

Incongruity-resolution humorous strategies in image macro memes

Francisco Yus

Internet Pragmatics 4(1): 131-149

(this is a draft and may differ slightly from the printed version)

ABSTRACT

In previous research (Yus 2016, 2017), a classification of jokes was proposed depending on how the humorous incongruity-resolution strategy was achieved. Twelve cases were isolated resulting from the combination of several parameters: (1) a differentiation between *discourse-based incongruities* and *frame-based incongruities*; (2) the location of the incongruity-triggering element (setup or punchline); and (3) three types of resolution: discourse-, frame- and implication-based. This paper proposes a similar taxonomy of incongruity-resolution patterns for a specific type of internet discourse: the image macro meme. The resulting seven-case taxonomy inherits some of the features of that proposed for verbal jokes, albeit exhibiting the extent to which the image plays a specific role in the successful meme-centred humorous strategy.

Keywords: image macro meme, incongruity-resolution strategy, internet humour.

1 Incongruity-resolution strategy in verbal jokes

This theory pictures humour as the outcome of two phases: one in which the hearer comes across some form of incongruity while interpreting the text; and another where the hearer finds a cognitive rule that reconciles the incongruity in question and triggers humorous effects, resolution becoming necessary to accomplish the humorous strategy (otherwise puzzlement would be generated instead of humour). The most outstanding representative of this theory is Suls (1972, 1977, 1983), for whom the key to humour lies in an initial stage during which a hearer detects an incongruity. Then, while the hearer tries to solve the incongruity (and make sense of the joke as a whole), a search for a cognitive rule takes place which reconciles the incongruous parts. Upon finding a resolution to the incongruity, the hearer will be relieved and perhaps also humorously entertained:

the theory is that humor results when the perceiver meets with an incongruity (usually in the form of a punch line or a cartoon) and then is motivated to resolve the incongruity either by retrieval of information in the joke or cartoon or from his/her own storehouse of information. According to this account, humor results when the incongruity is resolved; that is, the punch line is seen to make sense at some level with the earlier information in the joke. (Suls 1983: 43)

In order to account for the different kinds of incongruity-resolution (henceforth IR) patterns that might arise in humorous communication, three sets of variables were proposed in previous research (Yus 2016, 2017):

Firstly, two *types of incongruity*. Many instances of humorous texts fit conflicts that hearers encounter while processing the verbal content of the joke, subsequently leading to a discourse-centred resolution. By manipulating some of the hearer's inferential strategies such as disambiguation, reference assignment, concept adjustment, among others, the speaker may puzzle the hearer in their relevance-seeking interpretation of the joke. The speaker then has to alter the joke discourse inference so as to find the reconciling rule that leads to humour (e.g. by selecting a different sense for an ambiguous word, or replacing a referent for an indexical that, despite being initially relevant, eventually turns out to be invalidated).

However, the incongruity in other jokes does not rely on manipulations of the hearer's

inferential strategies to interpret the text of the joke. Instead, the hearer's construction of an initially appropriate mental situation (called frame, schema and script in the bibliography, among other terminological proposals) constructed to make sense of what is happening in the story depicted within the joke, and which is eventually invalidated. In a nutshell, inferring the intended interpretation of a joke (and of any utterance for that matter) entails retrieving general information both about the world and about everyday situations that is stored as accessible chunks of encyclopaedic information (more specifically as "I conceptualize X as *p*" or as a more factual "I believe that *p*"). This information is often activated almost unconsciously in order to make sense of the intended *scenario* for the comprehension of jokes and other such utterances. It is an information storing process that allows us to frame situations and save mental effort when facing a new input such as a joke, or to anticipate different stages in a communicative situation, as happens with scripts.

Therefore, a differentiation was proposed in Yus (2016) between *discourse-based incongruities* and *frame-based incongruities*. The former relate to the inferential strategies that the hearer applies to the coded input (the joke) seeking to turn it into a fully contextualised proposition (reference assignment, disambiguation, concept adjustment, free enrichment, the interface between explicit and implicated interpretations...). As for the latter, they focus on the situation that the hearer constructs to make sense of what is happening in the joke (the label *make-sense frame* arose as an umbrella term for the terminology on this activation of prototypical mental situations). In Yus (2013), (1) below was proposed as an example of discourse-centred incongruity specifically based on one of the typical inferential strategies meant to turn the schematic coded utterance into a meaningful interpretation: reference assignment of the pronoun in *do that*. Besides, joke (2) was offered as an example of *frame-based incongruity*, since the situation (*make-sense frame*) initially built up by the hearer is contradicted upon at the end of the joke, forcing them to backtrack and re-interpret the joke differently:

- (1) Eleven-year-old Lucy was walking down the village street leading a cow by a rope. She met the vicar who said, "Little girl, what are you doing with that cow?" "It is my father's cow and I am taking her to the bull," said Lucy. "Disgusting," said the parson, "can't your father do that?" "No," said Lucy, "it has to be the bull."
- (2) Jake was on his deathbed. His wife Susan, was maintaining a vigil by his side. She held his fragile hand, tears ran down her face. Her praying awoke him. He looked up and his pale lips began to move slightly. "My darling Susan," he whispered. "Hush, my love," she said. "Rest. Shhh. Don't talk." He was insistent. "Susan," he said in his tired voice. "I have something I must confess to you." "There's nothing to confess," replied the weeping Susan. "Everything's all right, go to sleep." "No, no. I must die in peace, Susan. I slept with your sister, your best friend and your mother." "I know," she replied. "That's why I poisoned you."

Secondly, *types of resolution*. A three-fold classification of resolutions was suggested depending on what the hearer is expected to perform inferentially in order to find some element that can solve the incongruity and causes the intended humorous effects:

a. *Discourse-based resolution*, when the hearer needs to perform a supplementary inferential operation to make sense of the incongruity found in some portion of the joke text. By way of example, to select a previously undetected sense of a word, to locate a referent for an indexical other than the one initially chosen, or to opt for a different *ad hoc* concept for a word in a context.

b. *Frame-based resolution*, when the hearer is forced to alter the current make-sense frame constructed for the situation and replace it with a different frame that reconciles the

incongruity.

c. *Implication-based resolution*, which applies to (a) cases where the hearer has to find the resolution outside the text of the joke through the derivation of implicated premises and implicated conclusions (implicatures); and (b) cases in which the hearer looks for implications that allow for deriving humorous effects after a full comprehension of the joke has taken place, as in (3), where the wife's words result in a humour-connoted contextual implication:

- (3) The husband was at home, talking to his wife one evening over supper. "Get this..." he chuckled, "That ridiculous janitor of ours claims he's made love to every woman in the building except one." "Hmm," said his wife, assuming a thoughtful faraway type expression, "must be that stuck-up Mrs. Stewart on the eighth floor."

Finally, *location of the incongruity-triggering element*. A distinction was drawn between incongruities whose source is located at the initial¹ part of the joke (the *setup*) and those whose source of incongruity stems from the final part (the *punchline*). The former often demand some kind of inferential backtracking from the hearer, that is, returning to an already processed part of the joke and re-interpreting it in the light of the new evidence extracted from the processing of a subsequent stretch of the joke text. The latter (incongruity-triggering elements placed at the end of the joke) often demand the hearer's participation through the derivation of implicatures required not only to make sense of the joke but also to obtain the desired humorous effects. Notice that, as highlighted in Yus (2016), the source of the incongruity and its actual identification may be located in different parts of the joke; the following example/example (4) below provides clear evidence thereof:

- (4) Tom and Jim talking to each other, Tom says: "You know, Jim, my wife and I were very happy for 25 years." His friend asks, "And then... what happened?" "We met," replies Tom.

The hearer activates a typical scenario of the married couple's loving relationship in the preceding joke. Its last part (punchline) invalidates this frame previously constructed and forces a resolution that involves replacing it with a different frame, this time regarding unsatisfactory marriage. Instead of being located in the punchline, the source of the incongruity (i.e., expressed differently, what makes the incongruity possible, what triggers it), is located in the setup, (more) specifically in the phrase *very happy for 25 years* —inevitably understood as "very happy [together] for 25 years" in the initially activated make-sense frame, and later re-interpreted as "very happy for 25 years [before meeting]." Therefore, the joke exhibits a frame-based incongruity whose source lies in the setup and provides a frame-based resolution (frame replacement from happy marriage to unhappy marriage).

The combination of these three parameters (type of incongruity; type of resolution; and location of the incongruity-triggering element) yielded the twelve cases listed in Table 1. The aim of this paper is to analyse the extent to which a similar taxonomy may also be proposed to account for the IR strategy in verbal-visual memes of the image-macro kind.

2 (Image macro) memes

1. *Initial part* does not mean the very beginning of the joke. The setup of the joke may take up most of its space, to such an extent that sometimes the location of the incongruity-triggering element close to the end of a joke does not (necessarily) prevent it from belonging to its initial part (setup).

Memes stand out among the most pervasive multimodal discourses on the Net nowadays. Despite having been initially created on websites, now there are specific meme-generation smartphone *apps*; and other smartphone *apps* such as social networking *apps* and mobile messaging *apps* typically act as channels through which these memes spread and are disseminated across mass populations. A particularly abundant type of meme at present is the *image macro*, a text-image multimodal discourse made up of one or two text lines at the top and/or the bottom of the meme complemented by an image in the middle with several possible interpretive combinations. This is the type of meme around which our analysis will revolve.

	Type of incongruity	Location	Type of resolution
1	frame-based	setup	discourse-based
2	frame-based	punchline	discourse-based
3	frame-based	setup	frame-based
4	frame-based	punchline	frame-based
5	frame-based	setup	implication-based
6	frame-based	punchline	implication-based
7	discourse-based	setup	discourse-based
8	discourse-based	punchline	discourse-based
9	discourse-based	setup	frame-based
10	discourse-based	punchline	frame-based
11	discourse-based	setup	implication-based
12	discourse-based	punchline	implication-based

Table 1. Incongruity-resolution patterns

According to Tay (2015, 49), image macros tend to be easier to produce, since several computer applications facilitate meme creation and alteration. The use of the expression “image macro” is widely accepted to describe an image with a superimposed caption, usually for a humorous purpose. From this perspective, a macro can be portrayed as the set of automated processes that users go through to add captions to images following the community’s accepted style.

As stated in Yus (2018: 114-115), the qualities of internet memes have led some researchers to equate them to *virals*. This is understandable, since virals spread unchanged across digital media (Dyrel 2016: 662) and are forwarded from one person to others, within and across multiple weakly linked personal networks, ultimately causing a rapid increase in the number of people exposed to the meme message (Shifman 2014: 55). Nevertheless, the qualities of memes and virals only partially overlap in one respect: both of them can be intentionally transmitted unaltered, exactly like the image macros under analysis in this paper. By contrast, virals may spread beyond the user’s intention rather than with an overt intention to ensure the widest possible dissemination of the discourse. For example, a user may write an online review of a hotel which, for some unexpected reason, goes viral beyond that user’s initial intention. This is not possible in memes —intentionally created and transferred with an expectation for them to spread. Furthermore, memes may be/are often altered by the free software available on the Net, whereas virals tend to spread unaltered.

Crucially, a distinctive feature of internet memes is their invariable ascription to a “family,” part of their effective processing and an adequate derivation of cognitive effects depending on this initial family ascription (Segev et al. 2015: 418; Nissenbaum and Shifman 2017: 484). Part of the appeal involved in interpreting memes and possibly part of their success as humorous instances lies in the viewers’ ability to track the family where the meme belongs and to compare the current meme to others already produced within the same family. Users with no meme literacy will often be unable to capture the full range of effects that those aware of meme families can perceive. The same applies to general background knowledge on the information contained in the meme and possible implicatures that may have to be derived to obtain a full understanding of the intended effects, as in meme (5), which demands knowledge about Bill Clinton to be able to obtain a proper interpretation.

- (5) Top text: Just realized if Hillary wins...
 Image: Photo of former US president Bill Clinton.
 Bottom text: I get interns.

3 Incongruity-resolution patterns in image macro memes

A sample of 150 memes was collected from random searches on the internet to isolate possible patterns of IR in image macro memes. Those memes subsequently underwent thorough analysis and distinctions were drawn regarding the kind of incongruity and resolution used to generate humorous effects. This analysis yielded several types of IR patterns, some of which more or less resembled the twelve cases of IR proposed above for verbal jokes; other types were more meme-specific due to the multimodal semiotic configuration of these discourses. Specifically, the two types of incongruity (*discourse-* and *frame-based*) suggested for verbal jokes were also at work in memes, plus a third combinatory pattern arising from a humour-triggering combination of the text and the image in the meme: *discourse-image incongruity* —which involves an inferential clash when inferring meanings that need a convergence between the partial meanings of the text and the image. The same as in memes, the meaning of the image and the interpretation of the text may prove incongruous when put together in the same semiotic discourse.

The role performed by the image in the meme additionally provided us with three major cases: (a) image plays no role in IR; (b) image aids in IR; and (c) image is essential in IR. The cases ensuing from the combination of such parameters yielded the 7 patterns collected in Table 2, which also shows the percent occurrence rate corresponding to these types in the aforementioned corpus of 150 memes.

		Type of incongruity	Role of image
31.33%	8.66%	frame-based	image plays no role
	18.00%	frame-based	image aids
	4.67%	frame-based	image is essential
38.67%		discourse-image based	image is essential
21.34%	8.00%	discourse-based	image plays no role
	7.34%	discourse-based	image aids
	6.00%	discourse-based	image is essential

8.66%		No incongruity	
-------	--	----------------	--

Table 2. Incongruity-resolution patterns in memes

The most frequent pattern was the one where incongruity arises from clashes between the meanings of the text and the image in the meme (near 40%), which makes sense given the multimodal quality of memes. More importantly, this stands out as the only pattern in which no differentiation is made depending on the role of the image, the reason for it being that, if the image plays an important role in contrasting its information to the one provided by the accompanying text, its role necessarily has to be essential. As for the other categories —i.e. discourse- and frame-based— they reveal a threefold distinction of roles performed by images, namely: no role; aiding role; and essential role. These 7 categories are briefly described below. The explanations provided in footnotes about particular memes and the series they belong come from the famous website *Know your Meme* (<https://knowyourmeme.com/>).

Only 8.66% of the memes analysed did not exhibit any humorous IR patterns whatsoever, probably because they were not meant to be humorous in the first place. Examples include the meme in (6) and the two memes from the series of memes “Not sure if...”² in (7) and (8):

- (6) Top text: Maggie Smith³ battled cancer whilst filming the last Harry Potter Movie.
Image: A photo of actor Maggie Smith.
Bottom text: She carried on, not wanting to disappoint the fans.
- (7) Top text: Not sure if everything is expensive.
Image: A still of character Fry in animated series *Futurama*.
Bottom text: or I am just poor.
- (8) Top text: Not sure if I’m sleepy because I’m tired.
Image: A still of character Fry in in animated series *Futurama*.
Bottom text: or I’m tired because I slept too much.

3.1 Frame-based incongruity where the image plays no role in the IR strategy

In this case of our taxonomy, the viewer is forced to replace the frame previously constructed to interpret the meme with another frame so as to find a resolution and obtain the intended effects. Only 8.66% of the memes fitted this category.

The prototypical situation in this category simply shows a contrast during the cumulative inferential strategy from the top text to the bottom text —regarding anticipatory hypotheses created during the inference of the top text and what may come up in the subsequent bottom text— without the aid of the information from the image in the meme. Two possible examples can be found in (9) and (10) below.

- (9) Top text: A guy walks into a bar...
Image: An image of a hen.
Bottom text: He has a family of four and a drinking problem.
- (10) Top text: Why do people say the sky’s the limit.
Image: A drawing of a Velociraptor dinosaur.
Bottom text: When there are footprints on the moon?⁴

2. This series, which portrays an image of *Futurama* character Fry squinting his eyes, is typically paired with overlaid text using a phrasal template. The top line always reads “Not sure if X,” and “or just Y” appears at the bottom line —used to represent an internal monologue.

3. Famous actor featuring in many films and TV series including *Harry Potter* and *Downton Abbey*.

4. Meme (9) belongs to a series called *anti-joke chicken*, where an image of a hen appears accompanied by a top line

In (9), the viewer will normally dismiss the image as potentially irrelevant —to concentrate on the text instead. The top text generates some expectations that the bottom text will continue the typical beginning of a joke provided in the top text; those expectations are subsequently disconfirmed, though, which creates an incongruity. It may conversely happen that the viewer knows about this meme series, and part of their background knowledge includes assumptions about the typical texts that accompany these hen-based memes. In this case, even if the bottom text remains incongruous, the accessibility to background (manifest) information reduces the degree of surprise. (10) faces the viewer with an utterance whose figurative interpretation is more relevant (given its storage as a unit and its being easy to retrieve), although the bottom text forces a literal re-interpretation of the utterance, causing an incongruity with regard to the initial expectations of continuation in figurativeness.

3.2 *Frame-based incongruity where the image aids in the IR strategy*

Two main varieties may be distinguished in this category, which amounts to 18% of the corpus analysed. Firstly, some instances show the referent of image as being compatible with the top text but not with the bottom text. This increases the potential effectiveness of the incongruity, since the initial interpretation selected for the top text is corroborated by the validating interpretation of the referent of the image in the meme (these are often processed in succession, although not necessarily so). Two examples can be found in (10-12), where the image —of an angry-looking woman in (10-11), and of a hippy-looking girl in (12)— is fully compatible with the top text, but the bottom text becomes incongruous with both the top text and the image, thus creating an incongruity.⁵

- | | | |
|------|--------------|---|
| (11) | Top text: | I am a strong independent woman. I don't need a man to take care of me. |
| | Image: | Angry-looking woman, shouting, with her arms raised. |
| | Bottom text: | But he better pay for dinner, buy me flowers, open the car door, tell me I'm beautiful and buy me an engagement ring. |
| (12) | Top text: | I demand equality between men and women. |
| | Image: | Angry-looking woman, shouting. |
| | Bottom text: | Except when I get special treatment for being a woman. |
| (13) | Top text: | "We need to love and respect all cultures and religions." |
| | Image: | A hippy-looking girl. |
| | Bottom text: | Except Christianity. |

Secondly, the image may be compatible with the bottom text but not with the top text. The inferential pattern in this sub-category entails the identification of an incongruity when reading top and bottom text in succession, and the image somehow smoothens the contrast between these texts, not only making the inferential incongruity less radical but also allowing for

of text that sounds like the introduction of a joke, while the bottom line provides an unexpected anticlimax. The chicken is centred within a typical advice animal starburst background. In turn, meme (10) falls within a series known as *Philosoraptor*, an “advice animal image macro series” featuring an illustration of a Velociraptor paired with captions which depict the dinosaur as deeply immersed in metaphysical inquiries or unravelling quirky paradoxes.

5. Meme (12) is from the *Feminist Nazi* series and features a photograph of Australian reality TV star Layla Joyce Subritzky from the ninth season of *Big Brother Australia*. The captions typically depict a naive approach to stereotypical feminist ideas, similarly to *College Liberal*. As for meme (13), which forms part of the *College Liberal* series, it shows a photo of a young Caucasian female with dreadlocks wearing a knitted cap. The captions in this series typically portray the character as a naive and hypocritical left-wing political activist, referencing various clichés associated with the hippie subculture.

a number of implicated conclusions about the person depicted in the image. This is what happens in the memes below.⁶ (14) exemplifies a sharp contrast between the texts (if the neighbour does not return the tools he may be murdered). However, the bottom text fits the character in the film *Taken*, where the actor plays a desperate father performing violent actions in order to rescue his kidnapped daughter. Instead, the contrast in (15) refers to the typical objective of university life: not to repeat courses. Again, the image of the student is not incompatible with the bottom text; and it smoothens the contrast as well. The bottom text in (16) will be compatible with the actor's image if the viewer has the ability to garner background knowledge to retrieve the information about the TV series in which he used to star, including its drug-related plot. In (17) after reading the top text, the audience will make expectations of some kind of comforting message coming up for lonely people, which are disconfirmed, that bottom text being compatible with the photo of Grumpy Cat. Finally, (18) is a curious example of image-text compatibility, since the bottom text reproduces the famous activist Greta Thunberg's words that she uttered exactly with the nonverbal behaviour portrayed in the photo of the meme while she was delivering a speech (viewers who are able to make that connection between the meme and her actual words will derive more humorous effects). This time, it is the top text that is completely incompatible.

- (14) Top text: Borrowing our tools? If not returned.
Image: Photo of actor Liam Neeson.
Bottom text: I will find you and kill you.
- (15) Top text: Why make college the 4 best years of your life.
Image: Photo of university student drinking a beer.
Bottom text: When you can make it 6.
- (16) Top text: I just met you and this is crazy.
Image: Photo of actor Aaron Paul.
Bottom text: But do you want some meth.
- (17) Top text: If you're sad about being alone on Valentine's day, just remember...
Image: Photo of the famous angry-looking "Grumpy cat".
Bottom text: No one loves you on the other days of the year either.
- (18) Top text: When you parents turn off the wifi.
Image: Photo of activist Greta Thunberg delivering a speech.
Bottom text: How dare you? You have stolen my dreams and my childhood.

3.3 Frame-based incongruity where the image is essential in the IR strategy

Only 4.67% of the samples analysed fell within this category. Here, the image becomes essential both to work out the incongruity detected in the frame constructed while processing the top text and when trying to find a possible resolution to the incongruity. In (19), the image of Jesus has a key importance for the proper comprehension of the accompanying texts. The same applies to (20), whose images are essential by providing visual information about the person to whom the top and bottom text are referring.⁷

6. Meme (14), also referred to as *The Taken meme* is a memorable quote said by Liam Neeson from the 2008 action-thriller movie *Taken*. This phrase has subsequently been adopted as a phrasal template for heckling comments, most typically in the form of reaction images. Meme (15) in turn belongs to the *Lazy College Senior* series and features a young man drinking a pint of beer. The images are typically captioned with unmotivated statements or behaviours associated with burnt-out college students during their final school year, a phenomenon commonly known as "senioritis." Finally, meme (16) refers to a series of memes that revolve around the famous TV series *Breaking Bad*, where this actor played one of the starring roles.

7. Meme (19) belongs to *Story Time Jesus*, an image macro series which features a painting of Jesus teaching a group of disciples with captions that contain gossipy, modern interpretations of famous Biblical narratives.

- (19) Top text: What is best in life?
Image: Jesus surrounded by his devoted followers.
Bottom text: Crush your enemies. See them driven before you. Hear the lamentations of other women.
- (20) Top text: Says he'll create 25 million new jobs.
Image: Donald Trump.
Bottom text: Went bankrupt 4 times.

3.4 Discourse-image based incongruity (*the image is essential*)

By contrast, this IR category was frequent, as attested by its presence in 38.67% of the samples analysed. This should come as no surprise, since image macros have a verbal-visual nature and authors are bound to exploit the interrelation of the partial text and image meanings when brought together in a single discourse for humorous purposes. Firstly, a commonly used strategy consists in creating a direct contrast between the meaning provided by the text and that of the image. Some examples are provided below.⁸

- (21) Top text: Join the marines they said.
Image: Marine holding an umbrella over President Obama.
Bottom text: You'll be a hero they said.
- (22) Top text: I want to go to Taco Bell.
Image: Woman weeping intensely.
Bottom text: But I'm on an all-carb diet.
- (23) Top text: I am hungry.
Image: Woman weeping intensely.
Bottom text: But I already brushed my teeth.

The sharp contrast portrayed in (21) between the expectations of the Marine job provided in the top and bottom text and the actual activity performed by a specific Marine becomes strengthened by the feeling that the text is being uttered by that very person as a complaint. Memes (22-23) instead generate incongruity by relating an intense nonverbal gesture (the act of crying) to the trivial situation provided by the top and bottom text.

A second sub-category involves the generation of ironic implicated conclusions by contrasting the image and the text. A clear example is the *Girls Laughing* series of memes, in which the image of two girls laughing and giggling invalidates a possible literal interpretation of the bottom text (the top text is always "And then I told him...") and triggers an ironic interpretation, as in (24-26) below. A similar series is the one in which a photo of several doctors and nurses laughing is portrayed in the meme, also invalidating the top and bottom texts in the meme, as in (27).

- (24) Top text: And then I told him...
Image: Girls laughing and giggling.
Bottom text: Size doesn't matter.
- (25) Top text: And then I told him...
Image: Girls laughing and giggling.

8. Memes (18-20) are examples of the *First World Problems* series, also known as *White Whine*, which depicts frustrations and complaints experienced only by privileged individuals in wealthy countries. They are typically used as a tongue-in-cheek comedic device to make light of trivial inconveniences.

- Bottom text: I'll be ready in 5 minutes.
- (26) Top text: And then I told him...
Image: Girls laughing and giggling.
Bottom text: I don't care about your money, I love you for who you are!
- (27) Top text: So then I told him...
Image: A group of doctors and nurses smiling and laughing.
Bottom text: "This won't hurt a bit."

Finally, an interesting source of text-image incongruity can be found in several instances that show images of actors in the new context of the meme, with an interpretation other than the ones that these images had in their initial context (film or TV series). The label *second-order incongruity* is proposed for this kind of incongruity, which exclusively applies to images taken from films or TV programs (e.g. stills) that are later re-used (and hence re-contextualised) in the meme. This time, the incongruity lies in the clash between the text and the image rather than in the clash between the image and the top-bottom texts *inside* the meme (in fact, very often texts and image are fully compatible). Expressed differently, the clash does not arise from the interpretation of the image within the meme; the clash actually takes place with that image as it had been initially produced and interpreted in the context of the film or the TV program. An example is provided in Figure 1, where a famous actor produces a certain facial expression provoked within some film-internal plot requirement. This expression is subsequently re-used in the meme with an accompanying text which, despite not being radically incongruous, does clash with the context and the first (i.e. *first-order*) interpretation in which the actor's gesture was originally appeared, and accordingly with the meaning of the image in that initial film-centred context too.

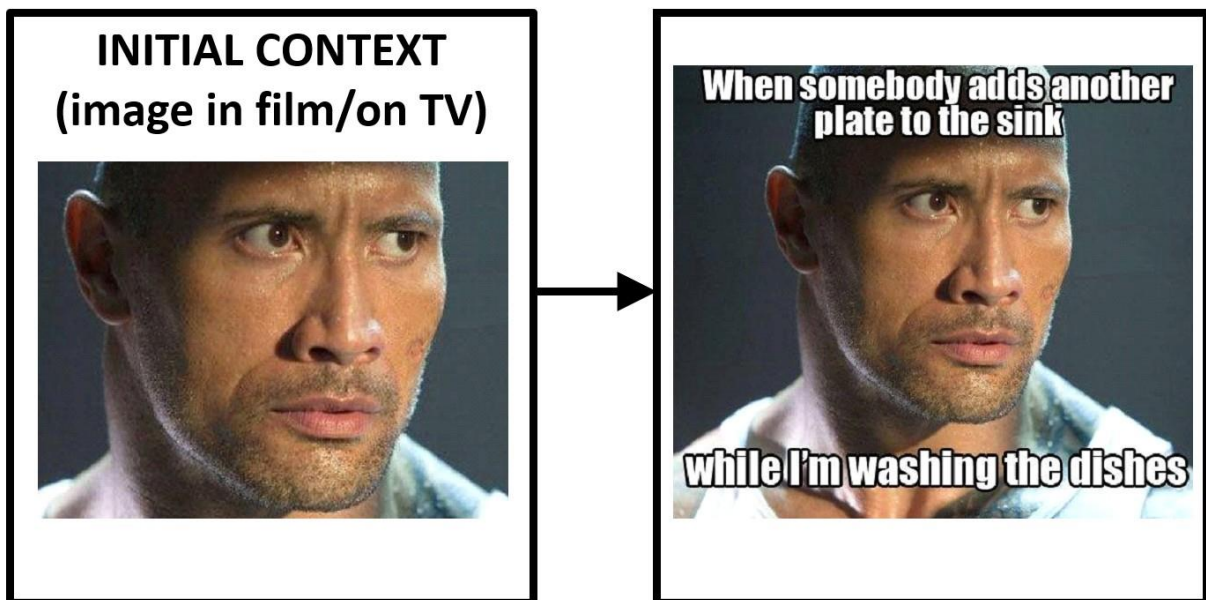


Figure 1. Second-order incongruity

3.5 Discourse-based incongruity where the image plays no role in the IR strategy

This category shows a discourse-based incongruity created without the aid of the meme image, and amounts to only 8% of the samples examined. This happens, for instance, with several memes whose IR strategy is based on the similar pronunciation corresponding to some stretch of discourse in the meme. By way of example, (28) below features a correlation between the pronunciation of *myself* and *my shelf*—the image of the actor seems to play no role whatsoever.

The same applies to (29), in which a similarity exists between *Al Gore rhythms* and *algorithms*, the drawing of a racoon being irrelevant in this IR strategy. Other memes play with different kinds of word ambiguity formats. Thus, meme (30) plays with the polysemy of *plain* and *plane*.⁹ The ambiguity may also extend to sentences and phrases, as illustrated by *get high* in (31), initially understood as “get high [on drugs]” (a relevant choice of an interpretation at first, then invalidated by the bottom text). The photo of a black man seems to play no significant role in this choice of an initially relevant interpretation that is subsequently invalidated, even though some biased viewers might find a connection between the race portrayed in the image and drug taking.¹⁰

- (28) Top text: A book fell on my head.
Image: Photo of actor Sean Connery.
Bottom text: I can only blame my shelf.
- (29) Top text: What do vice presidents listen to while doing math problems?
Image: Drawing of a racoon.
Bottom text: Al Gore rhythms.
- (30) Top text: Yo Dawg I heard you like planes.
Image: Photo of American hip hop artist Alvin Nathaniel Joiner.
Bottom text: So I bought you a plain full of plain planes so you can complain about yo plane or plain planes while you plane in yo plane of planes.
- (31) Top text: Let’s all get high.
Image: Photo of a black man.
Bottom text: Grades on our finals.

3.6 Discourse-based incongruity where the image aids in the IR strategy

This category, in which the accompanying image favours some kind of discourse-based incongruity typically triggered by some kind of ambiguity (homonymy, polysemy, homophony, similar pronunciation, to quote but a few/among others), accounts for 7.34% of the samples analysed. The similar pronunciations of *can’t* and *Kant* in (32), *handle* and *Handel* in (33), *Chopin* and *shopping* in (34), and the polysemy of *cabinet* in (35), are illustrative of this format/mechanism:

- (32) Top text: I want to stop philosophizing.
Image: A portrait of philosopher Kant.
Bottom text: But I Kant.
- (33) Top text: Bitches.
Image: A portrait of composer Handel.
Bottom text: Can’t Handel my oratorios.
- (34) Top text: Can’t we just make up...

9. Meme (26) belongs to the *Yo Dawg* (or *Sup Dawg*) series, based on portrait shots of American hip hop artist Alvin Nathaniel Joiner —better known by his stage name Xzibit. The humorous captions are composed around the recursive phrasal template “Yo Dawg, I herd you like (noun X), so I put an (noun X) in your (noun Y) so you can (verb Z) while you (verb Z).” Since rising to popularity in early 2007, this series has been considered one of the best-known and long-lasting examples of recursive humour on the Internet.

10. Meme (27) —taken from the *Successful Black Man* series— features a photo of a black man dressed in business attire superimposed upon a brown and beige colour background. The jokes in this series typically adopt a bait-and-switch format with the top caption appearing to set up a stereotype about African Americans followed by a bottom caption which renders the phrase innocuous, in a similar vein to what is done in the *Ordinary Muslim Man* series.

- (35) Image: Drawing of composer Chopin.
 Bottom text: And go Chopin.
 Top text: Hello, is that Ikea?...
 Image: Photo of British politician Theresa May.
 Bottom text: I need a new cabinet.

Frequent as word-based ambiguities may be, they abound at a phrase or utterance level too. This is typical in the *Insanity Wolf* series of memes, where some phrases and utterances could initially yield only one interpretation. However, the well-known image (especially among meme users with a certain *meme literacy*) leads viewers to expend more mental effort in entertaining a parallel interpretation of such phrases/utterances, with humorous effects compensating for that increased processing effort (having to accommodate several readings for the same chunk of text). Unlike the next category, where the image design inevitably aims at a single interpretation, here the image warns the viewer about a parallel interpretation, thus making it easier to obtain the intended dual-sense outcome when interpreting the meme.

Some examples are provided in (36-38). In (36), the viewer initially finds the metaphorical sense of the top text relevant enough to be selected as the intended interpretation, only to be invalidated by a more literal sense of the utterance (physical action). The viewer's expectations of relevance in (37) are not satisfied until two senses of *fast food* are simultaneously activated: one more accessible, and another which requires a connection with joggers and running (fast). Finally, (38) shows a simile in the top text which loses its figurative meaning (initially accessible and relevant) when the bottom text superimposes a more literal reading of the comparison between the source and target domains in that simile.

- (36) Top text: What's the fastest way to a woman's heart?
 Image: A fierce-looking wolf.
 Bottom text: Through her ribcage.
 (37) Top text: Joggers.
 Image: A fierce-looking wolf.
 Bottom text: Looks like I'm having fast food.
 (38) Top text: Friends are like trees.
 Image: A fierce-looking wolf.
 Bottom text: They fall down when you hit them multiple times with an axe.

3.7 Discourse-based incongruity where the image is essential in the IR strategy

Finally, this category —found in 6% of the sample memes— covers humorous strategies with an ambiguous top text where the image in the meme becomes essential in leading the viewer to select a single interpretation for that top text —subsequently invalidated when the bottom text is encountered. The viewers then realise that they have been fooled into selecting that initially relevant top text interpretation through the essential biasing role played by the image, which allows them to derive humorous effects out of the contrast between that initial interpretation and the one imposed on the viewer by the bottom text.

Examples include (39-41) below. Meme (39) —from the *Ordinary Muslim man* meme series¹¹— offers a combination of the word *covered* and the Muslim man's image that inevitably leads to select one of the senses corresponding to this polysemous word (i.e. to cover one's body

11. *Ordinary Muslim Man* is a series which features a smiling man wearing a Muslim taqiyah hat. The images are typically accompanied by bait-and switch captions with an initial statement that can be perceived as anti-American, followed by a phrase which renders the statement innocuous.

with clothes). However, the bottom text reminds the viewer of a parallel, initially less relevant sense of *covered* —to have insurance— which finally turns out to be the correct interpretation. As for (40), although the man’s image makes one interpretation of the ambiguous word *coming* (to reach orgasm) more likely to match the intended meaning, the bottom text ends up invalidating this initially relevant choice. The same happens in (41): albeit being presented as very accessible and initially relevant by the man’s image, the firstly assumed possible sense of *beating* is once again invalidated by the bottom text.

- (39) Top text: Unmarried women should be covered.
Image: A Muslim man.
Bottom text: By health insurance.
- (40) Top text: I am coming...
Image: A photo of a man with an ecstatic look on his face.
Bottom text: ...to help you with that spreadsheet.
- (41) Top text: I love beating women.
Image: A photo of a fierce-looking man.
Bottom text: To the door so I can hold it open for them.

Concluding remarks

This paper proposes seven IR strategy categories according to certain parameters: whether the incongruity is frame- or discourse-based and the role played by the image in the development of the incongruity-resolution that triggers a humorous outcome. With the exception of the category in which the incongruity results from combining the partial meanings of the top and bottom text and the image —and where the image necessarily has to play a key role in generating humorous effects— three basic roles of the image are suggested regarding the successful humorous outcome of both discourse-based and frame-based incongruities: (a) the image plays no role in the IR strategy; (b) the image aids in the IR strategy; and (c) the image is essential in the IR strategy.

The analysis of 150 sample memes taken randomly from the internet provided us with percentages corresponding to the ascription of memes to the different categories. The fourth one —i.e. IR derived from combining the partial meanings of texts and image— was the most frequent, which makes sense considering that, since memes are verbal-visual, users will find it easier to exploit the convergence of both semiotic modes for humorous purposes.

References

- Dynel, Marta. 2016. “‘I has seen image macros!’ Advice animal memes as visual-verbal jokes.” *International Journal of Communication* 10: 660-688.
- Nissenbaum, Asaf, and Limor Shifman. 2017. “Internet memes as contested cultural capital: The case of 4chan’s /b/ board.” *New Media & Society* 19(4): 483–501. doi: 10.1177/1461444815609313
- Segev, Elad, Asaf Nissenbaum, Nathan Stolero, and Limor Shifman. 2015. “Families and networks of Internet memes: The relationship between cohesiveness, uniqueness, and quiddity concreteness.” *Journal of Computer-Mediated Communication* 20: 417-433. doi: 10.1111/jcc4.12120
- Shifman, Limor. 2014. *Memes in Digital Culture*. Cambridge, MA: The MIT Press.
- Suls, Jerry M. 1972. “A two-stage model for the appreciation of jokes and cartoons: An information-processing analysis.” In *The Psychology of Humor: Theoretical Perspectives and Empirical Issues*, ed. by J.H. Goldstein, and P.E. McGhee, 81–100. New York: Academic Press.
- Suls, Jerry M. 1977. “Cognitive and disparagement theories of humour: A theoretical and empirical

- synthesis.” In *It’s a Funny Thing, Humour*, ed. by A.J. Chapman, and H.C. Foot, 41-45. Oxford: Pergamon Press. doi: 10.1016/B978-0-08-021376-7.50012-7
- Suls, Jerry M. 1983. “Cognitive processes in humor appreciation.” In *Handbook of Humor Research*, Vol. 1: *Basic Issues*, ed. by P.E. McGhee and J.H. Goldstein, 39-57. New York: Springer-Verlag. doi: 10.1007/978-1-4612-5572-7_3
- Tay, Geniesa. 2015. “Binders Full of LOLitics: Political Humour, Internet Memes, and Play in the 2012 US Presidential Election (And Beyond).” *European Journal of Humour Research* 2 (4): 46–73. doi: 10.7592/EJHR2014.2.4.tay
- Yus, Francisco. 2013. “An inference-centered analysis of jokes: The Intersecting Circles Model of humorous communication.” In *Irony and Humor: From Pragmatics to Discourse*, ed. by L. Ruiz Gurillo, and B. Alvarado, 59-82. Amsterdam and Philadelphia: John Benjamins. doi: 10.1075/pbns.231.05yus
- Yus, Francisco. 2016. *Humour and Relevance*. Amsterdam and Philadelphia: John Benjamins. doi: 10.1075/thr.4
- Yus, Francisco. 2017. “Incongruity-resolution cases in jokes.” *Lingua* 197: 103-122. doi: 10.1016/j.lingua.2017.02.002
- Yus, Francisco. 2018f. “Identity-Related Issues in Meme Communication.” *Internet Pragmatics* 1 (1): 113-133. doi: 10.1075/ip.00006.yus