The emotional impact of verbal irony: Eye-tracking evidence for a two-stage process

CrossMark

Ruth Filik [a,*], Emily Brightman [a], Chloe Gathercole [a], Hartmut Leuthold [b]

[a] School of Psychology, University of Nottingham, UK
[b] Department of Psychology, University of Tübingen, Germany

ARTICLE INFO

ABSTRACT

In this paper we investigate the socio-emotional functions of verbal irony. Specifically, we use eye-tracking while reading to assess moment-to-moment processing of a character's emotional response to ironic versus literal criticism. In Experiment 1, participants read stories describing a character being upset following criticism from another character. Results showed that participants initially more easily integrated a hurt response following *ironic* criticism; but later found it easier to integrate a hurt response following *literal* criticism. In Experiment 2, characters were instead described as having an amused response, which participants ultimately integrated more easily following ironic criticism. From this we propose a two-stage process of emotional responding to irony: While readers may initially expect a character to be more hurt by ironic than literal criticism, they ultimately rationalize ironic criticism as being less hurtful, and more amusing.

Introduction

Irony is a form of indirect language, used when the speaker or writer expresses one thing, but implies another (usually the opposite, e.g., Grice, 1975). A classic example would be uttering "What lovely weather!", when the conditions outside are somewhat disappointing. Sarcasm is a specific form of irony, which is used when the target of the comment is a person, with the intent to criticize (Kreuz & Glucksberg, 1989; Leggitt & Gibbs, 2000; Wilson, 2013). An example would be if you had just done something stupid, and your friend laughs and says, "That was clever!". Thus, it is clear that sarcasm is a form of ironic language that is strongly related to emotion (see also Shamay-Tsoory, Tomer, & Aharon-Peretz, 2005), and is likely to serve complex communicative functions that would not be achieved by speaking directly (e.g., Brown & Levinson, 1987; Colston, 1997; Dews & Winner, 1995; Leech, 1983). For instance, your friend might not only intend to criticize your behavior but also to induce a certain emotional response to his or her sarcastic comment (e.g., amusement). Experience shows, however, that as a recipient of sarcasm, one's emotional response might be quite different from what was intended by the speaker, highlighting the need to examine the socio-emotional functions of sarcasm from these different perspectives. To date, insights into the emotional impact of ironic language have been principally gained from questionnaire studies, thus, not much is known about the time course and underlying processing mechanisms. Here, we address these issues by analyzing participants' eye movements while they are reading, in order to examine their moment-to-moment expectations regarding the emotional impact of ironic versus literal criticism as described from the perspective of both the recipient and the speaker.

In tackling these questions, we take advantage of insights gained from text comprehension research, in

* Corresponding author at: School of Psychology, The University of Nottingham, University Park, Nottingham NG7 2RD, UK. Fax: +44 (0)115 951 5324.
E-mail address: ruth.filik@nottingham.ac.uk (R. Filik).

which it is generally assumed that readers construct a coherent mental representation of the people, objects, and events being described in the text, that is, a so-called situation model (e.g., Kintsch, 1988; Van Dijk & Kintsch, 1983). Situation models are held to encode various text dimensions such as time, space, causation, motivation, as well as information about the intentions and emotions of story characters (e.g., Zwaan & Radvansky, 1998). Crucially, this allows readers to generate inferences (e.g., McKoon & Ratcliff, 1992) and to thus anticipate the likely behavior or emotional responses of the characters that are being described. Early evidence to suggest that readers do in fact keep track of a character's likely emotional state comes from self-paced reading studies (e.g., de Vega, León, & Diáz, 1996; Gernsbacher, Goldsmith, & Robertson, 1992; Gernsbacher, Hallada, & Robertson, 1998), showing that reading times were longer for sentences that contained emotion words that were inconsistent rather than consistent with the emotional state implied by the context (e.g., someone enjoying the perfect end-of-year party feeling sad vs. happy, respectively). More recent neuroscientific studies support these findings (e.g., Ferstl, Rinck, & von Cramon, 2005) and furthermore, have demonstrated that readers can rapidly detect when characters do not show the anticipated emotional response to a situation (e.g., Leuthold, Filik, Murphy, & Mackenzie, 2012).

In relation to perspective effects, the nature of the situation model that is constructed may depend on the perspective adopted by the reader (cf. Zwaan & Rapp, 2006). For instance, it has been shown that the items recalled in a memory test can depend on whether participants read a story from the perspective of a potential homebuyer versus that of a burglar (Anderson & Pichert, 1978; Pichert & Anderson, 1977). Similarly, using a text-change detection paradigm, Bohan, Filik, MacArthur, and McClusky (2009) showed that changes to perspective-relevant words were more frequently detected than those that were perspective-irrelevant. Thus, findings from these separate strands of research on emotion, and on perspective, suggest that readers build a situation model that allows them to anticipate the likely emotional responses of story characters, and that the nature of the model depends on the perspective that is taken. However, it is unclear at present whether the perspective adopted by the reader also influences which and to what degree specific emotional responses are anticipated.

In relation to the current study, it seems likely that expectations regarding the emotional impact of criticism may be quite different depending on whether one takes the perspective of the victim of the comment, or of the protagonist. For example, Bowes and Katz (2011) found that although ironic and literal criticism were both judged as being relatively impolite, they were judged as more so when viewed from the victim's than from the protagonist's perspective. Perhaps underlying this difference is the observation that intent (e.g., intending to hurt) is central to the protagonist perspective, and emotional impact (e.g., feeling hurt) is central to the victim perspective (Toplak & Katz, 2000).

The results of similar previous studies of emotional responses to ironic versus literal criticism have led to a debate concerning whether ironic language is used to enhance or to mute the positive or negative nature of a message, compared to literal language. For example, Dews and Winner (1995) found in their rating study that ironic criticism (e.g., *That was just terrific*) was judged as less critical than a corresponding literal statement (e.g., *That was just awful*), suggesting that an attack becomes less negative when delivered ironically (see also Dews, Kaplan, & Winner, 1995; Filik et al., 2016; Harris & Pexman, 2003; Jorgensen, 1996; Matthews, Hancock, & Dunham, 2006). From this, Dews and Winner developed the *Tinge Hypothesis*, which states that in the case of ironic criticism, the (negative) ironic meaning is 'tinged' with the (positive) literal meaning of the expression (e.g., of *terrific*), thereby reducing the perceived negativity of the statement.

An alternative view is that irony (in particular, sarcasm) may actually *enhance* the negative emotions felt by the recipient of the criticism (e.g., Blasko & Kazmerski, 2006; Bowes & Katz, 2011; Colston, 1997; Kreuz, Long, & Church, 1991; Leggitt & Gibbs, 2000; Toplak & Katz, 2000). Specifically, it has been argued that the use of irony conveys information relating to the speaker's attitude towards the recipient, being especially appropriate if the speaker wishes to convey a hostile attitude towards the addressee (Lee & Katz, 1998). Thus, in contrast to the tinge hypothesis, this view suggests that being on the receiving end of ironic compared to literal criticism is likely to provoke an enhanced negative emotional response (e.g., be more hurtful).

Most previous studies of emotional responses to criticism have principally involved participants rating how the recipient of such a comment would feel, that is, the task draws attention to the emotional content of the stimuli and allows time for reflection. In the current study, we are interested in how readers process a character's emotional response to criticism 'on-line'. Some recent research has used on-line methodologies, such as eye-tracking while reading, to examine the time course of processes involved in computing the meaning of an ironic utterance (e.g., Au-Yeung, Kaakinen, Liversedge, & Benson, 2015; Filik, Leuthold, Wallington, & Page, 2014; Filik & Moxey, 2010; Kaakinen, Olkoniemi, Kinnari, & Hyönä, 2014; Olkoniemi, Ranta, & Kaakinen, 2016; Turcan & Filik, in press). These studies have principally focused on reporting reading times on the ironic comment itself compared to a non-ironic counterpart, and as a result, much has been learned about the on-line processing of ironic versus literal remarks. What is less clear, and therefore the focus of the current study, is how readers process a character's subsequent emotional response to such remarks. Thus, here we apply such on-line methods to study the time course of the emotional impact of using irony.

Specifically, to allow for a detailed examination of the moment-to-moment inferences regarding a character's emotional response that might be expected following ironic compared to literal criticism, from both the victim's and protagonist's perspective, we will monitor participants' eye movements while they are reading a series of short stories (see Table 1 for example scenarios). In Experiment 1, participants will be presented with scenarios in which one character criticizes another character, either literally or ironically, followed by a target sentence that describes either the recipient's hurt response