

## On Negation as Mitigation: The Case of Negative Irony

**Rachel Giora**

*Linguistics  
Tel Aviv University*

**Ofer Fein**

*Department of Behavioral Sciences  
The Academic College of Tel Aviv Yaffo*

**Jonathan Ganzi, Natalie Alkeslassy Levi, and Hadas Sabah**

*Linguistics  
Tel Aviv University*

Four experiments support the view of negation as mitigation (Giora, Balaban, Fein, & Alkabetz, 2004). They show that when irony involves some sizable gap between what is said and what is criticized (*He is exceptionally bright* said of an idiot), it is rated as highly ironic (Giora, 1995). A negated version of that overstatement, (*He is not exceptionally bright*), is also rated as ironic, albeit to a lesser extent. Indeed, rather than eliminating the stance, the negation marker only tones it down. Less ironic than both is a version that involves both a negation marker and a nonoverstatement (*He is not bright*). In contrast, an approximate opposite of the overstatement (*He is stupid*) is rated as nonironic, because it involves no considerable gap between what is said and what is referred to (Experiments 1–2). These results are replicated with other modifiers such as “looks like” (Experiment 3). In addition, negated overstatements are recognizable as ironic even when no explicit context is specified (Experiment 4).

## THE PHENOMENON

Consider the following, naturally occurring (originally Hebrew) examples:

1. "He is a little...somewhat...**not exceptionally bright**" (US, 19.10.02).  
(When asked why he hesitated before finally describing the person in question as "not exceptionally bright," US said he had wanted to say that that person was an idiot but changed his mind and eventually rephrased it more politely.)
2. "DL [a sleazy businessman] **will not be elected chair of the Human Rights Association**" (AR, 2.7.02).
3. "...Indeed, we are a sort of a democracy, but the minority in this country (religious settlers, extremist rightists) dictates its will using methods that **aren't always the pinnacle of democracy ...**" (Sabina Zeidman, a letter to the editor, Ha'aretz Supplement, 11.10.02).

The intuition is that these negated utterances, although truthful, are interpreted ironically (and are referred to hereafter as *negative ironies*). They are certainly not the opposite of their affirmative, literally untruthful articulations (4–6), which probably make up more poignant instances of irony:

4. He is exceptionally bright (said of an idiot).
5. DL will be elected chair of the Human Rights Association (said of a sleazy businessman).
6. The pinnacle of democracy (said about a nondemocratic or fascist regime).

Consider, further, the following negative metaphors (from Hasson & Glucksberg, 2004, p. XX), which have indeed been found to also have an ironic reading (Giora, Aschkenazi, & Fein, 2004):

7. The Boston train is no rocket.
8. Some school teachers are not encyclopedias.

Irony's apparent resistance to negation effects seems to hold for other types of 'negation' such as when another 'negative' modifier ('small') is used (9) as opposed to its positive ('great') alternative (10). Here too, both articulations seem ironic:

9. This telephone is a small success (said on a newly bought telephone that isn't functioning, MA. 16.1.03).
10. This telephone is a great success (said on a newly bought telephone that isn't functioning).

In this article, we explore the impact of negation on degree of irony. We aim to show that, as predicted by the view of negation as mitigation (Giora, Balaban, et al., 2004), negation (1–3) would not sieve out the ironic stance projected by its affirmative equivalent (4–6). Rather, it would only hedge it. In Giora et al. (2004), we argued against a suppression view of negation, according to which deactivating the affirmative meaning of a negated constituent is obligatory. Note that ‘suppression’ pertains here to late processes (Gernsbacher, 1990), which deactivate meanings and responses that have been activated earlier. ‘Inhibition’ and ‘inhibitory’ effects pertain here to early processes that block access of meanings and responses.

According to the suppression hypothesis, under specific circumstances, negation reverses the meaning of a concept or a statement by shifting focus to a diametrically opposite alternative (e.g., Kaup & Zwaan, 2003; MacDonald & Just, 1989; Mayo, Schul, & Burnstein, in press). We proposed, instead, that, among other things, negation operates as a hedge, allowing features of the negated item to be retained so that the end product of the negated constituent is a toned down version of the affirmative rather than a total eradication of that affirmative (see also Clark & Clark, 1977; Giora, 1995; Horn, 1989). As done earlier (Giora, 2003; Giora & Fein, 1999b), and here too, we assume that suppression and retention are pragmatically motivated rather than operating automatically. Particularly, we argue that negation is often used as a mitigator rather than as a suppressor, whose end product is the opposite of the negated concept (for information on function-oriented suppression and retention, see Frost & Bentin, 1992; Giora, 2003; Giora & Fein, 1999).

### ON NEGATION AS MITIGATION

On viewing negation as mitigation (Giora et al., 2004), a negation marker is often an instruction from a communicator to an addressee to mitigate rather than eliminate the representation of the negated concept (see also Horn, 1989, pp. 236–240; for a similar but still different view, see Fraenkel, 2003). This should be particularly true when no denials or rejections are at stake, which require a specific context (Tottie, 1991).

According to the mitigation view of negation, ‘not warm’ communicates ‘lukewarm’ or ‘less than warm’ rather than ‘cold’—its available complement (see also Horn, 1989; for a different view, see Mayo et al., in press). For instance, when Bank of Israel governor, David Klein, said in an interview with *Ma’ariv* (an Israeli daily) that the collapse of a major bank in Israel “is not an imaginary scenario,” he was interpreted by *Ha’aretz* (an Israeli daily) as communicating that such a collapse was “possible” (31.12.02). The *Ha’aretz* headline opted for a mitigated interpretation of the negated constituent even though an opposite (‘realistic’) alternative was available. Indeed, according to the mitigation view of negation, negation need not reverse the meaning of a concept or a statement by shifting focus to an al-

ternative, diametrically opposite concept (see also Clark & Clark, 1977, p. 426). Often, negation transforms a statement to an understatement, because it hedges the negated concept.

In Giora et al. (2004), we ran four experiments that argued in favor of the mitigation hypothesis. Our studies showed, first, that negation did not have inhibitory effects: *not in not sharp* did not block access of the salient (coded and prominent) meaning of *sharp* (on meaning salience, see Giora, 1997, 2003). Rather, in a short ISI of 100 msec, ‘piercing’ was primed following both *This instrument is sharp* and *This instrument is not sharp* (Experiment 1; for similar findings see Hasson & Glucksberg, 2004, Experiments 1 and 2).

In addition, our studies showed that, following the initial access stage, salient meanings of negated concepts were not wiped out. Instead, they were retained and affected the ongoing discourse processing. For instance, lists including negated items behaved like lists of nonnegated items. Specifically, participants found that *What I bought yesterday was not a bottle but a jug* was acceptable whereas *What I bought yesterday was not a bottle but a closet* was not. Such results indicated that the acceptability of the next item on a list was sensitive to the affirmative meaning of the negated entity appearing previously on that list. They demonstrated that, at least, some features of negated items were preserved and affected the classifiability and accessibility of the next item in line (Experiment 2).

Such retained features might also induce a mitigated reading of the negated item. Indeed, our findings showed that negated items such as *not pretty* were distinguishable from their opposite (‘ugly’): They were perceived to be halfway between the polarities (‘pretty’–‘ugly’); that is, they received a ‘less than pretty’ interpretation. This was true regardless of whether they were unmarked (*pretty*) or marked (*ugly*) items. That is, in both cases, they were viewed as occupying a midposition on the (pretty–ugly) polarity scale. The ‘less than’ reading induced by negation, then, is a mitigation aimed toward the middle, neutral position on a scale (Experiment 3). Indeed, people do not treat affirmatives and their negated opposites as exchangeable. For instance, Holleman (2000) showed that, in public opinion surveys, respondents were more likely to answer “no” to questions containing the verb *forbid*, than “yes” to questions phrased with the verb *allow*, suggesting that people did not treat ‘not forbid’ as equivalent to ‘allow’ and vice versa.

No wonder speakers show sensitivity to the modifying effect of negation. When asked to describe an undesirable state of affairs (e.g., failing) politely or tactfully, participants showed a clear-cut preference for negated items (*not succeeding*) over their antonyms (*failing*). This applied across the board, regardless of whether the items were scalar (pretty–ugly) or nonscalar (succeed–fail), adjectives or nouns (for different findings, see Colston, 1999; Fraenkel, 2003). Such results are consistent with the view that negation does not suppress the positivity (‘pretty’) or negativity (‘ugly’) of the negated concept. Rather, affirmative sense of the negated

concept dilutes the negativity of the negation marker, resulting in a more positive or less negative account of an undesirable situation.

In this study, we aim to further test the mitigation hypothesis by looking into negative ironies.

### NEGATIVE IRONY

In previous studies, we proposed that irony hinges on some significant gap or contrast between what is said and what is referred to (Giora, 1995; Giora & Fein, 1999a; Giora, Fein, & Schwartz, 1998). The greater the gap or contrast, the easier it is to perceive the irony (Colston & O'Brien, 2000; Dukas, 1997; Gerrig & Goldvarg, 2000; Ivanko & Pexman, 2003; Utsumi, 2000). One can control for the gap between what is said and what is referred to by manipulating strength of context (as did Colston & O'Brien, 2000, and Ivanko & Pexman, 2003). In our studies, we affected degree of contrast by keeping the context constant while manipulating the strength of the targets, which ranged between opposite-ends-of-a-scale expressions (*stupid*–*exceptionally bright*).

Indeed, if context is kept constant and is about or supports one end of the scale (*idiot*), an affirmative overstatement such as *exceptionally bright*, which points to the opposite end of the scale (11a), will be rated as highly ironic. An extreme end of a scale expression will strongly bring out the difference between what is said and what is implied. Given the view of negation as mitigation, a negated version of such an overstatement (*not exceptionally bright*, 11b) would also be rated as ironic, albeit to a lesser extent, because, rather than eliminating the gap, the negation marker would only hedge it. Consequently, it will preserve some observable irony. Less ironic than both would be a version that involves a negated nonoverstated version of the affirmative overstatement (*not bright*, 11c). Such a statement would provide for a smaller gap between what is said (*not bright*) and what actually is (*idiot*). In contrast, some opposite of the affirmative (*stupid*, 11d) would be rated as nonironic, because it hardly involves any gap between what is said (*stupid*) and what is referred to (*idiot*). Thus, although (11a–c) would be rated as ironic, with (11a) being most ironic and (11c) being least ironic, (11d) would be rated as hardly ironic, that is, nonironic:

11. Although Max was working very hard preparing for his exams, he failed them all.
  - a. Max is exceptionally bright (affirmative overstatement).
  - b. Max is not exceptionally bright (negated overstatement).
  - c. Max is not bright (negated nonoverstatement).
  - d. Max is stupid (opposite of the affirmative).

Note that these predictions follow from both a view of irony as residing in some perceivable gap between what is said and what is implied and from a view of negation as mitigation. These predictions cannot follow from the traditional view of the standard pragmatic model (Grice, 1975), because these negative ironies are both truthful and not necessarily the opposite of what is said and, thus, not breaching the truthfulness maxim.

These predictions also do not follow from relevance theory (Sperber & Wilson, 1986/1995; Wilson & Sperber, 1992, 2004). According to Wilson and Sperber (2004, p. 622), “verbal irony consists in echoing a tacitly attributed thought or utterance with a tacitly dissociative attitude.” Given this implicitness assumption, negative ironies that involve *explicit* dissociative attitude cannot be accounted for by such a view. They do not seem to be captured by a view of irony as a variety of *implicit* echoic interpretive use, in which the communicator *tacitly* dissociates himself or herself from the opinion echoed (Curc2000 ,7; Wilson & Sperber, 1992). Although *exceptionally bright* in (4) can be viewed as an echoed opinion the speaker tacitly dissociates herself from, it is not quite clear what opinion the communicator implicitly echoes in *not exceptionally bright* (1) while *tacitly* dissociating herself from, unless she wishes to communicate ‘exceptionally bright’ which is not what (1) is all about. The negation marker (*not*) can, of course, be indicative of a dissociative attitude, but then this would make the attitude of dissociation explicit, which defies the relevance theoretic account of irony (Carston, 2002, p. 298).

The allusional pretense view (Kumon-Nakamura, Glucksberg & Brown, 1995) does not seem to fare better either. According to the allusional pretense view, irony involves insincerity and alludes to or reminds the addressee of what should have been—of an expectation or a norm that went wrong (Kreuz & Glucksberg, 1989). However, it is not clear what expectation negated concepts (*not exceptionally bright*) could allude to and whether there is any insincerity involved in negative ironies.

Can the joint pretense view (Clark, 1996; Clark & Carlson, 1982; Clark & Gerrig, 1984) come up with these predictions? According to the joint pretense view, the ironist pretends “to be an injudicious person speaking to an uninitiated audience; the speaker intends the addressee of the irony to discover the pretense and thereby see his or her attitude toward the speaker, the audience, and the utterance” (Clark & Gerrig, 1984 p. 12). Indeed, the speaker of (11a–c) can be taken to be ironic while pretending not to be. However, it is not clear that the joint pretense view can account for degree of ironiness. Alternatively, for it to predict the hierarchy assumed in (11a–c), it needs to further accept the view of negation as mitigation (Giora et al., 2004).

Negative ironies often get across as litotes or understatements. We claim here that this effect is due to the mitigating effect of negation. Note the following example, taken from *A Book About Death* (Du Bois, 2000, p. XX):

## ON NEGATION AS MITIGATION 87

12. PAMELA: ...(H) I mean I'm not,  
...I'm not,  
... I'm not all bent out of shape about [it].

Pamela is “not all bent out of shape” about *death* (the topic of the conversation). This negative irony, which, in fact, implicates that the speaker rejects death altogether, gets across as a form of understatement. We claim here that the understatement flavor of negative ironies is induced by negation. Negating *all bent*... triggers a ‘somewhat bent...’ implicature. Having derived this implicature, the comprehender now reasons that the speaker is understating her case, intending instead to convey that she is ‘not *at all* bent... and rather rejects death’. It is only by positing this mitigating, midstage process invited by negation that the understatement flavor of such examples can be accounted for (Mira Ariel, personal communication, January 31, 2004).

Using written Hebrew materials, we intended Experiments 1 and 2 to test the effect of negation on irony. If the anticipated hierarchy in (11a–d) is substantiated, this will support both the view of irony as residing in some considerable gap between what is said and what is implied (Giora, 1995) and to the view of negation as mitigation (Giora et al., 2004).

### EXPERIMENT 1

#### Method

*Participants.* Forty-eight undergraduates of Tel Aviv University served as volunteer participants.

*Materials.* Materials were 18 contexts such as (11) or (13), each followed by 1 of 3 target sentences (11a–c and 13a–c); in all, there were 54 target sentences. Three booklets were prepared, each containing 18 contexts and 1 target sentence. Each student saw all the contexts and 1 target sentence of the triplet presented in a random order. In addition, there were filler contexts that were quite entertaining.

13. Yossi and Roni were chatting during class. At a certain point the teacher got angry and said:

- a. You are very helpful (affirmative overstatement).
- b. You are not very helpful (negated overstatement).
- c. You are not helpful (negated nonoverstatement).

*Procedure.* Participants read the passages and were asked to rate each target (13a–13c) on a 7-point irony scale ranging from 1 (*nonironic*) to 7 (*highly ironic*).

*Results and discussion.* As demonstrated in the first row of Table 1, results obtained from subject (*t1*) and item (*t2*) analyses support the mitigation hypothesis. They show that, rather than eliminating the ironic stance, a negative irony sustains its irony. The negation marker did not do away with the ironic effect. Instead, it preserved it, albeit to a lesser extent. Thus, an affirmative overstatement (11a and 13a) was rated as most ironic, which is significantly more ironic than a negated overstatement,  $t2(17) = 15.09, p < .0001, t1(47) = 13.79, p < .0001$ , and a negated nonoverstatement,  $t2(17) = 16.98, p < .0001, t1(47) = 17.65, p < .0001$ . Next in irony was a negated overstatement (11b and 13b), which was rated as significantly more ironic,  $t1(17) = 3.75, p < .001, t2(47) = 6.78, p < .0001$ , than a negated nonoverstatement (11c and 13c), which was rated the least ironic.

Findings thus show that overstatements and their negated versions are both ironic—with the latter to a lesser extent. The view of negation as mitigation (Giora et al., 2004) can account for such results. Indeed, it predicts that negation of utterances that involve a considerable gap between what is said and what is referred to would not wipe out that gap and consequently would not wipe out the ironic stance derivable from that gap, but it would tone it down. Narrowing the gap via negation allows for these ironies to get across as understatements.

Although off-line measures such as those used here cannot be revealing about on-line processes such as suppression, the results, nonetheless, do not seem consistent with a view of negation as suppression. The view that a negation marker is an instruction from a speaker to an addressee to eliminate the negated concept from the mental representation (Kaup & Zwaan, 2003) predicts that the negated items tested here should not get across as ironies. Wiping out the negated concept should have resulted in no gap at all, and, therefore, in no ironic reading. However, the assumption that negation affects mitigation, thus preserving some gap between what

TABLE 1  
Mean Irony Ratings as a Function of Negation and Hedging  
in Experiments 1 Through 40

	<i>Affirmative Overstatement</i>	<i>Negated or Hedged Overstatement</i>	<i>Negated or Hedged Nonoverstatement</i>	<i>Opposite of Overstatement</i>
Experiment 1 (not)	6.02 (0.40)	3.57 (0.62)	2.81 (0.95)	
Experiment 2 (not)	5.65 (0.43)	3.36 (0.62)		1.73 (0.85)
Experiment 3 (looks like)	5.65 (1.26)	4.91 (1.14)	4.38 (1.44)	
Experiment 4 (not)	3.03 (1.14)	4.27 (1.38)	3.19 (1.26)	

*Note.* Standard deviations are shown in parentheses.



is said and the situation or opinion described, can account for the gradedness in irony found here.

In addition, this gradedness further argues in favor of the view of irony as involving a visible gap between what is said and what is referred to (Giora, 1995). It shows that explicitly negating an overstatement results only in narrowing the gap, thus allowing for an ironic stance to be derived. In contrast, these findings question the echoic mention view (Sperber & Wilson, 1986/1995; Wilson & Sperber, 1992), according to which such items should not be ironic because they make explicit the dissociative attitude. They also somewhat defy the Gricean and allusional pretense views according to which irony should either breach the truthfulness maxim or involve insincerity. (For a similar view regarding negative metaphors, see Carston, 2002, p. 345.)

## EXPERIMENT 2

The aim of Experiment 2 was to replicate the results of Experiment 1 using new materials and new participants, with the exception, however, of one target condition: Instead of a negated nonoverstatement, the targets now involved a statement that constituted an affirmative, sort of opposite of the overstatement. Thus, the items compared in Experiment 2 were an ironic overstatement (11a), a negated overstatement (11b), and an opposite of the overstatement (11d). The latter was assumed to be nonironic because it hardly involves any gap between what is said and the actual state of affairs.

### Method

*Participants.* Sixty graduates of Tel Aviv University (32 women and 28 men) between the ages of 21 and 40 served as volunteer participants.

*Materials.* Materials were 18 contexts such as (11) or (14) each followed by 1 of 3 target sentences (11a, 11b, and 11d; 14a–c); in all, there were 54 target sentences. Three booklets were prepared each containing 18 contexts and 1 target sentence. Each student saw all the contexts and 1 target sentence of the triplet presented in a random order. In addition, there were 6 filler contexts followed by targets that were related to their context via their literal interpretation.

14. Kineret, my friend, lives in the center of Tel Aviv. When I go to see her, I spend at least half an hour looking for a parking space.
  - a. It's the easiest thing on earth finding a parking space in Tel Aviv (affirmative overstatement).
  - b. It's not the easiest thing on earth finding a parking space in Tel Aviv (negated overstatement).

- c. It's an intricate thing finding a parking space in Tel Aviv (opposite of the overstatement).

*Procedure.* Participants read the passages and were asked to rate each target on a 7-point irony scale ranging from 1 (*nonironic*) to 7 (*highly ironic*).

*Results and discussion.* As demonstrated by the *second* row of Table 1, results obtained from subject (*t1*) and item (*t2*) analyses support the mitigation hypothesis. As in Experiment 1, here too a negated top-of-a scale constituent was rated as ironic, albeit to a lesser extent than its affirmative counterpart. Thus, affirmative overstatements (14a) were rated as most ironic, which were significantly more ironic than negated overstatements,  $t2(17) = 17.40, p < .0001, t1(59) = 10.05, p < .0001$ , and the opposite of the overstatements,  $t2(17) = 29.97, p < .0001, t1(59) = 15.35, p < .0001$ . Next in irony were negated overstatements (14b), which were rated as significantly more ironic,  $t2(17) = 8.71, p < .0001, t1(59) = 9.13, p < .0001$ , than the opposite of the overstatements (14c). The latter were rated as least ironic.

The view of negation as *mitigation* (Giora et al., 2004), which assumes that negation does not suppress the negated item but only hedges it, predicts that negating overstatements would not wipe out the ironic stance but only tone it down. The reduction in irony found for negated overstatements compared to affirmative overstatements was, therefore, anticipated. Similarly, the reduction in irony found for negated nonoverstatements compared to negated overstatements was also predicted. Compared to a (literal) statement that hardly exhibits any gap between what is said and what is referred to, all the negated versions were rated as more ironic, as predicted.

### EXPERIMENT 3

In Experiment 3, we intended to find further support of the mitigation hypothesis by comparing negation to other modifiers. If the gradedness in irony induced via negation (Experiments 1 and 2) can be replicated by the use of other mitigators, this would provide evidence in favor of the view of negation as mitigation. Therefore, in Experiment 3 we aimed to replicate the results of Experiments 1 and 2, but we replaced the negation marker with a hedge (e.g., 'looks like', see Caffi, 2001, p. 450; on other mitigating modifiers, see Caffi, 1990, 1999). We thus aimed to show that although *He is exceptionally bright* would be rated as most ironic, *Looks like he is exceptionally bright* would be rated as less ironic but still, however, more ironic than *Looks like he is bright*.

## Method

*Participants.* Forty-eight graduate students of Tel Aviv University (21 women and 27 men) between the ages of 21 and 38 served as volunteer participants.

*Materials.* As in *Experiment 1*, materials were 18 contexts such as (11 and 15) each followed by 1 of 3 target sentences (15a–c); in all, there were 54 target sentences.

15. Yossi and Roni were chatting during class. At a certain point the teacher got angry and said:

- a. You are very helpful (affirmative overstatement).
- b. Looks like you are very helpful (hedged overstatement).
- c. Looks like you are helpful (hedged nonoverstatement).

*Procedure.* As in *Experiment 1*, participants read the passages and were asked to rate each target on a 7-point irony scale ranging from 1 (*nonironic*) to 7 (*highly ironic*).

*Results and discussion.* As demonstrated by the third row of Table 1, results obtained from subject (*t1*) and item (*t2*) analyses support the mitigation hypothesis. They show that a hedge mitigates the ironic stance. Consequently, as in Experiments 1 and 2, here too a hedged top-of-a-scale expression (15b) was rated as ironic, albeit to a lesser extent than its affirmative counterpart (15a). Thus, affirmative overstatements (15a) were rated as most ironic and significantly more ironic than hedged overstatements,  $t1(47) = 3.73, p < .0005, t2(17) = 1.62, p = .06$ , and hedged statements,  $t1(47) = 5.40, p < .0001, t2(17) = 8.42, p < .0001$ . Next in irony were hedged overstatements (15b) that were rated as significantly more ironic statements,  $t1(47) = 4.74, p < .0001, t2(17) = 2.62, p < .01$ , than the hedged statements (15c), which were rated the least ironic.

As anticipated, these results replicate those *obtained* for negation of the same materials in Experiments 1 and 2, in which negated overstatements were perceived as ironic although less ironic than affirmative overstatements and more ironic than negated statements. Replicating the negation results with a hedge supports the view that negation is a modifier rather than a suppressor, which tone down expressions like other modifiers.

## EXPERIMENT 4

The aim of Experiment 4 was to show that, even when no explicit context is provided, negated overstatements (*He is not exceptionally bright*) get across as ironic.

The assumption is that although processing these utterances, addressees would recruit a possible, unmarked context, which, for these items, should be compatible with their ironic interpretation. Specifically, if a negation marker is indeed a mitigator, a negated overstatement should be viewed as an exaggeration that is toned down by the negation marker. However, given the availability of shorter options than a negated top-of-a-scale expression (such as ‘bright’), if all a speaker is after is some mitigation, using a longer form than necessary seems to flout the manner maxim (Grice, 1975) and invite an implicature<sup>1</sup>. A negative overstatement would thus indicate a sizable gap between what is said and what is referred to, inviting an ironic interpretation of the utterance even when no explicit context is provided. On the view of negation as mitigation, unlike affirmative top-of-a-scale expressions, a negation of an extremely strong expression is suspect. It should thus alert the addressee to a nonliteral interpretation, which, in this case, would tend to implicate an approximate low end of the scale.

Consider other examples that come to mind involving different negative markers (e.g., ‘short of,’ ‘shy of,’ ‘missing,’ ‘minus,’ and ‘without’), which metaphorically ironize the victim of the irony even in the absence of a specific context (on other negation markers, see Israel, in press):

16.
  - a. One chapter short of a novel.
  - b. A few tiles missing from his space shuttle.
  - c. A few birds shy of a flock.
  - d. A violin minus the bow.
  - e. Left the store without all of his groceries.

Negative ironies cannot be captured by the relevance theoretic account. Given that negation is an overt marker of dissociation, an account based on implicit dissociation cannot explain such an irony. In contrast, negative utterances that negate a top-of-the-scale constituent can be accounted for by the view of irony as highlighting a gap between what is said and what is (or could be) alluded to.

Given the view of negation as mitigation, we anticipated that negated overstatements would be rated as more ironic than either their affirmative version (*He is exceptionally bright*) or their negated, non-top-of-a-scale variation (*He is not bright*).

---

<sup>1</sup> Negated expressions could also be viewed as flouting the informativeness requirement (Grice, 1975; for information regarding how negation fails to be informative enough, see Glenberg, Robertson, Jansen, & Johnson-Glenberg, 1999). For instance, a toned down top-of-a-scale expression such as *not exceptionally bright* seems to be underspecific. Although the speaker could be sufficiently informative as to whether the referent in question is either ‘bright,’ ‘mediocre,’ or ‘stupid,’ she chose not to.

## ON NEGATION AS MITIGATION 93

### Method

*Participants.* Forty-eight undergraduates of Tel Aviv University served as volunteer participants.

*Materials.* Materials were the same 54 target sentences used in Experiment 1. However, this time they were presented without their prior contexts. As in Experiment 1, each student saw only one target of each triplet previously associated with a specific context:

17.
  - a. He is exceptionally bright (affirmative overstatement).
  - b. He is not exceptionally bright (negated overstatement).
  - c. He is not bright (negated nonoverstatement).

*Procedure.* As in Experiment 1, participants read the targets (17a–17c) and were asked to rate each of them on a 7-point irony scale ranging from 1 (*nonironic*) to 7 (*highly ironic*).

*Results and discussion.* As demonstrated by the fourth row of Table 1, results obtained from subject (*t1*) and item (*t2*) analyses support the mitigation hypothesis (Giora et al., 2004) and the indirect negation view of irony (Giora, 1995). They show that, as predicted, rather than eliminating the negated concept, a negation marker only hedges it. Hedging a top-of-the-scale constituent when a shorter hedged alternative is available breaches the manner maxim (Grice, 1975) and calls for an ironic interpretation. Consequently, when a top-of-the-scale concept was negated (17b), ratings were highest—higher than affirmative overstatements (17a),  $t2(17) = 4.92, p < .0001, t1(47) = 5.30, p < .0001$ , and higher than negated statements (17c),  $t2(17) = 7.82, p < .0001, t1(46) = 7.77, p < .0001$ . Negated statements and affirmative overstatements did not differ significantly,  $t2(17) = 0.77, p = .23, t1(47) = 0.88, p = .19$ .

From an explicit context, items involving a negated exaggeration were rated as more ironic than items that did not involve such a top-of-the-scale constituent marked for negation. Although compared to their ratings in Experiment 1, they achieved higher irony ratings, this might be a result of the much less ironic competitors in the out-of-context condition. According to the view of negation as mitigation, however, negating such a strong expression should tone it down. Thus, given that other, shorter, or less strong alternatives are available, such a toned down expression would be suspect and would invite a nonliteral reading, implicating the almost opposite pole of the specified scale.

Indeed, indirect negation differs from explicit negation. Explicit negation modifies the negated concept, aiming toward the neutral, midpoint position on the scale

(see Giora et al. 2004, Experiment 3). Irony, which is a form of indirect negation, fleshes out the large gap between what is said and what is referred to. Although in a biasing context condition, affirmative overstatements could be more ironic than their negated versions (see Experiments 1 and 2), in the out-of-context condition, they do not invite such an interpretation, because the affirmatives involve no overt violation or implicit gap and thus need not trigger any such implicature.

Relevance theory (Sperber & Wilson 1986/1995; Wilson & Sperber, 1992, 2004) cannot account for the aforementioned results because they suggest that ironiness can be induced by an explicit dissociative attitude (indicated by the negation marker). In contrast, viewing negation as mitigation (while assuming rule violation, see Grice, 1975) can account for why, even out of an explicit context, negated top-of-a-scale utterances can have an ironic reading.

## GENERAL DISCUSSION

In Giora et al. (2004), we showed that a negation marker basically keeps intact the concept it negates; whereas when negating the concept, a negation marker does not discard it from the mental representation but only mitigates it. In this study, we pursue this line of research further, testing the mitigation hypothesis with respect to irony. Irony might be particularly amenable to mitigation effects because it hinges on a substantial gap between what is said and the reality referred to (Giora, 1995). If negation is indeed a mitigator, it will hedge the sense of the negated concept ('exceptional brightness' in 11a) and affect some narrowing of the gap in question while preserving it. The result should be a toned down irony (compared to the affirmative, nonmitigated overstatement). The mitigation hypothesis thus predicts that, when keeping the context constant, an affirmative top-of-a-scale expression (*exceptionally bright*) will be evaluated as highly ironic, because the gap between what is said and the reality referred to is considerable. Second in degree of ironicity will be a negated version of that overstatement (*not exceptionally bright*), which also sustains a noticeable gap between what is said and what is referred to. Third in ironicity will be a negated, less than a non-top-of-the scale expression (*not bright*), because even such an expression might allude to some noticeable gap between what is said and the reality described. A close opposite of the affirmative overstatement (*idiot*) will be evaluated as nonironic, because it hardly sustains any noticeable gap between what is said and what is referred to.

Findings from Experiments 1 and 2 support the hierarchy predicted by the view of irony as hinging on a substantial gap between what is said and what is criticized (Giora, 1995) and by the view of negation as mitigation (Giora et al., 2004). They show that, as predicted, in an irony-inviting context, an overstatement (11a and 14a) is most ironic, apparently because it involves the widest gap. Its negated articulation (11b and 14b) is a mitigated version of it, an understatement, which, given

the same context, nonetheless retains the ironic stance (for similar findings, see Colston, 1997, Experiment 2). Even a negated version of a nonoverstatement (11c) is still somewhat ironic—even more ironic than an expression that does not obviously exhibit a gap between what is said and what is referred to (11d and 14c). Admittedly, because the findings were obtained in two different experiments, the latter conclusion is not directly warranted. Still, on the basis of the similar ratings found in the two experiments regarding both ironic overstatements (6.02 and 5.65), on the one hand, and their negated versions (3.75 and 3.36), on the other, it is safe to conclude that the difference found between the negated nonoverstatements (2.81) and the affirmative opposites (1.73) is indeed considerable.

In all, our findings exhibit some irony gradedness that supports the view that negation induces mitigation. Indeed, these results have been replicated in other languages such as Russian and French (Chicheportiche & Rabits, 2003). Although our studies do not involve on-line measures, the results, nonetheless, seem hardly consistent with the view of negation as suppression. If indeed suppression had been involved (as might be deduced from, e.g., MacDonald & Just, 1989; Mayo et al., in press.), the gap between what is said and what is described should have been diminished entirely, resulting in nonironic reading of the items in question, which was not the case. The mitigation hypothesis gains further support from the results of Experiment 3, which show that negation is no different from a hedge. Thus, when the negation marker used in Experiments 1 and 2 was replaced by a conventional mitigator ('apparently' or 'looks like'), the irony gradedness found in Experiments 1 and 2 was replicated.

Note that, contrary to appearances, the view of negation as mitigation does not assume that negative ironies (*He is not exceptionally bright*) are always pragmatically derivative (although they might be syntactically and semantically derivative). That is, there is no need to assume that their ironiness is a result of the ironiness of their affirmatives. Comparing affirmative and negative ironies here only served as a means to make a point about negation as mitigation. It is quite plausible to assume that negative top-of-a-scale utterances would give rise to an ironic reading regardless of whether the affirmative counterparts do. Consider, for instance, the negative metaphoric statements in Hasson and Glucksberg's (2004) studies, which, we contend, lend themselves to an ironic reading (Giora, Aschkenazi, et al., 2004). These items involve negated top-of-the scale concepts, whereas their affirmative counterparts do not. Thus, although *The Boston train is no rocket* invites an ironic reading even when no specific context is given, its affirmative version, *The Boston train is a rocket*, will invite an ironic reading only in a context in which it is literally inappropriate (e.g., when the train is slow). Out of a biasing context, however, the affirmative version does not invite an ironic interpretation. Although negative ironies might be syntactic and even semantic derivatives, pragmatically they need not derive from their affirmative counterparts.

Indeed, this intuition gained support in Experiment 4, where, in the absence of a specific context, negated overstatements (*He is not exceptionally bright*) were rated as significantly more ironic than their affirmative counterparts (*He is exceptionally bright*). In Experiment 4, this was viewed as a result of the inappropriateness of using a negation marker to tone down a top-of-a-scale expression when shorter alternatives (*bright*) were available. In the absence of an explicit context, such inappropriateness may trigger an ironic interpretation.

Indeed, for top-of-a-scale expressions to be read nonironically, they require a specific context either where they would be interpreted as a rejection of a specific assumption or where a literal description is invited. For instance, in a context in which an interlocutor states “He is exceptionally bright” and the addressee begs to differ, he might reject this assumption by saying “He is not exceptionally bright” and mean it literally. Or, consider a context that makes it clear that the oncoming top-of-a-scale expression (shown here in bold) is to be interpreted literally:

18. <1682> [Well and she] doesn’t seem to be a person that has, she doesn’t come from wealth per se, and I said, well did you work? <mimicking> Well, yeah, I worked in education</mimicking> and then she worked as a social worker.  
<1691> Mhm  
<1682> **And those don’t make exceptionally good incomes.** I mean it’s [[not, maybe fifty thousand]] a year (J. Du Bois, P.C., retrieved from *The Longman Spoken American Corpus*, 2003 on 5.3.04).

Other than in such biasing contexts, negated top-of-a-scale expressions would most probably be interpreted as ironic.

The evidence brought here to bear on negation as mitigation also has implications with regard to the view of irony as sustaining some significant difference between what is said and what is referred to (Giora, 1995). The question is whether these findings are also accountable by the echoic mention view of irony (Sperber & Wilson, 1986/1995; Wilson & Sperber, 1992).

Some negative ironies are indeed echoic: They echo a negated phrase while dissociating from it. Consider the following examples (in bold) in the following text:

19. Where Have All the Nuances Gone? (original Hebrew title: Axis of European Evil) By Adar Primor  
These are difficult times for “old Europe’s” diplomats in Israel. Their foreign ministries read reports saying that in Israeli eyes, the French are hypocritical Israel-haters, the “great whores” of the international arena (Ma’ariv). The Belgians are a vain people with a record of colonialist crimes, anti-Semites from birth who, by the way, brought pedophilia to the



world. The Germans? Well, there's no need to even mention them—the history books say it all.

We've attached labels to everyone. Not only the French and Germans and Belgians. We've labeled the Flemish, Dutch-speaking Belgians, and the Walloons, the French speakers. And we haven't forgotten the British, either. They're good guys. **It doesn't matter** that more than a million of them demonstrated in the streets against going to war. **It doesn't matter** that 55 percent of them reject their prime minister's policies, while only 35 percent support him. *We won't let the facts confuse us*. The English speak English, after all. They were also the good guys in World War II, and despite the Mandate, they're kosher in our eyes nowadays, whether they like it or are embarrassed that we think so. (Haaretz.com, English edition, February 24, 2003, p. XX)

Indeed, it is possible to assume that Israelis could entertain the thought (echoed in the *It doesn't matter* phrases) that, regardless of some facts about the British people, the British should still be favored, because their government is on our side. Moreover, *It doesn't matter*... can be a thought to be attributable to people even though it is negative (and, therefore, nonnormative according to the relevance theoretic account), because it is a conventional phrase and way of thinking. Consequently, echoing such negated statements while dissociating from them can still be considered ironic & la relevance theory account.

Some negative ironies, however, seem to defy the echoic mention view. For instance, it is implausible to assume that *We won't let the facts confuse us* could be a thought attributable to anyone; no one in her or his right mind would entertain such a thought about herself or himself, not even the Israelis who the journalist is criticizing here<sup>2</sup>. In addition, this phrase is conventionally used ironically. Even for this reason, it is hard to see how this specific thought can be attributed to anyone while disengaging from it.

---

<sup>2</sup> A near counterexample to the negative irony discussed here has been brought to our attention by Arnon Kehat (personal communication, March 10, 2003), suggesting that some people can be thought of as entertaining a close associate to *We won't let the facts confuse us* and mean it literally. In (13), which is an excerpt from *The Simpsons* ("Lisa the Skeptic," season 9, episode 8; <http://www.snpp.com/episodes/5F05>), Homer responds to Lisa who is trying to test a theory by looking at the facts:

Lisa: I took a piece of the skeleton for scientific analysis, soon we'll have all the facts.

Crowd: You did what?! [mutterings]

Homer: Facts are meaningless. You can use facts to prove anything that's even remotely true.

Facts schmacts ...

It would thus take a simpleton such as Homer to admit that one can ignore facts when coming to test theories.

Note further that it is not the case that irony cannot be echoed with a dissociative attitude. The following example is a case in point. It shows that one can be ironic by dissociating herself from an explicitly echoed irony while endorsing what she says, that is, intending it literally (for a different view, see Curc2000 ,5):

20. I want to go for a picnic. My friend G does not. I convince him, and we drive out. It's very cloudy and windy out. My friend says, mockingly: "What a lovely day for a picnic!" By the time we get to our destination, the weather changes. It is now an extraordinarily beautiful day. I turn to my friend and say in a mocking tone: **"What a lovely day for a picnic!"** (Ronnie Brosh, personal communication, March 14, 2003)

Along the same lines, it is hard to see how the ironic reading assigned to the negative overstatements used in our studies can be accounted for by the echoic mention view. The only condition under which these ironies can be viewed as echoic—that is, as attributable to another speaker or state of mind—is when there is an explicit prior mention of them, which the speaker can dissociate from while endorsing some aspects of their affirmative meaning. For instance, in a context when one has mentioned that a certain person "is not exceptionally bright" and when later on it becomes clear that that person is doing much better than assumed, it is possible to dissociate from this utterance by a negative irony: "he is not exceptionally bright, ha?". However, there is no explicit mention of these negated utterances in our contexts. In addition, being negative reduces the possibility that these utterances can make up an implicit echo that the speaker is tacitly dissociating herself from. For these reasons, it is hard to see how the echoic mention view can account for their observed irony. Recall further that, in our contexts (Experiments 1 and 2), these ironies do not implicate some affirmative meaning of the negated utterances as they would if they were echoic. Rather, they convey a mitigated interpretation of their nonnegated ironic version.

In sum, because a negation marker serves to mitigate a statement rather than imply its opposite, negative ironies differ from affirmative ironies only quantitatively. Qualitatively, however, they are the same.

## ACKNOWLEDGMENTS

The studies reported here were supported by a grant to Rachel Giora by Adams Super Center for Brian, Tel Aviv University. We thank Mira Ariel, Ronnie Brosh, Claudia Caffi, Herb Clark, Jack Du Bois, Uri Hasson, Art Graesser, Carita Paradis, Noa Shual, Willibald Ruch, Andrew Wedel, Dana Zimmerman, and two anonymous reviewers for very helpful comments on an earlier version of this article.

## REFERENCES

- Caffi, C. (1990). Modulazione, mitigazione e litote. In M.-E. Conte, P. Ramat, & A. Gialone Ramat (Eds.), *Dimensioni della linguistica* (pp.169–199). Milan, Italy: Angeli.
- Caffi, C. (1999). On mitigation. *Journal of Pragmatics*, 31, 881–909.
- Caffi, C. (2001). *La mitigazione: Un approccio pragmatico alla comunicazione nei contesti terapeutici*. Münster: Lit.
- Carston, R. (2002). *Thoughts and utterances*. Oxford, England: Blackwell.
- Chicheportiche, S., & Rabits, R. (2003). *Negation as mitigation: The case of Russian and French*. Unpublished manuscript, Tel Aviv University.
- Clark, H. H., & Clark, V. E. (1977). *Psychology and language: An introduction to psycholinguistics*. San Diego: Harcourt Brace Jovanovich.
- Clark, H. H., & Gerrig, R. (1984). On the pretense theory of irony. *Journal of Experimental Psychology: General*, 113, 121–126.
- Colston, H. L. (1997). Salting a wound or sugaring a pill: The pragmatic functions of ironic criticism. *Discourse Processes*, 23, 25–45.
- Colston, H. L. (1999). "Not good" is "bad," but "not bad" is not "good": An analysis of three accounts of negation asymmetry. *Discourse Processes*, 28, 237–256.
- Colston, H. L., & O'Brien, J. (2000). Contrast and pragmatics in figurative language: Anything understatement can do, irony can do better. *Journal of Pragmatics*, 32, 1557–1583.
- Curc , C. (2000). Irony: negation, echo and metarepresentation. *Lingua*, 110, 257–280.
- Dukas, G. (1997). *Irony aptness: The case of visual and cinematic discourse*. Unpublished master's thesis, Tel Aviv University.
- Fraenkel, T. (2003). *The meaning of negated adjectives and the preservation of the negation in communication*. Doctoral dissertation, manuscript in preparation.
- Frost, R., & Bentin, S. (1992). Processing phonological and semantic ambiguity: Evidence from semantic priming at different SOAs. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 58–68.
- Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Gerrig, R. J., & Goldvarg, Y. (2000). Additive effects in the perception of sarcasm: Situational disparity and echoic mention. *Metaphor and Symbol*, 15, 197–208.
- Giora, R. (1995). On irony and negation. *Discourse Processes*, 19, 239–264.
- Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, 7, 183–206.
- Giora, R. (2003). *On our mind: Salience, context and figurative language*. New York: Oxford University Press.
- Giora, R., Aschkenazi, K., & Fein, O. (2004). *Suppression is not obligatory but functional: On negative metaphors*. Manuscript in preparation.
- Giora, R., Balaban, N., Fein, O., & Alkabets, I. (2004). Negation as positivity in disguise. In H. L. Colston, & A. Katz (Eds.), *Figurative language comprehension: Social and cultural influences* (pp. 233–258). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Giora, R., & Fein, O. (1999a). Irony: Context and salience. *Metaphor and Symbol*, 14, 241–257.
- Giora, R., & Fein, O. (1999b). On understanding familiar and less-familiar figurative language. *Journal of Pragmatics*, 31, 1601–1618.
- Giora, R., Fein O., & Schwartz, T. (1998). Irony: Graded salience and indirect negation. *Metaphor and Symbol*, 13, 83–101.
- Glenberg, M. A., Robertson, D. A., Jansen, J. L., & Johnson-Glenberg, M. C. (1999). Not propositions. *Journal of Cognitive Systems Research*, 1, 19–33.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Speech acts: Syntax and semantics* (Vol. 3, pp. 41–58). New York: Academic.

- Hasson, U., & Glucksberg, S. (2004). *Does understanding negation entail affirmation?* Manuscript submitted for publication.
- Holleman, B. (2000). *The forbid/allow asymmetry: On the cognitive mechanisms underlying wording effects in surveys*. Amsterdam/Atlanta, GA: Rodopi.
- Horn, L. R. (1989). *A natural history of negation*. Chicago: The University of Chicago Press.
- Israel, M. (in press). *The rhetoric of grammar: Scalar reasoning and polarity sensitivity*. Cambridge, England: Cambridge University Press.
- Ivanko, L. S., & Pexman, M. P. (2003). Context incongruity and irony processing. *Discourse Processes*, 35, 241–279.
- Kaup, B., & Zwaan, R. A. (2003). Effects of negation and situational presence on the accessibility of text information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 439–446.
- MacDonald, M. C., & Just, M. A. (1989). Changes in activation levels with negation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15, 633–642.
- Mayo, R., Schul, Y., & Burnstein, E. (in press). "I am not guilty" versus "I am innocent": The associative structure activated in processing negations. *Journal of Experimental Social Psychology*.
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition*. Oxford, England: Blackwell. (Original work published 1986)
- Tottie, G. (1991). *Negation in English speech and writing: A study in variation*. San Diego: Academic.
- Utsumi, A. (2000). Verbal irony as implicit display of ironic environment: Distinguishing ironic utterances from nonirony. *Journal of Pragmatics*, 32, 1777–1806.
- Wilson, D., & Sperber, D. (1992). On verbal irony. *Lingua*, 87, 53–76.
- Wilson, D., & Sperber, D. (2004). Relevance theory. In G. Ward & L. Horn (Eds.), *Handbook of pragmatics* (pp. 607–632). Oxford, England: Blackwell.