1 Introduction

In their position paper, Veale, Feyaerts & Forceville (this volume) argue that creativity can be found in duality. They analyze different examples of creative duality in the domains of word, image and sound. This paper further expands on their contribution by looking at the creative ways in which two of these domains – the verbal and the visual – can be combined to create meaning. As a case in point, this chapter focuses on the concept of verbal irony. Verbal irony may be particularly suited for this analysis, because an ironist always has two meanings to his or her disposal: the literal and the intended evaluation of the ironic remark. In fact, some of the examples discussed by Veale et al. (this volume) in the light of creative duality such as the gestalt-switching in the image of the hare and the duck (see Figure 1 in Veale et al. [this volume]) have also been discussed in the context of irony (see Hutcheon [1994]). Therefore, this chapter presents an exploratory corpus study that aims to investigate how creative duality can be found in images related to ironic remarks and how images can serve to make an ironic utterance more creative. Some scholars refer to this kind of image as a visual marker of verbal irony (see Attardo, Eisterhold, Hay & Poggi [2003]).

In order to investigate how these visual irony markers may work, one approach could be to investigate the rich literature on verbal markers of irony (e.g., Attardo [2000a], Kreuz [1996], Seto [1998]). These authors have identified a wide array of verbal irony markers ranging from tropes such as hyperbole (e.g., Kreuz & Roberts [1995]) and rhetorical questions (e.g., Muecke [1978]) to topographic signals such as the use of quotation marks (e.g., Attardo [2001]) or emoticons (e.g., Kreuz [1996]). A simple solution to the question which visual markers can be identified would be to transplant these verbal markers directly to the visual modality. Some scholars would agree with this notion and argue that nothing special separates visual from verbal communication. McQuarrie & Mick (2003) for instance claim that:

“there is nothing special about the visual modality. Pictures can be signs, and semiotic theory explains the communicative function of pictures using the same constructs that explain the function of words. [...] In short, at the level of scientific theory, whether something is visual or verbal, pictorial or auditory, may be of little consequence” (McQuarrie & Mick [2003: 215–216])
At a first glance, this observation seems to hold for the irony markers identified by authors such as Attardo (2000a), Kreuz (1996) and Seto (1998). After all, markers such as hyperboles may be transplanted to the visual mode (see Schilperoord & Maes, 2010) and could thus also work as visual irony markers (cf., El Refaie [2005]). The same may be true for the use of quotation marks. In the visual domain, it is possible to use “air quotes” to visually represent the use of quotation marks (cf. Muecke [1978]). Other irony markers, however, do not fare as well. How is it possible to visually represent a rhetorical question? Since these issues are hard to solve, this chapter takes a different perspective.

In a later article, Phillips & McQuarrie (2004) argued against McQuarrie & Mick (2003) and claimed that “[b]ecause pictures are not speech, we shall argue that existing taxonomies designed for verbal rhetorical figures (...) do not adequately capture important differentiations within the visual domain” (Phillips & McQuarrie [2004: 114]). This means that images may help a reader in detecting irony in different ways from verbal texts. This claim can be well illustrated by the fact that verbal irony itself cannot be completely transplanted to the visual modality (e.g., Hornikx & Van Mulken [2004], Kennedy [2008]). Kennedy (2008: 458) argues that

“irony in pictures is surely rare. (...) In language, the irony is often accompanied by a special tone of voice, but alas, no manner of portrayal has yet been invented that is the tip-off for pictorial irony. New Yorker cartoons are often ironic, but their irony lives in a caption’s fit to the picture”.

Of course, the question if and how verbal irony itself can be transplanted to the visual modality also depends on the definition of verbal irony that is used. The definition of verbal irony is a difficult matter that has led to much scholarly debate (cf. Attardo, 2000b). In another article, we compared definitions of irony from different scholarly traditions such as a Gricean and (Neo-)Gricean perspective (e.g., Grice [1978], Attardo [2000b]), speech act theory (e.g., Haverkate [1990]), relevance theory (e.g., Wilson & Sperber [2002]), pretense theory (e.g., Clark & Gerrig [1984]; Currie [2006]), mental space theory (e.g., Coulson [2005]; Kihara [2005]) and the view of irony as indirect negation (e.g., Giora [1995]). We then analyzed these definitions and looked at the aspects these definitions had in common. Based on these aspects, we defined irony as “an evaluative utterance, the valence of which is implicitly reversed between the literal and intended evaluation” (Burgers, Van Mulken & Schellens [2011]).
This definition implies that if the literal evaluation of the ironic utterance is positive, the intended evaluation is negative (or the other way round). As such, the valence of the literal evaluation of the ironic remark (positive or negative) is reversed in the intended evaluation of the ironic remark (negative or positive). This definition also shows that it is difficult to portray irony solely in the visual domain. A purely visual irony would require an image to implicitly negate its own literal evaluation. Following Kennedy (2008), it may be hypothesized that the image itself is not ironic, but it may sometimes be important in deciding whether a verbal utterance is ironic or not. An image can thus help a reader in detecting verbal irony.

The issue how an image serves to detect irony has received scant scholarly attention. Most scholars who looked at the relationship between verbal irony and images limited themselves to studying visual markers (e.g., Attardo et al. [2003], El Refaie [2005], Rockwell [2001]). A visual marker is a marker that is included in a text outside of the ironic utterance under discussion. A visual marker may thus work in the same way as a marker in the verbal co-text (cf. Burgers, Van Mulken & Schellens [in press]), which implies that a visual marker
sets up a co-textual support strategy that may alert a reader to the possible use of irony. In doing so, the literal evaluation is displayed and mocked at the same time, an example of which can be seen in the Dutch Lotto advertisement in Figure 1:

(1) **Lotto: the greatest risk of becoming a millionaire as well.**

Utterance (1) is ironic, because participation in the Lotto lottery is literally portrayed as something negative. Lotto’s contestants literally run the risk of becoming millionaires. Of course, anyone who is willing to participate in such a lottery is glad to run such a risk; they only dream of winning the jackpot. The advertisement’s image, however, illustrates the supposed risk of participation: an astronaut – representing somebody who won Lotto’s jackpot – goes to the moon to plant the Dutch flag, only to discover that many Dutchmen had been there before. The reader should then infer that these other Dutch visitors were previous Lotto winners who also went to the moon. The astronaut’s mission thus ends in disappointment. The image of Figure 1 serves to show one of the imagined risks of becoming a millionaire. At the same time, people who view this advertisement are supposed to recognize the image as a hyperbole. As of 2012, only two Dutchmen (Wubbo Ockels and André Kuipers) have ever gone into space, none of whom had landed on the moon. It is thus up to the reader to recognize the supposed risk of becoming a millionaire as ludicrous, opening up an ironic reading of utterance (1). In this utterance, the image serves as an irony marker by illustrating the literal evaluation of an ironic utterance while showing the absurdity of that literal evaluation at the same time.

Besides serving as an irony marker by illustrating a literal meaning, a visual may also exhibit incongruence with the literal evaluation of an ironic utterance. In other words, an image may display information that does not endorse a literal evaluation of the ironic utterance. An example of this strategy is seen in an advertisement for the mobile phone provider Libertel in Figure 2:

(2.1) **Sure, hedge-clippers are nice.**
(2.2) But it so difficult to call someone with them.
(2.3) The nicest gifts are iZi: calling on your mobile without a pay-monthly plan.
Utterance (2.1) is ironic; the man in this Libertel advertisement is not happy at all with the gift of hedge-clippers he just received. Instead, he would have preferred a mobile phone with a pay-as-you-go plan from mobile phone provider Libertel. In this advertisement, the man explicitly shows his displeasure with receiving hedge-clippers instead of a mobile phone. This implies that the image in Figure 2 highlights the incongruence between the literal and intended evaluation of the ironic utterance.

Of course, it should be noted that not every image in an advertisement or a cartoon with an ironic utterance automatically aids the addressee in finding the irony. An example of an advertisement in which the image does not help can be found in Figure 3:

(3.1) **Shameful!**
(3.2) Now only 1 Euro.
(3.3) Computer Idee, does not make computers difficult
Figure 3 contains an advertisement for the Dutch computer magazine *Computer Idee*. The advertisement has the ironic tagline of “Shameful” to advertise the magazine’s low price of 1 Euro. The image in the advertisement shows a cover of one of the issues of the magazine, which does not help the reader to arrive at an ironic interpretation of the advertisement’s tagline. The image does not help in detecting the irony.

Summarizing, if an ironic utterance is accompanied by an image, this particular image can have three relations to that ironic utterance. Firstly, the image may not help a reader in detecting the irony at all, like the image in Figure 3. If, however, an image does help a reader in detecting irony, this may happen in two distinct ways. A first way in which an image can help in detecting the irony is by serving as a visual irony marker by illustrating the literal meaning (see Figure 1). This means that an image highlights and shows the absurdity of the literal evaluation of an ironic utterance at the same time. A second way in which an image can help a reader in detecting the irony is related to incongruence with the literal evaluation of the irony: an image contradicts the literal evaluation of the ironic utterance (see Figure 2). For reader convenience, the first way in which an image can help to detect irony is labeled as a “visual marker”, while the second way is
labeled as a “visually incongruent image” from this point on. This also brings us to the first research question of this chapter:

RQ1. How are different types of visuals used in relation to irony in written discourse?

In answering this question, this chapter investigates which type of visual (i.e., a visual that does not help, a visual marker or an incongruent image) is most common in multimodal genres in relation to verbal irony.

Previous authors who have dealt with the relationship between visuals and verbal irony did not differentiate between visual markers and incongruent images. Instead, these scholars typically analyze pictorial elements that help a reader in detecting irony by focusing on one of these elements. Attardo et al. (2003), for instance, investigate the way in which facial expressions can help a reader to detect irony. Based on an analysis of humorous material from American sitcoms, Attardo et al. (2003) demonstrate how inappropriate facial expressions can serve to accomplish this goal. If a sitcom character for instance winks or smiles while making an ironic remark, a viewer may see the wink or smile as a clue that the speaker is ironic. At the same time, a blank expression (referred to as “blank face” by Attardo et al. [2003]) while uttering an (explicitly) evaluative ironic remark may also help in detecting the irony, because the absence of a facial emotion implicitly negates much of the literal evaluation.

Rockwell (2001) analyzed the facial expressions of respondents who used irony in a conversation. She divided the face into three areas (i.e., (1) the eyebrows and forehead, (2) the eyes, lids and upper part of the nose and (3) the mouths, cheeks, lower part of the nose, chin and jaw) and asked two coders to analyze “intentional movements” in these three facial areas when a speaker made an ironic remark. She found that only the mouth area (area 3) differed significantly between irony and non-irony; irony speakers made “mouth movements” to distinguish between irony and non-irony. Unfortunately, Rockwell (2001) does not elaborate on what these specific mouth movements entail. She also does not explain how it is possible to distinguish an intentional from an unintentional facial movement.

Like Rockwell (2001), Caucci & Kreuz (2012) investigated the facial cues that were used by ironic speakers in spontaneous conversation. They also found that movements of the mouth like tip tightens could indicate ironic intent. Furthermore, they also found empirical evidence that movements of the head (e.g., slow nods) were more associated with irony than with literal language.

In contrast to Attardo et al. (2003), Rockwell (2001) and Caucci & Kreuz (2012), El Refaie (2005) focuses on newspaper cartoons rather than face-to-face interactions. The case study of irony discussed by El Refaie (2005) shows that
visual hyperbole can be considered as a marker of verbal irony. This also demonstrates that irony markers in the verbal and visual domain can sometimes overlap; hyperbole can be a marker in the ironic utterance (e.g., Burgers, Van Mulken & Schellens [2012]) and an irony marker in the verbal (Burgers et al. [in press]) and visual co-text (El Refaie [2005]).

The studies conducted by Attardo et al. (2003), Rockwell (2001), Caucci & Kreuz (2012) and El Refaie (2005) analyzed the ways in which visuals may help a reader in detecting irony by focusing on specific pictorial elements. Although their approach has garnered interesting insights on how these particular pictorial elements help a reader in detecting verbal irony, their approach is still a bit ad hoc; a specific pictorial element was brought to the attention of researchers who decide to investigate it. The question remains which other pictorial elements that help a reader to detect irony may be distinguished. To try to capture this phenomenon as broadly as possible, we use a bottom-up approach in which all elements from the visual domain are analyzed separately to see if they help a reader to detect irony or not.

The “elements from the visual domain” in the bottom-up approach are derived from Verstraten’s (2006) visual narratology, an extension of Bal’s (1997) theory of narratology to the visual realm. Verstraten’s (2006) visual narratology shows a great overlap with other visual “grammars” (e.g., Bordwell & Thompson [2004], W. Phillips [2002]). According to Verstraten (2006), two types of elements play a role in the process of meaning-giving in the visual domain of static images. These include the mise en scène and cinematographic techniques. Both the mise en scène and cinematography consist out of five pictorial elements each.

The mise en scène is concerned with the question “who and/ or what is shown” (Verstraten [2006: 59]). The first element of the mise en scène is the choice of characters. The choice for a specific character or person may influence the way in which an image is processed. A second element is the position, body language and/ or facial expression of these characters. The way in which characters are represented can also affect the interpretation of an image. The facial expressions mentioned by Attardo et al. (2003) and Rockwell (2001) can for instance be assigned to this category. A third element is the clothing of the characters, which may also influence the ways in which such a character is perceived. Another aspect of the mise en scène includes objects visible in the picture. Like the choice of character, the choice of objects may also demonstrate important information about the image. The last element of the mise en scène is the location or setting of the image; the question of where a picture is situated may also influence the interpretation of the image (Verstraten [2006: 59–66]).

Cinematography deals with the question “how” something is shown (Verstraten [2006: 59]). Following Verstraten (2006), cinematographic techniques
include the use of color, framing, depth and sharpness, camera angle and focalization. The cinematographic technique of color focuses on the ways color is used (or not) to emphasize elements in the image. Framing is related to the frame of the image; which elements are included in the image and where in the image are they located? Depth and sharpness are concerned with the clarity with which certain things are visually represented. The camera angle deals with the angle from which a photo is taken. Focalization, finally, asks the question if the camera also takes somebody’s perspective or not.

The second research question of this chapter is thus:

RQ2. Which pictorial elements contribute to the identification of verbal irony?

The examples of visuals used earlier in this section were comprised of several pictorial elements. Figure 1 showed the supposed danger of winning the jackpot of the Dutch lottery Lotto. The first pictorial element that comprises the visual marker is the astronaut’s body language. The low shoulders and the head held low indicate the astronaut’s disappointment. The multitude of Dutch flags planted on the moon give the reason for his disappointment; numerous Dutch astronauts were on the moon before him. Finally, the shot is made from a bird’s-eye perspective, which makes it seem as if the entire surface of the moon is filled with these Dutch flags. This visual marker is thus comprised out of four pictorial elements: (1) body language (the astronaut’s shoulders and head), (2) objects (the multitude of Dutch flags) in combination with (3) location (the moon), and (4) camera angle (bird’s-eye perspective).

Figure 2 showed the man who was unhappy with his gift of hedge-clippers. The man showed his displeasure with his gift by means of his facial expression and body language; he looked puzzled and held the hedge-clippers as if they were a mobile phone. The red-purple tone of the advertisement suggests that the picture is old, possibly indicating that the gift of hedge-clippers is old-fashioned. Finally, the man in the image directly addresses the camera as if the camera was the person he is speaking to. This implies that the camera takes the perspective of the person who gave the hedge-clippers as a gift. This incongruent image is thus comprised out of four pictorial elements as well: (1) body language and facial expression (indicating disappointment), (2) object (hedge-clippers as substitute for a mobile phone), (3) color (tone of an old-fashioned photo) and (4) focalization (camera takes the perspective of the person who gave the present).

This bottom-up approach allows each image to be analyzed on each of the elements of the mise en scène and each of the cinematographic techniques. It is then possible to argue which pictorial elements help a reader to detect irony. Nevertheless, it should be taken into account that this is the first analysis to use
visual grammar to identify pictorial elements that help a reader in detecting irony. This analysis should thus be seen as a first, exploratory analysis in the ways in which visuals can help a reader to detect irony.

2 Method

2.1 Material

The analysis reported in this chapter was mainly exploratory. Fifty multimodal texts were randomly selected from a corpus of ironic utterances. In this corpus, the ironic utterances were identified with use of the Verbal Irony Procedure (VIP; see Burgers et al. [2011]), which consists out of four steps. Coders first read the entire text to get a general idea of its meaning and the position taken by the text’s author. Coders subsequently analyze each individual clause and determine whether it is descriptive or evaluative. If the utterance is evaluative, the third step of the VIP entails that coders determine if the literal evaluation is congruent with the co-text or not. If the literal evaluation is incongruent with the co-text, a coder can proceed to the fourth step, which determines if the literal evaluation can be contrasted with an intended evaluation about the same object. When a plausible reading can be found in which the valence of the intended (ironic) evaluation (negative or positive) is in the opposite domain from the literal evaluation (positive or negative), the utterance is considered ironic. In total, the fifty texts of this sub-corpus contained a total of sixty-five ironic utterances. Out of these fifty texts, twelve were cartoons, twenty-one were commercial advertisements and seventeen were non-commercial advertisements.

2.2 Procedure and reliability

The coding process of the visuals consisted out of two rounds of coding. For the first round, coders – the first author as well as two advanced BA students in Business Communication Studies at Radboud University Nijmegen – were given an instruction that informed them when a visual could help in detecting irony. They were subsequently asked to analyze each ironic utterance separately in relation to the image present in the advertisement or cartoon. In doing so, coders had to analyze several variables.

Firstly, they were asked if the visual actually helped in decoding the ironic utterance or not (compare Figures 1 and 2 to Figure 3). Texts in which a visual did
not help were not analyzed further. Then, coders had to decide whether the visual could be classified as a visual marker (see Figure 1) or as an incongruent image (see Figure 2).

The second part of the analysis aimed to identify as many pictorial elements as possible that indicated the use of irony. In order to do this, an exploratory analysis was executed where coders had to analyze the image on aspects of mise en scène (choice of characters, body language and facial expression, clothing, objects and location and setting) and cinematography (the use of color, framing, depth and sharpness, camera angle and focalization).

In the coding process, coders had to argue whether each of these pictorial elements would help in detecting the irony. This means that, say, the use of characters was only marked if the coder felt that this use of characters helped in detecting or solving the irony. The instruction provided coders with an example in which one of the elements from the mise en scène or one of the cinematographic techniques helped to decode the irony. Of course, these examples were not included in the actual corpus.

<table>
<thead>
<tr>
<th>Round</th>
<th>Variables</th>
<th>$\kappa$</th>
<th>% Agreement</th>
<th>$p_1$</th>
<th>$p_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Image helps in decoding irony or not$^1$</td>
<td>.30</td>
<td>75.3</td>
<td>.84</td>
<td>.45</td>
</tr>
<tr>
<td></td>
<td>Image is a visual marker or visually incongruent$^2$</td>
<td>.39</td>
<td>72.1</td>
<td>.59</td>
<td>.79</td>
</tr>
<tr>
<td>2</td>
<td>Image helps in decoding irony or not$^1$</td>
<td>.65</td>
<td>88.9</td>
<td>.93</td>
<td>.72</td>
</tr>
<tr>
<td></td>
<td>Image is a visual marker or visually incongruent$^2$</td>
<td>.69</td>
<td>86.5</td>
<td>.79</td>
<td>.90</td>
</tr>
</tbody>
</table>

Note: All scores are average scores based on the three possible coder pairs. $\kappa$ lists the Cohen’s Kappa; %Agreement indicates the total percentage of cases that coders agree on.

1 For the question whether an image helps in decoding the irony or not, $p_1$ indicates the probability that if a coder believes that the image helps in decoding a specific ironic utterance, at least one other coder would agree; $p_2$ indicates the probability that if a coder believes that the image does not help in decoding a specific ironic utterance, at least one other coder would agree.

2 For the question whether an image can be classified as a visual marker or as an incongruent image, $p_1$ indicates the probability that if a coder states that an image is a visual marker, at least one other coder would agree; $p_2$ indicates the probability that if a coder states that an image is visually incongruent, at least one other coder would agree.

Table 1: Cohen’s kappa, the percentage of agreement between coders, $p_1$ and $p_2$ in the first and second round of coding of visuals.

In the second round, coders were confronted with the coding and motivation of another rater on cases on which they disagreed, in a similar way as has been done in other studies (see e.g., Van Enschot [2006]). Since the coding in the second round was partially independent (coders saw the other ratings of
another coder), the scores of the kappas after the second round of coding should be treated with care.

Table 2 shows reliability scores for the elements that give meaning in the visual domain. These scores show a pattern that is similar to the other reliability scores. After the first round of coding, only two elements (depth and sharpness, and focalization) reach an acceptable Cohen’s kappa of at least .60, even though overall agreement is at least 67% for any of the variables. The second round of coding paints a different picture. After this round, only three elements

<table>
<thead>
<tr>
<th>Round</th>
<th>Pictorial element</th>
<th>( \kappa )</th>
<th>% Agreement</th>
<th>( p_{\text{present}} )</th>
<th>( p_{\text{absent}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Mise en scène:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>– Choice of characters</td>
<td>.37</td>
<td>67.3</td>
<td>.64</td>
<td>.69</td>
</tr>
<tr>
<td></td>
<td>– Position, body language, facial expression</td>
<td>.43</td>
<td>72.3</td>
<td>.76</td>
<td>.66</td>
</tr>
<tr>
<td></td>
<td>– Clothing</td>
<td>.53</td>
<td>87.1</td>
<td>.60</td>
<td>.92</td>
</tr>
<tr>
<td></td>
<td>– Choice of objects</td>
<td>.20</td>
<td>67.5</td>
<td>.65</td>
<td>.65</td>
</tr>
<tr>
<td></td>
<td>– Location or setting</td>
<td>.37</td>
<td>73.3</td>
<td>.52</td>
<td>.81</td>
</tr>
<tr>
<td>1</td>
<td><strong>Cinematography:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>– Color</td>
<td>.42</td>
<td>87.9</td>
<td>.47</td>
<td>.93</td>
</tr>
<tr>
<td></td>
<td>– Framing</td>
<td>.36</td>
<td>90.5</td>
<td>.41</td>
<td>.95</td>
</tr>
<tr>
<td></td>
<td>– Depth and sharpness</td>
<td>.65</td>
<td>96.5</td>
<td>.67</td>
<td>.98</td>
</tr>
<tr>
<td></td>
<td>– Camera angle</td>
<td>.22</td>
<td>85.1</td>
<td>.28</td>
<td>.92</td>
</tr>
<tr>
<td></td>
<td>– Focalization</td>
<td>.77</td>
<td>98.4</td>
<td>.78</td>
<td>.99</td>
</tr>
<tr>
<td></td>
<td><strong>Mise en scène:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>– Choice of characters</td>
<td>.73</td>
<td>86.3</td>
<td>.87</td>
<td>.86</td>
</tr>
<tr>
<td></td>
<td>– Position, body language, facial expression</td>
<td>.72</td>
<td>87.4</td>
<td>.90</td>
<td>.82</td>
</tr>
<tr>
<td></td>
<td>– Clothing</td>
<td>.65</td>
<td>88.1</td>
<td>.72</td>
<td>.92</td>
</tr>
<tr>
<td></td>
<td>– Choice of objects</td>
<td>.50</td>
<td>75.8</td>
<td>.80</td>
<td>.69</td>
</tr>
<tr>
<td></td>
<td>– Location or setting</td>
<td>.69</td>
<td>85.9</td>
<td>.80</td>
<td>.89</td>
</tr>
<tr>
<td>2</td>
<td><strong>Cinematography:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>– Color</td>
<td>.83</td>
<td>96.0</td>
<td>.86</td>
<td>.98</td>
</tr>
<tr>
<td></td>
<td>– Framing</td>
<td>.46</td>
<td>93.8</td>
<td>.49</td>
<td>.97</td>
</tr>
<tr>
<td></td>
<td>– Depth and sharpness</td>
<td>.66</td>
<td>97.4</td>
<td>.67</td>
<td>.99</td>
</tr>
<tr>
<td></td>
<td>– Camera angle</td>
<td>.37</td>
<td>91.3</td>
<td>.40</td>
<td>.96</td>
</tr>
<tr>
<td></td>
<td>– Focalization</td>
<td>.71</td>
<td>98.0</td>
<td>.72</td>
<td>.99</td>
</tr>
</tbody>
</table>

Note: All scores are average scores based on the three possible coder pairs. \( \kappa \) indicates Cohen’s Kappa; %Agreement indicates the total percentage of cases coder pairs on average agree on; \( p_{\text{present}} \) reports the probability that if a coder believes that a certain pictorial element is (partly) responsible for guiding an ironic interpretation, another coder agrees; \( p_{\text{absent}} \) reports the probability that if a coder believes that a certain element is not responsible for guiding an ironic interpretation, another coder agrees.

Table 2: Cohen’s kappa, the percentage of agreement between coders, \( p_{\text{present}} \) and \( p_{\text{absent}} \) in the first and second round of coding of pictorial elements that guide an ironic interpretation.
(i.e., choice of objects, framing and camera angle) score a Cohen’s kappa below the .60 threshold, even though overall agreement for any of these variables is at least 75%. More positively, seven elements (i.e., choice of characters; position, body language and facial expression; clothing; location and setting; color; depth and sharpness; focalization) are now coded with a Cohen’s kappa of at least .60.

The figures reported in Tables 1 and 2 show that the visual elements are difficult to code reliably with only one round of coding. After a second round of coding in which coders are confronted with the analysis of another coder, however, agreement can reach a substantial level. This makes it acceptable to report the coding of the author after the second round of coding. Results related to variables that scored kappa reliability lower than .60 should be treated with care.

3 Results

The first research question of this chapter dealt with the ways in which visuals are used in relation to irony in multimodal discourse. The first research question was:

RQ1. How are different types of visuals used in relation to irony in written discourse?

Table 3 shows that, for 52 ironic utterances in the multimodal texts (or 80% of ironic utterances from the selected sub-corpus), the image helped a reader to detect irony (see e.g., Figures 1 and 2). This means that for 13 ironic utterances from the multimodal texts (or 20% of ironic utterances from the selected sub-corpus), the image did not help a reader to detect irony (see Figure 3). These latter ironic utterances were not included further in the analyses in this chapter. Table 3 also shows that, in a number of cartoons, the image does not help in detecting the irony. These cartoons are comics in which the irony can be solved based on the text between the characters in the comic. Table 3 also demonstrates that, when the image helps a reader to detect irony, it is more often visually incongruent (36 times or 69.2% of ironic utterances from the corpus in which a visual helped in detecting the irony) than a visual marker (16 times or 30.8% of ironic utterances from the corpus in which a visual helped to detect irony).
In detecting irony, the image:

<table>
<thead>
<tr>
<th></th>
<th>Comm. ads</th>
<th>Non-comm. ads</th>
<th>Cartoons</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>helps</td>
<td>22</td>
<td>21</td>
<td>9</td>
<td>52</td>
</tr>
<tr>
<td>does not help</td>
<td>2</td>
<td>7</td>
<td>4</td>
<td>13</td>
</tr>
</tbody>
</table>

Image is:

<table>
<thead>
<tr>
<th></th>
<th>Comm. ads</th>
<th>Non-comm. ads</th>
<th>Cartoons</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>a visual marker</td>
<td>7</td>
<td>6</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>visually incongruent</td>
<td>15</td>
<td>15</td>
<td>6</td>
<td>36</td>
</tr>
</tbody>
</table>

Table 3: The number of ironic utterances for which the image helps to detect the irony and the number of ironic utterances for which the image is a visual marker or visually incongruent, by genre.

The second research question of this chapter was:

RQ2. Which pictorial elements contribute to the identification of verbal irony?

Table 4 shows how often individual elements from the mise en scène and cinematography were said to help a reader to detect irony. The most striking aspect of Table 4 is the difference in use between elements from the mise en scène and cinematography. Positioning of characters, body language and facial expression are used most often to help a reader to detect irony (34 times), followed by the choice of objects (32 times) and characters (25 times). Location and setting are used 16 times and clothing is used 13 times to help a reader in detecting irony. All cinematographic elements are used less than ten times to alert a reader to irony usage. While elements from mise en scène are thus used relatively often, elements from cinematography are hardly used to guide a reader towards an ironic interpretation.

Table 4: Frequency with which pictorial elements are used to guide a reader towards an ironic interpretation.
4 Conclusion and discussion

This chapter was concerned with the ways in which visuals could help in detecting verbal irony. Its first research question asked how different types of images were used in relation to irony in written discourse. Results demonstrate that, for over 80% of ironic utterances in the multimodal texts, the image helped a reader to detect irony. These results suggest that images are an important aspect of verbal irony in multimodal genres. In addition, results demonstrate that images that help a reader to detect irony can more often be labeled as visually incongruent than as visual markers. This implies that images that help in decoding irony often work with a contrast to the literal evaluation of the ironic utterance.

The division in visually incongruent images and irony markers also shows a duality which some authors see as central to the process of creativity. Veale at al. (this volume) for instance introduce the so-called CRIME model (Creative Integration Mechanism) which posits that creativity can be found in blends that combine two input spaces in a creative way. This study shows that this CRIME framework can also be extended to irony. After all, in an ironic utterance, the literal evaluation should be reversed to the intended evaluation. Studies on irony processing show that, even when the ironic utterance is solved and addressees come to the intended meaning, the original (literal) meaning remains just as active in memory (e.g., Giora & Fein [1999], Giora, Fein & Schwartz [1998]) which shows that both the literal and intended space are activated and that irony may be processed as a blend.

Given that, during processing of ironic utterances, both the literal and intended evaluation remain active in working memory, the division in visually incongruent images and irony markers may have implications for irony processing in multimodal contexts. After all, in the case of visual irony markers, the image reinforces and mocks the literal evaluation of the irony at the same time. During processing of visual irony markers, readers are thus forced to think about the literal evaluation and why this evaluation is presented as absurd. In other words, the focus during processing may be more on the absurdity of literal evaluation than on the issue why the intended evaluation is the better alternative. Visually incongruent images, in contrast, focus on the intended meaning, because they are incongruent with a literal reading of the irony. In other words, the focus during processing of visually incongruent images may be more on the intended evaluation and solving what the speaker actually means. Thus, since visual irony markers force readers to think about and reconsider the literal evaluation, it may be possible that, even when the irony is solved, the literal evaluation is still activated more strongly in working memory than when readers solve an ironic utterance that is marked with a visually incongruent image.
In addition to the types of visuals, the second research question dealt with the pictorial elements that help to signal irony. In this second question, we take on one of the challenges presented by Veale et al. (this volume) who argue that scholars should identify variables to analyze creative examples that are independent of the object under investigation. Our results show that elements from the mise en scène are more important in this matter than cinematographic techniques, because the number of elements from the mise en scène that alert a reader to the use of irony is much higher than the number of cinematographic techniques that do so. The pictorial elements of body language and facial expression, and the choice of characters and objects were used most often to direct a reader towards an ironic reading. The importance of body language and facial expression support the observations of Attardo et al. (2003) and Rockwell (2001) who also believed that facial expressions could be important in helping a reader to detect irony.

In contrast to elements from the mise en scène, elements from cinematography are hardly used to help a reader to detect irony. Previous research showed that cinematographic elements of images in advertisements can produce attitudinal effects (see e.g., Meyers-Levy & Peracchio [1992] for camera angle; Meyers-Levy & Peracchio [1995] for color and Peracchio & Meyers-Levy [1994] for framing). Nevertheless, in the corpus, these cinematographic elements were not found to help a reader in detecting the irony in a verbal utterance. Since the number of images included in this analysis was relatively low, future research may want to confirm these findings.

The results from this chapter suggest that images may help a reader to detect irony in two different ways. Firstly, the use of body language and facial expression strongly resembles kinetic cues that are used in face-to-face interaction (e.g., Attardo et al. [2003], Rockwell [2001]). This implies that an image that marks irony may resemble a freeze-frame shot from a conversation. The image conveys a critical moment in an (imagined) conversation; the moment at which the speaker “produced” a visual clue to help a reader identify the use of irony. The second way in which an image may help a reader to detect irony depends upon what is shown; the choice of characters or objects. These visuals may then be considered as a replacement of the situational context (cf. the ironic environment introduced by Utsumi [2000]). Imagine that two people planned to go on a picnic, but cannot go because it rains at the designated time. In frustration, one of them may utter the ironic remark “Great weather, eh?”. In a multimodal text, then, an image might show that it was raining at the moment the ironic utterance of “Great weather, eh?” was uttered. In case of such a multimodal text, a sender has of course the possibility to manipulate the situational context represented in the image.
This chapter distinguished the pictorial elements that help a reader to identify verbal irony. However, in many images, several pictorial elements are used to help a reader identify verbal irony. This study has not investigated the effect of these pictorial elements put together. It may be that some pictorial elements typically work together to help a reader in detecting irony. Visual metaphors and visual hyperboles are examples of a combination of various pictorial elements that together serve as a visual marker or to create visual incongruence. An example of a visual hyperbole that serves as a visual marker is the image of the disappointed Dutch astronaut on the moon who discovers that many Dutchmen had planted a flag on the moon before him (Figure 1). Future research may set up a coding instruction to identify visual markers and incongruent images that are comprised out of multiple pictorial elements more systematically.

Future research may also want to investigate if the use of visuals in relation to irony differs across various multimodal genres. Since this study was explorative in nature and the number of images that was included in the analysis was relatively low (50 images divided over three genres), it was difficult to do statistical tests to investigate whether this use also differed across genres. Future research may thus want to extend these findings with a bigger corpus. This analysis may also reveal whether verbal text and visuals work differently in alerting a reader to an ironic utterance. While the author of a verbal text may use a wide variety of written clues to help a reader detect irony, the number of visual clues that are at the “writer’s” exposal may be relatively small (only ten pictorial elements are identified in the Verstraten [2006] framework, for instance). In order to investigate these issues, future research may thus seek to replicate these results with a bigger sample of multimodal texts.

This chapter presented a first systematic analysis of the ways in which visuals may help a reader of multimodal texts to determine whether a specific verbal utterance is ironic. Results reveal that images may work in different ways to accomplish this goal, which means that ironists can have various creative options to alert a reader (or not) to the use of irony. This analysis thus shows that modalities may work together to form a creative text. This means that research into creativity within specific modalities (cf. Veale et al. [this volume]) may also be supplemented with research that cuts across various modalities.
5 References


