217. https://doi.org/10.1207/S15326950DP3303_1
- Pexman, P. M., & Zvaigzne, M. (2004). Does irony go better with friends? *Metaphor and Symbol*, 19(2), 143–163. <https://doi.org/10.1207/s15327868ms1902>
- 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 Research*, 1311, 121–135. <https://doi.org/10.1016/J.BRAINRES.2009.10.077>

- Regel, S., Gunter, T. C., & Friederici, A. D. (2011). Isn't It Ironic? An Electrophysiological Exploration of Figurative Language Processing. *Journal of Cognitive Neuroscience*, 23(2), 277–293. <https://doi.org/10.1162/jocn.2010.21411>
- Regel, S., Meyer, L., & Gunter, T. C. (2014). Distinguishing Neurocognitive Processes Reflected by P600 Effects: Evidence from ERPs and Neural Oscillations. *PLoS ONE*, 9(5), e96840. <https://doi.org/10.1371/journal.pone.0096840>
- Regenbogen, C., Schneider, D. A., Finkelmeyer, A., Kohn, N., Demtl, B., Kellermann, T., ... Habel, U. (2012). The differential contribution of facial expressions, prosody, and speech content to empathy. *Cognition and Emotion*, 26(6), 995–1014. <https://doi.org/10.1080/02699931.2011.631296>
- Rigoulot, S., Fish, K., & Pell, M. D. (2014). Neural correlates of inferring speaker sincerity from white lies: An event-related potential source localization study. *Brain Research*, 1565, 48–62. <https://doi.org/10.1016/J.BRAINRES.2014.04.022>
- Rigoulot, S., Jiang, X., Vergis, N., Pell, M. D. (2020). Neurophysiological correlates of sexually evocative speech. *Biological Psychology* 154:107909
- Rockwell, P. (2000). Lower, Slower, Louder: Vocal Cues of Sarcasm. *Journal of Psycholinguistic Research*, 29(5), 483–495. <https://doi.org/10.1023/A:1005120109296>
- Sally, D. (2003). Risky speech: Behavioral game theory and pragmatics. *Journal of Pragmatics*, 35(8), 1223–1245. [https://doi.org/10.1016/S0378-2166\(02\)00170-4](https://doi.org/10.1016/S0378-2166(02)00170-4)
- Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends in cognitive sciences*, 10(1), 24–30
- Schmitz, C. (2012). Limesurvey: an open source survey tool. *Limesurvey Project, Hamburg, Germany*.
- Schwoebel, J., Dews, S., Winner, E., & Srinivas, K. (2000). Obligatory Processing of the Literal Meaning of Ironic Utterances: Further Evidence. *Metaphor and Symbol*, 15(1–2), 47–61. <https://doi.org/10.1080/10926488.2000.9678864>
- Searle, J. R. (1965). What is a speech act. *Perspectives in the philosophy of language: a concise anthology*, 2000, 253–268
- Slugoski, B. R., & Turnbull, W. (1988). Cruel to be Kind and Kind to be Cruel: Sarcasm, Banter and Social Relations. *Journal of Language and Social Psychology*, 7(2), 101–121. <https://doi.org/10.1177/0261927X8800700202>
- Sperber, D., & Wilson, D. (1981). Irony and the Use-Mention distinction. *Radical Pragmatics*.
- Spotorno, N., Cheylus, A., Van Der Henst, J.-B., & Noveck, I. A. (2013). What's behind a P600? Integration Operations during Irony Processing. *PLoS ONE*, 8(6), e66839. <https://doi.org/10.1371/journal.pone.0066839>
- Stewart, J. L., Silton, R. L., Sass, S. M., Fisher, J. E., Edgar, J. C., Heller, W., & Miller, G. A. (2010). Attentional bias to negative emotion as a function of approach and withdrawal anger styles: An ERP investigation. *International Journal of Psychophysiology*, 76(1), 9–18. <https://doi.org/10.1016/J.IJPSYCHO.2010.01.008>
- Thierry, G., Berkum, J. J. A. Van, Brouwer, H., & Crocker, M. W. (2017). On the Proper Treatment of the N400 and P600 in Language Comprehension. *Frontiers in Psychology | Www.Frontiersin.Org*, 1, 1327. <https://doi.org/10.3389/fpsyg.2017.01327>
- Van Lancker Sidsit, D., Pachana, N., Cummings, J. L., & Sidsit, J. J. (2006). Dysprosodic speech following basal ganglia insult: Toward a conceptual framework for the study of the cerebral representation of prosody. *Brain and Language*, 97(2), 135–153. <https://doi.org/10.1016/j.bandl.2005.09.001>
- Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, 83(2), 176–190. <https://doi.org/10.1016/j.ijpsycho.2011.09.015>
- Vergis, N., & Pell, M. D. (2020). Factors in the perception of speaker politeness: the effect of linguistic structure, imposition and prosody. *Journal of Politeness Research*, 16(1), 45–84. <https://doi.org/10.1515/pr-2017-0008>
- Vergis, N., & Terkourafi, M. (2015). The Role of the Speaker's Emotional State in Im/politeness Assessments. *Journal of Language and Social Psychology*, 34(3), 316–342. <https://doi.org/10.1177/0261927X14556817>
- Wabnitz, P., Martens, U., & Neuner, F. (2015). Written threat: Electrophysiological evidence for an attention bias to affective words in social anxiety disorder. *Cognition and Emotion*, 9931(April 2015), 1–23. <https://doi.org/10.1080/02699931.2015.1019837>
- Wang, L., Bastiaansen, M., & Yang, Y. (2015). The influence of emotional salience on the integration of person names into context. *Brain Research*, 1609, 82–92. <https://doi.org/10.1016/J.BRAINRES.2015.03.028>
- Weissman, B., & Tanner, D. (2018). A strong wink between verbal and emoji-based irony: How the brain processes ironic emojis during language comprehension. *Plos One*, 13(8), e0201727. <https://doi.org/10.1371/journal.pone.0201727>
- Wichmann, A. (2000). The attitudinal effects of prosody, and how they relate to emotion. In *ITRW on Speech and Emotion, Newcastle, Northern Ireland, UK* (pp. 143–148).
- Wilson D. (2017) Irony, Hyperbole, Jokes and Banter. In: Blochowiak J., Grisot C., Durrleman S., Laenzlinger C. (eds) *Formal Models in the Study of Language*. Springer, Cham. https://doi.org/10.1007/978-3-319-48832-5_11
- Yu, R., Hu, P., & Zhang, P. (2015). Social distance and anonymity modulate fairness consideration: An ERP study. *Scientific Reports*, 5(1), 13452. <https://doi.org/10.1038/srep13452>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.