

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Do people understand irony from computers?

Permalink

<https://escholarship.org/uc/item/07g0p0p7>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 35(35)

ISSN

1069-7977

Authors

Utsumi, Akira
Watanabe, Yu
Wakayama, Yusuke

Publication Date

2013

Peer reviewed

Do people understand irony from computers?

Akira Utsumi (utsumi@inf.uec.ac.jp)

Yu Watanabe (yuwata@utm.inf.uec.ac.jp)

Yusuke Wakayama (ywganjin@utm.inf.uec.ac.jp)

Department of Informatics, The University of Electro-Communications
1-5-1, Chofugaoka, Chofushi, Tokyo 182-8585, Japan

Abstract

In this paper, we empirically investigate whether people understand irony from computers in order to test the recent argument for an egocentric tendency in irony comprehension. In the experiment, participants took a timed math test comprising 10 questions of 3-digit by 2-digit multiplication. After that, they received a feedback comment on their performance (including potentially ironic sentences) from either an intelligent evaluation system with an AI engine (AI condition), a non-intelligent automatic evaluation system (Auto condition), or a human judge connected via the network (Human condition). The result was that the participants in the AI and Auto conditions understood the comment as ironic as those in the Human condition, and the participants in the AI condition perceived more sarcasm than other participants. Because people know that computers cannot think just as humans do, these results can be regarded as evidence for the egocentric tendency in irony comprehension, indicating that participants understood irony egocentrically from their own perspective without taking into account the mental state of the ironic speaker. These findings are also consistent with the “media equation” theory, from which we can suggest implications for the media equation, anthropomorphism, and computer-mediated communication of irony.

Keywords: Irony; Egocentric interpretation; Theory of mind; Media equation; Computer-mediated communication; Anthropomorphism

Introduction

Verbal irony is a kind of nonliteral language that implicitly conveys the opposite of the literal meaning.¹ To interpret irony, people must infer the speaker’s beliefs and intentions including not only the first-order belief that the speaker does not think that the utterance is literally true, but also the second-order belief that the speaker thinks that the hearers do not think so. For example, imagine that you are rushing to a movie theater, where your friend is waiting for you so that you can see the movie together. When you arrive there about 30 minutes late and you miss the first part of the movie, your friend says to you, “You are always so punctual!” You can easily understand that this utterance is ironic, but you have to know beforehand many things about your friend’s belief and intention. First of all, you must be sure your friend does not think that you are always punctual because you are late for the movie. Furthermore, in order to recognize the speaker’s ironic intention, you must know that your friend thinks that you think this utterance is literally false, because your friend does not intend to convey his/her criticism using irony unless he/she is convinced that you can understand the utterance is literally false.

This property of irony leads to the widely accepted assumption that irony interpretation requires a “Theory of Mind” (henceforth, ToM), which refers to the ability to infer the mental states of others (e.g., Happé, 1993; Sperber & Wilson, 2002).

A large number of recent empirical studies have demonstrated the validity of this assumption. For example, developmental studies have revealed that typically developing children below 5-years of age, who do not completely acquire ToM, cannot understand irony (e.g., Creusere, 2000; Filippova & Astington, 2008; Pexman, 2008). It has also been found that people with pervasive developmental disorder (PDD) such as autism and Asperger syndrome have difficulty understanding irony, and this difficulty has been attributed to impaired ToM (Happé, 1993; Kaland et al., 2002; Wang, Lee, Sigman, & Dapretto, 2006). Hence, irony has been used as a benchmark for testing PDD and discriminating PDD from attention deficit hyperactivity disorder (ADHD) (Adachi et al., 2004). Furthermore, recent neuroimaging studies have demonstrated that, as compared to literal sentences, ironic sentences elicited higher activation in the medial prefrontal cortex, which is known to play a central role in ToM (Rapp, Mutschler, & Erb, 2012; Shibata, Toyomura, Itoh, & Abe, 2010).

On the other hand, some opposing evidence has been reported suggesting that irony comprehension does not always require ToM; people can interpret irony without considering the speaker’s beliefs and intentions (e.g., Akimoto, Miyazawa, & Muramoto, 2012; Keysar, 1994). Keysar (1994) demonstrated that people perceive an utterance as ironic even when it is obvious to them that the speaker of the utterance does not know the discrepancy between an utterance and reality (and thus, the speaker has no ironic intention). Akimoto et al. (2012) also found that people perceive irony by first attributing their own belief to the speaker automatically and subsequently by adjusting it through an effortful ToM process. These findings suggest that irony can be interpreted automatically by the *egocentric* process, and when the egocentric interpretation should be checked for errors and time allows², it is checked according to, or made consistent with, the speaker’s belief by the *allocentric* ToM process. For example, in the case of “late arrival” example presented above, you recognize the utterance “You are always so punctual!” as ironic first by considering the discrepancy between the content of the utterance and your own belief that you arrived late and thus you are not punctual. If you have enough time, you may then consider what your friend really thinks in order to check or confirm your own egocentric interpretation. It must be noted that this egocentric view of irony is also supported by a theoretical study of irony, i.e., Utsumi’s (2000) implicit display theory of irony. Note also that the egocentric interpretation is not specific to

¹This “folk” definition of irony has been recognized as problematic by irony researchers, but it is sufficient for the present purpose.

²Indeed, Epley, Keysar, Boven, and Gilovich (2004) found that ironic interpretation was more egocentric in the time-limited circumstance than in the leisurely circumstance.

irony; recent empirical studies have revealed that communication in general proceeds in a relatively egocentric manner, with addressees routinely interpreting what speakers say from their own perspective (Birch & Bloom, 2007; Keysar, 2007).

From these recent empirical findings of the egocentric nature of irony understanding, we also argue that people perceive irony in an egocentric way without resorting to the allocentric process of considering other's beliefs. In this paper, to obtain further evidence for this claim, we propose a different experimental methodology, namely an experiment with "irony generated by computers." The research question to be answered here is: Do people understand irony from computers? We know that computers cannot think just in the same way as humans do, and thus computers do not, or even cannot, intentionally say irony. Hence, it is highly reasonable to assume that people do not attempt to infer the "mental" states of computers, even when they see or hear a potentially ironic utterance generated by computers. It follows that, if irony interpretation essentially involves the allocentric process of inferring the mental state of the speaker, then people do not perceive irony in computer generated utterances. On the other hand, if irony interpretation does not always require the consideration of the speaker's mental states and is governed by the egocentric process, people may see the irony when they are given potentially ironic utterances by computers. In sum, by empirically examining whether people perceive irony in the computer-generated statements, we can obtain the evidence for or against the claim that irony is understood in an egocentric fashion without or before the allocentric ToM process. This is what this study aims to accomplish.

To examine people's reactions to irony from computers, we conducted a laboratory experiment. In this experiment, participants took a timed math test on computer comprising 10 questions of 3-digit by 2-digit multiplication (e.g., 768×59). They were instructed that the computer system not only provides a math test, but also (1) evaluates their overall performance on math calculation by an AI engine taking into account multiple information such as the test score, the time it took to calculate, and their behavioral data during calculation collected through Web cameras (AI condition); or (2) evaluates their overall performance on math calculation automatically from the test score and the time for calculation (Auto condition); or (3) displays their overall performance on math calculation evaluated by a human judge who observed their behavior during calculation through Web cameras (Human condition). After finishing the test, they received a highly positive comment on their performance from the computer. This positive comment can be ironically interpreted if participants could not get a satisfactory score.

As we mentioned above, our argument for the egocentric nature of irony interpretation predicts that people understand irony from computers just as they understand irony from humans. Therefore, we can predict that people's understanding of irony does not differ among these three conditions of this experiment. Specifically, supposing that the difference between the AI and Auto conditions may lie in the attributability of humanlike mental states to computers (e.g., people may be more likely or easier to attribute the mental state to intelligent computers with AI technology than non-intelligent computers), no difference between these two "computer irony" conditions

also suggests the egocentric nature of irony comprehension.

Our prediction can also be supported by the "media equation" theory (or "Computers Are Social Actors (CASA)" theory) for human-media interaction (Nass & Yen, 2010; Reeves & Nass, 1996). The media equation theory argues that people tend to unconsciously treat computers and other media (e.g., automobiles, cellphones, robots) as if they were real people. For example, people behave politely and cooperatively to computers, and attribute personality characteristics to computers. The media equation has been empirically supported by a number of studies demonstrating that the social rules and heuristics guiding human-human communication apply equally well to human-media interaction. Among these studies, Fogg and Nass's (1997) study on computers that flatter is most relevant to our study. They demonstrated that, when receiving a "flattery" feedback from a computer, people reported the same effects of flattery (e.g., more positive affect and evaluations on computers) as flattery from humans. Likewise, it is reasonable to assume that the media equation predicts that people understand irony from computer just as they understand irony from others, thus suggesting that the answer to the question "Do people understand irony from computers?" is yes. It must be noted here that our study can also be regarded as an empirical study on the media equation, and we can point out the relationship between the media equation and the egocentric communication, which will be discussed later in this paper.

Method

Participants

Fifty-three undergraduate and graduate students participated as volunteers. Note that the recruitment of participants continued until valid data were obtained from 45 participants (i.e., 15 participants for each of the three conditions).

Design

This experiment had three conditions: AI, Auto, and Human conditions. These conditions were manipulated by the instruction given to the participants and the time it took to provide feedback to them (i.e., to display the truth of their answer for each math question, and to display a final comment on their performance), except for which the three conditions were identical.

Procedure

The experiment was conducted using a computer system comprising a Windows PC, an LCD monitor, and two Web cameras. After arriving at the laboratory, participants seated in front of the computer system, and were given an explanation of the purpose of the experiment and an overall instruction of the task they had to perform. Specifically, participants were instructed that the purpose of the experiment was to test and evaluate a computer system in front of you that we were developing, and that this system not only would provide a math test, but also (1) would evaluate their overall performance on math calculation by an intelligent AI engine taking into account multiple information such as the test score, the time it took to calculate, and their behavioral data during calculation collected through Web cameras (AI condition); or (2) would evaluate their overall performance on math calculation automatically from the test score and the time for calculation (Auto

condition); or (3) would display their overall performance on math calculation evaluated by a human judge who observed their behavior during calculation through Web cameras (Human condition). Note that actually the computer system made no evaluation in all the conditions and no human judge evaluated the participants in the Human condition; the system simply displayed the same comment we prepared beforehand, independently of participants' performance. Note also that no humanlike agents were displayed on the monitor.

After the instruction, participants took a timed multiplication test comprising 10 questions of 3-digit by 2-digit multiplication (e.g., 768×59). Multiplication questions were randomly generated so that they did not differ in complexity. In order to make more errors and thus to be more likely to perceive irony, participants were requested to calculate as quickly as possible and complete each multiplication within 30 seconds. If 30 seconds passed since they started each question, they received a warning from the system. Multiplication questions appeared on the monitor one at a time and remained there until participants typed the answer. Participants calculated a given multiplication on paper and typed the answer. The truth of the answer was then presented on the monitor one second (in the AI and Auto conditions) or three to five seconds (in the Human condition) after the answer was typed.

After finishing the test, participants received from the system a highly positive comment on their performance, together with summary statistics including the number of correct and incorrect answers, the mean answering time, and the number of questions they took more than 30 seconds to answer. The comment was that "You have a perfect calculation performance. You were very careful not to make a mistake. The time you took to calculate is also excellently fast." In the comment, the first and last sentences were potentially ironic, if participants' performance is not satisfactory.

After receiving the comment from the computer system, participants were asked to answer a questionnaire. The questionnaire comprised six questions with 7-point Likert scales. The questions that we use in the analysis were: "Do you perceive irony in the comment?" (irony rating; 7 = *ironic*, 1 = *not at all ironic*), "Do you perceive sarcasm in the comment?" (sarcasm rating; 7 = *sarcastic*, 1 = *not at all sarcastic*), and "Do you think the comment is intentional?" (intentionality rating; 7 = *intentional*, 1 = *not at all intentional*). Other questions were: "Does the comment literally include praise or criticism?," "Is your performance satisfactory?," and "Is the evaluation given by the system appropriate?" At the close of the experiment, participants were told the true purpose of the experiment and debriefed. None of the participants were suspicious of the true purpose of the experiment.

Result

Whether the comment presented to the participants was understood as ironic greatly depends on their performance on the calculation test. Therefore, in the analysis, we did not use the data of eight participants who correctly answered all the multiplication questions within 30 seconds, because they were very unlikely to perceive irony in the comment. In other words, in order to collect the valid data of 15 participants per condition (and thus a total of 45 participants), we had to recruit 53 participants.

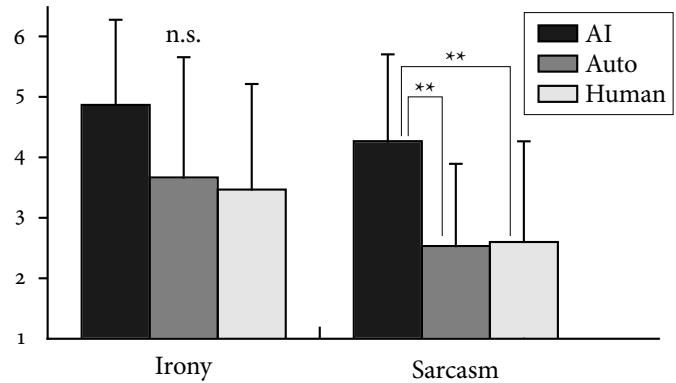


Figure 1: Mean irony and sarcasm ratings for the three conditions

In order to confirm that the likelihood of perceiving irony in the comment did not differ among three groups of participants, we analyzed the difference in the number of incorrect answers and answering time. The total numbers of incorrect answers were 31, 20, and 17 for the AI, Auto, and Human conditions, respectively; the difference did not reach the level of statistical significance, but was close to it, $\chi^2(2, N = 450) = 5.65$, $p = .059$. This result suggests that we must regress out the effect of incorrect answers in the following analysis. On the other hand, the mean answering times per question were 26.8, 26.1, and 26.8 seconds for the AI, Auto, and Human conditions, respectively; they did not significantly differ, $F(2, 42) = 0.17$, $p > .80$.

Irony and Sarcasm Ratings

First of all, we examined whether the mean irony and sarcasm ratings differ among the three conditions, as shown in Figure 1. Concerning irony ratings, the participants in the AI condition appeared to perceive the comment as more ironic than other participants. A one-way, between-participants ANOVA showed that the difference among the three conditions was marginally significant, $F(2, 42) = 2.08$, $p = .08$. However, an ANCOVA with the number of incorrect answers as the covariate revealed that this difference was no longer significant, $F(2, 41) = 1.55$, $p = .22$. This means that, when the number of incorrect answers was statistically controlled, the adjusted irony ratings did not differ among the three conditions; as predicted, people understood irony from computer just as they understood irony from humans.

The mean sarcasm rating was also higher in the AI condition than in the other two conditions, and this difference reached the level of statistical significance, $F(2, 42) = 6.06$, $p < .01$. An ANCOVA with the number of incorrect answers as the covariate also revealed that this main effect was reduced, but remained significant, $F(2, 41) = 4.91$, $p < .05$. Pairwise comparisons ($p < .05$) confirmed that the participants in the AI condition perceived the comment as significantly more sarcastic than the participants in the other two conditions. In particular, it is surprising that the mean sarcasm ratings in the Auto and Human conditions were very low, suggesting that the participants in these conditions did not perceive sarcasm. Possible reasons of this result will be discussed in the section of discussion.

In addition, we analyzed the correlation between irony and sarcasm ratings. In general, “blame-by-praise” irony, such as ones used in this experiment, accompanies a sarcastic effect (Kreuz & Glucksberg, 1989). Therefore, we can confirm from this analysis that participants’ judgment on irony and sarcasm was not arbitrary, and consequently the above result was reliable. The correlations between irony and sarcasm ratings were .51 (AI condition), .83 (Auto condition), and .77 (Human condition), and they were all significant ($p < .05$). This result clearly indicates that the participants who interpreted the comment as ironic also perceived sarcasm in the comment, and thus the obtained result on irony and sarcasm ratings was reliable.

Intentionality Rating

First of all, we analyzed the mean intentionality ratings for the three conditions. The mean intentionality ratings were 4.00 (SD = 1.73) for the AI condition, 3.93 (SD = 1.75) for the Auto condition, and 3.40 (SD = 1.96) for the Human condition. They did not significantly differ both in the ANOVA analysis, $F(2, 42) = 0.49$, and in the ANCOVA analysis with the number of incorrect answers as the covariate, $F(2, 41) = 0.58$. The participants perceived the same low degree of intentionality involved in the comment, regardless of whether the speaker was a computer or a human. This result suggests that people may interpret the comment egocentrically.

Next, we examined the correlations between the intentionality and irony ratings. If people interpret irony by allocentrically thinking about the speaker’s beliefs and intentions, they recognize the intentionality of the speaker in an ironic comment. Meanwhile, if people interpret the comment literally, they do not (or do not have to) recognize the intentionality behind the literal comment because it simply states the fact. Therefore, the allocentric view of irony understanding predicts a positive correlation between the irony and intentionality ratings. On the other hand, if people interpret irony egocentrically, they do not have to recognize the intentionality of the speaker, and thus the egocentric view of irony understanding predicts no correlation between the irony and intentionality ratings. The correlations between irony and intentionality ratings were $r = .11$ (AI condition), $r = -.05$ (Auto condition), and $r = .08$ (Human condition). All these correlations were not at all significant, thus providing additional evidence for the egocentric view of irony understanding.

Discussion

Anthropomorphism and Egocentric Comprehension

The design of the experiment in this paper is premised on the assumption that people consciously know that computers do not have minds and thus cannot think as humans do. However, many researchers criticize this assumption on the empirical grounds that people tend to attribute human characteristics, beliefs, intentions, or emotions to nonhuman agents and objects (e.g., Epley, Waytz, & Cacioppo, 2007). This tendency is known as *anthropomorphism*. According to the anthropomorphic explanation, people do not think of computers as mindless, and as a result our finding that people understand irony from computers does not imply the egocentric view of irony comprehension. Against this criticism, we defend our position as follows. A number of researchers have discussed a

variety of anthropomorphic experiences, which can be classified into two types: a strong, mindful anthropomorphism and a weak, mindless anthropomorphism (Kim & Sundar, 2012). Considering a number of existing empirical findings on anthropomorphism, we deny the possibility that the participants of our experiment anthropomorphized computers mindfully (i.e., in a strong, mindful sense). A weak, mindless anthropomorphism might occur, but it implies that people’s reasoning about the “mental states” of computers is quite egocentric.

Mindful anthropomorphism refers to the tendency to infer the mental states of nonhuman agents or objects from an allocentric perspective. For example, some pet owners perceive and speak of pets as being thoughtful and considerate. This anthropomorphic process is often conscious and seems to require ToM. Recent research demonstrates that whether people mindfully anthropomorphize nonhuman agents and objects depends on two properties, i.e., agency (the capacity to plan and act) and experience (the capacity to sense and feel) (Gray, Gray, & Wegner, 2007). Indeed, Krach et al. (2008) demonstrated through a fMRI experiment that activation of the brain regions (i.e., the medial frontal cortex and the right temporo-parietal junction) known to be associated with ToM was correlated with the degree of agency or humanlikeness (i.e., a human partner, a highly humanlike robot, a functional robot, and a non-humanlike computer, in descending order of agency). The computer system used in our experiment had no humanlike appearance and displayed no humanlike characters, and thus is very low in both agency and experience. Hence, we can safely say that the participants of our experiment did not anthropomorphize the computer system mindfully; this indicates that our assumption that people think of computers as mindless holds true for the experiment.

Mindless anthropomorphism refers to the tendency to automatically attribute human mental states to nonhuman agents or objects without much consideration of whether nonhuman targets have mental states (Kim & Sundar, 2012). Many of the studies on anthropomorphism have used the term “anthropomorphism” to refer to this mindless version. For example, Epley et al. (2007) state, “Using one’s own mental states and characteristics as a guide when reasoning about other humans is egocentrism. Using one’s own mental states and characteristics as a guide when reasoning about nonhuman agents is anthropomorphism (*ibid.*, p.868).” Their notion of anthropomorphism clearly indicates that mindless anthropomorphism is egocentric. Hence, even if the participants of our experiment mindlessly anthropomorphized computers during the experiment, they did not directly infer the “mental states” of computers. It follows that the finding that they perceived irony from computers implies that they did so egocentrically, as we argue in this paper.

It must be noted that the media equation is consistent with mindless anthropomorphism; in other words, the media equation is primarily due to the egocentric nature of communication. Nass and Moon (2000) have argued that the notion of mindlessness provides a robust explanation for the media equation. As have been observed in a variety of social situations, people mindlessly apply social rules and expectations to computers. This phenomenon completely coincides with mindless anthropomorphism. Although they reject an anthropomorphic explanation of the media equation, but the

anthropomorphism they rejected is the mindful version of anthropomorphism. Egocentric, mindless anthropomorphism is a main cause of mindless behavior observed in a variety of media equation studies.

Egocentric and Allocentric Comprehension of Irony

This paper has provided empirical evidence for the egocentric view of irony processing in a novel approach of using computers as ironists. People understand irony from their own perspective, and this leads to the obtained result that people understand irony from computers in the same way as they understand irony from humans. However, the egocentric view seems to be inconsistent with the strong relationship between the ability to understand irony and the ToM ability, which has been justified by a large number of empirical studies. Of course, the egocentric view does not imply that ToM (and the allocentric process) is unnecessary for irony processing, but the argument that people can understand irony without considering the mental states of others from the speaker's perspective seems to be inconsistent with the empirical findings that people with a ToM deficit cannot understand irony. How can the egocentric view explain these incompatible findings?

One possible explanation would be given from a developmental perspective; ToM is a prerequisite for acquiring the concept of irony (i.e., what is irony), but once people (i.e., children) know what is irony, they increasingly do not take into account the mental state of the speaker. The concept of irony essentially involves the speaker's intention of being ironic, which is motivated by a certain situational setting (which is referred to as *ironic environment* by Utsumi's (2000) implicit display theory) where the speaker's expectation has not been fulfilled and the speaker has a negative attitude toward it. Therefore, to acquire the concept of irony, children must be able to infer the mental state of the speaker. In general, children below 5-years of age cannot distinguish between what they know and what others know, and behave egocentrically as if their own beliefs are shared by others, from which it naturally follows that they cannot be aware of irony. Typically developing children at around 5 years of age can increasingly distinguish what others know from what they know, and they come to understand some aspects of irony. Children's appreciation of irony continues to develop into adolescence. As demonstrated by a number of developmental studies on irony, in this development period children's performance on irony understanding is correlated with their (allocentric) ToM ability, because they are acquiring the concept of irony with the help of their developing ToM ability. Adults, who completely acquired the concept of irony and have experienced a number of ironic communication, develop the egocentric tendency again, and increasingly do not take into account the mental state of the speaker (Keysar, 2007), mainly to shortcut the burdensome process of allocentric comprehension.

Irony in Computer-Mediated Communication

The experiment of this paper was conducted through computer-mediated communication. One may argue that our findings are specific to computer-mediated communication and should not be generalized to irony understanding in ordinary face-to-face communication. In other words, it may be pointed out that our finding is an artifact of computer-mediation communication and allocentric comprehension is

always required in face-to-face communication. We basically reject this possibility, but at the same time point out that there may be some truth in it.

Some empirical evidence against this possibility was obtained. Hancock (2004) found that comprehension of irony did not differ between computer-mediated communication and face-to-face communication. More important is their finding that a misunderstanding rate of irony did not differ between these two communication modes, and it was equal to the estimate (i.e., approximately 5%) given by Gibbs (2000). Considering Keysar's (2007) argument that egocentric understanding can provide a systematic reason for misunderstanding, this finding may suggest that people understand irony in computer-mediated conversation as egocentrically as in face-to-face conversation; this clearly rejects the specificity of computer-mediation communication, thus indicating that the finding of this paper is not an artifact of computer-mediated communication. Note also that, in almost all empirical studies of irony, participants of the experiment were asked to understand irony from the addressee's perspective, but they were not the addressees of irony. On the other hand, our participants were literally the addressees of irony, and thus we may safely say that the experiment of this paper was conducted in a more realistic setting, which is more similar to face-to-face communication.

At the same time, some positive arguments for the specificity of computer-mediated communication can be pointed out concerning irony production. Hancock (2004) also revealed that speakers in computer-mediated conversation produced more irony than face-to-face speakers. Furthermore, it is pointed out that sentences created via social media such as blogs and Twitter include more irony, which motivates recent NLP studies on automatic recognition of irony (e.g., Reyes, Rosso, & Veale, 2013). These findings are concerned with the production of irony, but appear to suggest that irony comprehension in computer-mediated conversation may be somewhat different. For example, a younger generation, who is familiar with blogs and Twitter, may be likely to interpret utterances in computer-mediated conversation as ironic. In addition, micro-bloggers in Twitter often use the hashtag #irony or #sarcasm to clearly indicate their ironic intention. It is a very characteristic property of irony in microblogs, because in general ironic intention should not be explicitly expressed so that irony does not lose its effect. This specific property of Twitter may possibly induce the younger generation (including the participants of our experiment) to understand irony more egocentrically in computer-mediated communication. This may be a potential reason for the result that the mean sarcasm rating was higher in the AI condition, assuming that the perceived agency of the AI condition is most similar to that of micro-bloggers. On the other hand, the participants in the Auto condition might perceive little agency. The participants in the Human condition might be aware of more humanity in an imaginary human judge than in micro-bloggers because they were told that the human judge was observing their behavior throughout the math test.

Concluding Remarks

The experiment reported in this paper demonstrated that people perceived the comment as ironic regardless of whether the

speaker of irony is a computer or a human. The experiment was based on the “Computers Are Social Actors” or “CASA” paradigm, which is a novel approach for the study of irony. The obtained finding provided empirical evidence in favor of the egocentric tendency in irony comprehension, because if people understand irony by routinely considering the mental state of the speaker, they could not perceive irony when a computer is the speaker of irony. Through the study of this paper, we have also discussed some features of irony comprehension in computer-mediated communication and the relationship among egocentric communication, anthropomorphism, and media equation. It is worth pursuing these issues for further research.

Acknowledgments

This study was supported in part by a Grant-in-Aid for Scientific Research B (No.23300098) from the Japan Society for the Promotion of Science.

References

- Adachi, T., Koeda, T., Hirabayashi, S., Maeokae, Y., Shiota, M., & Wada, E. W. A. (2004). The metaphor and sarcasm scenario test: A new instrument to help differentiate high functioning pervasive developmental disorder from attention deficit/hyperactivity disorder. *Brain & Development*, 26, 301–306.
- Akimoto, Y., Miyazawa, S., & Muramoto, T. (2012). Comprehension process of verbal irony: The effects of salience, egocentric context, and allocentric theory of mind. *Metaphor and Symbol*, 27, 217–242.
- Birch, S., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science*, 18, 382–386.
- Creusere, M. (2000). A developmental test of theoretical perspectives on the understanding of verbal irony: Children’s recognition of allusion and pragmatic insincerity. *Metaphor and Symbol*, 15(1&2), 29–45.
- Epley, N., Keysar, B., Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology*, 87(3), 327–339.
- Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4), 864–886.
- Filippova, E., & Astington, J. (2008). Further development in social reasoning revealed in discourse irony understanding. *Child Development*, 79(1), 126–138.
- Fogg, B., & Nass, C. (1997). Silicon sycophants: The effects of computers that flatter. *International Journal of Human-Computer Studies*, 46, 551–561.
- Gibbs, R. (2000). Irony in talk among friends. *Metaphor and Symbol*, 15(1&2), 5–27.
- Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, 315, 619.
- Hancock, J. (2004). Verbal irony use in face-to-face and computer-mediated communications. *Journal of Language and Social Psychology*, 23(4), 447–463.
- Happé, F. (1993). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition*, 48, 101–119.
- Kaland, N., Moller-Nielsen, A., Callesen, K., Mortensen, E., Gottlieb, D., & Smith, L. (2002). A new ‘advanced’ test of theory of mind: Evidence from children and adolescents with Asperger syndrome. *Journal of Child Psychology and Psychiatry*, 43(4), 517–528.
- Keysar, B. (1994). The illusory transparency of intention: Linguistic perspective taking in text. *Cognitive Psychology*, 26, 165–208.
- Keysar, B. (2007). Communication and miscommunication: The role of egocentric processes. *Intercultural Pragmatics*, 4, 71–84.
- Kim, Y., & Sundar, S. S. (2012). Anthropomorphism of computers: Is it mindful or mindless? *Computers in Human Behavior*, 28, 241–250.
- Krach, S., Hegel, F., Wrede, B., Sagerer, G., Binkofski, F., & Kircher, T. (2008). Can machines think? Interaction and perspective taking with robots investigated via fMRI. *PLoS ONE*, 3(7), e2597.
- Kreuz, R., & Glucksberg, S. (1989). How to be sarcastic: The echoic reminder theory of verbal irony. *Journal of Experimental Psychology: General*, 118(4), 374–386.
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81–103.
- Nass, C., & Yen, C. (2010). *The man who lied to his laptop: What machines teach us about human relationships*. Penguin/Current.
- Pexman, P. (2008). It’s fascinating research: The cognition of verbal irony. *Current Directions in Psychological Science*, 17(4), 286–290.
- Rapp, A., Mutschler, D., & Erb, M. (2012). Where in the brain is nonliteral language? a coordinate-based meta-analysis of functional magnetic resonance imaging studies. *NeuroImage*, 63, 600–610.
- Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge University Press.
- Reyes, A., Rosso, P., & Veale, T. (2013). A multidimensional approach for detecting irony in Twitter. *Language Resources and Evaluation*, 47(1), 239–268.
- Shibata, M., Toyomura, A., Itoh, H., & Abe, J. (2010). Neural substrates of irony comprehension: A functional MRI study. *Brain Research*, 1308, 114–123.
- Sperber, D., & Wilson, D. (2002). Pragmatics, modularity and mindreading. *Mind & Language*, 17(1/2), 3–23.
- Utsumi, A. (2000). Verbal irony as implicit display of ironic environment: Distinguishing ironic utterances from non-irony. *Journal of Pragmatics*, 32(12), 1777–1806.
- Wang, A., Lee, S., Sigman, M., & Dapretto, M. (2006). Neural basis of irony comprehension in children with autism: The role of prosody and context. *Brain*, 129, 932–943.