

The mental space structure of verbal irony

YOSHIHIKO KIHARA*

Abstract

This article presents a unified theory of irony which claims, with the help of Fauconnier's (1985) mental space theory, that an ironical utterance refers to the mental space of a mutually manifest expectation. According to this view, what a typical ironical speaker does is to say without any distinct space builders that something is the case in the mental space of expectation in order to make it mutually manifest that it is not so in the initial reality space. This expectation space theory of irony integrates the explanatory power of Sperber and Wilson's (1986) echoic interpretation theory with the descriptive power of the view of irony as relevant inappropriateness.

Keywords: irony; mental space; relevance; expectation space; verbal irony.

1. Introduction

The literature on irony is so extensive that I cannot but limit myself to the analysis of verbal irony in this paper, deliberately ignoring other types of irony such as Socratic irony, dramatic irony, situational irony and so on. The definition of irony in *The Concise Oxford Dictionary* as “an expression of meaning, often humorous or sarcastic, by the use of language of a different or opposite tendency” will do as a working definition, since the examples of what are called *ironical* utterances supplied by various scholars discussed below will circumscribe the category of their own accord. Also since “[t]here is no consensus on whether sarcasm and irony are essentially the same thing, with superficial differences, or if they differ significantly” (Attardo 2000: 795), I do not attempt to distinguish between irony and sarcasm here, regarding the latter as a subcategory of the former.

There have been two mainstream families of approaches to verbal irony so far; one is based on the traditional view of irony as a trope and the other on a more recent view of irony as echo or pretense.¹

The traditional view of irony as a kind of negation (i.e., saying something while meaning the opposite) or as a figure of speech (i.e., saying something while meaning something else) was radically refined by Grice (1975, 1978) who sees irony as an effect of implicature produced by the violation of the maxim of quality. According to him, for instance, an ironical speaker means something is bad by saying “good” when it is evidently bad. This Gricean pragmatic analysis of irony has been recently modified by such theorists as Giora (1995), whose view involves indirect negation and the graded salience hypothesis, and Attardo (2000), who regards irony as *relevant inappropriateness*.

On the other hand, as is well known, another important approach to irony comes from the *echoic mention* theory first proposed by Sperber and Wilson (1981), who deal with verbal irony as a subtype of echoic use of language, although the term *echoic* is very confusing as will be clarified below. According to them, roughly speaking, when an ironical speaker says “good”, she is repeating an antecedent utterance by someone, while dissociating herself from the opinion echoed. This theory has been revised as the *echoic interpretation* theory (Sperber and Wilson 1986; Wilson and Sperber 1992) and has at the same time given rise to a family of variations, namely Clark and Gerrig’s (1984) *pretense* theory, Kreuz and Glucksberg’s (1989) *echoic reminder* theory, and the *allusional pretense* theory of Kumon-Nakamura et al.’s (1995).

Although each of these revisions seems to add to the explanatory and descriptive power of their precedents, neither family of approaches seems to explain the coexistence of the following pair of ironical utterances.² Consider a situation in which two people approach a door. The first person opens the door and lets it swing shut behind her. The second person who happens to be carrying a heavy box says to the first person:

- (1) a. Thanks for holding the door.
- b. Thanks for shutting the door.

Here is a striking but previously unnoted fact; that both (1a) and (1b) can come across as irony. The speaker of (1a) may be echoing an imaginary utterance in a counterfactual expected situation, but such an analysis obviously fails in the case of (1b); thus the echo theory fails here. In the meantime, both (1a) and (1b) are clearly *pragmatically insincere* and *relevantly inappropriate* but neither insincerity nor inappropriateness can account for the compatibility of these two similarly ironical but propositionally contradictory remarks; therefore the Gricean theory also fails.

This is not an isolated example. Consider also a situation in which a driver comments, when the car ahead abruptly turns left without signaling:

- (2) a. I love people who signal when turning.
- b. I love people who don't signal when turning.

As in the case of (1a) and (1b), (2a) comes across no less ironical than (2b). The speakers of (1a) and (2a) refer to events that did not take place while the speakers of (1b) and (2b) refer to events that *did* take place. The ironical effect, however, seems to remain virtually the same. How is it possible that one utterance should be as ironical as another that propositionally contradicts it? An integral theory of irony must be one that can account for such phenomena.

I present a theory that verbal irony is a reference to a mutually manifest expectation space (that is, in short, an expectation held by someone that the speaker assumes is recognizable by the hearer) without any distinct space builders. The account offered will explain how irony achieves its effect, how the same content can be conveyed by both an ironic positive assertion **A** and its negative counterpart **-A**, why irony is most often used to criticize and not to praise, and why negative sarcastic utterances do not require explicit antecedents while positive ones do.

2. Mental space theory

A similar strange phenomenon of the parallel coexistence of seemingly contradictory representations has been studied as an opaque/transparent ambiguity in semantics. When Max is planning to marry a green-eyed woman who he thinks has blue eyes, the following two sentences are compatible with each other.

- (3) a. Max believes he will marry the woman with green eyes.
- b. Max believes he will marry the woman with blue eyes. (Sweetser and Fauconnier 1996: 14)

To solve the puzzle, let us specify in what context such an oddity comes about.

- (4) Max believes the woman with green eyes has blue eyes. (Sweetser and Fauconnier 1996: 13)
- (5) In the picture, the woman with green eyes has blue eyes.

Although a woman with green eyes cannot, in a normal context, have blue eyes at the same time, the contradiction in the clause *the woman with green eyes has blue eyes* disappears completely in (4) and (5) when it

is put in what is called a *belief-context* or an *image-context* (i.e., when it is preceded by *Max believes*, *in the picture*, and so on). The question is what the mechanism underlying such sentences is.

Fauconnier's (1985) *mental space* theory, built and developed on what is called *scope theory* (Russell 1905; Quine 1956) and then Jackendoff's (1975, 1983) *opacity principle*, provided a general model for studying the diverse and rich cognitive phenomena that involve domain connection in human thought and language.

The main idea of the mental space theory can be shown quite simply in the case of (4) and (5) as follows (Sweetser and Fauconnier 1996: 10):

When such sentences appear at some point in a discourse they open a new domain (set up to contain structured information and inferences “about” beliefs or time periods or pictures, etc.). Phrases such as *in 1952* or *in the picture* or *Max believes* are all thus space builders—overt mechanisms which speakers can use to induce the hearer to set up a new mental space. They provide in themselves very little explicit information about that new domain, or what it purports to refer to. And for that reason, any additional structure that may be needed in the domain for reasoning purposes will typically be inherited according to default mechanisms from other domains, and ultimately often from background knowledge. In the same fashion, counterparts for elements in existing domains will be created in the new domains. Connectors link domains, and domains may be linked in more than one way.

So in the case of (4), the space builder *Max believes* opens a new mental space **M** (i.e., Max's belief) in which “a woman with green eyes” in the initial base space **B** (i.e., reality) has her counterpart “a woman with blue eyes”, and the two women are linked to each other by the identity connector. There are thus two ways of accessing the woman in question, either through **B** or through **M**. So if Max believes he will marry the woman, there are two ways to say it as in (3a), accessed from **B**, and (3b), accessed from **M**. Mental space access explains how these two seemingly contradictory sentences are compatible.

This account of opaque/transparent ambiguities by mental space theory sheds a new light on verbal irony where we find similar oddities like (1) and (2). Irony can best be analyzed as a mental space phenomenon.

3. Irony as reference to an expectation space

According to Sperber and Wilson's echoic interpretation theory of irony (the revised version of echoic mention theory), the recovery of ironical implicatures depends on three factors: (a) a recognition of the utterance

as echoic, (b) an identification of the source of the opinion echoed, and (c) a recognition that the speaker's attitude to the opinion echoed is one of rejection or dissociation (1986: 240). Sperber and Wilson (1986: 238) call those cases *echoic* in which the "interpretations achieve relevance by informing the hearer of the fact that the speaker has in mind what so-and-so said, and has a certain attitude to it". The term *echoic* is, however, very confusing, and often misleading, because it does not require either that an echoic ironical utterance should represent an attributable belief or that there should be an original utterance to be echoed ironically.³ That is, in their words:

Not all ironical echoes are as easily recognisable. The thought being echoed may not have been expressed in an utterance; it may not be attributable to any specific person, but merely to a type of person, or people in general; it may be merely a cultural aspiration or norm. (Wilson and Sperber 1992: 60)

Therefore it makes little sense any longer to call it echoic, so that Kreuz and Glucksberg propose to rename Sperber and Wilson's account the *echoic reminder theory*, and redefine ironical remarks as allusive reminders of some antecedent event, or as "allusions to prior occurrences or states of affairs" (1989: 375).

Kumon-Nakamura et al. (1995), casting aside these shortcomings of the echoic mention theory and integrating part of Clark and Gerrig's (1984) pretense theory and Kreuz and Glucksberg's (1989) echoic reminder theory, propose the *allusional pretense theory of irony*. The allusional pretense theory has two tenets (Kumon-Nakamura et al. 1995: 5): one is that "a necessary property of discourse irony is an allusion to some prediction, expectation, preference or norm that has been violated"; the other is that "all ironically intended utterances involve pragmatic insincerity, in that they violate one or more of the felicity conditions for well-formed speech acts". In their view, in other words, recognition of the violation of conversational felicity conditions gives the hearer the clue for detecting irony in an utterance that alludes to a failed expectation. But they neither ask nor answer the question why irony is characterized by these two properties (i.e., allusion and insincerity) instead of one such thing, nor what the relationship between them is. Attardo's view that "irony is essentially an inappropriate utterance which is nonetheless relevant to the context" (2000: 823) is open to a similar question: why does verbal irony have to involve relevant inappropriateness at all?

I therefore make only one major claim, not two: that ironical remarks have their effects by referring to a counterfactual mental space of expectation without any distinct space builders.⁴ I contend that this is the only

necessary and at the same time the sufficient condition for an utterance to be ironical.

3.1. *Counterfactual assertives*

Let us begin by analyzing simpler cases of discourse:

- (6) a. Peter: It's a lovely day for a picnic.
[They go for a picnic and the sun shines.]
Mary (happily): It's a lovely day for a picnic, indeed.
- b. Peter: It's a lovely day for a picnic.
[They go for a picnic and it rains.]
Mary (sarcastically): It's a lovely day for a picnic, indeed.
(Sperber and Wilson 1986: 239)

Sperber and Wilson call these utterances by Mary echoic since they echo Peter's preceding remark. So far so good. But Sperber and Wilson (1998: 284) begin to lose their way when they take examples in which they find the speaker "echo[ing] general norms or universal desires":

- (7) Oh great. That's nice.
- (8) A: Bob has just borrowed your car.
B: Well, I like that.

Sperber and Wilson (1998: 285) claim that "someone who says emphatically that things are great when they patently are not does achieve her ironical effect by echoing a representation of what is always desirable". Similarly, they claim that *I like that* in (8) "echoes the universal desire for things to be such that we can like them" (Wilson and Sperber 1998: 285). In other words, what is being echoed may be merely an expectation or a norm (Wilson and Sperber 1992: 60).

We do not need such a forced extension of the key concept of *echo* if we see (6b) as reference to the mental space of expectation that is mutually manifest to the speaker and the hearer. By a *mutually manifest* mental space I mean a mental space which the speaker can have reasonable confidence that the hearer will be able to identify.⁵ The parallelism of (6a) and (6b) is a natural outcome because it makes little or no superficial difference whether Mary echoes Peter's antecedent utterance or whether she refers to the expectation space built by the utterance. What is not clear at all in Sperber and Wilson's argument is what kind of relationship exists among (a) preceding specific utterances, (b) imaginary attributed thoughts and utterances, and (c) standard expectations or norms that are echoed in various ways. The answer to the question must be that what they call echoic utterances in (6b) through (8) all refer to a mutually

manifest expectation, or more strictly speaking, a mutually manifest expectation space.

Then the sarcastic speaker of (7) is not speaking about something **a** that is far from great in the base reality space **R**, but about its counterpart **a'** that is great in the (imaginary) expectation space **E** or, in other words, that is *expected* to be great. Except some elements in **E** that are to be computed (i.e., implicated by the speaker and inferred by the hearer) as different from their counterparts in **R**, the other elements in **E** are by default linked by the identity connector to their counterparts in **R**. We can suppose the existence of a connector that links **a** to its ironical (therefore frequently opposite) counterpart **a'** just as we have customer-dish connectors in restaurants (which enable one to refer to a customer as, say, *a ham sandwich*) or actor-character connectors (which enable one to refer to a movie character as, say, *Tom Cruise*).

Similarly the speaker **B** in (8) is not speaking about her liking Bob borrowing her car in reality. One way of making sense of the utterance is to suppose that *I* refers to herself in **E** (i.e., her counterpart who is expected to be far more generous than her real self and therefore likes Bob borrowing her car), with a (probably jocular) self-critical implicature that she is not generous enough in reality. In this case the speaker seems virtually pretending to be someone else, with the result that the whole picture looks like Clark and Gerrig's pretense theory. But the pretense-theoretical approach does not exhaust the possible interpretation of the utterance.⁶ The second way we can make sense of (8) is to suppose that the speaker likes it that the cautious counterpart of Bob in **E** (who is not like the real-life rough Bob) borrows her car, in which case the speaker implicitly criticizes Bob's roughness. Either way, the speaker is not commenting about something in the base reality space **R**, but is speaking about something in the expectation space **E**.

According to the relevance theory (Sperber and Wilson 1986: 179), when a person hears an utterance, she knows that there is a set of assumptions which the speaker intends to make manifest to her by saying it. So when a person hears an ironical utterance, she knows that there is a set of assumptions which the speaker intends to make manifest to her by saying that something is the case in **E**. Although the communicated assumptions necessarily involve the discrepancy between **E** and **R**, it does not follow that the hearer is able to recover all the precise assumptions, apart from that of the discrepancy itself, which the speaker can be said to intend the hearer to share. This is why (8) can be interpreted in more than one way.

Here is a Necker-cube-like flipping of the view on irony. Whereas all the major theories on irony so far have dealt with ironical remarks as

referring critically to something in reality by means of a kind of lie, I propose that ironical remarks always refer to something in the counterfactual expectation space that is linked by a connector to its ironical counterpart in the base reality space. In other words, the ironical speaker is pretending that **R** is **E**.⁷

Roughly speaking, what a typical ironical speaker does is to say without any distinct space builders that something is the case in **E** in order to make it mutually (more) manifest that it is not so in **R**. If there should be any distinct space builders in an ironical utterance, the utterance will only be a plain statement about some failed expectation. For example, Mary's utterance in (6b) would be almost equivalent to "I had expected a situation where I could say 'It's a lovely day for a picnic, indeed'".

I propose that the main effect of irony lies in the surprise you feel when you hear a sentence like (9), which indubitably involves two mental spaces without any distinct space builders:

(9) The woman with green eyes has blue eyes.

The fact noted by Grice (1978: 125) that one cannot prefix the employment of irony with *to speak ironically* is therefore a natural consequence; it would completely spoil the surprise. Although it is possible to suppose that an ironical tone of voice operates as a covert (non-lexical) space builder, I regard it as optional since we often find deadpan ironical remarks.⁸ An ironical tone is necessary only when it is not manifest enough that the utterance refers not to reality but to an expectation space.

Now, by way of experiment, let us contrive a context in which (9) can be a consistent ironical utterance without any space builders in it. Suppose a film director puts an advertisement in the newspaper for a green-eyed actress for his next film, but then he receives only one application from a woman, who at the interview with the assistant director turns out to have blue eyes. Now the assistant can say (9) to the director felicitously. The assistant is saying that the woman who is expected to have green eyes has blue eyes, implicating that the advertisement was a failure. In other words, in the case of (9), the green-eyed woman in **E** has blue eyes in **R**.

We now have adequate conceptual tools to analyze (1a), (1b), (2a), and (2b). The speaker of (1a) is expressing gratitude for the expected (though failed) act of the hearer's kindly holding the door. The kindly act **a'** in **E** is connected to its ironical counterpart **a** in **R**, namely the hearer's unkind act, thus making it possible to access the act **a'** through both **R** (i.e., as **a**) and **E** (i.e., as **a'**).

The case of (2a) and (2b) is a little more complicated, but the basic analysis remains the same. There is an element **a'** (i.e., people who signal

when turning) that has the property $P'(_)$ (i.e., the speaker likes $_$) in **E**, while in **R** there is a counterpart element **a** (i.e., people who do not signal when turning) that has a counterpart property $P(_)$ (i.e., the speaker hates $_$), so that it is possible to ironically say either (2a) or (2b), implicating that the speaker hates people who do not signal when turning.⁹ Either of the utterances inevitably contains a manifest reference to **E** (i.e., **a'** and $P'(_)$ in [2a], and $P'(_)$ in [2b]), which is sufficient to give the hearer the signal to detect the ironical intent.¹⁰

Now ironical understatements like (10), which Kumon-Nakamura et al. (1995: 3) claim the echoic interpretation theory cannot handle, can also be analyzed in a similar way.

(10) [In a very noisy place:] It is a little noisy here.

The speaker expects the noise to be less than it is in **R**, so that she refers to **E** where it is not very noisy but *a little noisy*. In the same fashion, when something fails to measure up to some expectation, one can readily use ironical overstatements to refer to the failed expectation as in (11):

(11) [About a lukewarm hamburger:] This hamburger is really sizzling hot. (Seto 1998: 245)

The speaker expects the hamburger to be hotter than it is in **R**, so that she refers to **E** where it is not lukewarm but *sizzling hot*.

(10) incidentally shows that the expectation space does not have to exist prior to the recognition of the present circumstances; the expectation can be constructed posteriorly. This is why I call the mental space in question not a *failed expectation space* but simply an expectation space. The wording *failed expectation* necessarily connotes that something posterior falls short of a prior expectation.

3.2. *Asymmetry of irony and the unexpected*

The mental space analysis of irony throws a new light on an aspect of irony that has been given various explanations so far; that is, the asymmetry of affect. It has often been noted that a strange asymmetry exists concerning irony. Why do negative sarcastic utterances require explicit antecedents while positive ones do not, or, to put in a slightly different but related way, why is irony most often used to criticize, not to praise?¹¹ Let us consider (12):

(12) [The mother says to her dirty child ironically:] You're such a clean child!

As Sperber and Wilson see "the discrepancy between the norm of cleanliness that the child is supposed to meet and his actual appearance" (1990:

152), so we see the discrepancy between the two mental spaces and find the child's dirty appearance in **R** blamable as compared to her clean appearance in **E**, hence blame by praise. In this case, the speaker's expectation is in accord with the standard expectation. In addition, needless to say, what is praiseworthy very often coincides with what is expected. So (12) needs no antecedent utterance, and achieves blame by praise.

Now consider a seemingly opposite case of (12) in (13):

- (13) [The mother says to her clean child ironically:] You are such a dirty child!

It is not impossible to suppose a comparatively rare situation where the child is expected to be far dirtier than she is and the irony can work: for example, when the mother tells the child to play outdoors to her heart's content without worrying about dirtying her clothes and hands. This is one of the rare cases in which a usually blamable state of things receives a positive expectation. The specific expectation in this context contradicts the usual expectation, so that it requires an antecedent utterance or context. It is noteworthy that this ironical utterance still comes across not as praise but as blame. Since verbal irony always involves reference to an expectation space, it *usually* implies a disappointment in the reality space.

However, this is not always the case because there are times when something goes *unexpectedly* well. In these cases, subtle praise-by-blame irony works successfully. Sperber and Wilson (1998: 288) take up such an example. Kyoko learns that her husband Jiro has fiddled his traveling expenses and bought her a nice present. Kyoko says:

- (14) You're so naughty.

In (14) Kyoko is referring to either of the following two expectation spaces. The first is the mental space that Kyoko expects **E_k**, where Jiro's fraud is judged nothing more than a piece of childish *naughty* mischief; in this case, Kyoko is in effect saying that she is ready, like a forgiving mother, to acquit him in **E_k**, while critically meaning that it is not a permissible act in **R**. The second is the mental space expected by the social norm **E_s**, where Jiro's fraud is judged bad; in this case, Kyoko is in effect saying that in **E_s** she is supposed by social rules to criticize his fraud, while favorably meaning that she does not have the intension to do so in **R**. In the first interpretation, she means that, though she is happy, his act is still bad, whereas in the second interpretation she means that, though his act is bad, she is happy.

Now we find little difficulty in analyzing the following example cited by Attardo (2000: 796):

- (15) [Your stock broker says on calling for the third time to announce unexpected dividends:] Sorry to keep bothering you like this.

The broker is evidently not speaking in the present situation **R** but is talking in a normally expected situation **E**, where a customer will get irritated by frequent phone calls from a stock broker. As is evident from this case, what is referred to in **E** does not have to be a thing, a property, or an event; it can be an expected situation or expected circumstances in which the utterance is made. Thus the utterance (15) highlights the disparity between the usually expected formality and the current good news that exceeds it.

In short, by referring to something or some state that is the case not in **R** but only in **E**, the ironical speakers of these utterances are implying that something that is expected to be the case is not the case, setting each space off. But even if something runs counter to the expectation, it does not always mean disappointment, because there are times when a personal expectation runs counter to the social expectation or when something goes successfully in reality against all expectations.

3.3. *Mutually manifest expectation*

(16), cited in Kumon-Nakamura et al. (1995: 20), needs a similar but a slightly different analysis:¹²

- (16) [A parent, while trying to concentrate on an important piece of work, says to a teenager who is watching a very loud television program:] Would you mind if I asked you perhaps to consider turning off the TV?

The degree of politeness of requests is determined by such factors as the difference in social status, and the degree of familiarity, between addresser and addressee. The teenager's inconsiderate and impudent behavior is such that it looks as if she were someone to whom the parent has to ask in an extremely humble manner. The discrepancy between what the things are supposed to be in **R** and what the teenager seems to expect the things to be gives birth to a new space (the teenager's supposed expectation space **E**) where the parent has to address the child in such a manner. The connector links **a** (i.e., the social state differential and the familiarity between the parent and the child) in **R** to its counterpart **a'** in **E**, highlighting the preposterous human relations lurking behind the child's arrogant behavior.

Why is it that *under*-polite ironical requests are impossible though under-politeness has essentially the same "relevant inappropriateness"

and same “pragmatic insincerity” as *over*-politeness does? The expectation space theory of irony claims as follows: the theory has it that *over*-polite irony can be used when the subordinate (or coordinate) hearer behaves as if she were superordinate (i.e., when she is superordinate in **E**), and it is safe because it superficially sounds polite; on the other hand, *under*-polite irony can be used only when the superordinate (or coordinate) hearer behaves as if she were subordinate (i.e., when she is subordinate in **E**), and therefore it is quite risky because it sounds impolite to the polite superordinate hearer.

(17) also involves the hearer’s expectation:

(17) Go ahead and break my vase.

It goes without saying that the speaker does not want to have the vase broken, but the situation is such that the hearer is behaving as if she were given leave to do what she wants. In other words, in the hearer’s expectation space, the speaker is supposed to be ready to allow, or to have allowed, the hearer’s outrageous behavior.

When an annoyed listener says (18) to someone who is arrogantly showing off knowledge, it also refers to the hearer’s expectation space:

(18) You sure know a lot. (Kumon-Nakamura et al. 1995: 4)

The speaker is referring to the hearer’s expectation space in which the hearer is not a know-it-all but an unaffected knowledgeable person and admired by others as such, with the implicature that it is not the case in reality.

(16), (17), and (18) are utterances in which the speaker implicates that something that the *hearer* expects to be the case is not the case. The expectation in question, however, does not have to be entertained either by the speaker or the hearer; it just has to be a mutually manifest expectation entertained by some party. Therefore it is possible to speak ironically of a failed weather forecast even when neither the speaker nor the hearer has believed in it:

(19) [Looking at the pouring rain, the speaker says:] What lovely weather, as the forecast said!

Generally speaking, an ironical utterance **U** that refers to the speaker’s, or a mutually shared, expectation space can be translated into a formulaic sentence “I (or we) had expected a situation where I could say **U** (without irony)”, or in short, “I expect such a situation”. When the hearer’s expectation is referred to, the formulaic sentence will be “Had you expected a situation where I should say **U** (without irony)?” or in short, “Do you expect such a situation?” When the expectation is not shared either by the

speaker or the hearer, then the formula will be “So-and-so had expected a situation where I could say U (without irony)”, or in short, “They expect such a situation”.

3.4. Further applications of the expectation space theory

(20), cited by Hamamoto (1998), is a good challenge to the view of irony as relevant inappropriateness as well as to the echoic interpretation theory.¹³ Taro and his wife Hanako are environmental activists who work very hard on environmental issues, spending all their time away from home. Their son Jiro says:

(20) Our home is an environment.

I disagree with Hamamoto (1998: 261) who comments that Jiro does not dissociate himself from the echoed opinion. Jiro does dissociate from the opinion, as is evident in the wording of the utterance, namely the use of the word *environment* and the implied importance attached to it. Jiro expects his parents to recognize that their home is part of what they call environment, while he himself does not think much of environmental issues. So the ironical disparity can be detected in the word *environment*. It lies between the current definition given to *environment* by the parents that excludes their home and the definition that Jiro expects his parents to give the word so that it will include their home. In other words, the word *environment* *e'* in Jiro's expectation space is different from its counterpart word *environment* *e* in **R**, or what the parents call environment. The subtlety of the irony in (20) is that the ironical disparity does not appear either as pragmatic insincerity or as relevant inappropriateness but as a so-to-speak double-bottomed meaning of a word.

The expectation space theory can deal not only with ironical utterances but also with other related tropes, such as oxymoronic mini-phrases of the kind cited in Seto (1998: 248–249): *open secret*, *cruel kindness*, and others. *Open secret* can be analyzed as something that is expected to be a secret but is publicly known, or in our terms, it is something that is a *secret* in **E** but is *open* in **R**. Similarly, *cruel kindness* is something that is *kindness* in **E** but is *cruel* in **R**. Such applicability proves this approach to be more explanatory than other approaches.

4. Mention, pretense, pragmatic inappropriateness, and the expectation space

In some cases, the mental space theory of irony looks similar to Clark and Gerrig's (1984) pretense theory, because the ironical speaker seems to be pretending to be someone who mistakes an expectation space for

the reality space. The important difference is that, in the mental space theory, the ironical speaker do not have to “view the world through rose-colored glasses” (Clark and Gerrig 1984: 122) as in the case of the pretense theory, but she can view the world as expected by anyone. Clark and Gerrig do not explain in the first place why one has to pretend to be an injudicious person, instead of a genius or a scholar, to be ironical.

Also in some cases, the expectation space theory looks similar to Sperber and Wilson’s (1981) echoic mention theory, Sperber and Wilson’s (1986) echoic interpretation theory, and Kreuz and Glucksberg’s (1989) echoic reminder theory, because the ironical speaker sometimes echoes a previous real or imaginary utterance that contains some implicit or explicit expectation. Therefore all the experimental results along the line of echoic theories presented by Jorgensen et al. (1984), Kreuz and Glucksberg (1989), and Kumon-Nakamura et al. (1995) consistently support my view.¹⁴

The expectation space theory can account for the fact that verbal irony often appears as “relevant inappropriateness” (Attardo 2000) or as “pragmatic insincerity” (Kumon-Nakamura et al. 1995), because it claims that an ironical utterance refers, without any markers, not to the state of affairs that is the case in **R** but to a situation that is the case in **E**. The resultant utterance cannot but contain a heterogeneous element in some way.

The family of echoic theories does not account for the existence of pragmatic insincerity or relevant inappropriateness that is usually found in verbal irony, while the family of irony-as-inappropriateness theories do not account for the fact that a majority of ironical utterances involve an echo of some prior utterance or expectation. The expectation space theory of irony integrates the explanatory power of the family of echoic theories with the descriptive power of the family of irony-as-inappropriateness theories.

5. Conclusion

All the major theories on verbal irony have so far concerned themselves with such so-to-speak epistemological questions as why the speaker says “good” about something **a** while **a** is patently not good. The theory of irony as reference to an expectation space, on the other hand, is concerned with such ontological questions as what is the relationship between something **a** that patently is not good and another thing **a’** that the ironical speaker says is good.

Verbal irony is, in conclusion, a reference to an expectation space **E** without any distinct space builders. In spite of the absence of any space builders, it is recognized as ironical because the utterance contains at least

one element **a'** which exists only in **E** (in a mutually manifest manner for the speaker and the hearer) and which will form a marked contrast with its counterpart **a** in the initial reality space **R**.

Received 10 June 2003

Osaka University, Japan

Revision received 27 June 2004

Notes

* I would like to thank Professor Adele Goldberg for her helpful comments and suggestions. I hold MA and PhD in Literature from the Graduate School of Letters at Kyoto University, Japan. I currently teach at the Department of Language and Culture at Osaka University. E-mail <y-kihara@lang.osaka-u.ac.jp>.

1. For a more detailed theoretical survey, see Attardo's (2000) careful review of more than a dozen previous theories of irony and his extensive and insightful discussion, and also Gibbs' chapter on irony (1994: chapter 8).
2. The situation and the utterances (1a) and (2a) are taken from Kumon-Nakamura et al. (1995: 4, 19) with slight modification, and I added (1b) and (2b).
3. Similar criticisms can be found in Kreuz and Glucksberg (1989: 375) and Giora (1995: 246).
4. Since Kumon-Nakamura et al. (1995) regard verbal irony as allusion to a failed expectation, my proposal that verbal irony is reference to the counterfactual mental space of expectation is in part a thorough recast of Kumon-Nakamura et al.'s proposal. Hamamoto's (1998) approach that sees irony as mention to a "prior cognition" (266), which can be interpreted as an expectation or an ideal, is an insightful modification of echoic interpretation theory, but Hamamoto presupposes only ironical propositional assertions. Though Martin's (1992) theoretical framework based on the notions of possible world and universe of belief has given me some hint, my expectation space theory differs considerably from his approach.
5. Mutual manifestness is one of the most important key tenets of relevance theory and is discussed thoroughly in Sperber and Wilson (1986: 38–46). They define *manifest* as follows: "A fact is *manifest* to an individual at a given time if and only if he is capable at the time of representing it mentally and accepting its representation as true or probably true" (Sperber and Wilson 1986: 39). Then as regards *mutually manifest*:

Any shared cognitive environment in which it is manifest which people share it is what we call a *mutual cognitive environment*. In a mutual cognitive environment, for every manifest assumption, the fact that it is manifest to the people who share this environment is itself manifest. In other words, in a mutual cognitive environment, every manifest assumption is what we call *mutually manifest* (Sperber and Wilson 1986: 41–42).

6. The expectation space theory of irony allows such multiple interpretations of an ironical utterance. This unique characteristic is an advantage rather than a weakness. Such possibility of multiple interpretations of an ironical utterance has a lot to do with the analogical counterfactual phenomenon treated in detail in Fauconnier (1997: chapter 4).
7. This is a kind of pretense theory turned upside down. Whereas the usual pretense theory claims that the speaker is pretending to be someone else (i.e., an "unseeing" and "injudicious" person [Clark and Gerrig 1984: 122]), my theory claims that she is pretending that the surrounding world is different from **R** (i.e., **E**).

8. For a similar account of ironical tone, see Jorgensen et al. (1984: 116) and Clark and Gerrig (1984: 122). Seto (1998: 248) lists a variety of irony markers, including prosodic ones.
9. Note that the ironical speaker means by (2a) not that she really loves people who signal when turning but that she hates people who do not signal when turning.
10. See Sperber and Wilson (1986: 232) who summarize the four aspects of verbal communication in a tentative diagram which shows that the propositional form of an utterance is an interpretation of a mental representation of the speaker which can be entertained as an interpretation of (a) an actual representation or (b) a desirable representation, or as a description of (c) an actual state of affairs or (d) a desirable state of affairs. They say that irony falls into category (a), but in my view it roughly fits into category (d) or some other category (assumed to exist next to [c] and [d]) that involves *an expected state of affairs*. See also Curco (2000: 271–273) who goes into the details of the relevance-theoretical process of retrieving a final ironical interpretation from a contradiction between the context and the propositional content of the utterance; she spells out how verbal irony differentiates itself from other cases where propositional clash is encountered (i.e., mistakes, lies, arguments, and deception).
11. For the various answers given to these questions, see Jorgensen et al. (1984: 115), Clark and Gerrig (1984: 122), Kreuz and Glucksberg (1989: 376), Myers Roy's (1977) argument reported in Haverkate (1990: 90), and Giora (1995: 255).
12. Kumon-Nakamura et al. (1995: 20) do not carefully analyze (16) but seem to regard (15) as alluding to the parent's failed expectation (though they do not make it clear what the expectation is). But their view fails to account for the impossibility of an under-polite request to ironically allude to the same expectation.
13. Sperber and Wilson (1998: 288) reply to Hamamoto's critical comments, saying that "Jiro is echoing approvingly a thought that his parents should have". But this account far deviates from their echoic interpretation theory, first in that the speaker's attitude toward the opinion echoed is not one of rejection or dissociation, and secondly in that what is echoed is not a thought that is attributable to some real or imaginary person but a thought that is *expected* to be attributable to some specific person.
14. Jorgensen et al. (1984: 118) assert that "[t]he results tend to support the claim that people do not perceive an implausible non-normative utterance as ironic unless it echoes some antecedent use, which is the outcome predicted by the mention theory of irony". Kreuz and Glucksberg (1989: 374) claim that "Positive statements do not require explicit antecedents because such statements may implicitly allude to societal norms and expectations.... Negative statements, however, cannot implicitly allude to such positive norms, and so they should require explicit antecedents if they are to be understood". Kumon-Nakamura et al. (1995: 5) show that "a necessary property of discourse irony is an allusion to some prediction, expectation, preference, or norm that has been violated". See Curco (2000: 264–267) for a detailed evaluation of a number of experiments looking at the psycholinguistic processing of ironical remarks.

References

- Attardo, Salvatore
 2000 Irony as relevant inappropriateness. *Journal of Pragmatics* 32, 793–826.
- Clark, Herbert H. and Richard J. Gerrig
 1984 On the pretense theory of irony. *Journal of Experimental Psychology: General* 113 (1), 121–126.

- Curco, Carmen
2000 Irony: Negation, echo and metarepresentation. *Lingua* 110, 257–280.
- Fauconnier, Gilles
1985 *Mental Spaces: Aspects of Meaning Construction in Natural Language*. Cambridge, MA: MIT Press.
1997 *Mappings in Thought and Language*. Cambridge, MA: Cambridge University Press.
- Gibbs, Raymond W.
1994 *The Poetics of Mind*. Cambridge: Cambridge University Press.
- Giora, Rachel
1995 On irony and negation. *Discourse Processes* 19, 239–264.
- Grice, H. Paul
1975 Logic and conversation. In Cole, P. and J. L. Morgan (eds.), *Syntax and Semantics*, volume 3: *Speech Acts*. New York: Academic Press, 41–58.
1978 Further notes on logic and conversation. In Cole, P. (ed.), *Syntax and Semantics*, volume 9: *Pragmatics*. New York: Academic Press, 113–127.
- Hamamoto, Hideki
1998 Irony from a cognitive perspective. In Carston, Robyn and Seiji Uchida (eds.), *Relevance Theory: Applications and Implications*. Amsterdam: Benjamins, 257–270.
- Haverkate, Henk
1990 A speech act analysis of irony. *Journal of Pragmatics* 14 (1), 77–109.
- Jackendoff, Ray
1975 On belief-contexts. *Linguistic Inquiry* 6, 53–93.
1983 *Semantics and Cognition*. Cambridge, MA: MIT Press.
- Kreuz, Roger J. and Sam Glucksberg
1989 How to be sarcastic: The echoic reminder theory of verbal irony. *Journal of Experimental Psychology: General* 118 (4), 374–386.
- Kumon-Nakamura, Sachi, Sam Glucksberg, and Mary Brown
1995 How about another piece of pie: the allusional pretense theory of discourse irony. *Journal of Experimental Psychology: General* 124 (1), 3–21.
- Martin, Robert
1992 Irony and universe of belief. *Lingua* 87, 77–90.
- Myers Roy, Alice
1977 Towards a definition of irony. In Fasold, Ralph W. and Roger Shuy (eds.), *Studies in Language Variation*. Washington, DC: Georgetown University Press, 171–183.
- Quine, Willard Van Orman
1956 Quantifiers and propositional attitudes. *Journal of philosophy* 53.
- Russell, Bertrand
1905 On denoting. *Mind* 14, 479–493.
- Seto, Ken-ichi
1998 On non-echoic irony. In Carston, Robyn and Seiji Uchida (eds.), *Relevance Theory: Applications and Implications*. Amsterdam/Philadelphia: Benjamins, 239–255.
- Sperber, Dan and Deirdre Wilson
1981 Irony and the use-mention distinction. In Cole, P. (ed.), *Radical Pragmatics*. New York: Academic Press, 295–318.
1986 *Relevance: Communication and Cognition*. Oxford: Blackwell.

- 1990 Rhetoric and relevance. In Wellbery, David and John Bender (eds.), *The Ends of Rhetoric: History, Theory, Practice*. Stanford, CA: Stanford University Press, 140–155.
 - 1998 Irony and relevance: A reply to Seto, Hamamoto and Yamanashi. In Carston, Robyn and Seiji Uchida (eds.), *Relevance Theory: Applications and Implications*. Amsterdam/Philadelphia: Benjamins, 283–293.
- Sweetser, Eve and Gilles Fauconnier
- 1996 Cognitive links and domains: Basic aspects of mental space theory. In Fauconnier, Gilles and Eve Sweetser (eds.), *Spaces, Worlds, and Grammar*. Chicago, IL: Chicago University Press, 1–28.
- Wilson, Deirdre and Dan Sperber
- 1992 On verbal irony. *Lingua* 87, 53–76.