Journal of Memory and Language 99 (2018) 35–48

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Context, facial expression and prosody in irony processing

Gaétane Deliens [a], Kyriakos Antoniou [a,1,2], Elise Clin [a,b], Ekaterina Ostashchenko [a], Mikhail Kissine [a,*]

[a] ACTE at LaDisco and Neuroscience Institute, Université libre de Bruxelles, CP 175, 50 avenue F.D. Roosevelt, 1050 Brussels, Belgium
[b] F.R.S.-FNRS, Fund for Scientific Research, Wallonia-Brussels Federation of Belgium, Belgium

ARTICLE INFO

Article history:
Received 28 April 2017
Revision received 13 October 2017

Keywords:
Irony
Figurative language
Prosody
Facial expression
Context
Eye-tracking

ABSTRACT

While incongruence with the background context is a powerful cue for irony, in spoken conversation ironic utterances often bear non-contextual cues, such as marked tone of voice and/or facial expression. In Experiment 1, we show that ironic prosody and facial expression can be correctly discriminated as such in a categorization task, even though the boundaries between ironic and non-ironic cues are somewhat fuzzy. However, an act-out task (Experiments 2 & 3) reveals that prosody and facial expression are considerably less reliable cues for irony comprehension than contextual incongruence. Reaction time and eye-tracking data indicate that these non-contextual cues entail a trade-off between accuracy and processing speed. These results suggest that interpreters privilege frugal, albeit less reliable pragmatic heuristics over costlier, but more reliable, contextual processing.

Introduction

Imagine that, as you announce that you will not attend a crisis meeting because of a party, your boss replies 'I love your sense of responsibility!'. Most likely, the incongruity of her comment with the conversational context – broadly understood as shared background knowledge or beliefs (in the classic sense of Stalnaker, 2002) – will (correctly) prompt you to interpret it as ironic. While such ironic utterances pervade our daily conversations, irony is notoriously difficult to define in precise terms (e.g. Gibbs, 2000; Gibbs & Colston, 2012, p. 52) and surfaces under many different guises (such as sarcasm, jocularity, hyperbole, rhetorical question, and understatement). Nevertheless, in one sense or another, all ironically intended messages deliberately mismatch the utterance literal content.[3]

Incongruence with the background context, of the kind just illustrated, is a powerful cue for irony (Gerrig & Goldvarg, 2000; Kreuz & Link, 2002). However, there are indications that a statement may still be interpreted as ironic in the absence of such contextual incongruity, provided that other cues are available (e.g. Kowatch, Whalen, & Pexman, 2013; Jacob, Kreifelts, Nizielski, Schutz, & Wildgruber, 2016). In particular, spoken ironic utterances are often associated with a specific facial expression and a distinctive prosody (e.g. Attardo, Eisterhold, Hay, & Poggi, 2003; Rankin et al., 2009). To the extent that such cues to irony do not directly rely on background context, in what follows we will dub them 'non-contextual', as opposed to contextual incongruity.

The precise role non-contextual cues play in irony processing remains ill understood. On one hand, there is some evidence that a global ironic prosody can be correctly discriminated from a non-ironic one (Bryant & Fox Tree, 2005), provided that the statement is uttered in a familiar language (Cheang & Pell, 2011). And, in fact, many experimental designs implicitly presuppose that ironic prosody is efficient, as they use a distinctive prosody to contrast between ironic and literal stimuli (e.g. Chevallier, Noveck, Happé, & Wilson, 2011; Colich et al., 2012; Kowatch et al., 2013). On the other hand, Bryant and Fox Tree (2005) report that a prosodic contour that is successfully discriminated as ironic is also perceptually associated with other dimensions, such as anger or inquisitiveness. Furthermore, the perception of a given prosodic contour as ironic or not may be influenced by the contextual availability of an ironic interpretation (Voyer, Thibodeau, & Delong, 2016).

We submit that while ironic tone of voice and/or ironic facial expression may be correctly *discriminated*, these cues are not necessarily efficient in a genuine process of irony *comprehension*. Arguably, successful social interactions do not reduce to tagging statements as literal or not (viz. discrimination), but require the identification of the speaker's discourse goals, and the selection of

* Corresponding author.
E-mail addresses: gaetane.deliens@ulb.ac.be (G. Deliens), ka353@cam.ac.uk (K. Antoniou), elisclin@ulb.ac.be (E. Clin), ekaterina.ostashchenko@ulb.ac.be (E. Ostashchenko), mkissine@ulb.ac.be (M. Kissine).
[1] Department of Theoretical and Applied Linguistics, University of Cambridge, English Faculty Building, 9 West Road, Cambridge CB3 9DP, United-Kingdom.
[2] School of Humanities, Hellenic Open University, 18 Parodos Aristotelous, Patra 26335, Greece.
[3] Of course, contextual incongruence does not necessarily boil down to manifest falsity; for instance, hyperbolic, but nevertheless literally true statements may be ironic (Kreuz & Roberts, 1995; Sperber & Wilson, 1981).

https://doi.org/10.1016/j.jml.2017.10.001
0749-596X/

an appropriate reaction (viz. comprehension; see Kreuz, 2000). Studies in brain-damaged patients suggest a dissociation between these two processes: some patients fail to understand the speaker's intent when contextual and prosody cues are available, even though they are able to identify the tone of voice as sarcastic (McDonald, 2000; McDonald & Pearce, 1996). Yet, irony processing is usually investigated through tasks in which participants have to judge as quickly as possible if statements are ironic or not, thus measuring only the discrimination component. For instance, Bryant and Fox Tree (2002) found that participants successfully discriminate ironic vs. non-ironic utterances based on their prosody.[4] However, making decisions in a binary, forced-choice task is very different from interpreting a message as would its actual addressee. The precise role of prosody within irony comprehension is further blurred by the fact that Bryant and Fox Tree (2002) found context to be a more powerful cue for ironic judgements than prosody.

A notable exception to such metalinguistic decision paradigms is the study by Kowatch et al. (2013), who designed an innovative 'shopping task' that positions participants as active interpreters. In this experimental design, a puppet faces food items (e.g. an apple and an orange) and utters literal or ironic statements about what it wants to buy (e.g. 'I just love apples'). Only the puppet's tone of voice allows to disentangle ironic criticisms (e.g. 'I just love apples'), literal criticisms (e.g. 'I just hate apples') and literal praise (e.g. 'I just love oranges'). Participants are asked to put in a shopping cart the food item the puppet really wants. In this way, participants' response mirrors their interpretation of the discourse goals of the speaker. The results of Kowatch et al. (2013) display an interesting asymmetry between accuracy and reaction time. The rate of correct responses for ironic items is low (less than 60%), and significantly so relative to literal items. At the same time, the authors report no difference in processing time or in frequencies of first looks to the correct object for ironic and literal criticisms. It could be the case, then, that while ironic prosody and/or facial expression are not very reliable for accurately grasping an ironic communicative intention, they still prompt a rapid, cognitively shallow attribution of ironic intentions to the speaker.

Importantly, Kowatch et al. (2013) did not compare ironic prosody relative to the role of context, so it is unclear whether interpreters still use prosody when context is available, and if yes, whether non-contextual cues merely complement context-based processing or whether they may take precedence over it. There is ample evidence that mastery of irony presupposes complex mental-state attribution skills (e.g. Akimoto, Miyazawa, & Muramoto, 2012; Bryant, 2012; Spotorno & Noveck, 2014). Such mentalising processes require inferring the speaker's intention by assessing the utterance content against the background context. Some theorists hold that any type of pragmatic processing involves complex, context-based inferences about the speaker's communicative intentions (Sperber & Wilson, 2002). Consistently with this idea, in 5- to 7-year-old children, it is the capacity to attribute multilayered mental states, and not ironic prosody, that predicts correct discrimination between irony and white lies (Wimmer & Leekam, 1991; see also Filippova & Astington, 2010).

However, it is also plausible that conversationally experienced interpreters sometimes rely more on salient non-contextual cues than on context. For instance, Deliens, Antoniou, Clin, and Kissine (2017) recently found that in the presence of salient ironic prosody, participants do not engage in context-based perspective-shifting to gauge the sarcastic nature of a message. According to the parallel-constraint-satisfaction account (Katz, 2005; Pexman, 2008), all cues are processed in parallel and activate a certain – possibly ironic – interpretation. However, as acknowledged by Pexman (2008) herself, this model does not currently provide any indication as to the relative weight of different cues. A more radical idea, to which we subscribe, is that the presence of salient, albeit perhaps less reliable, non-contextual cues prompts interpreters to disregard costlier contextual processing. This hypothesis is consistent with the Direct Access view (e.g. Gibbs, 2002), which predicts that interpreters do not always need to analyse literal meaning in full to form a hypothesis about the meaning communicated by the speaker. It is also in line with a model of pragmatics according to which interpreters are driven by considerations of cognitive economy, and do not necessarily engage in extensive context-driven reasoning about speaker's intentions (Kissine, 2016; see also Ferreira & Patson, 2007).

By contrast, Giora's Graded Salience theory (Giora, 2003; Giora, Givoni, & Fein, 2015) holds that, unless the sentence form bears a conventional or by-default association with irony,[5] utterance literal, compositional meaning will necessarily be activated first before being rejected in favor of a contextually computed ironic interpretation. On different grounds, authors like Sperber and Wilson (2002), who hold that any pragmatic processing involves context-based inference of speaker's intentions, would also have to predict that non-contextual cues can supplement, but not replace context in irony comprehension.

Summing up, two related research questions clearly emerge from the current state of the art: one about the reliability of non-contextual cues, and the other about the relative processing roles of contextual and non-contextual cues. In Experiment 1 of this paper we assess the discrimination of ironic prosody relative to neutral prosody, as well as to positive or negative literal prosody; we also test, in the exact same way, the discriminability of ironic facial expression. (While the discrimination of ironic prosody has been previously investigated, to the best of our knowledge no such evidence is available for ironic facial expression.) In Experiments 2 and 3 we assess how the same prosody and facial cues, as well as contextual information impact irony comprehension, using an act-out task inspired by Kowatch et al. (2013). Our Hypothesis 1 is that in a categorization task ironic prosody and ironic facial expression should allow correct discrimination of ironic items. In line with the model put forward by Kissine (2016), as well as with the Direct Access view (Gibbs, 2002), we predict that in the act-out tasks of Experiments 2 and 3 the presence of salient – albeit potentially less reliable – non-contextual cues should prompt interpreters to bypass costlier contextual processing. That is, our Hypothesis 2 is that ironic prosody and facial expression are privileged in irony comprehension at the expense of costlier, but more accurate assessment of the utterance literal content relative to the context. Accordingly, one should expect non-contextual cues to be associated with faster responses; furthermore, if, as we predict, the processing of ironic prosody or facial expression does not supplement context-based assessment of the compositional meaning, non-contextual cues should not entail any accuracy gain relative to contextual incongruence.

Our Hypothesis 2 may also be seen as one possible implementation of the parallel-constraint-satisfaction model of irony interpretation (Katz, 2005; Pexman, 2008). As we already mentioned, this model predicts that contextual and non-contextual cues are processed in parallel. If parallel processing of all cues must be completed before the outputs are weighted and the final interpretation reached, then, contrary to our predictions, the presence of non-contextual cues along with contextual incongruence should

[4] There are many other experimental studies that approach irony exclusively through discrimination; see, for instance, Kreuz and Roberts (1995), Climie and Pexman (2008), Epley, Keysar, Van Boven, and Gilovich (2004), Chevallier et al. (2011), Colich et al. (2012).

[5] So far, evidence for such by-default ironic meanings, outside conventionally ironic constructions, is limited to negative statements of the form 'X is not the most Y' (Giora et al., 2015).