Saying What You Don’t Mean
Social Influences on Sarcastic Language Processing
Albert N. Katz,1 Dawn G. Blasko,2 and Victoria A. Kazmerski2
1University of Western Ontario, London, Ontario, Canada, and 2Pennsylvania State University at Erie

ABSTRACT—In recent years, an increasingly large body of research has examined the common situation in which one thing is said in order to express another. Although research has examined the understanding of figurative language such as metaphor in some depth, sarcasm has been less studied. Understanding sarcasm requires considering social and cultural factors, which are often ignored in models of language. We report diverse experiments that point to the same conclusions: Sarcastic interpretation occurs early in processing, with gender and social-cultural factors associated with class playing an important role. These data support interactive models of nonliteral language processing, in which social and cultural factors serve as early-acting constraints on interpretation.

KEYWORDS—figurative language; sarcasm; social-cultural factors; ERPs

Consider a statement such as “all men are animals.” The intent and meaning of this simple statement is quite different if its author is a professor teaching logic, a biologist speaking on human evolution, or a woman, being sarcastic, telling her friend about her passive and boring boyfriend. But in the latter case, how would the listener recover the sarcastic intent from any of the plausible alternatives?

Although computational models of language processing have begun to address the basic question of how a language comprehender can understand the nonliteral sense of a statement, they have yet to be extended to the processing of sarcasm, nor have they acknowledged the important role of social factors in language comprehension. We argue here that in order to fully understand the processes of language processing, it is critical to consider the nature of the person who makes a statement, the nature of the person who receives it, and the context in which this social interaction occurs.

SARCASM AND ITS RELATION TO OTHER FORMS OF NONLITERAL LANGUAGE

Although figurative usage is ubiquitous in natural language, most theories focus only on literal language. In recent years, cognitive scientists have begun to focus on metaphor, simile, and idioms, but there has been much less work on sarcasm and irony. This neglect is unfortunate because, relative to metaphor and idioms, sarcasm is especially dependent on both the discourse and the social contexts in which it is embedded. Irony is typically understood as conveying the opposite meaning from that expressed. Sarcasm, a variant of irony, has a caustic element in its usage and is, unlike irony in general, directed at a specific victim (Lee & Katz, 1998). Indeed, Winner (1988) made the developmental claim that whereas understanding of metaphor is usually present by about age 4, a child’s understanding of irony is not manifested until a few years later because irony, unlike metaphor, cannot be understood from the words alone, but depends on an understanding of social factors that does not develop until later in life.

There is a long but nonexperimental tradition that makes the human actor and the social context central to the act of irony in general, and to sarcastic irony specifically. About 2000 years ago, Quintilian (trans. 1920) wrote that “[irony] is made evident to the understanding either by the delivery, the character of the speaker, or the nature of the subject. For if any one of these three is out of keeping with the words, it at once becomes clear that the intention of the speaker is other than what he actually says” (p. 333). Or, as Hutcheon (1994) put it,

It seems to me that it is precisely the mutual contexts that an existing community creates that set the scene for the very use and comprehension of irony. . . . I want to define these “discursive communities” in general by the complex configuration of shared knowledge, beliefs, values, and communicative strategies. (p. 91)

Although there has been some examination of these communities in the sociolinguistic tradition, these studies have been largely naturalistic observations regarding the frequency, type, or perceived function of language use in different communities.

In contrast, we present here experimental approaches to the study of sarcasm. The studies described typically used short passages (“textoids”) describing the interactions of two or more people in a common social situation, such as moving to a new house or going to the beach. In each passage, one character makes an everyday statement, such as “You are sure a good friend.” We have gained insight into how sarcasm is understood by comparing reactions to these statements when they are produced under different conditions that make them more or less likely to be understood as literally true. For example, we have varied factors such as the context preceding the critical statement, the nature of the person making the statement (for
instance, designating the speaker as John or Jane), and the nature of the person to whom the statement is directed.

**USING AND UNDERSTANDING SARCASM**

Figurative language plays an important role in conveying emotion and modulating emotional intensity (Gibbs, Leggitt, & Turner, 2002). The emotional effects can be different for the speaker and the listener, and a sarcastic comment is seen as more caustic and less funny by the victim than by the speaker. In reading tasks, people perceive characters who use a statement sarcastically as more verbally aggressive, and yet more humorous, than characters who use the same statement literally (Toplak & Katz, 2000). A sarcastic message is also perceived as more insincere, impolite, noninstitutional, and ambiguous than a literal statement.

The classic theoretical position holds that sarcastic language (and other forms of nonliteral language) is processed by obligatory processes during the initial stages of language comprehension, and that contextual and pragmatic effects come into play only later. For instance, Searle (1979) claimed that one must initially process the literal sense of a trope and only when that fails, find a plausible (and nonliteral) interpretation. These classic models suggest that nonliteral language processing should be optional if a literal interpretation is possible and that nonliteral understanding should take more time than literal understanding, because additional processing is needed. In contrast, several more recent models suggest that, with appropriate contextual support, the nonliteral (sarcastic) meaning of a statement is made available as rapidly as the literal sense and that accessing the nonliteral interpretation is not dependent on a failure to achieve a plausible literal interpretation. Both classes of models have been asocial inasmuch as social factors, such as the nature of the speaker or audience, have not been considered. Moreover, both classes of models have lumped sarcasm in with other forms of nonliteral language.

The majority of psychological studies of language comprehension have employed off-line procedures. In comparison with on-line studies, in which investigators attempt to assess language during comprehension, off-line studies involve inferring what occurs in the mind during comprehension by analyzing readers’ interpretations or ratings of an utterance some time after they first encounter it. Such off-line studies have identified some of the social and cultural factors important in producing a sarcastic reading. For example, the effect of a verbal barb is perceived differently depending on the closeness and social status of the speaker and listener. Participants are particularly likely to remember statements that are incongruous with social status, such as when a low-status person talks in an impolite manner to a high-status person. Moreover, perception of a statement as being ironic varies as a function of audience privilege; an interpretation of irony is especially salient when the listener not only perceived a speaker to be making an incongruent statement (perceived the speaker as rejecting its truth value), but also recognizes that someone else in the conversation (the nonprivileged listener) was unaware that the speaker was not, in fact, endorsing the position he or she just espoused.

In addition to contextual constraints engendered by the degree of relatedness and shared knowledge of the participants in a conversation, there are contextual constraints engendered by social stereotypes, such as those associated with gender or socioeconomic class. For instance, in everyday usage, males make sarcastic remarks almost twice as often as do females. Consequently, when the gender of the speaker is manipulated in a textoid, the same comment is rated as more sarcastic when made by a male than when made by a female.

One marker of social class, occupation, has been shown to serve as a contextual marker of ironic and sarcastic language. Members of certain occupations (e.g., clergy and teachers) are perceived stereotypically as likely to use metaphor, whereas members of other occupations (e.g., comedians and factory workers) are perceived as more likely to use irony. Statements such as “that child is a precious gem” are seen as being more sarcastic when made by speakers in high-irony occupations than when made by speakers in high-metaphor occupations (and are also more likely to be recalled correctly). The effect of the speaker’s occupation on creating a sarcastic reading depends on the discourse context and the salience of the statement itself, indicating that multiple sources of information are weighed conjointly when a listener attempts to understand the possibly sarcastic intent behind a given statement.

When participants are asked to read textoids one word at a time, we find evidence for the emergence of sarcasm at a very early stage of processing—by the last word of a target, potentially sarcastic, statement. This effect is produced by having a character in the textoid make a statement incongruent with events unfolding in the story. We have now demonstrated that the information that this statement is made by a comedian (rather than a priest, for instance) further predisposes people to a sarcastic interpretation while they read it. The evidence is based on a convergence of reading-time data with ratings of the subjective sense produced by reading the text. With respect to reading time, we find that when the discourse context is congruent with a sarcastic interpretation, the target sentence is read more rapidly if it is made by a person from a high-irony occupation rather than a person from a low-irony occupation; in addition, if the statement is made by a person from a high-irony occupation, there is no slowdown in reading the first words of the next sentence in the textoid, indicating that the target statement’s meaning is integrated very shortly after the statement is read. However, when there is an incongruity created by having a high-irony speaker make a comment in a context that invites a nonsarcastic interpretation, this incongruity is noted by the beginning of the next sentence, as evidenced by a spike in reading times. In addition, we find that ratings of the degree to which the statement seems sarcastic are correlated reliably with differences in how long it takes to read the last word of the target sentence. Thus, a sense of sarcasm, which is associated with the time taken to read the potentially sarcastic sentence, is present by the last word of the sentence.

There is also emerging data suggesting on-line influences of the speaker’s gender: Participants are slower reading the last words of a sarcastic statement when it is made by a female character than when it is made by a male character, an effect that is particularly marked when the comment is directed at a female. Because sarcasm is more likely to be associated with males than females, comprehension of noncanonical usage is delayed as people attempt to integrate the text they are reading with their stored “knowledge” (stereotypes) of men and women.

These results have implications for theoretical models of language comprehension. Models that posit obligatory processing of the literal or expressed sense of statements predict that evidence for the processing of sarcasm should not, on average, occur until some time after a critical statement is encountered. More interactive, context-driven models predict that the evidence for processing of sarcasm...
should occur very early, during the initial processing of the critical statement. Our evidence favors the latter class of theories and indicates that context involves stored social-cultural stereotypes, such as those relating to gender and class.

We have recently begun investigating these same theoretical questions with event-related potentials (ERPs), which are electrical recordings of the ongoing electrical activity of the brain that are directly linked in time to the onset of the presentation of a specific stimulus or event. ERPs provide a true on-line measure of the brain’s processing of specific information and are recorded with millisecond accuracy. ERP waveforms have series of positive and negative peaks, relative to a baseline, that are correlated with sensory, motor, and cognitive processes. Although often used to study literal language processing, ERPs have seldom been employed for the study of non-literal language processing. We have shown that ERPs are a sensitive means of identifying individual differences in automatic activation of metaphorical meaning (Kazmerski, Blasko, & Dessalegn, 2003) and have now expanded this methodology to the examination of how sarcasm is comprehended.

Our first goal was to examine whether the ERPs in response to a given statement (e.g., “Bob, you’re a really good driver”) would differ depending on whether the preceding context biased toward a literal or sarcastic interpretation. Figure 1 shows the ERPs, from a central scalp site, evoked by the final word of a critical utterance. The ERPs for the sarcastic context, compared with those for the literal context, showed a greater negativity beginning at 500 ms and peaking at 650 ms, followed by a larger positivity at 900 ms (P900). Figure 2 shows activation across all 64 scalp sites at 650 and 900 ms after the initial presentation of the final word of a critical utterance. It is clear that the difference between the conditions is broadly distributed across the scalp.

A negative peak at 400 ms (N400) is often associated with sentence-level semantic integration. The greater negativity for our sarcastic than for our literal stimuli may reflect a greater difficulty in integrating the critical utterance with the ongoing mental representation of the discourse; that is, this integration process may be more resource intensive for the sarcastic than for the literal condition. Alternatively, we know that ERP amplitude has been correlated with emotional intensity (Carretie, Mercado, Tapia, & Hinojosa, 2001), and so the difference between the two sentence types may reflect greater emotional processing associated with sarcastic than literal sentences. Finally, the late P900 we found is somewhat novel but may be related to the processing of humor (Coulson, 2001). Recall that humor is highly associated with the use of sarcasm and, indeed, in an earlier

![Fig. 1. Event-related potentials (ERPs) evoked by the last word of a key statement, in both literal and sarcastic contexts. These tracings show activity at a central parietal scalp site, Pz. Onset of the last word is indicated by an arrow.](image1)

![Fig. 2. Scalp distribution of the mean amplitude of event-related potentials (ERPs) 650 and 900 ms after presentation of the last word of a key statement. ERPs were recorded in both sarcastic (top) and literal (bottom) contexts.](image2)
study, we found that people rated our sarcastic passages as more humorous than the literal passages.

CONCLUSIONS

The data presented here support the following conclusions. First, the nature of the person who makes a statement and the context in which he or she makes it play a role in influencing whether the statement is interpreted as sarcastic. The fact that social and cultural factors influence the processing of sarcasm opens the door for a wide range of studies identifying processing similarities and differences cross-culturally. Such studies might compare sarcasm with other forms of verbal aggression, such as ridicule, or with other forms of nonliteral language. Second, it is the combination of social, cultural, and other factors that determines the likelihood that a statement is interpreted as sarcastic. Gender, occupation, familiarity of the statement, and whether the discourse invites a literal, metaphorical, or sarcastic reading of the statement have all been shown to play a role in producing a sense of sarcasm. In principle, other factors, such as whether there is a victim targeted in the statement, may also prove to be effective. Third, there are implications for the roles of humor and emotion in the ERP data that must be pursued. Finally, the on-line data indicate that at least some of these sources of information come into play very early in the comprehension process, and this time course is inconsistent with the theoretical models positing initial obligatory processing of the literal sense of the expressed meaning. The reading-time data indicate effects that occur by the time the last word of a statement is read, though we have reason to believe that a more sensitive methodology may isolate processing differences between literal and nonliteral readings even earlier. The ERP data may provide clues to when these differences emerge. By analyzing the ERPs occurring prior to the last word of key sentences, we may be able to discover the exact point in the sentences where the brain responses begin to diverge.

How can we integrate the early processing of sarcasm with the fact that a number of factors interactively influence the likelihood that a statement will be perceived as sarcastic? We propose that a constraint-satisfaction approach is a viable option to pursue. In this type of model, different sources of information (i.e., constraints) provide immediate probabilistic support for competing interpretations (e.g., literal vs. nonliteral or metaphorical vs. ironic interpretations). These constraints operate in parallel over time. In general, the duration of the competition between interpretations (and thus reading time) is itself a function of the strength of the various alternatives; competition is resolved rapidly if the constraints all point to the same interpretation, whereas settling on an interpretation is delayed if support for different alternatives is nearly equal. This type of model can explain many of the seeming contradictions in the literature on nonliteral language processing, such as the fact that some studies have found that literal meaning is assessed more rapidly than nonliteral meaning, whereas other studies have found either that there are no differences in processing or that the nonliteral sense is processed more rapidly. Aspects of the constraint-satisfaction explanation of this discrepancy have been supported experimentally for proverbs, an instance of nonliteral language (Katz & Ferretti, 2003).

For this model to be successful, it must also be linked to neurological evidence. We are only now beginning to explore how sarcasm is processed in the brain. Now that we have validated the use of ERPs as a sensitive tool for investigating the processing of sarcastic language, we are ready to begin using this tool to investigate the many social factors that may influence sarcastic understanding.

Recommended Reading


REFERENCES


