Full citation: Caffarra, S., Motamed, A.H., Michell, E., & Martin, C.D. (2019). When is irony influenced by communicative constraints? ERP evidence supporting interactive models. *European Journal of Neuroscience*, 00, 1–12. doi: 10.1111/ejn.14503

When is irony influenced by communicative constraints?

ERP evidence supporting interactive models

Sendy Caffarra[a], Arman Motamed Haeri[a], Elissa Michell[a], & Clara D. Martin[a, b]

a BCBL, Basque Center on Cognition, Brain and Language, Paseo Mikeletegi 69, 20009, Donostia, Spain

b Ikerbasque, Basque Foundation for Science, Maria Diaz de Haro 3, Bilbao, 48013, Spain

Corresponding author:

Sendy Caffarra

BCBL, Basque Center on Cognition, Brain and Language, Donostia, Spain

Paseo Mikeletegi 69, 20009 Donostia-San Sebastian, Spain.

tel.: +34 943 309 300

Email: s.caffarra@bcbl.eu

Abstract

Distinct theoretical proposals have described how communicative constraints (contextual biases, speaker identity) impact verbal irony processing. Modular models assume that social and contextual factors have an effect at a late stage of processing. Interactive models claim that contextual biases are considered early on. The constraint-satisfaction model further assumes that speaker's and context's characteristics can compete at early stages of analysis. The present ERP study teased apart these models by testing the impact of context and speaker features (i.e., speaker accent) on irony analysis. Spanish native speakers were presented with Spanish utterances that were ironic or literal. Each sentence was preceded by a negative or a positive context. Each story was uttered in a native or a foreign accent. Results showed that contextual biases and speaker accent interacted as early as 150 ms during irony processing. Greater N400-like effects were reported for ironic than literal sentences only with positive contexts and native accent, possibly suggesting semantic difficulties when non-prototypical irony was produced by natives. A P600 effect of irony was also reported indicating inferential processing costs. The results support the constraint-satisfaction model and suggest that multiple sources of information are weighted and can interact from the earliest stages of irony analysis.

Introduction

Verbal irony is a trope in which the intended meaning is usually the opposite of what is literally said. Given this gap, interpreting irony represents a complex inferential process and identifying the correct figurative meaning requires consideration of multiple sources of information. Who is producing irony can easily influence the way we understand the utterance. If a person in front of a broken plate says loudly “Great!”, knowing that the speaker is a clown instead of a waiter would likely change our interpretation. In addition, the type of context in which irony is produced (e.g., having a plate broken in hundreds of pieces or simply chipped) can influence how ironically the comment is perceived. When and how each of these communicative constraints (i.e., speaker identity and context) plays a role in verbal irony interpretation is yet to be defined and it will be the focus of the present event-related potential (ERP) study.

Distinct theoretical models describe the time course of verbal irony analysis and its relative interaction with social and contextual variables (for a review see Gibbs, 2001). Serial modular models claim that literal meaning is accessed first (Grice, 1975; Searle, 1979). Contextual information is taken into account only at later stages, once the literal meaning has already been accessed. At this point, if a conflict arises between the literal interpretation and the communicative context, the literal meaning will be rejected, leading to re-analysis and figurative interpretation (Grice, 1975; Searle, 1979). Other similar modular proposals theorize an early and context-free access to the salient meaning (i.e., the familiar and conventional meaning of the lexical input, corresponding to the literal meaning in the case of novel ironic phrases), which can never be preceded (or inhibited) by non-salient figurative meanings (Giora, 1997; 2002; Giora, Fein, & Schwartz, 1998). Importantly, all these theoretical models describe a two-step analysis where literal meaning is initially encapsulated and information regarding social and contextual factors is lately taken into account.

On the other hand, parallel interactive models assume that figurative meaning can be accessed as early as literal meaning. There is no principled delay for the processing of ironic meaning and its relevance depends on multiple discourse factors (Sperber & Wilson, 1995). A set of interactive models have highlighted the role of contextual biases, such as the presence of ironic cues in the context preceding the target sentence (direct access model; Gibbs, 1986; 1994; 2002). According to this view, the type of context can have an early impact on sentence interpretation, supporting or discouraging figurative readings. One typical situation in which context would have an early impact on sentence interpretation is when sentence context has a positive or negative valence: In fact, many studies have described qualitative distinctions between the type of irony produced in negative contexts (i.e., ironic criticism) and the one produced in positive contexts (i.e., ironic praise), with the first type being more prototypical, more frequently used, and easier to be learned, detected, produced and understood as compared to the second type (Averbeck, 2015; Bromberek-Dyzman, 2015; Bruntsch & Ruch, 2017; De Groot, Kaplan, Rosenblatt, Dews, & Winner, 1995; Gibbs, 1986).

An even more extreme version of the interactive framework comes from a recent theoretical model claiming that contextual biases are only one of the factors that have an early impact on irony interpretation (constraint-satisfaction model; Katz, 2005; Katz, Blasko & Kazmerski, 2004; Pexman, 2008). Other communicative constraints would interact and compete early on in order to support the most likely sentence interpretation. Among these additional constraints, there are sociocultural aspects related to the identity of the speaker (e.g., stereotypical information about the speaker; Katz & Pexman, 1997; Katz et al., 2004; Pexman & Olineck, 2002).

The abovementioned theoretical proposals have been operationalized and translated into experimental predictions that are focused on the temporal dimension of irony interpretation (i.e., when literal and ironic interpretations become available). Behavioral evidence supporting classic modular models showed that the ironic interpretation requires more time and processing costs as compared to the literal/salient interpretation (Dews & Winner, 1999; Fein, Yeari, & Giora, 2015; Filik & Moxey, 2010; Giora et al., 1998; and Turcan & Filik, 2016 for unfamiliar ironic phrases). On the other hand, behavioral findings supporting interactive models showed that, with a facilitating context (usually describing negative events and triggering ironic criticism), the ironic meaning can become available as fast as the literal interpretation (Gibbs, 1986; Ivanko & Pexman, 2003). Finally, the authors that proposed the constraint-satisfaction model highlighted that, not only the type of context, but also the identity of the speaker (e.g., being a comedian instead of a priest) influences whether and how people perceive verbal irony and can potentially interact with contextual cues (Climie & Pexman, 2008; Katz & Pexman, 1997; Pexman, Ferretti, & Katz, 2000; Pexman & Olineck, 2002). Interestingly, in a recent behavioral study both context and speaker identity were manipulated in the same experiment (Caffarra, Michell, & Martin, 2018). Spanish native listeners were presented with Spanish ironic and literal utterances and had to rate their degree of irony. Each sentence was preceded by either a positive or a negative context. Speaker identity was manipulated through speech accent so that each story could be uttered in a native (Spanish) or foreign (English) accent. The results showed a triple interaction between irony, contextual bias and speaker's accent. The degree of perceived irony dropped when the stories were produced in a foreign accent. This was true only when the contextual bias was positive, which corresponds to the less prototypical type of irony. This suggests that deriving pragmatic inferences was particularly problematic when unusual ironic comments were produced by foreigners. Note that all the studies described above collected behavioral measures, which can be influenced by strategic behavior and do not provide a measure of brain correlates as irony analysis unfolds over time. As a result, behavioral findings cannot definitely solve the theoretical debate. For instance, the presence of a triple interaction in behavioral studies (irony, contextual biases and speaker identity) can be compatible with both modular and interactive predictions since behavioral measures cannot detect at which stage of analysis this interactive effect emerges.

High-temporal resolution techniques are ideal to precisely identify when this interaction arises and, thus, they are able to test the predictions of the abovementioned models. Event-related potentials are evoked brain responses time-locked to the presentation of a stimulus of interest (e.g., a spoken word embedded in a sentence). They can provide information about how brain processes unfold over time on a millisecond-by-millisecond basis even when there is no behavioral response. Previous ERP experiments compared ironic and literal sentences that were preceded by short stories (e.g., the sentence “These artists are gifted” after the description of a bad or a good performance; Cornejo, Simonetti, Aldunate, Ibañez, López, & Melloni, 2007; Filik, Leuthold, Wallington, & Page, 2014; Regel, Coulson, & Gunter, 2010; Regel, Gunter, & Friederici, 2011; Regel, Meyer, & Gunter, 2014; Spotorno, Cheylus, Van der Henst, & Noveck, 2013). The use of this paradigm follows the assumption that the ERP differences observed after comparing ironic and literal interpretations (when all the other linguistic variables are held constant) must be related to the way the brain specifically treats the feature under investigation (e.g., irony) and to the availability of a certain type of interpretation over time. The most consistent finding across all studies has been a greater posterior positivity for ironic as compared to literal sentences about

500 ms after the target word presentation (e.g., “gifted”). This effect, commonly reported as a P600 effect, seems to be related to late controlled processes of sentence comprehension (Hahne & Friederici, 1999) and increasing complexity at the discourse level (Kaan & Swaab, 2003). Within the domain of verbal irony, the P600 effect has been functionally interpreted as reflecting late inferential processing costs that are necessary in order to derive the final intended meaning (Regel et al., 2011). Importantly, the P600 was usually reported in the absence of an N400 effect (but see Cornejo et al., 2007; Filik et al., 2014). The N400 is a centro-posterior negativity elicited around 400 ms after stimulus onset and it is typically related to lexical-sematic processes (Kutas & Federmeier, 2011). The lack of an N400 effect has been taken as suggesting that irony might not imply early semantic processing difficulties (Balconi & Amenta, 2008; Regel et al. 2011; 2014; Spotorno et al., 2013). In line with this interpretation, it should be noted that the only two studies that did report an N400 effect (followed by a P600) in response to irony adopted paradigms where ironic sentences might have been initially misinterpreted and treated as semantic anomalies (e.g., by mixing semantic violations and ironic sentences: Cornejo et al., 2007; e.g., by using unfamiliar ironic phrases: Filik et al., 2014). Despite the wide interest on verbal irony in the ERP literature, less attention has been given to the role of communicative constraints on the time course of irony processing.

The ERP studies conducted so far examined only one type of verbal irony (i.e., ironic criticism). For instance, the ironic sentences were always preceded by a context describing negative events (e.g., a bad performance by the artists). To the best of our knowledge, there is no available ERP study that tested how and when different contextual biases (i.e., describing positive and negative events) would affect the time course of irony processing. Similarly, the role of other types of communicative constraints, such as speaker identity, has not been widely investigated in the ERP literature on irony interpretation. A few ERP studies focused on literal sentences showed that information inferred from the speaker's voice can influence the utterance analysis within the N400 time window (200-700 ms; Van Berkum, van den Brink, Tesink, Kos, & Hagoort, 2008, see also Bornkessel-Schlesewsky, Krauspenhaar, & Schlesewsky, 2013; Caffarra & Martin, 2019) or even earlier on (100-400 ms; Hanulíková & Carreiras, 2015). In addition, a German ERP study manipulated the communicative style of the speaker (i.e., being ironic or non-ironic) by varying the amount of ironic sentences produced by each speaker during the entire experiment (70% or 30% of the times; Regel et al., 2010). The results seem to be in line with interactive models, as the pragmatic knowledge about the speaker had an impact on early stages of processing (200 ms after stimulus onset). Greater early positivities (P200) were observed when the target sentence matched with the speaker's communicative style (i.e., being ironic for the ironic speaker and literal for the non-ironic speaker), suggesting that the speaker's style constrained the interpretation early on. Late stages of analysis (500 ms after stimulus onset) were also affected by the speaker's characteristics. A P600 irony effect was observed only for the non-ironic speaker, suggesting greater inferential processing costs when the irony was not expected. This P600 pattern was reversed (P600 effect only for the ironic speaker) in a subsequent experimental section where participants' expectations about the speaker's ironic habits were constantly disconfirmed (e.g., the ironic speaker started to produce irony less often), suggesting that late inferential processes are flexible and can be rapidly adjusted based on new pragmatic information. Although this study provides supporting evidence for the interactive view, it manipulated only one communicative constraint (speaker's style, but not contextual biases) and it could not check for early interactions among social and contextual factors, which is specifically predicted by the constraint-satisfaction model.

In the present ERP study we manipulated both contextual biases and speaker identity to increase the chances of teasing apart multiple theoretical models on irony processing.

The present study

In the present electroencephalogram (EEG) study we presented a series of literal and ironic utterances embedded in stories. Contextual biases were manipulated, so that the target sentence could be preceded by a negative or a positive context. Speaker identity was manipulated through speech accent (as in Caffarra et al., 2018), so that each story could be uttered by a native or a foreign speaker. The present ERP results have a significant theoretical contribution as they can discriminate among multiple predictions based on the presence of interactive effects between speaker and context factors and their temporal localization (Pexman et al., 2000).

According to modular serial models any effect of context or speaker identity should emerge at a late stage of processing (Giora, 1997; 2002; Giora et al., 1998; Grice, 1975; Searle, 1979), presumably resulting in a modulation of the P600 effect (Regel et al., 2011; Regel et al., 2014; Spotorno et al., 2013). According to classical interactive models (Gibbs, 1986; 1994; 2002) contextual biases should have an earlier effect on irony interpretation. Finally, an updated and more extreme version of the interactive view (Katz, 2005; Katz et al., 2004; Pexman, 2008) would predict an interaction between contextual biases and speaker identity at an early stage of processing (as early as 200 ms after stimulus onset, according to Regel et al., 2010).

Methods

Participants

Thirty-nine native Spanish listeners participated in the experiment in exchange for payment (10 € per hour). Three participants were excluded because of insufficient number of artifact-free trials (less than 60%, final sample size: 36; 24 women, mean age: 24, SD: 4). None of the participants reported neurological disorders, psychiatric disorders or hearing problems. All participants signed an informed consent form before taking part in the study, which was done in accordance with the World Medical Association Declaration of Helsinki and was approved by the Basque Center on Cognition, Brain and Language ethics committee.

Materials

The experimental materials were taken from Caffarra et al. (2018). One hundred-twenty Spanish stories were created (of six sentences each). The target sentence was always the second-to-last sentence and it ended with either an ironic or a literal target word (these two groups of words were matched for lexical features; Caffarra et al. 2018). In addition the pre-target context could describe a positive or a negative event[1]. This led to four versions of the same story (see Table 1): negative context-irony, negative context-literal, positive context-irony, positive context-literal. In addition, each version of the stories was recorded in a native (i.e., Spanish) and in a foreign accent (i.e., British English). There were three female speakers for each accent (three Spanish native speakers and three British English native speakers with a clear and highly intelligible

[1] The distinction between sarcasm and ironic criticism is not always straightforward, and potential overlaps are possible (Averbeck, 2015). It is beyond the aim of this study to distinguish between these two categories. However, if we consider sarcasm as being a positive comment able to convey "a negative ironic argument that is directed to the addressee and generally malevolent" (Averbeck, 2013), only 17% of our stories corresponded to this definition.

accent). Pre-target contexts and target utterances were recorded separately, so that the speakers were not aware of which meaning (literal, ironic) would have been finally attributed to the sentence. In addition, the same target utterance was used for the ironic and the literal condition and all story recordings were obtained by cross-splicing pre-target and target utterances in a fully-crossed design. Story recordings were normalized with respect to average root mean squared amplitude. Acoustic features (intonation, duration, speech rate, and pitch) of the target sentence and target word were matched across conditions (Caffarra et al., 2018) so that the distinction between story types (irony vs. literal, positive vs. negative context) was mainly based on lexical cues (rather than prosodic cues). Importantly, differences in speech rate, pitch, and duration could not explain any significant interaction in the ERP results since all interactions involving these acoustic features were not significant (all Fs<1.2; all ps>0.20; Caffarra et al., 2018). This procedure might have reduced the naturalness of the recorded sentences, but it allowed us to control for potential confounds (Anolli, Ciceri, & Infantino, 2002; Rockwell, 2007; but see Bryant, 2010), and clearly isolate the effects of irony, context and speech type. Note that previous behavioral data on the same material confirmed the reliability of our experimental manipulations showing that the intelligibility of the stories was high and matched across conditions, that the foreign accent was easily detectable, and that the ironic stories were clearly perceived as ironic while their literal counterparts were literally interpreted (Caffarra et al. 2018). Four experimental lists of 240 trials (30 items per condition) were created so that each version of the pre-target context appeared twice, once followed by an ironic statement and once followed by a literal one. For each list, half of the trials were in a native accent and the other half in a foreign accent.

Table 1. Examples of experimental materials (the English translation is provided in the bottom part). Target sentences are in italics. Target words are underlined and the underlining style marks the EEG comparisons of interest.

	Negative context	Positive context
Literal	Compré un billete de lotería de mi pueblo.	Compré un billete de lotería de mi pueblo.
	Leí más detalles en internet.	Leí más detalles en internet.
	El primer premio era una chistorra de 5 €.	El primer premio era un viaje de 10.000 €.
	Le dije a mi novio:	Le dije a mi novio:
	¡Vaya premio más triste!	*¡Vaya premio más tentador!*
	Al final no me tocó ningún premio.	Al final no me tocó ningún premio.
Ironic	Compré un billete de lotería de mi pueblo.	Compré un billete de lotería de mi pueblo.
	Leí más detalles en internet.	Leí más detalles en internet.
	El primer premio era una chistorra de 5 €.	El primer premio era un viaje de 10.000 €.
	Le dije a mi novio:	Le dije a mi novio:
	¡Vaya premio más tentador!	*¡Vaya premio más triste!*
	Al final no me tocó ningún premio.	Al final no me tocó ningún premio.

Literal	I bought a lottery ticket in my town.	I bought a lottery ticket in my town.
	I read more details online.	I read more details online.
	The first prize was a 5 € sausage.	The first prize was a 10.000 € trip.
	I said to my boyfriend:	I said to my boyfriend:
	What a sad prize!	*What a tempting prize!*
	In the end I didn’t win anything.	In the end I didn’t win anything.
Ironic	I bought a lottery ticket in my town.	I bought a lottery ticket in my town.
	I read more details online.	I read more details online.
	The first prize was a 5 € sausage.	The first prize was a 10.000 € trip.
	I said to my boyfriend:	I said to my boyfriend:
	What a tempting prize!	*What a sad prize!*
	In the end I didn’t win anything.	In the end I didn’t win anything.

Procedure

Participants were seated in a dimly lit and sound-attenuated room. They were asked to carefully listen to each story and reply to a yes/no question appearing after 25% of the trials. They were informed at the beginning of the study that the Spanish stories were uttered by Spanish and British native speakers. This was done to make sure that the accent difference was clear to all participants and that listening to accented speech did not come as a surprise. Each trial began with the symbol *.* that stayed at the center of the screen for one second, during which the participants were encouraged to blink. After a 300-ms blank a story was auditorily presented through two speakers (approximate mean duration: 19 sec, SD: 2). A fixation cross was displayed on the screen during the auditory presentation. The participants were asked to minimize their eye movements and keep their gaze on the cross. After a quarter of the sentences a yes/no comprehension question was displayed on the screen for a maximum of 3 seconds. The participants replied by pressing one of the two response buttons. The questions focused on the content of the story (e.g., "Did they talk about a lottery ticket?" after the story displayed in Table 1) and they were used in order to make sure participants payed attention to the auditory material. The questions never focused on the experimental manipulations under interest (e.g., ironic/literal utterances) in order to discourage strategic behavior (e.g., paying attention only to the target utterance). At the beginning of the EEG session, the participants listened to a short audio from each speaker in order to familiarize themselves with the six different voices. As the overall experiment lasted more than three hours, it was divided into two sessions (120 items each, approximate session duration: two hours) and the participants came twice to the laboratory (mean time distance between the two sessions: 15 days). Each story (i.e., specific combination of pre-target and target utterance) was presented only once in the whole experiment (i.e., two sessions). Each target sentence was presented once per session (once with an ironic meaning and once with a literal meaning). The order of each session was counterbalanced across participants.

EEG recording and analyses

The EEG was recorded from 27 channels placed in an elastic cap: Fp1, Fp2, F7, F8, F3, F4, FC5, FC6, FC1, FC2, T7, T8, C3, C4, CP5, CP6, CP1, CP2, P3, P4, P7, P8, O1, O2, Fz, Cz, Pz. There were six external electrodes: two on the mastoids, two on the ocular canthi, one above and one below the right eye. The online reference was to the left mastoid and the sampling rate was 500 Hz. Impedance was kept below 5 KΩ for the electrodes on the scalp and below 10 KΩ for the external channels. EEG data were re-referenced offline to the average activity of the two mastoids. A bandpass filter of 0.01–30 Hz (12 dB/oct) was applied. Vertical and horizontal eye movements were corrected using the Independent Components Analysis (ICA). The EEG of each subject was decomposed into independent components. The components that explained the highest percentage of the variance in the Veog and Heog channels (recorded as the voltage difference between electrodes placed around the eyes) were identified. The time course and the topographic distribution of these components were visually inspected to ensure they represented real artifacts, and then subtracted from the original data. Residual artifacts exceeding ±100 μV in amplitude were rejected. On average, 15% of trials were excluded, with no difference across conditions ($F(7,280)<1$, $p= 0.71$). For each target noun, an epoch of 1700 ms was obtained including a 200 ms pre-stimulus baseline. For each condition average ERPs were computed time locked to the onset of the target noun and baseline corrected to -200 to 0 ms.

Statistical analyses were carried out in the following time windows defined on visual inspection: 150-300, 500-1000, 1000-1500 ms. Three topographic factors were included in the statistical analyses: Hemisphere (left and right), Distance to midline (DML, two levels, close to midline: F3, F4, FC1, FC2, C3, C4, CP1, CP2, P3, P4, far from midline: F7, F8, FC5, FC6, T7, T8, CP5, CP6, P7, P8), and Anterior-Posterior factor (AP, five levels, frontal; F7, F3, F4, F8, fronto-central: FC5, FC1, FC2, FC6, central: T7, C3, C4, C8, centro-posterior: CP5, CP1, CP2, CP6, parietal: P7, P3, P4, P8)[2]. A repeated measures analysis of variance (ANOVA) was performed for each time window including: Type (Irony, Literal), Accent (Foreign, Native), Context (Positive, Negative), and the three topographic factors as within-subject factors. Data acquired from midline electrodes (Fz, Cz, Pz) were separately analyzed and included in an ANOVA with Type, Accent, Context, and AP (three levels: frontal, central, parietal) as within-subject factors. The Greenhouse-Geisser procedure was applied when the sphericity assumption was violated. Post-hoc *t*-tests were corrected for multiple comparisons (Hochberg, 1988). Effects of topographic factors are reported only when they interacted with the experimental factors.

Results

Participants accurately responded to the comprehension questions, suggesting that they paid attention to the content of the auditory material along the experiment (mean: 85%, with no differences across conditions, Fs<1).

150-300 ms

The effect of Accent was significant (Accent x DML x AP: $F(4,140)=3.32$, $p<.05$, $\eta^2_p=0.09$), with native accent eliciting more negative waveforms than foreign accent

[2] Additional analyses using different topographic factors (Anteriority, Laterality) showed interactive effects similar to those reported here.

especially over central sites (central: $t(35)=3.91$, $p<.01$; other sites: $ps>.05$). In addition the factor Context was significant (main effect: $F(1,35)=13.88$, $p<.001$, $\eta^2_p=0.28$; Context x DML x AP: $F(4,140)= 3.87$, $p<.05$, $\eta^2_p=0.10$), reflecting greater centro-posterior negativity for positive as compared to negative contexts (fronto-central sites: $ts>2.40$, $p<.05$; centro-posterior: $ts>3$, $p<.01$). Finally, the factor Type did not reach significance ($Fs<1.4$, $ps>.05$) but it interacted with Accent and Context (Type x Accent x Context x DML: $F(1,35)=4.47$, $p<.05$, $\eta^2_p=0.11$). Additional ANOVAs were run for each context (negative and positive). For negative contexts no effect of Type were observed ($Fs<1.3$, $ps>.05$ including interactions with other experimental factors, see Figure 1), while for positive contexts the effect of Type was significant (Type x Hemisphere: $F(1,35)=4.48$, $p<.05$, $\eta^2_p=0.11$) and it also interacted with Accent (Type x Accent: $F(1,35)=4.61$, $p<.05$, $\eta^2_p=0.12$; Type x Accent x DML: $F(1,35)=8.30$, $p<.01$, $\eta^2_p=0.19$, see Figure 2). This suggests a greater central negativity for ironic as compared to literal utterances only in the case of native accent and positive context (close to midline, native: $t(35)=2.18$, $p<.05$; foreign: $t(35)=1.36$, $p=.18$; far from midline: $ps>.05$). The general ANOVA on the midline showed a main effect of Context ($F(1,35)=13.88$, $p<.001$, $\eta^2_p=0.28$).

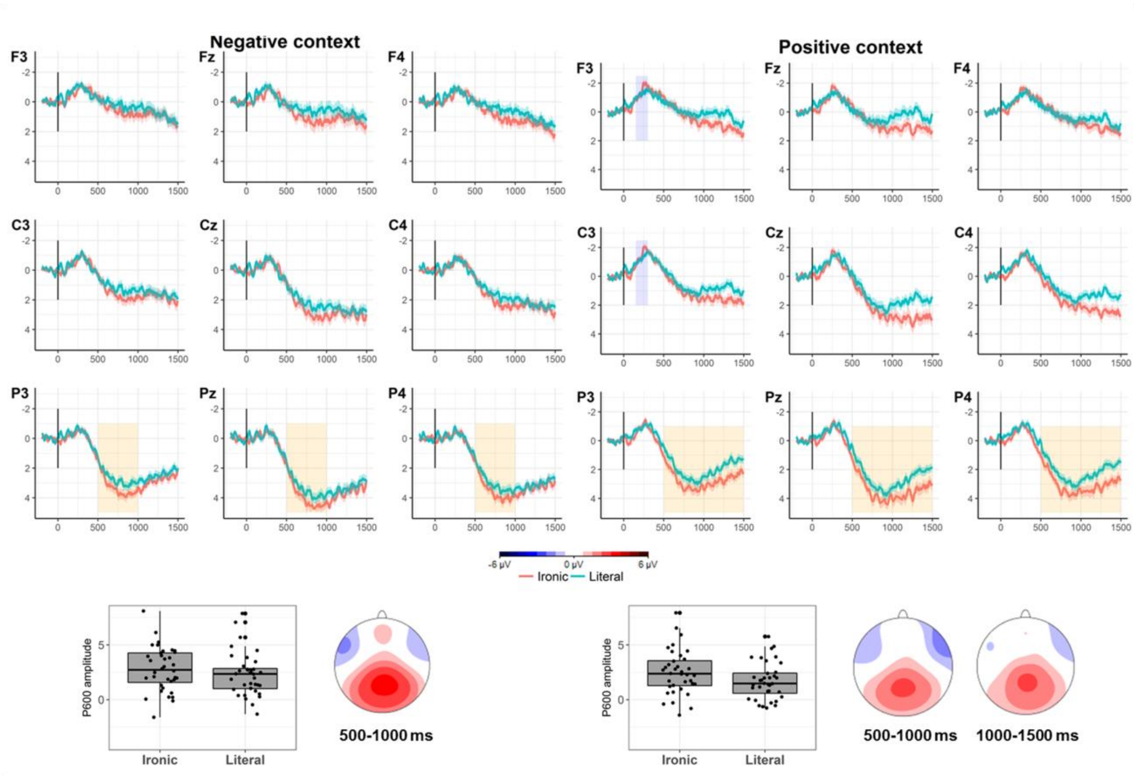

Figure 1. Top panel: Grand-average waveforms for each context. Negativity is plotted upwards. Shaded areas represent one standard error from the average waveform. The x-axis reports time (in ms) and the y-axis reports brain response amplitude (in µV). The time windows at F3 and C3 mark the time range where Type interacted with Context and Accent. The time windows at posterior electrodes mark significant P600 effects for irony (500-1000 ms for negative context, 500-1000 and 1000-1500 ms for positive context). **Bottom panel:** The box plots showed the P600 amplitude for the ironic and the literal condition (calculated as the average activity from centro-posterior sites: C3, C4, Cz, P3, P4, Pz) in the time window of 500-1000 ms for the negative context, and of 500-1500 ms for the positive context. Beside each box plot, the topographic distribution of the P600 effect (calculated based on the difference between ironic and literal utterances) is displayed for the time windows of interest.

500-1000 ms

The factor Accent was significant ($F(1,35)=15.74$, $p<.001$, $\eta^2_p=0.31$; Accent x DML: $F(1,35)=17.63$, $p<.001$, $\eta^2_p=0.34$; Accent x AP: $F(4,140)=4.32$, $p<.05$, $\eta^2_p=0.11$; Accent x hemisphere x DML x AP: $F(4,140)=3.62$, $p<.05$, $\eta^2_p=0.10$) with more positive waveforms for native as compared to foreign accent over fronto-central sites (*ts*>4.4; *ps*<.001; other sites *ts*>2.5, *ps*<.05). The factor Type was also significant (Type x DML: $F(1,35)=7.81$, $p<.01$, $\eta^2_p=0.18$; Type x AP: $F(4,140)=5.93$, $p<.01$, $\eta^2_p=0.15$; Type x DML x AP: $F(4,140)=2.99$, $p<.05$, $\eta^2_p=0.08$) with ironic utterances eliciting more positive waveforms than literal utterances over centro-posterior sites (*ts*>2.8, *ps*<.05; other sites: *ps*>.05, see Figure 1). No other significant interaction was observed. Similarly, the ANOVA on the midline showed an effect of Accent ($F(1,35)=15.74$, $p<.001$, $\eta^2_p=0.31$; Accent x AP: $F(4,140)=4.32$, $p<.05$, $\eta^2_p=0.11$), and Type (Type x AP: $F(4,140)=5.93$, $p<.01$, $\eta^2_p=0.14$).

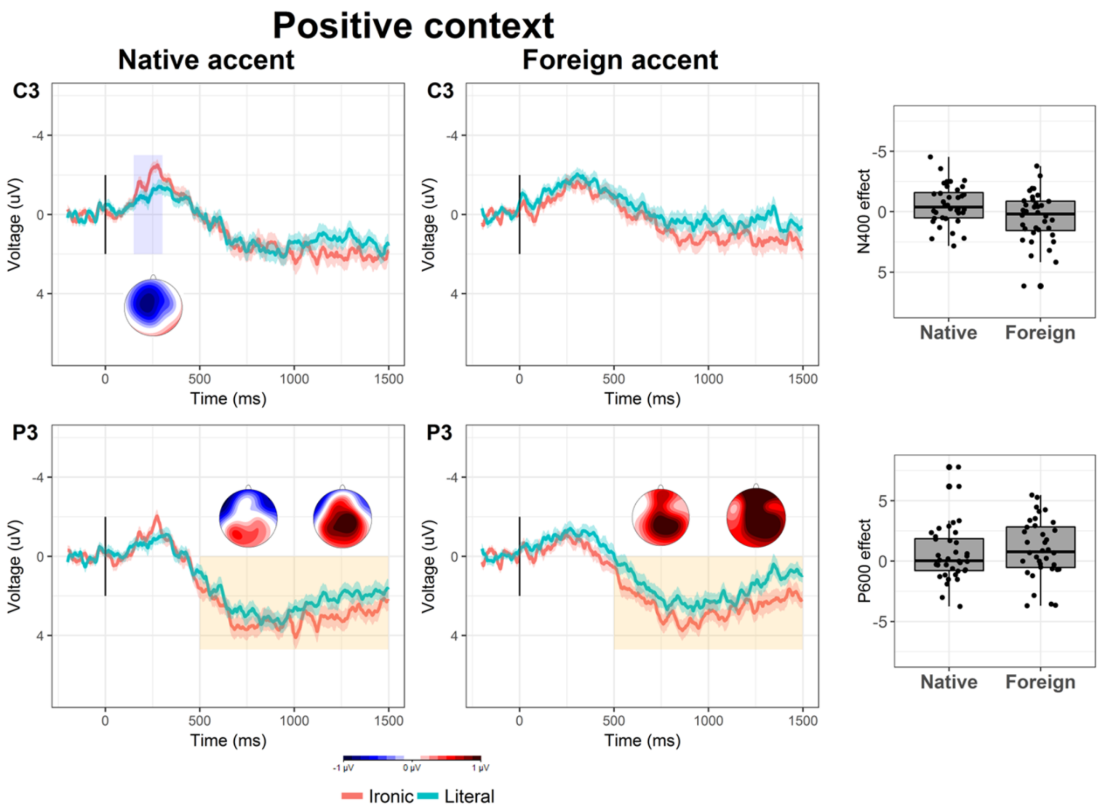

Figure 2. Grand-average waveforms for positive contexts. Negativity is plotted upwards. Shaded areas represent one standard error from the average waveform. The time window at C3 marks the significant effect of irony in native accent. The time windows at P3 mark the P600 irony effects for each accent. Each time window is displayed together with the topographic distribution of the corresponding effect (calculated based on the difference between ironic and literal utterances). The box plots on the right side show the early negative effect (calculated as the ERP difference between ironic and literal sentences at the channels of the midline and close to the midline: F3, F4, Fz, C3, C4, Cz, P3, P4, Pz) and the P600 effect (calculated as the ERP difference between ironic and literal sentences at the centro-posterior channels: C3, C4, Cz, P3, P4, Pz) for native and foreign accent.

1000-1500 ms

The effect of Accent was still significant (Accent x DML: $F(1,35)=4.84$, $p<.05$, $\eta^2_p=0.12$; Accent x hemisphere x DML x AP: $F(4,140)=6.26$, $p<.001$, $\eta^2_p=0.15$), but the follow-up analyses did not show significant differences across accents (all *p*s>.05). There was an effect of Type ($F(1,35)=6.65$, $p<.05$, $\eta^2_p=0.16$; Type x DML: $F(1,35)=10.41$, $p<.01$, $\eta^2_p=0.23$), with a greater positivity for ironic as compared to literal utterances over centro-posterior sites (*t*s>2.6; *p*s<.05; other sites: *p*s>.05). In addition the factor Type interacted with Accent, Context, and topographic factors (Type x Accent x Context x DML x hemisphere: $F(1,35)=4.22$, $p<.05$, $\eta^2_p=0.11$). Follow-up ANOVAs for each context showed that in negative contexts the factor Type was not significant (*F*s<1.7, *p*s>.05 including interactions with other experimental factors), while for positive contexts there was a significant effect of Type ($F(1,35)=7.35$, $p<.05$,

η^2_p=0.17; Type x DML: $F(1,35)=7.29$, $p<.05$, η^2_p=0.17; Type x AP: $F(4,140)=4.57$, $p<.05$, η^2_p=0.12) suggesting a greater positivity for ironic as compared to literal sentences over centro-posterior sites (*t*s>2.7, *p*s<.05; other sites: *p*s>.05, see Figure 1). In addition, for positive contexts the factor Type marginally interacted with Accent and topographic factors (Type x Accent x Hemisphere x DML: $F(1,35)=3.91$, $p=.06$, η^2_p=0.10), suggesting that the P600 effect for irony was stronger for foreign accent over right sites (right electrodes, foreign: $t(35)=2.61$, $p<.05$; native: $t(35)=1.74$, $p=.09$; left electrodes, all *p*s>.10, see Figure 2). The ANOVA on the midline showed an effect of Type ($F(1,35)=6.65$, $p<.05$, η^2_p=0.16).

In summary, a clear P600 effect (500-1000 ms) was observed for ironic as compared to literal utterances. In addition, both context and accent had an impact on utterance interpretation at an early stage of processing (150-300 ms), showing a greater negativity for ironic as compared to literal sentences only in the case of positive context and native accent. Finally, a triple marginal interaction was observed at a later stage of processing (1000-1500 ms), suggesting that the P600 effect was longer-lasting in the case of positive contexts, especially for foreign accent.

Discussion

The present ERP study aimed at teasing apart modular and interactive theoretical predictions on verbal irony processing. It explored the impact of contextual biases and speaker identity on the time course of ironic utterance analysis. ERPs were recorded in response to ironic and literal sentences, which could be preceded by negative or positive contexts. Each sentence was embedded into a story uttered by a native speaker speaking in her native language, or a foreign speaker speaking in her second language with a strong foreign accent. The ERP results showed that ironic sentences consistently elicited a greater P600 as compared to literal utterances, in line with previous ERP studies (Cornejo et al., 2007; Filik et al., 2014; Regel et al. 2010; 2011; 2014; Spotorno et al., 2013). This ERP effect has been interpreted as reflecting late inferential processes based on high-level information (e.g., pragmatic conventions, conversations rules, expectations about the interlocutor) that are necessary to successfully derive the final intended meaning (Regel et al., 2011). The present ERP findings also showed that the processing of ironic utterances was modulated by communicative constraints at an early (150-300 ms) and a late (1000-1500 ms) stage of analysis. These results are compatible with previous ERP findings showing early and late speaker-related effects on irony analysis (Regel et al., 2010) and on literal utterances (Bornkessel-Schlesewsky et al., 2013; Caffarra & Martin, 2019; Hanulíková & Carreiras, 2015; Van Berkum et al., 2008). The presence of early effects of context is difficult to reconcile with modular models (Giora, 1997; 2002; Giora et al., 1998; Grice, 1975; Searle, 1979), and it is rather in line with interactive models (Gibbs, 1986; 1994; 2002; Katz, 2005; Katz et al., 2004; Pexman, 2008). Importantly, the present study showed a triple interaction between irony, contextual biases and speaker's characteristics as early as 150 ms after stimulus presentation. This finding represents the first ERP evidence fully supporting the constraint-satisfaction model (Katz, 2005; Katz et al., 2004; Pexman, 2008). It is compatible with the idea that social and contextual sources of information are quickly available and interact from the earliest stage of processing in order to reach the most likely interpretation of the utterance (Katz, 2005; Katz et al., 2004; Pexman, 2008).

Follow-up tests of the triple interaction can also inform us about what types of communicative constraints are particularly influential in the on-line processing of irony. For instance, positive contexts seem to heavily impact the interpretation of the subsequent target comment, in line with previous behavioral evidence (Caffarra et al., 2018). In this case, the processing of irony started to differ from the processing of literal meaning 150 ms after stimulus onset. This was observed only in the case of native speakers, with irony eliciting a greater centro-posterior negativity as compared to literal sentences. Based on its polarity and distribution this ERP effect can be categorized as an N400-like effect (Kutas & Federmeier, 2011). However, the latency appeared to be shorter as compared to the typical N400 effect. It should be noticed that there is a certain degree of variability in the N400 time window boundaries reported in the literature (Kutas & Federmeier, 2011). This is particularly true within the auditory domain, where the incremental nature of the auditory signal together with the uniqueness point position can reduce the latency of the N400 effect (O'Rourke & Holcomb, 2002). For instance, previous ERP studies on spoken sentences reported early N400-like effects, with a time range similar to the present study (Hanulíková & Carreiras, 2015; Shen, Staub, & Sanders, 2013; Ye, Luo, Friederici, & Zhou, 2006). The early latency of the negative effect here reported might be due to the fact that most of the target words were relatively short (mean duration: 635 ms, SD: 168 ms; average number of syllables: 3.21, SD: 1.08) and had an early uniqueness point (between the second and third syllable; O'Rourke & Holcomb, 2002). N400-like effects have not often been reported in previous ERP studies on verbal irony (Balconi & Amenta, 2008; Regel et al. 2011; 2014; Spotorno et al., 2013), but it should be noted that previous studies only focused on negative contexts. As far as we know, there are only two ERP studies using negative contexts that did report a greater N400-P600 pattern for ironic as compared to literal sentences (Cornejo et al., 2007; Filik et al., 2014). In contrast to other studies, Cornejo et al. (2017) adopted a paradigm where irony was presented together with semantically implausible sentences and participants had to explicitly judge whether each sentence made sense. The authors concluded that the experimental design (having irony only 1/3 of the times and having irony and semantic anomalies presented together) and the task (being focused on the semantic plausibility of the sentence) made irony low-predictable and led participants to initially treat ironic sentences as semantic anomalies. In Filik et al. (2014), a biphasic N400-P600 response was reported in the case of unfamiliar irony (i.e., phrases not typically used to convey irony, which were identified based on an offline rating). Here, the N400 effect (together with early effects in eye-tracking measures) was considered to reflect initial difficulties in constructing the meaning of the target word in relation to the previous context. The N400-like effect observed in the present study might reflect similar semantic difficulties. Ironic comments inserted in positive contexts express compliments (i.e., ironic praise) and they are much less frequent, familiar and prototypical than ironic comments in negative contexts (i.e., ironic criticism; Averbeck, 2015; Bromberek-Dyzman, 2015; Bruntsch & Ruch, 2017; De Groot et al., 1995; Gibbs, 1986). This difference between contexts might result in different expectations and/or integration costs (Bruntsch & Ruch, 2017). While a negative context echoes irony and can trigger predictions about possible ironic comments, positive contexts do not often provide anticipatory cues to irony and the detection of the final intended meaning might require greater cognitive demands (Bruntsch & Ruch, 2017; Kreuz & Glucksberg, 1989; Nakassis & Snedeker, 2002). In this last case where irony is not much expected, nonliteral comments might be initially treated as a logical contradiction, leading to an N400 effect reflecting semantic processing costs (similar to Cornejo et al., 2007 and Filik et al., 2014). In other words, the N400 irony effect observed in native accent reflects the integration costs of the target word meaning in a positive context. This effect suggests that, in this particular case, irony has been initially misunderstood as a literal comment, probably because listeners did not expect irony after a positive context.

The present data suggest that contextual cues have important effects in the early stage of verbal irony analysis (Filik et al., 2014; see also Bambini, Bertini, Schaeken, Stella, & Di Russo, 2016 for similar results in metaphor interpretation). Importantly, this data showed for the first time that this is true for native accented speech, but not for foreign accented speech. This might suggest that ironic compliments were initially misinterpreted only when produced by native speakers.

In the case of foreign accent no clear semantic integration difficulties (indexed by N400-like effects) were reported, and this was true for both positive and negative context. Different explanations can be proposed to account for this finding. One possibility is that adverse listening conditions (such as foreign accented speech) might disrupt anticipatory processes, with a concomitant reduction of N400 effects (as observed in Strauβ, Kotz, & Obleser, 2013). Another possibility is related to the stereotypical knowledge associated with the speaker. Knowing that second language speakers may not reliably convey their intended meanings and that their second language production is prone to errors (Bergen & Grodner, 2012; Lev-Ari, 2015; Fairchild & Papafragou, 2017), listeners might have kept their initial interpretation relatively open, with no specific semantic expectations (Caffarra et al., 2018). Further investigation is needed in order to confirm these tentative explanations.

Overall, the ERP data on the early stage of analysis (N400 time window) reflect lexico-semantic integration of the target word into the context. Our findings suggest that while in native accent listeners put in place semantic expectations which can lead to additional processing costs, in foreign accent semantic expectations seems to be overall reduced. In the second phase of analysis listeners' brain responses (indexed by the P600 effect) suggest that in both accents high-level information is considered in order to infer the correct intended meaning. Interestingly, this late stage of processing (1000-1500 ms) is modulated by communicative constraints, with longer-lasting P600 effects for irony in the case of positive contexts. This result further extends our knowledge on irony processing showing P600 effects not only with negative but even with positive contexts. This is in line with previous studies suggesting that the late P600 effect is modulated by communicative constraints (Regel et al., 2011). It is also compatible with the idea that positive contexts make irony more difficult to be understood and require greater inferential processing costs as compared to negative contexts (Gibbs, 1986; Ivanko & Pexman, 2003; Woodland & Voyer, 2011). Finally, the marginal triple interaction in this late time window seems to suggest that the P600 irony effect observed for positive contexts was greater for foreign accented speech. This might indicate that greater pragmatic effort is required to interpret low-frequency irony produced by foreigners. This result fits well with recent behavioral findings showing that non-prototypical irony produced by foreigners is perceived as less ironic than non-prototypical irony produced by native speakers (Caffarra et al., 2018). Taken together, these findings are generally in line with previous behavioral studies showing that producing figurative inferences becomes problematic when interlocutors do not share the same linguistic and cultural background (Averbeck, 2015; Averbeck & Hample, 2008; Chaeng & Pell, 2011; Kaufer & Neuwirth, 1982).

Finally, it is important to note that this ERP experiment reported an interaction between contextual biases and speech accent in an experimental situation where prosodic cues to irony were minimized and acoustic features were matched across conditions. The role of prosodic cues in irony perception and processing is still highly debated, with some studies reporting irony-specific acoustic patterns (Anolli et al., 2002; Rockwell, 2007), and others showing no evidence for a phonological/prosodic characterization of irony (Bryant, 2010; Bryant & Fox Tree, 2005) and no clear neural effects due to prosodic cues to irony (Regel et al., 2011). The present results cannot speak to this debate, as acoustic and prosodic features were controlled across conditions, and additional studies are needed in order to better specify the behavioral and neural effects of prosody on irony analysis. However, we think that the presence of an early interaction between contextual biases and speaker accent in such a "prosodically controlled" situation further strengthens our conclusions, showing that early interactive effects of communicative constraints can be observed even when irony is not cued by prosody.

To conclude, the present ERP study examined the role of different communicative constraints on the time course of verbal irony processing in order to tease apart distinct theoretical proposals. The present findings are not compatible with modular models (Giora, 1997; 2002; Giora et al., 1998; Grice, 1975; Searle, 1979), but they are rather in line with interactive models (Gibbs, 1986; 1994; 2002; Katz, 2005; Katz et al., 2004; Pexman, 2008). Specifically, the early interactive effects reported here fully support the constraint-satisfaction model and they suggest that multiple communicative constraints influence irony analysis as early as 150 ms after stimulus onset (Katz, 2005; Katz et al., 2004; Pexman, 2008).

Acknowledgements:

The research was supported by the Spanish State Research Agency (Severo Ochoa excellence accreditation SEV-2015-0490) and the Basque Government (BERC 2018-2021). The authors were funded by the Spanish Ministry of Economy and Competitiveness (SC and CM: PSI2014-54500, PSI2017-82941-P; SC: IJCI-2016-27702), the Basque Government (SC and CM: Eusko Jaurlaritza; PI_2015_1_25), and the European Research Council (SC: H2020-MSCA-IF-2018-837228). The authors wish to thank Rodrigo Loredo and Leah Gosselin for their insightful comments on the manuscript. The funding source had no involvement in the study preparation.

Conflict of Interest Statement: There is no conflict of interest.

Author Contributions: SC designed the experiment, analyzed the data and drafted the paper. AM and EM collected the data. CM designed the experiment and drafted the paper.

Data Accessibility Statement: The materials and the EEG data are available on the OSF platform (https://osf.io/m8ynw/).

Abbreviations:

EEG (electroencephalogram)

ERP (event-related potential)

References

Anolli, L., Ciceri, R. & Infantino, M. G. (2002) From "Blame by praise" to "praise by blame": Analysis of vocal patterns in ironic communication. International *J. Psychol.*, **37**, 266–276. doi: 10.1080/00207590244000106

Averbeck, J. M. (2013) Comparisons of Ironic and Sarcastic Arguments in Terms of Appropriateness and Effectiveness in Personal Relationships. *Argumentation and Advocacy*, **50**, 47–57. doi: 10.1080/00028533.2013.11821809

Averbeck, J. M. (2015) Irony in interpersonal conflict scenarios. Differences in ironic argument endorsement and suppression across topics. *J. Argumentation in Context*, **4**, 87–109. doi: 10.1075/jaic.4.1.05ave

Averberk, J. M. & Hample, D. (2008) Ironic message production: how and why we produce ironic messages. *Commun. Monogr.*, **75**, 396–410. doi: 10.1080/03637750802512389

Balconi, M. & Amenta, S. (2008) Isn't it ironic? An analysis on the elaboration of ironic sentences with ERPs. *The Open Applied Linguistics Journal*, **1**, 9–17. doi: 10.2174/1874913500801010009

Bambini, V., Bertini, C., Schaeken, W., Stella, A. & Di Russo, F. (2016) Disentangling Metaphor from Context: An ERP Study. *Front. Psychol.*, 3, 7–559. doi: 10.3389/fpsyg.2016.00559

Bergen, L. & Grodner, D. J. (2012) Speaker knowledge influences the comprehension of pragmatic inferences. *J. Exp. Psychol. Learn.*, **38**, 1450–1460. doi: 10.1037/a0027850

Bornkessel-Schlesewsky, I., Krauspenhaar, S. & Schlesewsky, M. (2013) Yes, you can? A speaker's potency to act upon his words orchestrates early neural responses to message-level meaning. *PLoS One*, **24**, e69173. doi: https://www.doi.org/10.1371/journal.pone.0069173

Bromberek-Dyzman, K. (2015) Irony processing in L1 and L2: same or different? In Heredia, R. R. (ed), *Bilingual figurative language processing*. Cambridge University Press, Cambridge, pp. 268–298.

Bruntsch, R. & Ruch, W. (2017) Studying irony detection beyond ironic criticism: Let's include ironic praise. *Front. Psychol.*, **8**, 606. doi: 10.3389/fpsyg.2017.00606

Bryant, G. A. (2010) Prosodic contrasts in ironic speech. *Discourse Process.*, **47**, 545–566. doi: 10.1080/01638530903531972

Bryant, G. A. & Fox Tree, J. E. (2005) Is there an ironi tone of voice? *Lang. Speech*, **48**, 257–277. doi: 10.1177/00238309050480030101

Caffarra, S., Michell, E. & Martin, C.D. (2018) The impact of foreign accent on irony interpretation. *PLoS One*, **13**(8). e0200939. doi: 10.1371/journal.pone.0200939.

Caffarra, S. & Martin, C. (2019) Not all errors are the same: ERP sensitivity to error typicality in foreign accented speech perception. *Cortex*, **116**, 308–320. doi: 10.1016/j.cortex.2018.03.007.

Cheang, H. S. & Pell, M. D. (2011) Recognizing sarcasm without language: A cross-linguistic study of English and Cantonese. *Pragmat.Cogn.*, **19**, 203–223. doi: 10.1075/bct.55.02che

Climie, E. A. & Pexman, P. M. (2008) Eye Gaze Provides a Window on Children's Understanding of Verbal Irony. *J. Cogn. Dev.*, **9**, 257–285. doi: 10.1080/15248370802247939

Cornejo, C., Simonetti, F., Aldunate, N., Ibañez, A., López, V. & Melloni, L. (2007) Electrophysiological evidence of different interpretative strategies in irony comprehension. *J. Psycholinguist. Res.*, **36**, 411–430. doi: 10.1007/s10936-007-9052-0

De Groot, A., Kaplan, J., Rosenblatt, E., Dews, S. & Winner, E. (1995) Understanding versus discriminating nonliteral utterances: evidence for a dissociation. *Metaphor Symb. Act.*, **17**, 255–273. doi:10.1207/s15327868ms1004_2

Dews, S. & Winner, E. (1999) Obligatory processing of literal and nonliteral meanings in verbal riony. *J. Pragmatics*, **31**, 1579–1599. doi: 10.1016/S0378-2166(99)00005-3

Fairchild, S. & Papafragou, A. (2017) Flexible expectations of speaker informativeness shape pragmatic inference. *Proceedings of the 40th Annual Penn Linguistics Conference*, **23**, 49–56.

Fein, O., Yeari, M. & Giora, R. (2015) On the priority of salience-based interpretations: The case of sarcastic irony. *Intercult. Pragmat.*, **12**, 1–32. doi: 10.1515/ip-2015-0001

Filik, R., Leuthold, H., Wallington, K. & Page, J. (2014) Testing theories of irony processing using eye-tracking and ERPs. *J. Exp. Psychol. Learn.*, **40**(3), 811–828. doi: 10.1037/a0035658

Filik, R. & Moxey, L. M. (2010) The on-line processing of written irony. *Cognition*, **116**, 421–436. doi: 10.1016/j.cognition.2010.06.005

Gibbs, R. W. (1986) On the psycholinguistics of sarcasm. *J. Exp. Psychol. Gen.*, **115**, 3–15. doi: 10.1037/0096-3445.115.1.3

Gibbs, R. W. (1994) Figurative thought and figurative language. In Gernsbacjer, M. A. (ed), *Handbook of psycholinguistics*. Academic Press, New York, pp. 411–446.

Gibbs, R. W. (2001) Evaluating contemporary models of figurative language understanding. *Metaphor Symbol*, **16**, 317–333. doi: 10.1080/10926488.2001.9678900

Gibbs, R. W. (2002) A new look at literal meaning in understanding what is said and implicated. *J. Pragmatics*, **34**, 457–486. doi: 10.1016/S0378-2166(01)00046-7

Giora, R. (1997) Understanding figurative and literal language: The graded salience hypothesis. *Cogn. Linguist.*, **7**, 183–206. doi: 10.1515/cogl.1997.8.3.183

Giora, R. (2002) Literal vs. figurative language: different or equal? *J. Pragmatics*, **34**, 487–506. doi: 10.1016/S0378-2166(01)00045-5

Giora, R., Fein, O. & Schwartz, T. (1998) Irony: graded salience and indirect negation. *Metaphor Symbol*, **13**, 83–101. doi: 10.1207/s15327868ms1302_1

Grice, H. P. (1975) Logic and conversation. In Cole, P. & Morgan, J.L. (eds), *Syntax and Semantics*. Academic Press, New York, pp. 41–58.

Hahne, A. & Friederici, A. D. (1999) Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *J. Cognitive Neurosci.*, **11**(2), 194–205. doi: 10.1162/089892999563328

Hanulíková, A. & Carreiras, M. (2015) Electrophysiology of subject-verb agreement mediated by speakers' gender. *Front. Psychol.*, **6**, 1396.doi: https://www.doi.org/10.3389/fpsyg.2015.01396

Hochberg, B. (1988) A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, **75**(4), 800–802. doi: 10.1093/biomet/75.4.800

Ivanko, L. S. & Pexman, M. P. (2003) Context congruity and irony processing. *Discourse Process.*, **35**, 241–279. doi: 10.1207/S15326950DP3503_2

Kaan, E. & Swaab, T. Y. (2003) Repair, revision, and complexity in syntactic analysis: an electrophysiological differentiation. *J. Cognitive Neurosci.*, **15**, 98–110. doi: 10.1162/089892903321107855

Katz, A. N. (2005) Discourse and sociocultural factors in understanding nonliteral language. In Colston, H. & Katz, A. N. (eds), *Figurative language comprehension: Social and cultural influences*. Erlbaum, Mahwah, pp. 183–207.

Katz, A. N., Blasko, D., G. & Kazmerski, V. A. (2004) Saying what you don't mean: social influence on sarcastic language processing. *Curr. Dir. Psychol. Sci.*, **5**, 186–189. doi: 10.1111/j.0963-7214.2004.00304.x

Katz, A. N. & Pexman, P. M. (1997) Interpreting figurative statements: Speaker occupation can change metaphor to irony. *Metaphor Symbol*, **12**, 19–41. doi: 10.1207/s15327868ms1201_3

Kaufer, D. S. & Neuwirth, C. M. (1982) Foregrounding norms and ironic communication. *Q. J. Speech*, **68**, 28–36. doi: 10.1080/00335638209383589

Kreuz, R. & Glucksberg, S. (1989) How to be sarcastic: the echoic reminder theory of verbal irony. *J. Exp. Psychol. Gen.*, **118**, 374–386. doi: 10.1037/0096-3445.118.4.374

Kutas, M. & Federmeier, K. D. (2011) Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.*, **62**, 621–647. doi: 10.1146/annurev.psych.093008.131123.

Lev-Ari, S. (2015) Comprehending non-native speakers: theory and evidence for adjustment in manner of processing. *Front. Psychol.*, **5**, 1546. doi: 10.3389/fpsyg.2014.01546

Nakassis, C. & Snedeker, J. (2002) Beyond sarcasm: intonation and context as relational cues in children's recognition of irony. In Greenhill, A., Hughs, M., Littlefield, H. & Walsh, H. (eds), *Proceedings of the Twenty-sixth Boston University Conference on Language Development*. Cascadilla Press, Somerville, pp. 429–440.

O'Rourke, T. B. & Holcomb, P. J. (2002) Electrophysiological evidence for the efficiency of spoken word processing. *Biol. Psychol.*, **60**, 121–50. doi: 10.1016/S0301-0511(02)00045-5

Pexman, P. M. (2008) It's fascinating research: the cognition of verbal irony. *Curr. Dir. Psychol. Sci.*, **17**, 286–290. doi: 10.1111/j.0963-7214.2004.00304.x

Pexman, P. M., Ferretti, T. R. & Katz, A. N. (2000) Discourse factors that influence online reading of metaphor and irony. *Discourse Process.*, **29**, 201–222. doi: 10.1207/S15326950dp2903_2

Pexman, P. M. & Olineck, K. M. (2002) Understanding irony: How do stereotypes cue speaker intent? *J. Lang. Soc. Psychol.*, **21**, 245–274. doi: 10.1177/0261927X02021003003

Regel, S., Coulson, S. & Gunter, T. C. (2010) The communicative style of a speaker can affect language comprehension? ERP evidence from the comprehension of irony. *Brain Res.*, **1311**, 121–135. doi: 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. *J. Cognitive Neurosci.*, **23**, 277–293. doi: 10.1162/jocn.2010.21411

Regel, S., Meyer, L. & Gunter, T. (2014) Distinguishing processes reflected by P600 effects: evidence from ERPs and neural oscillations. *PLoS One*, **9**, e96840. doi: 10.1371/journal.pone.0096840

Rockwell, P. (2007) Vocal features of conversational sarcasm: A comparison of methods. *J. Psycholinguistic Res.*, **36**, 361–369. doi: 10.1007/s10936-006-9049-0

Searle, J. (1979) Metaphor. In Ortony, A. (ed), *Metaphor and thought*. Cambridge University Press, New York, pp. 92–123.

Shen, E. Y., Staub, A. & Sanders, L. D. (2013) Event-related brain potential evidence that local nouns affect subject-verb agreement processing. *Lang. Cognitive Proc.*, **28**, 498–524. doi: 10.1080/01690965.2011.650900

Sperber, D. & Wilson, D. (1995) *Relevance: Communication and cognition*. Oxford, UK: Blackwell.

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*, **6**, e66839. doi: 10.1371/journal.pone.0066839

Strauβ, A., Kotz, S. & Obleser, J. (2013) Narrowed expectancies under degraded speech: revisiting the N400. *J. Cognitive Neurosci.*, **25**, 1383–1395. doi: 10.1162/jocn_a_00389

Turcan, A. & Filik, R. (2016) An eye-tracking investigation of written sarcasm comprehension: The roles of familiarity and context. *J. Exp. Psychol. Learn.*, **42**(12),1867–1893. doi: 10.1037/xlm0000285

Van Berkum, J. J., van den Brink, D., Tesink, C. M., Kos, M. & Hagoort, P. (2008) The neural integration of speaker and message. *J. Cognitive Neurosci.*, **20**, 580–91. doi: https://www.doi.org/10.1162/jocn.2008.20054

Woodland, J. & Voyer, D. (2011) Context and intonation in the perception of sarcasm. *Metaphor Symbol*, **26**, 227–239. doi: 10.1080/10926488.2011.583197

Ye, Z., Luo, Y., Friederici, A. D. & Zhou, X. (2006) Semantic and syntactic processing in Chinese sentence comprehension: evidence from event related potentials. *Brain Res.*, **1071**, 186–196. doi: 10.1016/j.brainres.2005.11.085