

Context Incongruity and Irony Processing

Stacey L. Ivanko and Penny M. Pexman

*Department of Psychology
University of Calgary*

In this research we investigated the role of context in interpretation (Experiments 1 and 2) and relative processing (Experiment 3) of literal and ironic statements. As such, we manipulated degree of situational negativity (using strongly negative, weakly negative, and neutral contexts) and found that, in strongly negative situations, reading times for ironic statements were slower than for literal statements, whereas in weakly negative situations reading times for ironic statements were faster than or equivalent to reading times for literal statements. When these reading time data were related to interpretation data, we found that differences in processing time could be predicted by certain aspects of interpretation. We argue that these results support a direct access model of figurative language processing (e.g., Gibbs, 1994), and also support Colston's (2002) theory of contrast and assimilation in verbal irony.

Imagine the following situation: Joe has agreed to give John a ride to school. Joe is 1 hr late to pick John up and apologizes. John says "you are so punctual." Is this statement sarcastic (ironic)? Probably, but the interpretation of John's intent in making the statement might be easier if there had been even stronger incongruity between the events and the literal meaning of John's statement. For instance, if Joe had never arrived to pick up John and had never apologized, then the event could be perceived as even more negative, and would have contrasted more sharply with the positive literal meaning of the statement. Research has shown that irony can be perceived as a result of incongruity between the context and the statement (e.g., Colston, 2002; Colston & O'Brien, 2000; Gerrig & Goldvarg, 2000; Katz & Lee, 1993; Katz & Pexman, 1997; Kreuz & Glucksberg, 1989; Pexman, Ferretti, &

Katz, 2000; Pexman & Olineck, 2002). The purpose of this research was to investigate how degree of incongruity influences interpretation and processing of ironic statements.¹

In pursuing this research question our hope was that, by examining the role of context incongruity, we might shed light on a long-standing debate in the verbal irony literature. That debate concerns the manner in which irony (compared to nonirony) is processed. We also sought to shed light on this debate by combining offline ratings and decision data with precise moment-by-moment online reading time measures. By combining these measures we had the opportunity to detect where processing differences occur, and to what aspects of interpretation those differences are related.

INCONGRUITY

Verbal irony serves many communicative purposes. One of these is to highlight disparity between expectations and reality. For instance, ironic statements like “you are so punctual” often convey failed expectations. The statement expresses what the speaker expected (punctuality), and because the statement is out of keeping with events (the target was actually late) there is incongruity between the speaker’s attitude (negative, disappointment) and their actions (a positive statement). As noted, this incongruity is a reliable cue to ironic intent. Further, the degree of incongruity influences the extent to which ironic intent is perceived. That is, Gerrig and Goldvarg (2000) examined the effect of the degree of situational disparity (e.g., 5 min late vs. 50 min late) on the perception of irony. They found that greater situational disparity led to a higher perception of irony, compared to a situation with less disparity. Thus, there is clearly an effect of degree of disparity on the perception of irony.

Similarly, Colston and O’Brien (2000) manipulated the degree of contrast between a context situation and an ironic statement (e.g., *Sheila was looking forward to her boyfriend Walter’s visit. When Walter arrived, he was in a terrible mood and was snapping and yelling at Sheila and her housemates. Sheila turned to him and said:* (a) weak verbal irony: “Aren’t you in an agreeable mood?”; (b) strong verbal irony: “Aren’t you in a magnificent mood?”; or (c) literal: “Aren’t you in a bad mood?”). Colston and O’Brien examined pragmatic functions of both strong and weak ironic statements. When there was a high degree of difference between the strong and weak version of statements, the speakers of strongly ironic statements were rated to be more condemning, more humorous, and more self-protecting than the speakers of weakly ironic statements. These results suggest that a perception of

¹We refer to the items used in this study as “ironic,” but they could be described, somewhat more narrowly, as instances of *sarcasm*.

irony is dependent on disparity or contrast, which can be created by the strength of the statement (as in the Colston & O'Brien study) or by the strength of the context (as in the Gerrig & Goldvarg, 2000, study).

Colston (2002) expanded on the straightforward notion of contrast in verbal irony comprehension to suggest that the perception of verbal irony and appreciation of its pragmatic functions are subject to *contrast effects*. Contrast effects are observed in many contexts (e.g., perception, judgment, interpretation) and describe the situation where biasing information is presented and influences perception or interpretation in a direction away from the biasing information. This occurs when the biasing information is considerably different from the target, along the dimension on which judgments or perceptions are being made. Colston argued that this is what often happens with interpretation of verbal irony. A strongly positive statement (the biasing information) presented in a negative situation can make the situation (the target) appear *more* negative. Colston also argued that interpretation of verbal irony can be subject to *assimilation effects*. These occur when the biasing information is only slightly different from the target and results in a perception that the target is more similar to the biasing information. This occurs in verbal irony interpretation when a statement is only slightly positive. In that case, the statement makes the situation appear *less* negative. To summarize, with an ironic (positive) statement, if the discrepancy between the negative context situation and the statement is large, a contrast effect could emerge and, consequently, the situation would be judged as being more negative than it would with a literal (negative) statement. On the other hand, if the discrepancy is smaller, because the ironic statement is less positive, an assimilation effect would be observed and the situation would be judged to be less negative than it would have been with a literal statement.

In support of these predictions, Colston (2002) found that, when extremely positive statements followed moderately negative situations, the situations were considered more negative, and the speakers more condemning, than when moderately positive statements were presented following the same situations. Further, speakers of ironic statements were considered to be more condemning than speakers of literal statements. In this work, Colston examined how contrast and assimilation effects influence the interpretation of irony's pragmatic functions, but did not examine how contrast influences processing of irony. It is not clear how a situation involving contrast effects might be processed differently from a situation involving assimilation effects. One might predict that situations involving contrast effects require additional processing, because the contrast effect alters the perception of the situation to be more negative, and there is some evidence that negative information takes longer to process than positive information (Matlin & Stang, 1975; Temple & Geisinger, 1990). We will consider the issue of contrast and processing time in this work but next consider previous research on irony processing.

IRONY PROCESSING

There are several theories of how irony is processed, two of which are in opposition. Gibbs (1986, 1993, 1994, 2002) suggested that irony can be processed without first activating and assessing the literal interpretation of the statement. However, Giora (1995, 1997, 2002; Giora, Fein, & Schwartz, 1998) suggested that the salient (usually literal) interpretation of the statement must be examined first, and then is considered in conjunction with the ironic meaning.

The Direct Access Model

Gibbs (e.g., 1994) argued that it does not necessarily take longer to process the figurative meaning of a statement than it does to process the literal meaning of a statement (though in some circumstances it can), and this is made possible by the contextual information that is available. Contextual information influences early processing of the utterance, allowing direct activation of contextually appropriate meanings. The relevant contextual information displays the speaker's attitude toward or beliefs about the situation described.

Gibbs (1994) stated that "recognition of the incongruity between what people say and what they do reflects the cognitive ability of people to think ironically" (p. 437). Thus, understanding irony requires no special cognitive processes, and does not need to be particularly effortful. Instead, understanding irony requires parallel activation of literal and figurative meanings. This seems evident because, in some situations, while we are speaking ironically, we may also be making a literal statement. For example, a driver may say, "I love people who signal" after being cut off by another driver. Although the speaker is being ironic in the sense that the other driver did not use his signal, this statement also reflects the speaker's literal belief.

To test these claims, Gibbs (1986) examined the processing time involved in reading and understanding ironic, nonironic, literal, and acknowledgment statements by presenting participants with paragraphs ending in one of these statement types. There was a negative context condition (e.g., *Harry was building an addition to his house. He was working real hard putting in the foundation. His younger brother was supposed to help. But he never showed up. At the end of a long day, Harry's brother finally appeared. Harry was a bit upset with him. Harry said to his brother*) in which both ironic statements (which were positive in surface meaning; e.g., *you're a big help*) and nonironic statements (which were negative; e.g., *you're not helping me*) were presented. There was also a positive context condition (e.g., *Greg was having trouble with calculus. He had a big exam coming up and he was in trouble. Fortunately, his roommate tutored him on some of the basics. When they were done, Greg felt he'd learned a lot. "Well," he said to his roommate*) in which the ironic statements were used literally (e.g., *you're a big help*) and acknowledgment statements (e.g., *thanks for your help*) were presented. Gibbs found that the

ironic statements were processed faster than the nonironic statements but equivalently to the literal statements, which suggested that a literal meaning does not have to be processed before an ironic meaning when the statements are presented in realistic social contexts.

In support of the direct access model, online studies of metaphor processing have shown that context influences relative processing of literal and figurative meanings. For instance, when the context preceding the metaphor is relatively long, metaphors are processed in the same time as their literal interpretations, but when the context is relatively short, metaphors take longer to process (e.g., Gerrig & Healy, 1983; Ortony, Schallert, Reynolds, & Antos, 1978; Shinjo & Myers, 1987). Similarly, Gerrig and Healy (1983) demonstrated that metaphor processing is also affected by the placement of a metaphor in the context. These authors placed the statements either before the context or after the context. They found that metaphors were processed in the same time as a literal statement when the context came before the metaphor, but that metaphors were processed more slowly than literal statements when the context appeared after the metaphor.

The Graded Salience Hypothesis

An alternative view of irony processing is that the literal interpretation of the statement is salient and must be processed first, before the ironic interpretation (e.g., Giora, 1997). Giora proposed the graded salience hypothesis of irony. According to this hypothesis, regardless of contextual fit, the salient literal meaning is interpreted first (except in cases of familiar irony). Then, if context suggests a figurative interpretation, the literal and figurative meanings are considered together. Giora (1995) argued that irony does not involve a cancellation of the indirectly negated message; "rather, it entertains both the explicit and implicated messages so that the dissimilarity between them may be computed" (p. 240). Therefore, irony involves what is said (the literal meaning) and what is implicated (the ironic meaning). Giora concluded that an ironic statement is more difficult to process than a literal statement and will consequently take longer to process. Thus, according to this hypothesis, it should take longer to read figurative statements than literal statements.

Giora et al. (1998) tested this hypothesis experimentally by presenting participants with paragraphs that were constructed of three or four sentences and a final statement. The final statement was either literal (a positive statement following a positive context; e.g., *Anna is a great student and very responsible. One day she called to tell me she did not know when she would be able to show up for my lecture. However, just as I was starting, she entered the classroom. I said to her: "You are just in time"*) or ironic (a positive statement following a negative context; e.g., *Anna is a great student, but she is very absent-minded. One day while I was well through my lecture, she suddenly showed up in the classroom. I said to her: "You*

are just in time"). Consistent with the graded salience hypothesis, Giora et al. found that the figurative statements took longer to read than the literal statements.

Dews and Winner (1999) examined the processing time of ironic and literal language by asking participants to judge whether a speaker's statement was intended to be positive or negative. Like Giora et al. (1998), Dews and Winner presented participants with positive contexts followed by positive, literal statements or negative contexts followed by positive (in surface meaning), ironic statements (with a negative underlying meaning). They found that judgments of the ironic statements took longer than judgments of the same statements used literally, which supported the graded salience hypothesis.

Effects of Context

The direct access model allows the possibility that context can influence processing of potentially ironic utterances. Gibbs (1994) argued that "the ease with which many figurative utterances are produced and comprehended is due in part to the context for linguistic understanding or, more specifically, common ground (i.e., the knowledge, beliefs, and attitudes that are recognized as being shared by speakers and listeners in any discourse situation)" (p. 413). This model of ironic processing is interactive, such that context information affects how an ironic statement is interpreted or processed. Giora (2002), however, stated that, "Even when a context is so highly constraining that it predicts the appropriate meaning very early, even before the linguistic stimulus is encountered, it will not restrict access of salient but contextually incompatible meanings" (p. 491). This model of irony processing is modular; contextual information does not influence how an ironic statement is initially processed.

Explicit predictions about context and irony processing were recently offered by Utsumi (2000). Utsumi proposed the implicit display theory, which involved three main criteria: (a) the presence of an ironic environment (i.e., context); (b) the ironic environment displayed implicitly; and (c) prototypicality of the ironic utterance. Utsumi agreed with some tenets of the graded salience hypothesis, but did not agree that literal language is always processed faster than irony. Instead, Utsumi argued that the ironic environment can allow for irony to be processed in the same time as literal language. The appropriate ironic environment (which involves the speaker having an expectation, this expectation not being met, and the presence of a negative emotional attitude, such as anger, toward the incongruity between expectation and outcome) can lead to the ironic statement being more prototypical and therefore processed in the same time as the same statement uttered literally. Utsumi explained the opposing results of Gibbs (1986) and Giora et al. (1998) as being a result of context. More specifically, Utsumi argued that the context situations used by Gibbs displayed a strong ironic environment, whereas those used by Giora et al. displayed a weaker ironic environment, one that is not

necessarily interpreted as being ironic. In this study we tested this possibility: Does the processing of ironic statements differ as a function of context strength?

THIS RESEARCH

In this study we examined the role of context in the processing of figurative (ironic) and literal² language. Based on previous research (Colston, 2002; Colston & O'Brien, 2000; Gerrig & Goldvarg, 2000), we hypothesized that the degree of incongruity between the context and the statement would influence the interpretation of speaker intent. We also tested predictions of theories of irony processing by examining whether degree of incongruity between the context and the statement influenced relative processing times. According to the direct access hypothesis (Gibbs, 1994) and implicit display theory (Utsumi, 2000), relative processing times should be influenced by context conditions. According to the graded salience hypothesis, relative processing times should not be influenced by context conditions.

In studies in which the goal is to compare processing times for literal and ironic statements, the researchers have two design options. The first is to make comparisons such that the statement is the same but the context varies, wherein one context is negative and invites an ironic interpretation of the statement and one context is positive and invites a literal interpretation of the statement. This was, for instance, the approach taken by Giora et al. (1998) and by Dews and Winner (1999). The second option is to make comparisons such that the context is the same (consistently negative) but the surface form of the statement varies, wherein one form of the statement is positive (ironic) and one form is negative (literal). This was essentially the approach taken by Gibbs (1986). In choosing between these options, we conducted an online pilot study in which we compared processing (self-paced reading) times for positive and negative contexts. We found that negative contexts were consistently read more slowly: As participants read about negative events, their reading times slowed, and they carried that slower processing into reading of a neutral target statement. We found a similar trend in a lexical decision task (LDT). We presented participants with a set of positive words (e.g., *useful*, *dependable*) and a set of negative words (e.g., *useless*, *irresponsible*). Although the positive and negative word sets were matched for word length and frequency (Kučera & Francis, 1967), we found that LDT responses were significantly slower for the negative words. This bias, involving slower processing times for negative informa-

²We refer to these statements as "literal," but this is not strictly correct. Many of these statements involve a figurative element (e.g., *Sam is a rotten friend*). Although these statements could perhaps be more accurately described as "nonironic," we use the term *literal* here for the sake of consistency with previous studies.

tion, was something we wanted to avoid in this study. As such, we chose the latter design option: using the same contexts (with equivalent negativity), but varying a single word in the target statements to produce ironic and literal versions of the statements. We used three context conditions that differed only in degree of negativity: strongly negative, weakly negative, and neutral. These contexts were presented with a pair of target statements that had similar form except that one word (the modifier) differed. When the statement had a positive modifier it was intended as an ironic statement (e.g., *Sam is a nice friend*) and when the statement had a negative modifier it was intended as a literal statement (e.g., *Sam is a rotten friend*). To control for item-specific effects, or possible differences between the two forms of each target statement, we collected ratings for a set of potential statements. We only chose those statements where the positive and negative forms were rated as equivalent in familiarity and where the positive form was rated just as positive as the negative form was negative.

In Experiment 1 we assessed how our stimuli were interpreted: The stimuli were rated by participants for sarcasm, mocking, politeness, and confidence in interpretation. Thus, the experiment served several purposes. First, it allowed us to test Colston's (2002) claims about contrast and verbal irony, and to extend those claims to other pragmatic functions; for example, were the ironic statements in the strongly negative contexts perceived to be more mocking and less polite than the literal statements (due to contrast effects), and were the same statements perceived to be less mocking and more polite than literal statements in the weakly negative contexts (due to assimilation effects)? Second, the experiment allowed us to attempt to replicate Gerrig and Goldvarg (2000): We hypothesized that the ironic statements following the strongly negative context would be perceived as more sarcastic compared to the ironic statements following the weakly negative contexts. A final purpose of Experiment 1 was to provide ratings of speaker intent for us to relate to reading times in Experiment 3. That is, to draw inferences from reading times in an online task we assessed how the statements were interpreted in an offline task.

The purpose of Experiment 2 was to collect additional information about interpretation. In this case, we examined participants' expectations as a function of context. Participants were presented with the context paragraphs and were asked which of the two statements (literal or ironic) they would be more likely to expect in each situation. This allowed us to assess how the context manipulation influenced participants' expectations for literal and ironic language. Further, by subsequently relating these expectations to reading times collected in Experiment 3, we were able to assess how and when these expectations influenced processing. This analysis provided a test of Giora's (2002) claim that context does not influence early processing, and a test of Gibbs' (1994) and Utsumi's (2000) claims that context does influence early processing.

In Experiment 3, we examined the word-by-word reading times for ironic and literal statements. The stimuli were presented in a moving windows paradigm and

we measured the time taken to read each word in each paragraph presented. In previous studies of irony processing, processing time was measured only in terms of an overall response time for each statement. The moving windows procedure used here allowed a more precise examination of processing. Reading times for words in the target statement were examined as well as reading times for words in a wrap-up sentence following the statement. Reading times for the wrap-up sentences were examined to assess processing that occurred after the statement was read. Pexman et al. (2000) found that processing times for words in wrap-up sentences were related to perceptions about ironic intent. They argued that participants continued to consider the meaning of potentially ironic statements even as they viewed the subsequent sentence. It seemed possible that, when participants moved on to the wrap-up sentence, they were checking whether it provided any additional information that would help them process the target statement. We therefore hypothesized that processing time differences observed for wrap-up sentences following literal and ironic statements would reflect differences in processing for literal and ironic statements.

EXPERIMENT 1

Purpose

The general purpose of this experiment was to determine whether context strength influences interpretation of literal and ironic statements.

Method

Participants. Participants in this experiment were 89 (65 female and 24 male) undergraduate students from the University of Calgary, ages 18 to 44 ($M = 21.66$, $SD = 4.66$). In exchange for participation they received bonus credit toward a psychology course. There were six versions of the stimuli in this experiment and 15 participants were assigned to each, with one version having only 14 participants (1 participant did not follow instructions and was excluded from the analysis).

Materials. This experiment involved a ratings task consisting of 12 items with three context sentences each and a final statement (see the Appendix for the items). These items were derived from a pilot study.

Pilot study. For the pilot study, we created paragraphs describing 29 context situations. For each situation, we constructed four potential statements (in total, 116) and five or six versions of the contexts (in total, 149). Sixty-five undergraduate students from the University of Calgary participated in the pilot study. They

were given two booklets: one contained 58 statements (one positive and one negative, from each situation) and one contained 87 context paragraphs (three from each situation). After each statement, 7-point rating scales for familiarity, plausibility, and positivity were presented. Based on the ratings, 12 statement pairs were chosen such that statements within each pair were similar in familiarity and plausibility (and were both moderately familiar and quite plausible) and were opposite in negativity (i.e., the positive statement had a mean positivity rating over 5 and the negative statement had a mean positivity rating below 3). Similarly, 7-point rating scales for plausibility and positivity were placed after each context paragraph. Based on the ratings, 12 sets of three context paragraphs were chosen. The sets were rated as equivalent in plausibility and the strongly negative contexts were rated as significantly more negative than the weakly negative contexts, which were rated as significantly more negative than the neutral contexts.

Experiment 1. The stimuli for Experiment 1 were 12 sets of statements and contexts, involving a 3 (context: strongly negative, weakly negative, neutral) \times 2 (statement: ironic, literal) combination of conditions. Each statement appeared in each context condition across six versions of the stimuli.

The stimuli were presented in booklets. Each context-statement paragraph was followed by four 7-point scales: (a) sarcasm—ranging from 1 (*not at all sarcastic*) to 7 (*very sarcastic*); (b) mocking—ranging from 1 (*not at all mocking*) to 7 (*very mocking*); (c) politeness—ranging from 1 (*not at all polite*) to 7 (*very polite*); and (d) confidence of ratings—ranging from 1 (*not at all confident*) to 7 (*very confident*).

Procedure

Participants were tested in small groups of 5 to 10 individuals.

RESULTS AND DISCUSSION

The mean ratings for each combination of statement and context condition were calculated (see Table 1) and were analyzed by subjects ($F1$ and $t1$) and by items³ ($F2$ and $t2$) using separate 2 (context: strongly negative, weakly negative) \times 2 (statement: ironic, literal) repeated-measures analyses of variance (ANOVAs). We also conducted planned comparisons to compare ratings for ironic and literal statements within each context condition, as well as to examine context effects for the ironic and literal statements. Bonferroni adjustments were applied. We excluded

³With only 12 critical items in the experiment the degrees of freedom in the items analyses were very small. Thus, in drawing our conclusions, we place more weight on the analyses by participants, although in many cases the effects are also significant by items.

TABLE 1
Mean Ratings and Standard Deviations by Condition for Experiment 1

	<i>Strongly Negative Context</i>					<i>Weakly Negative Context</i>					<i>Neutral Context</i>				
	<i>Ironic Statement</i>		<i>Literal Statement</i>		<i>Difference</i>	<i>Ironic Statement</i>		<i>Literal Statement</i>		<i>Difference</i>	<i>Ironic Statement</i>		<i>Literal Statement</i>		<i>Difference</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Sarcasm	5.74	1.21	1.87	0.98	3.87	4.85	1.48	2.18	1.20	2.67	1.85	1.02	2.84	1.51	-0.99
Mocking	5.17	1.20	3.64	1.64	1.53	4.45	1.49	3.97	1.30	0.48	1.78	0.97	4.18	1.50	-2.40
Politeness	3.35	1.37	2.82	1.26	0.53	4.13	1.45	2.71	1.01	1.42	6.14	0.83	2.52	1.03	3.62
Confidence	5.37	0.92	5.57	0.87	-0.20	5.19	0.96	5.35	0.90	-0.16	5.74	0.92	4.90	1.14	0.84

the neutral context condition from the analyses because it was not clear how the statements were being interpreted. The “ironic” statements in that context condition actually received lower sarcasm ratings than the “literal” statements. We assumed this was because the neutral contexts were perceived as slightly positive, which resulted in some of the literal statements being interpreted as ironic compliments (see Table 1).

Sarcasm Ratings

There was a significant Context \times Statement interaction, $F(1, 88) = 27.33, p < .001, MSE = 1.17$; $F(2, 11) = 15.08, p < .01, MSE = .30$. The nature of this interaction was revealed with planned comparisons. In the strongly negative context condition, the ironic statements were rated as significantly more sarcastic than the literal statements, $t(88) = 23.19, p < .001, SE = .17$; $t(11) = 17.14, p < .001, SE = .22$, and this difference was smaller, but still significant, in the weakly negative context condition, $t(88) = 13.55, p < .001, SE = .34$; $t(11) = 7.73, p < .001, SE = .34$. Planned comparisons also showed that the ironic statements were rated as significantly more sarcastic when placed after the strongly negative contexts than when placed after the weakly negative contexts, $t(88) = 4.63, p < .01, SE = .19$; $t(11) = 4.08, p < .01, SE = .23$. For the literal statements, sarcasm ratings were somewhat lower for statements in the strongly negative contexts than they were for the same statements in the weakly negative contexts, $t(88) = -2.48, p < .05, SE = .12$; $t(11) = -2.27, p < .05, SE = .13$. This showed that the level of incongruity between the contexts and the statements affected how sarcastic the ironic statements were perceived to be, in that as the incongruity increased the perception of sarcasm also increased, replicating Gerrig and Goldvarg (2000).

Not surprisingly, the main effects of context, $F(1, 88) = 6.41, p < .05, MSE = 1.16$; $F(2, 11) = 10.91, p < .01, MSE = .12$, and statement, $F(1, 88) = 527.24, p < .001, MSE = 1.81$; $F(2, 11) = 184.05, p < .001, MSE = .67$, were also significant.

Mocking Ratings

The Context \times Statement interaction was significant, $F(1, 88) = 18.99, p < .001, MSE = 1.28$; $F(2, 11) = 9.54, p < .05, MSE = .34$. In the strongly negative contexts, the ironic statements were rated as significantly more mocking than the literal statements, $t(88) = 6.92, p < .001, SE = .22$; $t(11) = 7.04, p < .001, SE = .21$, though this difference was only marginally significant in the weakly negative contexts, $t(88) = 2.51, p < .05, SE = .19$; $t(11) = 1.76, p > .05, SE = .25$. Additional planned comparisons showed that the ironic statements were rated as significantly more mocking when placed after the strongly negative contexts than when placed after the weakly negative contexts, $t(88) = 3.84, p < .01, SE = .19$; $t(11) = 3.38, p < .05, SE = .22$. Mocking ratings did not differ significantly for literal statements

placed after the strongly negative contexts and literal statements placed after the weakly negative contexts, $t_1(88) = -1.81, p > .05, SE = .18$; $t_2(11) = -1.50, p > .05, SE = .19$. As with the sarcasm ratings, this showed that the perceived level of mocking for ironic statements was affected by the degree of incongruity between the context and the statement. With stronger incongruity, the ironic statements were perceived to be more mocking than the literal statements, and this difference was much smaller for situations with less incongruity. In terms of Colston's (2002) framework, it appears that contrast effects were at work in the strongly negative context condition.

The main effect of context was not significant, $F_1(1, 88) = 2.00, p > .05, MSE = 1.72$; $F_2(1, 11) = 3.66, p > .05, MSE = .18$, but the main effect of statement was significant, $F_1(1, 88) = 35.38, p < .001, MSE = 2.54$; $F_2(1, 11) = 38.22, p < .001, MSE = .29$.

Politeness Ratings

The Context \times Statement interaction was significant, $F_1(1, 88) = 13.58, p < .001, MSE = 1.31$; $F_2(1, 11) = 25.85, p < .001, MSE = .10$. In the strongly negative context condition the ironic statements were rated as only slightly more polite than the literal statements, $t_1(88) = 2.55, p < .05, SE = .21$; $t_2(11) = 2.02, p = .07, SE = .25$, but this difference was larger, and significant, in the weakly negative context condition, $t_1(88) = 7.71, p < .001, SE = .18$; $t_2(11) = 4.47, p < .01, SE = .32$. Additional planned comparisons showed that the ironic statements were perceived to be less polite (more rude) when placed after the strongly negative contexts than when placed after the weakly negative contexts, $t_1(88) = -4.57, p < .001, SE = .17$; $t_2(11) = -3.74, p < .05, SE = .21$. Politeness ratings for literal statements, in contrast, did not differ across the two context conditions, $t_1 < 1$; $t_2 < 2$. This demonstrated that context strength also affects how polite an ironic statement is perceived to be. With stronger incongruity between the contexts and the statements, the ironic statements were perceived to be only marginally more polite than the literal statements. With weaker incongruity, the ironic statements were perceived to be considerably more polite than the literal statements. These results are suggestive of contrast effects (in the strongly negative context condition) and assimilation effects (in the weakly negative context condition).

There were also significant main effects of context, $F_1(1, 88) = 8.00, p < .01, MSE = 1.24$; $F_2(1, 11) = 4.39, p = .06, MSE = .29$ and statement, $F_1(1, 88) = 40.06, p < .001, MSE = 2.11$; $F_2(1, 11) = 12.56, p < .01, MSE = .91$.

Confidence Ratings

We should note, first, that across conditions participants expressed quite a high level of confidence in their interpretations of the speakers' intent. The Context \times

Statement interaction was not significant, $F(1, 88) < 1$; $F(2, 11) < 1$. None of the planned comparisons were significant.

The main effect of context was significant, $F(1, 88) = 9.16, p < .01, MSE = .38$; $F(2, 11) = 4.64, p = .05, MSE = .10$. Across statement types, participants were more confident in their ratings for the statements presented after the strongly negative contexts ($M = 5.47, SD = 1.08$) than they were for the statements presented after the weakly negative contexts ($M = 5.27, SD = 1.18$). The main effect of statement was also significant, $F(1, 88) = 5.84, p < .05, MSE = .49$; $F(2, 11) = 4.33, p = .06, MSE = .09$. Across context conditions, confidence ratings for the literal statements ($M = 5.47, SD = 1.08$) were somewhat higher than they were for the ironic statements ($M = 5.28, SD = 1.19$).

EXPERIMENT 2

Purpose

The general purpose of this experiment was to examine participants' expectations for literal or ironic language as a function of context incongruity.

Method

Participants. Participants in this experiment were 54 (34 female and 20 male) undergraduate students from the University of Calgary, ages 18 to 28 ($M = 20.74, SD = 2.74$). Participants received bonus credit toward a psychology course in exchange for their participation. There were three versions of the stimuli and 18 participants were presented with each version.

Materials. The stimuli were the same 12 sets of stimuli used in the previous experiment. These stimuli were presented in a paper-and-pencil task. The context paragraphs and statements were pre-recorded on an audiotape. A female narrator read each paragraph, then both statements in the pair, with either an ironic intonation or a literal intonation as appropriate.

Procedure

Participants were tested in groups of up to 15 individuals. They were presented with a booklet containing 12 items. One context type from each set was presented in the booklet, followed by both statements for that set. Contexts were presented in random order in the booklet. Participants were instructed to listen to the narration and to read along in the booklet. When the paragraph was finished and the state-

ment choices were read, they were asked to place an X beside the statement that they would expect to follow the paragraph.

RESULTS AND DISCUSSION

Each participant was presented with four contexts of each type. Thus, it was possible to calculate the mean number of expected ironic and literal statements (out of 4) for each context condition. For the strongly negative contexts, literal statements ($M = 2.52$, $SD = 1.14$) were expected more often than ironic statements ($M = 1.48$, $SD = 1.14$), $t(53) = 3.33$, $p < .01$, $SE = .31$. For the weakly negative contexts, literal and ironic statements were expected equally often ($M = 2.20$, $SD = 1.25$; $M = 1.80$, $SD = 1.25$, respectively), $t(53) = 1.20$, $p > .05$, $SE = .34$. For the neutral context conditions, literal and ironic statements were also expected equally often ($M = 2.00$, $SD = 1.36$; $M = 2.00$, $SD = 1.36$, respectively), $t < 1$, although, as in Experiment 1, it is difficult to know how the statements in the neutral context condition were interpreted. These results suggest that a literal criticism was somewhat more likely to be expected in the strongly negative context condition. Following Experiment 3, we examined how these expectations were related to reading times.

EXPERIMENT 3

Purpose

This experiment was designed to investigate reading times for literal and ironic statements, as a function of context incongruity.

Method

Participants. Participants in this experiment were 48 (27 female and 21 male) undergraduate students from the University of Calgary, ages 17 to 37 ($M = 20.79$, $SD = 3.59$). Participants received bonus credit toward a psychology course in exchange for their participation. As in Experiment 1, there were six versions of the stimuli. In this experiment, 8 participants were presented with each version.

Materials. The stimuli for this experiment consisted of 10 practice trials, 12 experimental trials, and 63 filler trials. The filler trials involved the same context paragraphs and literal statements used in Pexman et al. (2000). The 12 experimental trials involved the same 12 stimuli sets (context paragraphs and target statements) used in Experiments 1 and 2, with a wrap-up sentence added after each target statement. The wrap-up sentence was included to catch any processing that

occurred after the target statement had been read (i.e., a spillover effect). That is, if a reader has not resolved the meaning of the target statement and it requires more time to process, this may be evident in longer reading times in the wrap-up sentence. Each wrap-up sentence was eight words long. For example, for the stimulus set with the target statement, *Jason is a poor driver*, the wrap-up sentence was *April and Lynn talk while drinking their coffee*. Importantly, we tried to make the wrap-up sentence neutral in tone and did not provide information about how the listener reacted to the target statement. One wrap-up sentence was used for each stimulus set, so these sentences did not differ across statement types or context conditions.

To ensure that participants were attending to the stimuli they were reading, simple yes–no comprehension questions were asked after each trial. The questions did not involve assessment of speaker intent for the target statements.

The stimuli were presented on a 17-in. Sony monitor controlled by a Macintosh G3 and presented using PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993) in a one-word-at-a-time moving window format. A PsyScope button box was used for participants to advance the text and respond to the comprehension questions. The button box recorded the reading times for each word (the interval between successive button presses) as well as processing times for the space between each sentence with millisecond accuracy.

Procedure

Participants were tested individually. Each paragraph and context statement was presented on the computer screen, as a series of dashes for each word, with a space separating each word. The target statements were always presented on one line on the computer screen so that participants did not switch lines in the middle of reading the target statement.

Participants pressed the middle button on the button box to reveal the first word in the paragraph. Each subsequent button press revealed the next word, replacing the previous word with dashes. Participants were instructed to read at their own normal pace, as if reading a newspaper or magazine. Paragraphs were presented to each participant in a different random order.

RESULTS AND DISCUSSION

The data in this experiment were reading times for each word in the target statements and for each word in the wrap-up sentences. We examined 14 reading locations: the five words in each target statement, the space after each target statement, and the eight words in each wrap-up sentence (e.g., *Jason is a poor driver*. [space] *April and Lynn talk while drinking their coffee*). Mean reading times for all loca-

tions are presented in Figure 1. We included reading times for the neutral context condition in Figure 1 for the interested reader but note, again, that it was difficult to know how statements were interpreted in that condition. To test whether the context conditions affected the reading times of ironic and literal statements, $2 \text{ (context)} \times 2 \text{ (statement)}$ repeated-measures ANOVAs were conducted for each location. We also conducted planned comparisons (which were simple main effects for the interaction) for statement effects within each context condition. Before analyses were conducted, 12.5% of the data were excluded: 1.5% for reading times shorter than 100 ms or longer than 2,000 ms and 11% for incorrect answers on the comprehension questions. Of the 14 reading locations, four locations showed significant effects (see Table 2 for mean reading times for these four locations). Results for these four locations are presented next.

Fifth Word in Target Statement

There was a significant Context \times Statement interaction, $F(1, 47) = 7.37, p < .01, MSE = 24523.39$; $F(1, 11) = 2.81, p > .05, MSE = 14894.55$. In the strongly negative contexts, the reading times for this location in the ironic statements were marginally slower than they were for this location in the literal statements, $t(47) = 1.84, p = .07, SE = 32.58$; $t(11) = 1.32, p > .05, SE = 45.73$. The reading times for the weakly negative contexts at this location showed the opposite effect: Reading times for the ironic statements were significantly faster than those for the literal statements, $t(47) = 2.07, p < .05, SE = 30.39$; $t(11) = 1.46, p > .05, SE = 39.58$. The main effect of context was not significant, $F(1 < 1; F(2 < 1)$. The main effect of statement was also not significant, $F(1 < 1; F(2 < 1)$.

Sixth Word in Wrap-up Sentence

The Context \times Statement interaction was significant by items and marginally significant by subjects, $F(1, 47) = 3.59, p = .06, MSE = 9514.66$; $F(1, 11) = 7.17, p < .01, MSE = 1697.37$. In the strongly negative contexts, reading times at this location following the ironic statements were slower than they were for this location following the literal statements, $t(47) = 2.46, p < .05, SE = 25.40$; $t(11) = 3.19, p < .01, SE = 21.46$. In the weakly negative contexts there were equivalent reading times for this location following the ironic and literal statements, $t(1 < 1; t(2 < 1)$.

The main effect of context was not significant, $F(1 < 1; F(2 < 1)$, but the main effect of statement was significant, $F(1, 47) = 4.93, p < .05, MSE = 12566.28$; $F(1, 11) = 7.58, p < .05, MSE = 2123.01$.

Seventh Word in Wrap-up Sentence

The Context \times Statement interaction was significant by items only, $F(1, 47) = 2.24, p > .05, MSE = 11777.69$; $F(1, 11) = 5.89, p < .05, MSE = 1213.01$. In the

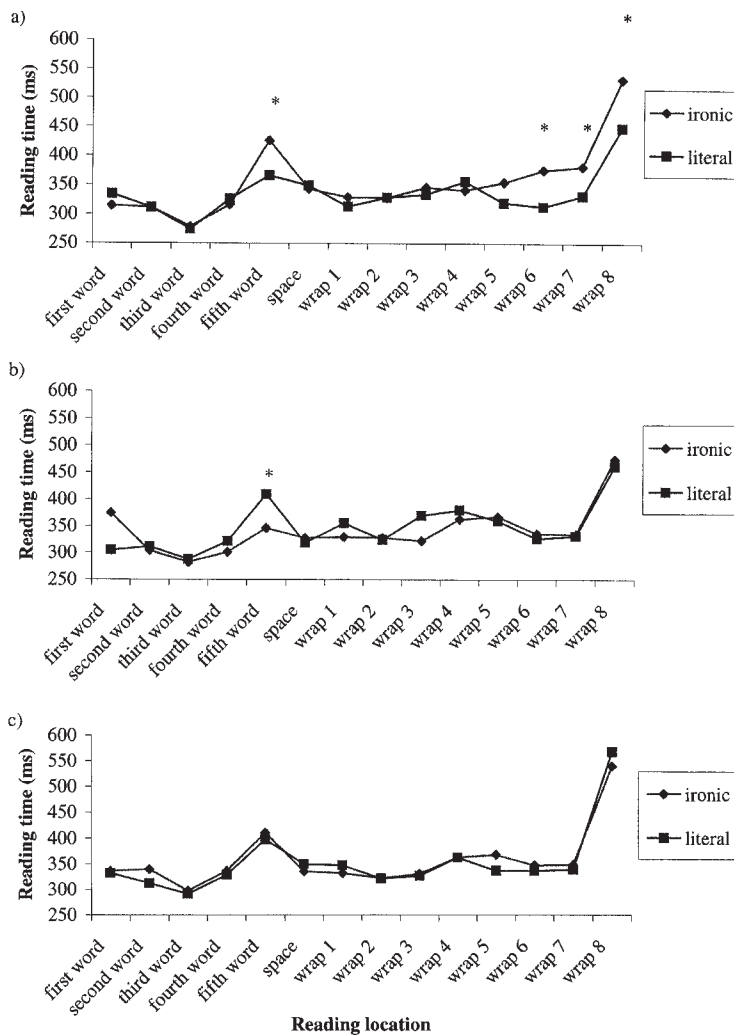


FIGURE 1 Mean reading times (in ms) for all reading locations for the (a) strongly negative context, (b) weakly negative context, and (c) neutral context. Asterisks represent the locations at which reading times for literal and ironic statements were significantly different ($p < .05$). First word = first word in target statement; second word = second word in target statement; third word = third word in target statement; fourth word = fourth word in target statement; fifth word = fifth word in target statement; space = blank space between target statement and wrap-up sentence; wrap 1 = first word in wrap-up sentence; wrap 2 = second word in wrap-up sentence; wrap 3 = third word in wrap-up sentence; wrap 4 = fourth word in wrap-up sentence; wrap 5 = fifth word in wrap-up sentence; wrap 6 = sixth word in wrap-up sentence; wrap 7 = seventh word in wrap-up sentence; wrap 8 = eighth word in wrap-up sentence; as in the example, *Jason is a poor driver. [space] April and Lynn talk while drinking their coffee.*

TABLE 2
Mean Reading Times and Standard Deviations in ms for Locations With Significant Effects in Experiment 3

	<i>Strongly Negative Context</i>					<i>Weakly Negative Context</i>					<i>Neutral Context</i>				
	<i>Ironic Statement</i>		<i>Literal Statement</i>		<i>Difference</i>	<i>Ironic Statement</i>		<i>Literal Statement</i>		<i>Difference</i>	<i>Ironic Statement</i>		<i>Literal Statement</i>		<i>Difference</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Fifth word in statement	425.89	313.20	365.99	194.11	59.90	346.35	183.59	409.17	247.35	-62.82	410.91	311.36	397.70	205.74	13.21
Wrap-up 6	376.37	165.99	313.80	139.44	62.57	335.65	103.24	326.40	104.63	9.25	347.55	109.06	336.70	92.27	10.85
Wrap-up 7	382.42	150.54	332.38	121.34	50.04	334.63	106.85	331.45	120.70	3.18	348.78	150.03	339.30	87.53	9.48
Wrap-up 8	533.12	281.58	449.37	167.37	83.75	473.38	309.96	460.84	229.26	12.54	540.50	331.92	568.86	367.17	-28.36

Note. Wrap-up 6 = sixth word in wrap-up sentence; wrap-up 7 = seventh word in wrap-up sentence; wrap-up 8 = eighth word in wrap-up sentence.

strongly negative contexts, reading times for this location following the ironic statements were longer than they were for this location following the literal statements, $t_1(47) = 2.03, p < .05, SE = 24.68$; $t_2(11) = 2.75, p < .05, SE = 20.17$. In the weakly negative context, there were equivalent reading times for this location following both the ironic and literal statements, $t_1 < 1$; $t_2 < 1$.

The main effect of context was marginally significant, $F_1(1, 47) = 3.33, p = .07, MSE = 8556.12$; $F_2(1, 11) = 2.48, p > .05, MSE = 1902.38$. Across statement conditions, the reading times for this location following the strongly negative contexts ($M = 351.72$ ms, $SD = 169.34$) were slower than they were for this location following the weakly negative contexts ($M = 329.77$ ms, $SD = 119.42$). The main effect of statement was only significant by items, $F_1(1, 47) = 2.78, p > .05, MSE = 12211.65$; $F_2(1, 11) = 4.70, p = .05, MSE = 2453.36$. Across context conditions, the reading times for this location following the ironic statements ($M = 355.79$ ms, $SD = 168.24$) were marginally slower than they were for this location following the literal statements ($M = 324.36$ ms, $SD = 116.67$).

Eighth Word in Wrap-up Sentence

The Context \times Statement interaction was not significant, $F_1(1, 47) = 2.33, p > .05, MSE = 26084.76$; $F_2(1, 11) = 1.98, p > .05, MSE = 8620.60$. However, in the strongly negative contexts, the reading times for this location following the ironic statements were longer than they were for this location following the literal statements, $t_1(47) = 2.20, p < .05, SE = 38.07$; $t_2(11) = 2.04, p = .07, SE = 39.43$. In the weakly negative contexts, there were equivalent reading times for this location following both the ironic and literal statements, $t_1 < 1$; $t_2 < 1$.

The main effect of context was not significant, $F_1 < 1$; $F_2 < 1$. The main effect of statement was marginally significant, $F_1(1, 47) = 3.39, p = .07, MSE = 32848.47$; $F_2(1, 11) = 1.82, p > .05, MSE = 11929.23$. Across context conditions, the reading times for this location following the ironic statements ($M = 493.39$ ms, $SD = 316.82$) were marginally slower than they were for this location following the literal statements ($M = 444.86$ ms, $SD = 242.18$).

The results of this experiment showed that ironic statements in the strongly negative context condition took somewhat longer to process than literal statements in the same context condition. In contrast, in the weakly negative context condition, ironic statements tended to be processed in equivalent time, or faster than, literal statements. These results show that context incongruity plays a role in how irony is processed relative to literal language. What is not clear, however, is why processing differences were observed. That is, although participants seemed to take longer to process ironic statements at certain locations following the strongly negative contexts, we do not know what that processing involved. Were participants taking longer to read the statements at those locations because the statement was inconsistent with their context-derived expectations? Were participants taking longer because certain

pragmatic implications, perhaps activated via contrast effects, result in longer processing times? We address these questions in the following analyses.

COMBINED ANALYSES FOR EXPERIMENTS 1, 2, AND 3

We examined the relations between results from Experiments 1 and 3, and Experiments 2 and 3, in separate sets of regression analyses. In both cases, these were essentially item analyses, because the criterion variables were reading times for items (Experiment 3 data). We used separate sets of analyses because the criterion variables were calculated slightly differently when the predictor variables were ratings for ironic and literal statements (Experiment 1 data) than when the predictor variables were expectations for literal (vs. ironic) statements (Experiment 2 data). The criterion variables in the first set of regression analyses were reading times for each literal statement and each ironic statement, and we examined these for strongly negative contexts separately from weakly negative contexts. We separated contexts in this analysis because we suspected that the relation of sarcasm ratings to reading times would differ in the two contexts. We included data for both literal and ironic statements in this set of analyses because we wanted to capture the effect of the full range of sarcasm ratings. Thus, the first set of analyses involved regression analyses for each context condition (2 conditions), using reading times (14 locations) for each literal and ironic statement (24 statements) as criterion variables, and ratings (sarcasm,⁴ politeness, and confidence) as predictors. This amounted to 28 regression analyses involving 24 cases each.

The criterion variables in the second set of regression analyses were reading times for each literal statement and, separately, for each ironic statement, and we examined these for strongly negative contexts separately from weakly negative contexts. We separated statement types (literal vs. ironic) in this set of analyses because the correlation data suggested that the expectancy variable (expectation for a literal vs. ironic statement) had different relations to reading times for literal and ironic statements. Thus, the second set of analyses involved regression analyses for each context condition (2 conditions), using reading times (14 locations) for each literal statement (12 statements) and then each ironic statement (12 statements) as criterion variables, and expectancy for a literal or ironic statement as the predictor variable. This amounted to 56 regression analyses involving 12 cases each.

Correlations for variables included in these regression analyses are presented in Table 3 (for the first set of analyses involving Experiment 1 and Experiment 3 data) and in Table 4 (for the second set of analyses involving Experiment 2 and Experiment 3 data). There were several significant correlations. The results presented

⁴We did not include mocking ratings as a predictor variable in these analyses to avoid multicollinearity: Mocking ratings and sarcasm ratings were highly correlated.

TABLE 3
Correlations Between Ratings From Experiment 1 and Reading Times From Experiment 3

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
(1) Sarcasm	—	.87**	.31	-.29	-.24	.10	.10	-.07	.26	.12	.32	.05	.21	-.10	.16	.51**	.42*	.32
(2) Mocking	.60**	—	.09	-.10	-.19	.01	.02	.01	.10	.30	.35	.15	.25	-.05	.13	.35	.28	.31
(3) Politeness	.42*	-.32	—	-.16	-.12	-.07	.30	-.13	.13	-.21	-.09	-.06	.06	-.23	.10	.33	.38	.16
(4) Confidence	-.26	-.39	-.03	—	-.26	-.33	-.49*	-.24	-.37	.14	-.27	-.20	-.33	-.48*	-.59**	-.16	-.04	-.02
(5) First word in statement	.07	-.11	.32	-.32	—	.20	.31	-.08	.41*	-.16	.36	.44*	.46*	.45*	.23	.07	-.22	-.22
(6) Second word in statement	-.05	-.04	-.12	-.28	.07	—	.50**	.18	.10	.24	.58**	.41*	.25	.30	.22	.06	-.05	-.13
(7) Third word in statement	-.01	-.01	-.01	-.19	.06	.55**	—	.21	.17	.12	.42*	.23	.33	.28	.32	.30	.19	-.01
(8) Fourth word in statement	-.14	-.04	-.39	-.06	.10	.20	.03	—	-.12	.18	-.07	-.37	-.14	.58**	.48*	-.04	-.15	-.18
(9) Fifth word in statement	-.17	.09	-.34	-.10	.17	.15	-.17	.37	—	-.17	.43*	.31	.05	.02	.04	.43*	.00	.03
(10) Space	.17	.20	-.11	-.30	.38	.47*	.42*	.36	.32	—	.28	.05	.11	-.23	-.11	-.08	-.17	.26
(11) First word in wrap-up	-.06	.03	-.23	-.13	.26	.22	.22	.47*	.52**	.65**	—	.47*	.54**	.09	.12	.28	.09	-.19

(12) Second word in wrap-up	.19	.29	-.23	-.30	.20	.34	.46*	.41*	.21	.81**	.69**	—	.40*	.06	-.13	-.23	-.33	-.06
(13) Third word in wrap-up	-.22	-.09	-.13	-.44*	.01	.19	.16	.16	.18	.12	.34	.30	—	.14	.05	.01	.15	-.07
(14) Fourth word in wrap-up	-.24	-.06	-.33	-.19	.26	.10	-.06	.65**	.10	.21	.46*	.31	.20	—	.70**	.05	-.03	-.32
(15) Fifth word in wrap-up	.08	.16	-.22	-.33	.35	.45*	.32	.40*	.04	.53**	.54**	.58**	.19	.62**	—	.27	.18	-.09
(16) Sixth word in wrap-up	-.13	-.46*	.32	-.27	.46*	.24	.19	.09	-.09	.32	.37	.37	.49*	.35	.42*	—	.70**	-.03
(17) Seventh word in wrap-up	.30	.29	-.16	-.05	-.07	.08	.08	.37	.17	.37	.50**	.54**	.19	.38	.31	.23	—	.12
(18) Eighth word in wrap-up	.01	.03	.01	-.25	.03	-.29	-.02	.25	.02	.05	.32	.24	.28	.40*	.17	.19	.43*	—

Note. The data for strongly negative contexts are presented above the diagonal and the data for the weakly negative contexts are presented below the diagonal.

* $p \leq .05$. ** $p \leq .01$.

TABLE 4
Correlations Between Expectancies From Experiment 2 and Ratings and
Reading Times From Experiments 1 and 3

	<i>Expectancy</i>			
	<i>Strongly Negative</i>		<i>Weakly negative</i>	
	<i>Ironiic Statement</i>	<i>Literal Statement</i>	<i>Ironiic Statement</i>	<i>Literal Statement</i>
Sarcasm	.36	.16	.24	.30
Mocking	.27	.05	-.08	-.46
Politeness	-.44	-.07	-.02	.20
Confidence	.13	.06	.24	.22
First word in statement	.02	.11	.61*	.27
Second word in statement	.34	-.13	.19	-.06
Third word in statement	.22	.50	.10	.06
Fourth word in statement	.63*	.21	.43	.29
Fifth word in statement	.13	-.59*	.29	-.66*
Space	.30	.52	.29	.50
First word in wrap-up	.40	.08	.21	.67*
Second word in wrap-up	.46	-.44	.31	-.32
Third word in wrap-up	-.14	-.05	.06	-.27
Fourth word in wrap-up	.35	.05	.16	.03
Fifth word in wrap-up	.00	.04	.04	.10
Sixth word in wrap-up	.40	.33	.05	.07
Seventh word in wrap-up	.23	.24	-.10	-.12
Eighth word in wrap-up	.05	.23	.14	.06

* $p < .05$.

next show that, in some cases, we were able to explain 30% or 40% of the variance in online reading times using offline interpretation data. We would argue that this is a striking result, given all of the sources of unexplained variance. For instance, the online and offline data were produced by different groups of participants, and reading skill and other individual differences were not measured.

Results—Regression Analyses Set 1

The regressions in this set were calculated using the backward method (with a criterion of excluding predictor variables with $p > .10$). For the strongly negative contexts, 5 out of the 14 reading locations obtained significant regression equations (third word in the target statement, fourth word in the wrap-up sentence, fifth word in the wrap-up sentence, sixth word in the wrap-up sentence, and seventh word in the wrap-up sentence; see Figure 2a), and for the weakly negative contexts there were two reading locations with significant regression equations (third word in the wrap-up sentence and sixth word in the wrap-up sentence; see Figure 2b).

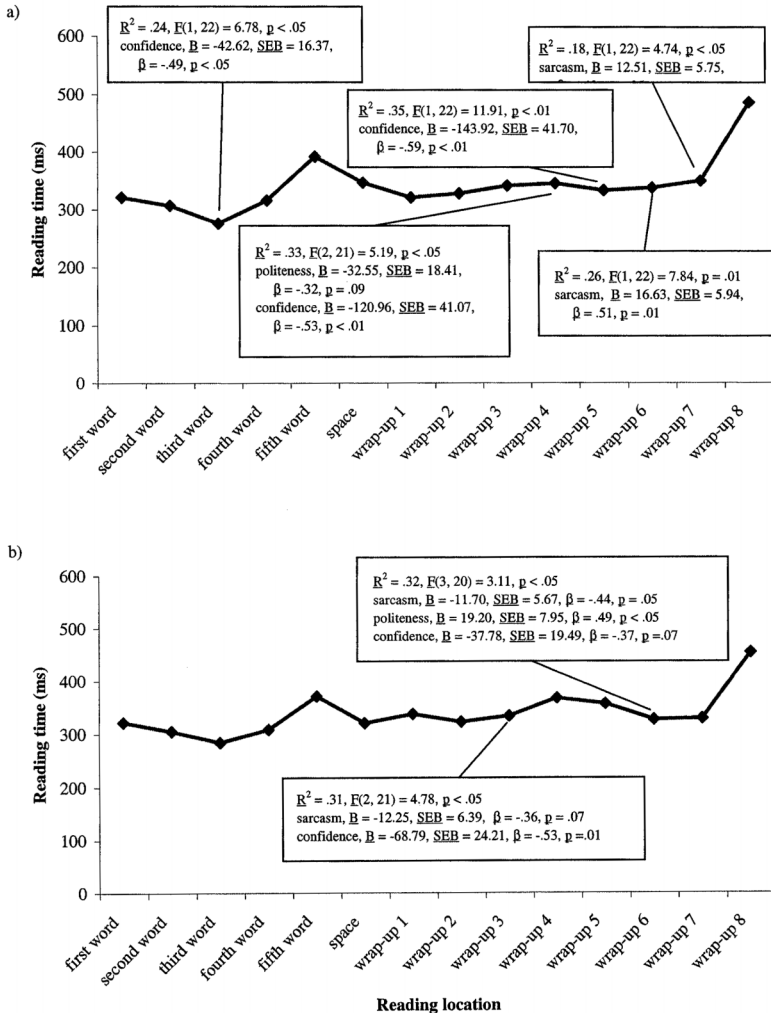


FIGURE 2 Mean reading times (in ms) for the combined analyses of Experiment 1 and Experiment 3, with significant regression equations for the (a) strongly negative contexts and (b) weakly negative contexts.

Confidence ratings alone accounted for a significant portion of the variance at the third word in the statement and the fourth and fifth words in the wrap-up sentence for the strongly negative contexts, and at the third and sixth words in the wrap-up sentence for the weakly negative contexts. This relation was negative in nature such that higher confidence ratings were related to faster reading times. The implication is that feelings of confidence about interpretation of a statement were asso-

ciated with faster processing of that statement and the subsequent sentence. Conversely, lack of confidence in the interpretation given to a statement was associated with slower processing of that statement and the subsequent sentence. Perhaps the most important aspect of this finding is the fact that these relations were observed for locations in the wrap-up sentence. These relations suggest that, quite often, participants were thinking about speaker intent for a target statement well into their processing of the next sentence.

Politeness ratings alone accounted for a significant portion of the variance for two reading locations: fourth word in the wrap-up sentence in the strongly negative contexts, and the sixth word in the wrap-up sentence in the weakly negative contexts. For the strongly negative contexts the relation was negative such that higher politeness ratings for a statement were related to faster reading times. In contrast, for the weakly negative contexts the relation was positive: Higher politeness ratings were related to slower reading times. This is a curious finding, but our hunch is that it was not actually an effect of context. That is, when one examines the pattern of correlations between reading times and politeness ratings (see Table 3), one sees very similar patterns across the context conditions: negative correlations at the fourth word in the wrap-up sentence, and positive correlations at the sixth word in the wrap-up sentence. It seems, however, that politeness only emerges as a significant predictor in the regression equations at the fourth word in the wrap-up sentence for the strong context condition and the sixth word in the wrap-up sentence for the weak context condition. The fact that these are different locations lends some support to our claim that context did not have a substantial effect on the relation between reading times and politeness ratings. Instead, in both contexts, statements that received more moderate politeness ratings (recall that few of these statements actually received high politeness ratings, as illustrated by the means in Table 1) tended to be read faster at the fourth word in the wrap-up sentence and then slower at the sixth word in the wrap-up sentence. We speculate that statements that received moderate or neutral politeness ratings were those for which participants were unsure about the statements' social impact. For these items, participants read the wrap-up sentence quickly at the fourth word because they were advancing through the sentence to see if any information was provided about the target statement's social impact, then read the wrap-up sentence slowly at the sixth word when they realized they would not receive any such information and must achieve comprehension based on little more than what they gleaned from the sentence to that point.

Most importantly, results showed that sarcasm ratings alone accounted for a significant portion of the variance for four reading locations: sixth and seventh words in the wrap-up sentence for the strongly negative contexts, and the third and sixth words in the wrap-up sentence for the weakly negative contexts. Interestingly, the relations were in opposite directions in the two context conditions: positive in the strongly negative context condition and negative in the weakly negative context condition.

The direction of these effects suggest that, in the strongly negative context condition, the statements that were rated as more sarcastic tended to be those with slower reading times. This is consistent with the results of planned comparisons in Experiment 3, where reading times at locations at the end of the wrap-up sentence were significantly slower following the ironic statements than they were following the literal statements. In the weakly negative context condition, however, statements that were rated as more sarcastic tended to be those with faster reading times. These opposite effects were observed at the same location (sixth word in the wrap-up sentence), which suggests that it really is the case that ironic statements take longer to process than literal statements in the strongly negative context condition, and they take less time to process than literal statements in the weakly negative context condition. These results indicate that the differences observed in reading times are in fact related to perceived speaker intent for the statement. Thus, the longer processing times observed following ironic statements in the strongly negative context condition were likely a function of additional processing required to determine speaker intent and pragmatic functions in that situation. In the weakly negative context condition, those functions took less time to process for ironic statements than for literal statements.

Results—Regression Analyses Set 2

The regression analyses in this set were run using the *enter* method, as there was only one predictor variable. For the strongly negative contexts presented with ironic statements, there was one reading location with a significant regression equation: the fourth word in the target statement (see Figure 3a). The relation here was positive such that a higher expectancy for literal statements was related to slower reading times at this location. There was also one reading location (the fifth word in the target statement) with a significant regression equation for the strongly negative contexts presented with literal statements (see Figure 3b). This relation was negative such that a higher expectancy for a literal statement was related to faster reading times at this location.

For the weakly negative contexts presented with ironic statements, there was one reading location with a significant regression equation: the first word in the statement (see Figure 3c). The nature of the relation was negative such that a higher expectancy for literal statements was related to faster reading times at this location. This is an odd result. The relation indicates that a preference for the literal statement (in Experiment 2) was related to slower reading times for the first word in the target statement (in Experiment 3). At the first word in the target statement, however, participants did not yet know that they were reading an ironic statement. Variability in reading time at that location must therefore be a function of the context they have just read. Yet for some reason the same context/first word combination in the weakly negative context/literal statement condition did not produce the same relation. As

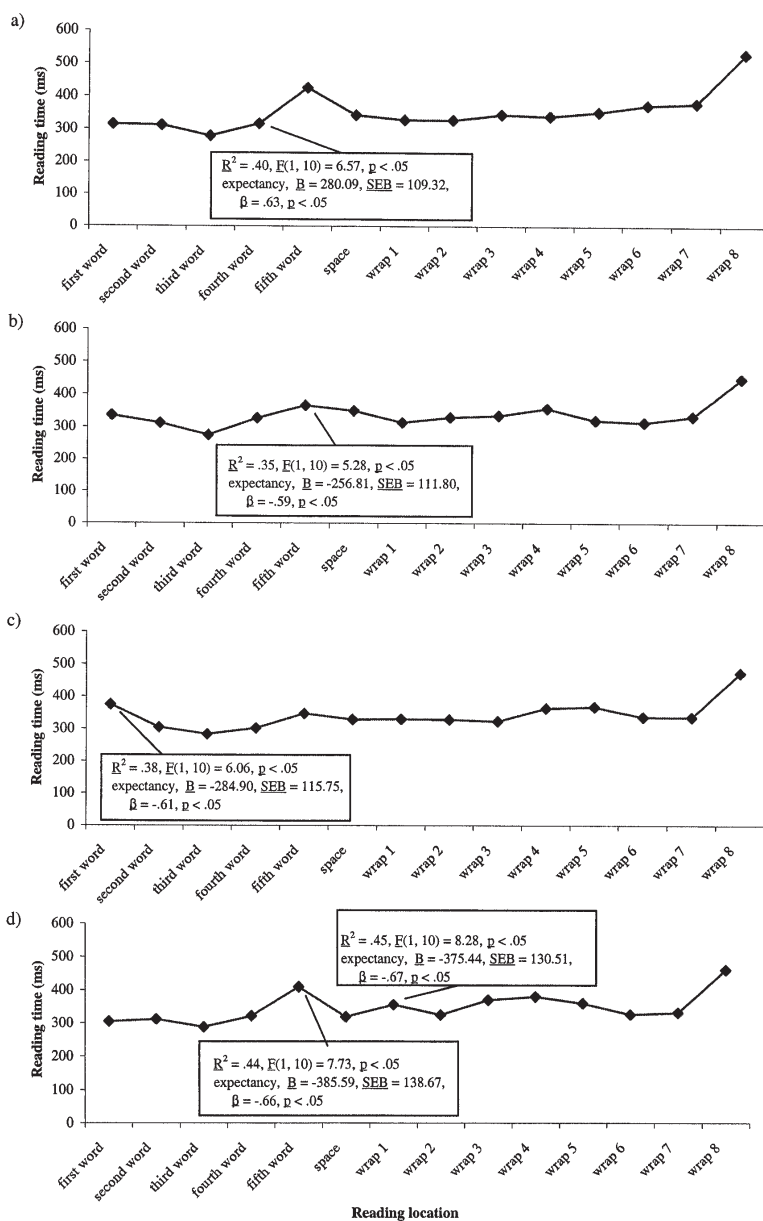


FIGURE 3 Mean reading times (in ms) for the combined analyses of Experiment 2 and Experiment 3, with significant regression equations for the (a) strongly negative contexts with the ironic statements, (b) strongly negative contexts with the literal statements, (c) weakly negative contexts with the ironic statements, and (d) weakly negative contexts with literal statements.

such, it seems likely that the relation observed for reading times at the first word of the target statement in the weakly negative context/ironic statement condition was caused by a spurious data point—there just happened to be a good deal of variability in reading times for the first word of the statement in that condition and some of that variability was related to expectancy for a literal statement. The standard deviations for reading times at that location back this up: 278.41 for the weakly negative context/ironic statement condition, and 100.30 for the weakly negative context/literal statement condition. This variability may have been a function of the way items and contexts were crossed in the reading time study, as well as individual differences in the way those items were perceived. Perhaps a few cautious readers found the weakly negative contexts they were presented with to be slightly ambiguous, and they contributed slow reading times for the first word in the target statement location.

There were two reading locations with a significant regression equation for the weakly negative contexts with the literal statements: the fifth word in the target statement and the first word in the wrap-up sentence (see Figure 3d). The nature of these relations was negative such that a higher expectancy for a literal statement was related to faster reading times at these locations.

It is difficult to know what participants were basing their choices on in Experiment 2. They were presented with the context paragraph and then the two target statements (one literal, one ironic) and had to choose the one that they would be more likely to “expect.” They could have compared the context to the statements and made their choice based on some type of appropriateness dimension. Thus, the fact that higher expectations of a literal statement were related to faster reading times for literal statements at the fifth word of the target statement or first word of the wrap-up sentence is perhaps not all that surprising. Participants read the more appropriate context/statement combinations somewhat faster. The most interesting finding was that, in situations in which participants tended to have a higher expectation for literal statements (strongly negative context condition), ironic statements tended to be read more slowly at the fourth word in the target statement. This suggests that, in these situations, context had created an expectation for literal language and a type of bias against ironic language. Participants seemed to do a quick “double take” when they read the fourth word (which was the positive modifier), because it was out of keeping with their expectations for a direct insult. Thus, we speculate that reading time differences around the end of the target statement are related to the extent to which the target statement fits with expectations derived from context.

GENERAL DISCUSSION

The purpose of this research was to investigate the effect of context incongruity on interpretation and processing of verbal irony. In Experiment 1, when participants

interpreted ironic and literal statements presented in strongly negative and weakly negative context conditions, we observed that ironic statements were perceived to be more sarcastic when presented in strongly negative contexts (replicating Gerrig & Goldvarg, 2000). We also observed that ironic statements in the strongly negative context condition were perceived to be more mocking than literal statements, and were only marginally more polite than literal statements. In contrast, in the weakly negative context condition, ironic statements were perceived to be considerably more polite than literal statements, and only marginally more mocking than literal statements. This pattern of results is consistent with Colston's (2002) contrast and assimilation theory of verbal irony. Contrast was more extreme in the strongly negative contexts, and this seemed to produce a contrast effect: There was a perception that ironic statements were more mocking and only slightly more polite than literal statements. With less contrast in the weakly negative contexts, an assimilation effect seemed to emerge: The same ironic statements were perceived to be more polite and only slightly more mocking than literal statements.

Colston (2002) argued that the mechanism for contrast and assimilation effects was that the statement, depending on the extent to which it contrasted with the situation, changed the perception of the *situation*. In this study we manipulated the contrast by modifying the situation, and by doing so altered the interpretation given to the statement. However, the contrast and assimilation account still seems to apply because, in one case (the strongly negative context condition), the contrast made the ironic statement seem more negative than the literal; in the other case (the weakly negative context condition), the contrast made the ironic statement seem more positive than the literal. Thus, we have contrast effects in the former case and assimilation effects in the latter.

One might argue that our results are not completely in line with Colston's (2002) theory because, in the strongly negative context condition, the ironic statements were still perceived to be somewhat more polite than the literal statements; in the weakly negative context condition, the ironic statements were still perceived to be somewhat more mocking than the literal statements. We would argue that these results are a function of the pragmatic functions we asked participants to rate. The mocking and politeness dimensions are not perfectly correlated with the negativity–positivity dimension (Pexman & Olineck, 2002). Our experiment might have been a better test of Colston's theory had we included the negativity–positivity dimension for ratings. If this had been the case, we would predict that the ironic statements would be rated as more negative than the literal statements in the strongly negative context condition, but more positive than the literal statements in the weakly negative context condition. Certainly, these predictions need to be tested empirically.

The purpose of Experiment 2 was to try to tap into participants' expectations for literal versus figurative language, as a function of the context manipulation. Results showed that participants had a somewhat higher expectation for literal utter-

ances in the strongly negative context condition. This may have been because those contexts, as intended, described situations that were extremely negative and because the speaker had been seriously offended by the behavior of the actor. It was clear in all of our items that the actor was not present when the statement was made; instead, a third party was the listener. Consequently, there was little reason for speakers to hide their criticism and, in the strongly negative context, the offense may have been serious enough that participants considered it socially appropriate for speakers to offer the literal criticism. Thus, they had a higher expectation for the literal statement in that context condition.

The results of Experiment 3 showed that relative reading times for literal and ironic utterances do vary as a function of context, resulting in reading time differences at the end of the target statement and also in the subsequent wrap-up sentence. To interpret those differences, we analyzed the relations between Experiment 1 ratings, Experiment 2 expectancies, and Experiment 3 reading times. We found that speaker intent and pragmatic functions (sarcasm, politeness) tended to predict reading times for locations in the wrap-up sentence. Importantly, sarcasm ratings were a negative predictor of reading times at the sixth word in the wrap-up sentence in the weakly negative context condition, and a positive predictor of reading times at that location in the strongly negative context condition. Statements that were perceived as relatively sarcastic in the weakly negative context condition (the ironic statements) were read relatively quickly, whereas statements that were perceived as relatively sarcastic in the strongly negative context condition (the ironic statements) were read relatively slowly. Another important finding was that expectancy for a literal statement in the strongly negative context condition predicted slower reading times for ironic statements at the fourth word in the target statement. The overall pattern suggests that strong contrast resulted in slower reading times as soon as there was any evidence that the statement was nonliteral (at the positive modifier of the ironic statements), and also resulted in slower reading times somewhat downstream: at the sixth and seventh words of the wrap-up sentence. Less strong contrast resulted in only a downstream effect: faster reading times at the third and sixth words of the wrap-up sentence.

Utsumi (2000) argued that processing times for irony would be faster when the environment and the utterance were more prototypical of irony. Prototypical irony involves a failed expectation on the part of the speaker and thus a negative attitude that is implicitly displayed. Explicit display, in contrast, signals nonirony. Events in the strongly negative context condition are so negative that implicit display of the speaker's disappointment is not possible: Perception of the disappointment would seem unavoidable. This makes the weakly negative context a better example of prototypical irony. The speaker's disappointment could be missed, making implicit display more likely. Thus, our results seem consistent with that theory, although it offers no obvious explanation for why both early effects (at the end of the target statement) and late effects (in the wrap-up sentence) were observed.

These results do not seem at all consistent with the graded salience hypothesis (Giora, 1995, 2002), because we observed effects of the context manipulation quite early in processing. It also does not seem possible for this account to explain faster processing times following ironic statements in the weakly negative context condition. Given that our stimuli were not salient ironies, there is no reason, by this account, why the ironic utterances should ever produce processing faster than or equivalent to the literal utterances.

Our results do seem consistent with the direct access model (Gibbs, 1994), because we observed that context modulated relative processing of literal and ironic statements. When the context led to neither a literal nor ironic bias or expectation (in the weakly negative context condition) we observed equivalent reading times for literal and ironic utterances (or faster reading times for the ironic utterance at the fifth word in the target statement). When the context led to a bias for a literal criticism, we observed slower reading times for ironic utterances, both initially (fourth word of target statement) and downstream. Again, there is the unresolved issue of why the effects occur downstream.

To explain the complete pattern of findings, we offer the following account. As Gibbs (1994) predicted, ironic utterances can sometimes be processed as quickly as literal utterances, and our results suggest that ironic statements are being interpreted as such immediately, without any delay in processing, in the weakly negative context condition. In the strongly negative context condition, however, there is some degree of bias against irony; in this case it appears that a literal interpretation is initially favored and results in a delay of processing at the end of the target statement when the words are inconsistent with that expectation. As such, it appears that the literal interpretation does sometimes dominate the ironic interpretation, but we would be very cautious about drawing conclusions from our data regarding the earliest phases of meaning activation. Our results suggest that processing of an ironic statement can take more time, but it can also take equal time, initially. The fact that the literal meaning is dominant as participants read the fourth word in the target statement in the strongly negative context condition does not necessarily mean that the literal meaning was activated *first*. Instead, the literal meaning could just be weighted more heavily. Katz and Ferretti (2001) argued that figurative language processing could be accounted for by a constraint satisfaction model in which “constraints interact to provide probabilistic evidence in support of the various alternatives with competition ending when one alternative ‘wins’” (p. 215). The constraints are all the sources of information that the comprehender can draw on in deriving their interpretation. In our strongly negative context condition, the literal statement could be favored as a result of pragmatic cues (constraints) and consequently the ironic statement is processed more slowly.

To explain the differences in processing time observed downstream, we suggest that these might be the result of contrast and assimilation effects. That is, results from Experiment 1 suggested that contrast effects might be influencing

perception of ironic statements in the strongly negative context condition, whereas assimilation effects might be influencing perception of ironic statements in the weakly negative context condition. The contrast effects cause a shift in interpretation to one that is even more negative, whereas assimilation effects cause a shift in interpretation to one that is less negative. When participants are reading the wrap-up sentence, evidence suggests that they are still considering speaker intent and pragmatic functions for the target statement. When their perception is more negative than for literal statements (in the strongly negative context condition), this may lead to longer processing times. When their perception is more positive than for literal statements (in the weakly negative context condition) this may lead to faster processing times. Our assumption here (supported by our pilot data, mentioned earlier) is that negative information is generally processed more slowly than positive information. Thus, there is a processing cost associated with a more negative interpretation, as well as a processing benefit associated with a more positive interpretation.

Although we believe that this account explains the pattern of results we observed, there are certainly many aspects of the account that are speculative and would need to be tested. In addition, we suspect that certain aspects of our findings were influenced by the stimuli and procedures we chose. For instance, as mentioned, in all of our statements the speaker utters the target statement when the actor is not present. Because of this, the speaker clearly has little fear of offending the actor. This may have decreased the salience of the face saving function of irony (Glucksberg, 1995; Jorgensen, 1996). Consequently, this aspect of our stimuli may have contributed to participants' expectations for a literal target statement in the strongly negative context condition. If our situations had been described such that the actor was present when the utterance was made, participants may have been more likely to expect irony. With a different expectation, we would anticipate different results, particularly in terms of early reading time differences, at the end of the target statement. Given our account, we would predict, further, that this different expectation should not affect the pattern of results we observed in the wrap-up sentence because we are attributing those to contrast and assimilation effects, which would still apply.

In addition, we suspect that our results were influenced by the length of the wrap-up sentences we used (eight words each). Processing differences tended to be observed later (approximately 1,800 ms after the last word in the target statement was presented) rather than earlier in these sentences. We do not, however, think that the timing of these effects should be interpreted as the necessary timeline for understanding irony. We suspect that reading time effects emerged toward the end of the wrap-up sentence because it was at this point that participants realized they would not be provided with any additional information relevant to their interpretation of speaker intent. Our prediction, therefore, is that these effects would be observed earlier in shorter wrap-up sentences.

We should reiterate that our processing time measure did not directly address the related issue of whether literal or ironic meanings were activated first. Giora et al. (1998), however, claimed to have addressed this issue. That is, they derived a procedure to more directly examine which meanings are activated in the time course of irony processing. In their Experiment 2, those authors examined LDT response times (RTs) to test words presented just after the target statement to determine which meaning was activated first. The test words were related either to the literal meaning or to the figurative meaning of the statements. The target sentences (e.g., *you are just in time*) were presented in a figurative or a literal context; the word following the sentences supported the literal (e.g., *punctual*) or figurative (e.g., *late*) meaning of the statement, or was a nonword. The words were presented at an interstimulus interval (ISI) of 150 ms or 1,000 ms.

In both ISI conditions, Giora et al. (1998, Experiment 2) found that the fastest RTs were for the test words that were literally related to the statements; this was true in both the literal and figurative contexts. When an ISI of 2,000 ms was used (Experiment 3), RTs were equivalent for literally related and figuratively related words. This supported their hypothesis that the more salient (literal) meaning is activated first, before the less salient (figurative) meaning is activated. According to these results, it is not until 2,000 ms that the figurative meaning is activated.

Gibbs (2002) has offered criticisms of this procedure, and we would argue that this procedure has methodological weaknesses that may cloud interpretation of the findings. Specifically, Graesser, Singer, and Trabasso (1994) suggested that certain controls should be applied when examining meaning inferences with test words. First, the test words should be generated by individuals from the participant population instead of the experimenters because this increases the likelihood that the majority of participants will associate the items with what was previously read. If the experimenters choose the items, it is not clear that the participants would generate the exact meanings captured by the test words. Second, a baseline RT should be computed for each test word. This baseline should then be used to calculate an inference activation score. This allows for differences between words to be subtracted out so that the differences in question can be examined. This control is particularly relevant to Giora et al.'s procedure because, as mentioned previously, in the pilot study we conducted for this research, we observed differences in RTs between positive and negative words in an LDT. Because all of Giora et al.'s "ironic" test words were negative, and all of their "literal" test words were positive, the faster response times they observed for literal test words may have been due, in part, to a response bias against negative words and not necessarily to activation of the literal meaning. Using an inference activation score would allow these baseline differences to be eliminated.

The results of this study clearly indicate that context plays a role in the processing and interpretation of irony. We also provide evidence that the processing of irony is interactive (and not modular). The degree of negativity of the context situation constrained relative processing of ironic and literal statements. These findings support a

model of irony processing that allows direct access of the ironic meaning. Processing of the ironic meaning is facilitated by appropriate context conditions. Appropriate context conditions involve, among other things, a speaker with a negative attitude and, importantly, a reason for expressing that negativity indirectly.

ACKNOWLEDGMENTS

This research was supported in part by a grant to Penny M. Pexman from the Social Sciences and Humanities Research Council (SSHRC) of Canada.

We thank Lorraine Reggin for assistance in programming Experiment 2, and Kara Olineck for providing the narration in Experiment 3. We are also grateful to Ray Gibbs and three anonymous reviewers for valuable feedback on an earlier version of this article.

The research presented here was part of an MSc thesis submitted by Stacey L. Ivanko at the University of Calgary. Portions of the research described in this article were presented at the meeting of the Society for Text and Discourse, Santa Barbara, CA, July 2001.

REFERENCES

- Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, & Computers*, 25, 257–271.
- Colston, H. L. (2002). Contrast and assimilation in verbal irony. *Journal of Pragmatics*, 34, 111–142.
- Colston, H. L., & O'Brien, J. (2000). Contrast of kind versus contrast of magnitude: The pragmatic accomplishments of irony and hyperbole. *Discourse Processes*, 30, 179–199.
- Dews, S., & Winner, E. (1999). Obligatory processing of literal and nonliteral meanings in verbal irony. *Journal of Pragmatics*, 31, 1579–1599.
- Gerrig, R. J., & Goldvarg, Y. (2000). Additive effects in the perception of sarcasm: Situational disparity and echoic mention. *Metaphor and Symbol*, 15, 197–208.
- Gerrig, R. J., & Healy, A. F. (1983). Dual processes in metaphor understanding: Comprehension and appreciation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 667–675.
- Gibbs, R. W., Jr. (1986). On the psycholinguistics of sarcasm. *Journal of Experimental Psychology: General*, 115, 3–15.
- Gibbs, R. W., Jr. (1993). Process and products in making sense of tropes. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed., pp. 252–274). New York: Cambridge University Press.
- Gibbs, R. W., Jr. (1994). Figurative thought and figurative language. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 411–446). New York: Academic.
- Gibbs, R. W., Jr. (2002). A new look at literal meaning in understanding what is said and implicated. *Journal of Pragmatics*, 34, 457–486.
- Giora, R. (1995). On irony and negation. *Discourse Processes*, 19, 239–264.
- Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, 7, 183–206.
- Giora, R. (2002). Literal vs. figurative language: Different or equal? *Journal of Pragmatics*, 34, 487–506.

- Giora, R., Fein, O., & Schwartz, T. (1998). Irony: Graded salience and indirect negation. *Metaphor and Symbol, 13*, 83–101.
- Glucksberg, S. (1995). Commentary on nonliteral language: Processing and use. *Metaphor and Symbolic Activity, 10*, 47–57.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*, 371–395.
- Jorgensen, J. (1996). The functions of sarcastic irony in speech. *Journal of Pragmatics, 26*, 613–634.
- Katz, A. N., & Ferretti, T. (2001). Moment-by-moment reading of proverbs in literal and nonliteral contexts. *Metaphor and Symbol, 16*, 193–221.
- Katz, A. N., & Lee, C. J. (1993). The role of authorial intent in determining irony and metaphor. *Metaphor and Symbolic Activity, 8*, 257–279.
- Katz, A. N., & Pexman, P. M. (1997). Interpreting figurative statements: Speaker occupation can change metaphor to irony. *Metaphor and Symbol, 12*, 19–41.
- Kreuz, R., & Glucksberg, S. (1989). How to be sarcastic: The echoic reminder theory of verbal irony. *Journal of Experimental Psychology: General, 118*, 374–386.
- Kučera, H., & Francis, W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Matlin, M. W., & Stang, D. J. (1975). Some determinants of word-frequency estimates. *Perceptual & Motor Skills, 40*, 923–929.
- Ortony, A., Schallert, D. L., Reynolds, R. E., & Antos, S. J. (1978). Interpreting metaphors and idioms: Some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior, 17*, 465–477.
- Pexman, P. M., Ferretti, T. R., & Katz, A. N. (2000). Discourse factors that influence on-line reading of metaphor and irony. *Discourse Processes, 29*, 201–222.
- Pexman, P. M., & Olineck, K. M. (2002). Does sarcasm always sting? Investigating the impact of ironic insults and ironic compliments. *Discourse Processes, 33*, 199–217.
- Shinjo, M., & Myers, J. L. (1987). The role of context in metaphor comprehension. *Journal of Memory and Language, 26*, 226–241.
- Temple, D. E., & Geisinger, K. F. (1990). Response latency to computer-administered inventory items as an indicator of emotional arousal. *Journal of Personality Assessment, 54*, 289–297.
- Utsumi, A. (2000). Verbal irony as implicit display of ironic environment: Distinguishing ironic utterances from nonirony. *Journal of Pragmatics, 32*, 1777–1806.

APPENDIX

Context and Statement Sets

Strongly negative, weakly negative, and neutral contexts appear, respectively, in italics.

1) Sam agreed to pick Christopher up after school. *Sam never arrived to pick up Christopher and never called to say why / Sam arrived 1 hour late and apologized / Sam and Christopher talked about the dance on Friday.* The next day Christopher is explaining to Jodi what happened. Christopher says:

Ironic statement: Sam is a nice friend.

Literal statement: Sam is a rotten friend.

Wrap-up sentence: Christopher and Jodi were walking home from school.

2) Terri agreed to help Joan with the toy drive on Saturday. *Terri spent the entire day socializing / Terri spent half the day socializing / Terri spent some of her time helping.* The next day Joan is explaining to Sara what happened. Joan says:

 Ironic statement: Terri is a super helper.

 Literal statement: Terri is a lazy helper.

 Wrap-up sentence: The toy drive lasted from dawn until dusk.

3) Meredith told Kevin that she could memorize a 20 line poem in 5 minutes. *Meredith started to recite the poem but forgot everything after the first 2 lines / Meredith started to recite the poem but only recited half before forgetting the rest / Meredith recited the poem and only forgot 2 lines.* The next day Kevin is talking to Harry about the Poem. Kevin says:

 Ironic statement: Meredith has a powerful memory.

 Literal statement: Meredith has a weak memory.

 Wrap-up sentence: Kevin and Harry were watching a funny movie.

4) Jason was driving April home from school. *Jason cut off another driver and barely avoided a collision / Jason drove very quickly and made tight turns / In the car, Jason and April made plans for the weekend.* The next day April is explaining to Lynn what happened. April says:

 Ironic statement: Jason is a good driver

 Literal statement: Jason is a poor driver.

 Wrap-up sentence: April and Lynn talk while drinking their coffee.

5) Cory entered a cooking competition, claiming he was the best chef in town. *Lauren tasted his main dish and bit into raw chicken / Lauren tasted his main dish which had too much salt / Cory and Lauren talked about all of the different dishes being made.* The next day Lauren is talking to Elsie about the food. Lauren says:

 Ironic statement: Cory cooked a delicious dinner.

 Literal statement: Cory cooked a terrible dinner.

 Wrap-up sentence: Lauren and Elsie were shopping for new clothes.

6) Tracy was in a dance recital and invited Steve to watch. *Tracy tripped many times during her performance / Tracy bumped into another dancer during her performance / Tracy and Steve talked about the recitals that Tracy still had to perform in.* The next day Steve is talking to Dawn about the recital. Steve says:

 Ironic statement: Tracy is a graceful dancer.

 Literal statement: Tracy is a clumsy dancer.

 Wrap-up sentence: Steve and Dawn went shopping with their friends.

7) Brad convinced Melissa to go out to a karioke bar. *Brad got up and sang and people started to throw things at him / Brad got up and sang and a light laugh*

could be heard / Brad and Melissa ordered their drinks and talked about their week at school. The next day Melissa is talking to Jeremy about the bar. Melissa says:

Ironic statement: Brad is a wonderful singer.

Literal statement: Brad is a horrible singer.

Wrap-up sentence: Melissa and Jeremy started studying for an exam.

8) Anne gave Krista tickets to a lecture where she was a guest speaker. *Before Anne was 20 minutes into her lecture, the audience started to leave / At the end of her lecture, Anne noticed that some people fell asleep / Before the lecture, Anne was telling Krista what the lecture was about.* The next day Krista is talking to Mary about the lecture. Krista says:

Ironic statement: Anne is a strong speaker.

Literal statement: Anne is a boring speaker.

Wrap-up statement: Krista and Mary went to attend their class.

9) Kristine and Marie wrote the same math exam. *Kristine hardly studied and received 20% on the exam / Kristine studied for 2 days and received 60% on the exam / Kristine and Marie talked about studying together for the rest of their exams.* The next day Marie is talking to Lane about the exam. Marie says:

Ironic statement: Kristine is a clever student.

Literal statement: Kristine is a poor student.

Wrap-up statement: Marie and Lane went to visit their parents.

10) Kim entered a 10k road race with Shelly, claiming she runs all the time. *Kim started running and fell to last place after 1k / Kim started running and fell to last place after 7k / Kim and Shelly talked about the activities after the race.* The next day Shelly is telling Aaron about the race. Shelly says:

Ironic statement: Kim is a speedy runner.

Literal statement: Kim is a slow runner.

Wrap-up sentence: Aaron went home and did his math homework.

11) Darren and Gordie had a camping trip planned with their sons. *Darren stayed at work late and arrived 1 hour late and missed the bus to the campsite / Darren stayed at work late and arrived 20 minutes late, almost missing the bus to the campsite / Darren and Gordie talked about what they were going to do with their sons on the trip.* At the campsite, Gordie is talking to Bob. Gordie says:

Ironic statement: Darren is a nurturing father.

Literal statement: Darren is a neglectful father.

Wrap-up sentence: Bob and Gordie collected wood for a campfire.

12) Sherri convinced Chuck to take her introductory psychology course. *During her first lecture, Sherri was unable to answer most of the questions asked / Dur-*

ing her first lecture, Sherri was only able to answer a few of the questions asked / Sherri and Chuck talked about what was involved in the course. The next day Chuck is talking to Sally about the lecture. Chuck says:

Ironic statement: Sherri is a knowledgeable professor.

Literal statement: Sherri is an ignorant professor.

Wrap-up sentence: Chuck went to eat dinner with his friends.