

Reflections on developing Multimodal Metaphor Theory into Multimodal Trope Theory

Charles Forceville, University of Amsterdam, dept. of Media Studies
(c.j.forceville@uva.nl)

Abstract

The publication of Lakoff and Johnson's pioneering *Metaphors We Live By* (1980) launched Conceptual Metaphor Theory, which located the essence of this trope in cognition. This model entails that metaphors in language are no less but also no more than verbal manifestations of what is in the last resort a cognitive process. Unsurprisingly, scholars studying other discourses than (exclusively) verbal ones began to research how metaphors could be, and were, expressed both in co-speech gestures and in visual media. In more recent years, cognitivist scholars have begun to theorize and analyse verbal manifestations of other tropes besides metaphor, such as metonymy, antithesis, hyperbole, and irony. A logical next step is examining if, and if so, how, classic tropes can assume visual and multimodal forms. This paper discusses work that has been done in this area, launches some new proposals, and sketches *desiderata* of a truly "multimodal trope theory."

Mit der Publikation von *Metaphors We Live By* entwickelten Lakoff und Johnson 1980 den Ansatz der kognitiven Metapherntheorie. Er betont im Gegensatz zu anderen Ansätzen die Relevanz des bildlichen Sprachgebrauchs für das menschliche Denken, Verstehen und Handeln. Aufbauend auf den Überlegungen von Lakoff und Johnson entwickelte sich eine umfassende Forschung zu Sprachbildern im Alltagsdiskurs und es ist nicht überraschend, dass Forscher:innen auch damit begannen, die Relevanz und Bedeutsamkeit von Metaphern in Interaktion, Gestik sowie für visuelle Medien zu analysieren. Hier entwickelte sich neben dem Fokus auf die Metapher in den vergangenen Jahren ein zunehmendes Interesse für Sprachfiguren wie Metonymien, Antithesen, Hyperbeln oder die Ironie. Deren multimodale Dimensionen und Formen stellen jedoch nach wie vor ein Forschungsdesiderat dar, mit dem sich der vorliegende Beitrag kritisch auseinandersetzt. Sein Ziel besteht darin, erste Denkanstöße zu geben, die für die Entwicklung einer adäquaten „multimodalen Theorie der Sprachfiguren“ wichtig sind.

1. Introduction

Discussing the command of language a good poet possesses, Aristotle famously wrote in his *Poetics* that "it is important to use aptly each of the features mentioned [...] but much the greatest asset is a capacity for metaphor. This alone cannot be acquired from another, and is a sign of natural gifts: because to use metaphor well is to discern similarities" (1999: 115). However, it took many centuries before metaphor studies became truly popular, mostly thanks to Black (1979), Ortony (1979) and Lakoff and Johnson (1980). Particularly Lakoff and Johnson (1980) emphasized that metaphor is primarily a matter of thought and

only derivatively a matter of language, and thereby pioneered the influential Conceptual Metaphor Theory (CMT). Scholars such as Whittock (1990), Carroll (1994), and Forceville (1996, heavily indebted to Black 1979) took this idea seriously by embarking on metaphor research involving other modes than language, mainly focusing on visuals. Research in this area is still in full swing, not least because robust analyses of metaphor (as of any other phenomenon in discourse) need to be cognizant of (1) the combination of modes deployed; (2) the genre to which the discourse belongs; and (3) the medium in which it occurs. There are still many mode combinations, genres, and media to be studied.

But research needs to expand into a different direction as well. If “metaphor” is first and foremost a matter of thought, then surely other tropes are, too. It then makes sense to systematically start investigating which other tropes may be usefully claimed to have visual and multimodal manifestations. Within CMT the awareness that metonymy, though less spectacular than metaphor, is no less crucial in meaning-making gained ground in the early years of the 21st century (Barcelona 2000; Dirven/Pöring 2002). This insight in turn spawned research on metonymy in co-speech gesturing (e.g., Mittelberg/Waugh 2009), and in discourse involving visuals and written language, such as advertising (Pérez-Sobrino 2017).

The next step is to examine if, and if so, how, any other non-verbal and multimodal constellations besides metaphor and metonymy can be claimed to constitute tropes. An affirmative answer would require on the one hand defining each candidate trope in a mode-independent, conceptual manner, and on the other hand demonstrating how this candidate trope could manifest itself. Systematically addressing these questions requires joint efforts by scholars with expertise in rhetoric and scholars knowledgeable about visual and multimodal analysis (cf. Tseronis/Forceville 2017a).

In this paper I cannot but scratch the surface of these issues, expanding on ideas in Forceville (2010, 2019). Examining examples (some of them discussed in my earlier papers), I will say something about the role of mode, genre, and medium in analysing visual and multimodal manifestations of metaphor, metonymy, antithesis, hyperbole, and irony, and sketch some of the problems that need to be addressed by scholars motivated to extend classic verbal rhetoric into “Multimodal Trope Theory.”

2. Some preliminary assumptions

First of all, I subscribe to the view that all communication is governed by the relevance principle as proposed in Sperber and Wilson's relevance theory/RT (e.g., Sperber and Wilson 1995; Wilson and Sperber 2012; Clark 2013). In my formulation, slightly adapted from the original version, the central claim of relevance theory is that "every act of communication comes with the presumption of optimal relevance to its envisaged addressee" (Forceville 2020a: 99). Informally phrased, this means that each communicator tries to the best of his/her/+s ability to convince the envisaged audience of the message conveyed (an utterance, a letter, an advertising billboard, a political cartoon, a film scene...) that it is worth the attention of that audience; that it expresses pertinent information, attitudes, and/or emotions; and that it is in the audience's interest to invest mental energy to understand and (hopefully) accept the message's contents. The presumption (not: guarantee!) that a message is relevant thus amounts to the promise that the envisaged audience (which can vary from an individual to millions of people) will, in a microscopic or life-changing way, benefit from processing and accepting the message. It is to be noted that subscribing to the RT model entails recognizing that, in the last resort, *all* communication is rhetorical in that it aims to attain an effect on the envisaged audience in such a way that this audience changes (the strength of) its ideas about something at least partly on the basis of the communicated message.

Secondly, I will assume that one of the strategies that communicators have at their disposal to persuade audiences of the correctness and/or validity of ideas, perspectives, and attitudes is the use of tropes (cf. Tseronis 2021, and references quoted therein). Inasmuch as modern communication becomes ever more visual and multimodal, it is correspondingly more important to further theorize not only verbal but also non-verbal and multimodal tropes.

Thirdly, I propose it is impossible to fruitfully analyse multimodal tropes in any discourse - actually, to analyse *anything* in discourse - without taking into account the genre to which the trope (or other phenomenon) examined belongs (cf. e.g., Altman 1999; Neale 2000; Busse 2014; Frow 2015). Genre is the single most important pragmatic principle governing the interpretation of mass-communicative messages (Forceville 2020a: Chapter 5).

Fourthly, it is necessary to specify what is meant by “mode”. Embarrassingly, multimodality scholars have hitherto not been able to agree on a definition of mode, and this issue evokes heated debate (cf. for instance Bateman et al.’s [2020] response to Forceville [2020b]). There is no space here to go further into this debate. For present purposes it will have to suffice to present my mode candidates for mass-communication: visuals; written language; spoken language; music; sound; and bodily behaviour (the latter including touch, gestures, postures, manner of movement, and facial expressions; for more discussion, cf. Forceville 2021).

In the fifth place, it makes good theoretical sense to retain the distinction between monomodal and multimodal discourse – even though there is growing consensus that completely monomodal discourse is rare. After all, even a book without any pictures features visual elements such as different fonts, signals for chapter divisions, and margins, while purely spoken language cannot help but draw on the sound mode (pitch, loudness, timbre). In the present paper a trope will nonetheless be considered monomodal if its key elements (more on which below) can be identified via information in *one* mode only; it will be called multimodal if its key elements can only be identified by accessing information conveyed in at least two different modes.

Some further comments are in order here. It is to be noted that I have made “identification” of the key elements partaking in a trope a sufficient criterion for deciding whether a trope is monomodal or multimodal. But it may well be that although a trope’s *identification* is possible by drawing on a single mode – thereby making the trope “monomodal” – its *interpretation* may be enriched by, or even require, the input from (an)other mode(s). If it were to be decided that both identifiability and interpretability of a trope are necessary criteria for distinguishing between monomodal and multimodal tropes, the number of monomodal tropes would be considerably smaller than under the broad definition adhered to here. Clearly, there is a continuum from monomodal to multimodal tropes. That said, it is important to remember that monomodal metaphors remain the norm in (pictureless) books and spoken language, while even if one should adopt the strict definition it is possible to have completely monomodal metaphors in visuals, as there are discourses in which the visual mode suffices for both the identification and the interpretation of a trope.

3. Identifying and analysing visual and multimodal tropes

3.1 Identifying and analysing visual and multimodal metaphor

In Forceville (1996) I wrote extensively on the identification and interpretation of a single type of visual and multimodal trope, namely metaphor, adapting the model developed by Black (1979) – surely the single most important modern work on metaphor predating Lakoff and Johnson’s (1980) CMT – so as to make it work at a conceptual level. I would now formulate the procedure for identifying and interpreting a phenomenon as a metaphor as follows:

- (i) A discourse expresses, or suggests, two phenomena that in the given context belong to different semantic domains in such a way that it invites (or forces) equating them, as if they were the same phenomenon. The incongruity or salience of the equation or similarity-relation invites the judgment that it should not be taken literally – or that it should not be taken *only* literally (Lankjær 2016: 119; Forceville 2016: 25-26). How the similarity is created depends crucially on the medium in which the (supposed) metaphor occurs: similarity between visuals is created by different means than similarity between musical themes or between sounds or between gestures. In multimodal metaphors, similarity is typically signalled by salient synchronous cueing of target and source (cf. for some discussion Forceville 2006: 384-385).
- (ii) On the basis of (i), decide which of the two phenomena is the one that is (part of) the subject about which something is predicated (= the metaphor’s target) and which is the one that predicates something about the target (= the metaphor’s source). Verbalize the metaphor (irrespective of the mode(s) in which it occurs) in a TARGET A IS SOURCE B form, or in its dynamic equivalent: TARGET A-ING IS SOURCE B-ING.
- (iii) Resolve what feature(s) is/are to be mapped from source to target on the basis of (a) the context within the discourse; (b) the supposed intention of the communicator of the metaphor (rooted in the relevance principle); and (c) your knowledge of the world. It is crucial to realize that the emotions and valuations conventionally associated with the source domain (which may differ from one (sub)culture to another) are typically co-mapped onto the target. Stage (iii) amounts to interpreting the metaphor.

Here are some examples. In figure 1a, belonging to the genre of public service advertising, we see five girls in swimwear, located in what appears to be a shower block. Three girls are lined up and extend their hands in an odd, unnatural pose toward a slightly overweight, cowering girl, while a girl in the background sticks up her hand. Even without the accompanying text, which begins “One shot is enough,” many viewers will construe a metaphor that can be verbalized as *TAKING PICTURES OF SOMEONE AGAINST HER WILL IS EXECUTING THAT PERSON BY A FIRING SQUAD*. In presenting this metaphor UNICEF warns children against taking unflattering or private pictures of each other (and subsequently sharing them on social media). Clearly, the negative emotions and attitudes pertaining to the source domain are co-mapped onto the target domain. It is worth observing that some, but not all, viewers will need the accompanying text to construe the metaphor. For the former the metaphor functions as a monomodal one, for the latter as a multimodal one. The metaphor’s interpretability is of course aided by the fact that in English one can “shoot” both bullets and photographs.



Fig. 1a: “One shot is enough.” Public service ad by UNICEF (2015)



Fig. 1b: Cartoon by Chen Song, *China Daily* 18.07.2019



Fig. 1c: Screenshot from “The Wound” (Anna Budanova, 2013)

Figure 1b, part of the corpus analysed in Zhang/Forceville (2020), shows a Chinese political cartoon providing a perspective on the Sino-US trade conflict, presenting the metaphor *TRADE CONFLICT IS PLAYING TWO DIFFERENT GAMES*. At the moment of the cartoon’s publication presumably most viewers would not need the help of the text “trade talks” (written on the table) to construe the metaphor (which would thus be monomodal for these viewers). What is minimally mapped is the awareness that two opponents playing different games will by definition not be able to agree on the rules of the game – yielding the interpretation that any negotiations to resolve the trade conflict are bound to fail. As in figure 1a, the negative valuations of the source domain are co-mapped onto the target.

Figure 1c, a screenshot from the short animation film “The Wound,” discussed by Forceville/Paling (2021), depicts a monster. From the narrative context of the film it is clear that this monster is to be construed as the source domain of the metaphor DEPRESSION IS A MONSTER. Mappable features are scariness, unwantedness, dangerousness – and again, the negative emotions the source domain evokes are a crucial part of the mapping.

3.2 Identifying and analysing visual and multimodal metonymy

A mode-independent description of metonymy, adapted with minor changes from Forceville (2009: 58), is the following:

- (i) A metonym consists of a source concept/structure, which via a cue in a specific mode (language, visuals, music, sound, gesture ...) allows the metonym’s addressee to infer the target concept/structure.
- (ii) Source and target are, in the given context, part of the same conceptual domain.
- (iii) The choice of metonymic source makes salient one or more aspects of the target that otherwise would not, or not as clearly, have been noticeable, and thereby makes accessible the target *under a specific perspective*. The highlighted aspect may have an evaluative dimension.

Whereas the short-hand formula to capture metaphor is A IS B or A-ING IS B-ING, the short-hand formula for metonymy is B STANDS FOR A: we are given access to a source B, from which we infer target A. Metonymies are ubiquitous in pictures of all kinds if only because many pictures present an element that is, in fact, part of a bigger whole. Instances of such PART FOR WHOLE metonymies are, in a given context, FACE FOR PERSON, FLOWERBED FOR GARDEN, and GENERAL FOR ARMY. These are relatively conventional metonymies, of whose highlighting dimension we are usually not even aware, but *that* there is such a dimension becomes clear when we realize that other options are available, such as FINGERPRINT FOR PERSON, GRASS FOR GARDEN, and SOLDIER FOR ARMY. These latter offer a different perspective on the target than the first three. In another variety of metonymy, a typical specimen stands for the class, category, or entity to which it belongs. In figure 1b, for instance, the Chinese Go-player stands for China, while Uncle Sam stands for America. The fact that we could also say that Uncle Sam *symbolizes* America, incidentally, reminds us that in symbolism, too, we use the B STANDS

FOR A formula: in certain contexts, a flag stands for a country, a cross for Christ/suffering, a rose for love.

Figure 2a is an advertising billboard promoting a bank, ABN-AMRO, recognizable via its logo (a metonym for the bank) and the tag line (“Making more possible”). To find the ad relevant, we need to be aware that the object depicted is a grape; that wine is made from grapes; and that the French phrase *grand cru* refers to high quality wines. There is thus a metonymic part-whole relation between the grape (in the visual mode) and *grand cru* (in the written verbal mode): GRAPE STANDS FOR *GRAND CRU* WINE. Of course the source GRAPE is not coincidentally chosen: clearly, it takes a lot of work, and investments, to transform grapes into a *GRAND CRU* WINE – and this is where ABN-AMRO presumably ‘makes more possible’ by providing loans to invest in a wine-making business.

Figure 2b is the cover of Shaun Tan’s wordless *The Arrival*. In the context of this graphic novel, the suitcase is a metonym for travelling, here specifically for immigrating: SUITCASE STANDS FOR TRAVEL. Of course the written title cues the source domain, but to the extent that the suitcase is a fairly context-independent metonym for travel, we could say that the suitcase in many contexts functions as a *symbol* for travel.



Fig. 2a: Billboard from an advertising campaign by ABN-Amro bank, The Netherlands, ±2009

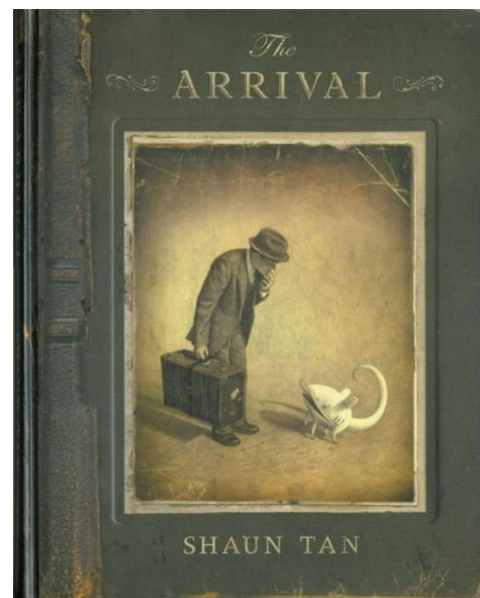


Fig. 2b: Cover of *The Arrival* by © Shaun Tan, Arthur A. Levine books/Lothian books 2006

3.3 Identifying and analysing visual and multimodal antithesis

On the basis of definitions by scholars of rhetoric, Tseronis and Forceville (2017b: 168) launch a proposal for criteria to identify a certain configuration as a visual or multimodal antithesis. Here is a slight rephrasing of that proposal:

- (i) Find a contrastive relation between two states of affairs, entities or persons ...
- (ii) that is conveyed by saliently presented stylistic means emphasizing both difference and similarity ...
- (iii) that, in the given context, gives rise to an awareness of diametrically opposed viewpoints, ideas, or interests associated with the two states of affairs, entities or persons.

I note in passing that Tseronis (2021) further pursues this line of thinking by making a distinction between antitheses (and metaphors, and allusions) that have (only) “rhetorical relevance, in the sense that they convey meaning which helps to frame the message for a particular audience and a particular situation,” and those that (also) provide “argumentative relevance,” namely “when the meaning conveyed by the figure contributes content that is somehow part of the argument (the claim and/or reasons) that may be recovered from the (multimodal) text” (2021: 378).



Fig. 3a: Screenshot from the documentary *Hospital*, Frederick Wiseman, USA 1969 © Zipporah Films



Fig. 3b: “Give the aids babies of Africa a chance.” Advertisement by Orange Babies, The Netherlands 2002

Figures 3a and 3b provide examples of antithesis. Figure 3a is a screenshot from a scene in the documentary *Hospital*, discussed in Tseronis and Forceville (2017b). A young black male prostitute tells his psychiatrist that his typical client looks like an average Wall Street worker with a suit and a tie, “and his hair combed to the side, looking like a billion dollars” (Tseronis/Forceville 2017b: 179). Toward the end of his utterance, the camera zooms out to reveal a poster of the then mayor of New York, hanging behind the black man – a striking exemplification of the latter’s stereotypical client. The antithesis could be phrased as something like “underprivileged low-status black male prostitutes typically have as their clients privileged high-status white men held up as role models for society”.

Figure 3b is a public service advertisement from the Orange Babies foundation that fights HIV in Africa. The main text runs, translated, “Give the AIDS babies of Africa a chance.” Building on the folk belief that babies are brought by storks, we can here construe an antithesis that can be formulated as “Whereas Western babies are auspiciously delivered by storks, African babies are ominously delivered by vultures.”

3.4 Identifying and analysing visual hyperbole

In order to define hyperbole in a mode-independent manner it is, as always, important to start with definitions and descriptions of its verbal manifestations. Aristotle asserts: “Effective hyperboles are also metaphors. [...] Hyperboles are adolescent, for they exhibit vehemence” (1991: 253). In her *Dictionary of Stylistics* Katie Wales provides the synonyms of “exaggeration” and “overstatement” for hyperbole, stating it is “often used for emphasis as a sign of great emotion or passion. Common phrases, often involving metaphor [...] at least imply an intensity of feeling, and add vividness and interest to conversation” (2001: 190). Burgers et al. (2016) adopt a cognitivist perspective, and begin by analysing and discussing various proposals to characterize hyperbole. Emphasizing that construing something as a hyperbole presupposes common knowledge of what is to be considered ‘normal’ in the everyday world – which pertains to knowledge of factual as well as of fictional events – they define hyperbole as “an expression that is more extreme than justified given its ontological [i.e. factual or fictional, ChF] referent” (2016: 166, emphasis in original). Peña-Cervel and Ruiz

de Mendoza-Ibáñez, although proposing some refinements, by and large accept this characterization of hyperbole:

We take sides with Burgers et al.'s (2016) claim that the clash with the context should be given primary status in the recognition of hyperbole. In our view [...] this is cognitively substantiated by postulating a cross-domain mapping from a hypothetical to a real scenario, which allows the hearer to pin down the nature of the speaker's emotional reaction including its intensity (2022: 188).

Although both Burgers et al. (2016) and Peña-Cervel/Ruiz de Mendoza-Ibáñez (2022) focus on the trope's verbal manifestations, their characterizations are mode-independent enough to help identify visual hyperboles.

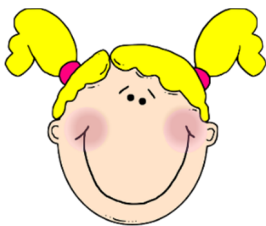


Fig. 4a: Hyperbolic smile¹

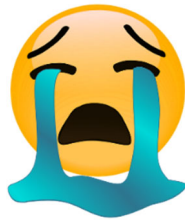


Fig 4b: Hyperbolic crying²



Fig 4c: Cartoon by Sempé. Provenance and year unknown

Figure 4a depicts a smile that cannot physically be procured. Similarly, the uninterrupted stream of tears of the emoji in figure 4b makes it hyperbolic. The Sempé cartoon suggests a degree of historical imagination that no tourist possesses. Examples 4a-4c also support the insight that hyperbole is “scalar” (Burgers et al. 2016: 164): the smile, the tear-flood, and the imagined scene could have been even bigger/larger/more detailed, but they could also have been smaller/less detailed – in the latter case crossing a border after which the expressions would no longer be hyperbolic. Moreover, all three emphasize, in line with Peña-Cervel and Ruiz de Mendoza-Ibáñez (2022), that the communicator aims to evoke an emotional response in the envisaged addressee, namely of joy, sadness, and humorous ridicule, respectively.

¹ Source: <https://pixabay.com/nl/vectors/meisje-vrolijk-glimlach-vrouwelijk-311674/> (21.08.2022).

² Source: Christian Dorn, <https://pixabay.com/nl/illustrations/smiley-huilend-rouw-verdrietig-5566743/> (21.08.2022).

3.5 Identifying and analysing visual and multimodal irony

The *Dictionary of Stylistics* characterizes irony as follows: “the words actually used appear to contradict the sense actually required in the context and presumably intended by the speaker” (Wales 2001: 224). Burgers et al. (2011), discussing and evaluating different approaches to verbal irony, propose that all of them agree on four aspects:

(1) irony is implicit, (2) irony is evaluative, and it is possible to (3) distinguish between a non-ironic and an ironic reading of the same utterance, (4) between which a certain type of opposition may be observed. Of course, an ironic utterance is also usually directed at someone or something; its target (Burgers et al. 2011: 189).

On this basis they define irony as “an utterance with a literal evaluation that is implicitly contrary to its intended evaluation” (2011: 190). Although Burgers et al. (2011) approvingly mention the relevance theory perspective on irony (Sperber/Wilson 1995: 237-243), they do not discuss it at length. In a later formulation Wilson and Sperber state that “irony [...] rests on the perception of a discrepancy between a representation and the state of affairs it purports to represent. [...] Ironical utterances [...] are a loosely defined sub-class of echoic utterances” (Wilson/Sperber 2012: 94).

Peña-Cervel/Ruiz de Mendoza-Ibáñez “take sides with Wilson and Sperber (2012) and support their claim that echoing is key to explaining irony” (2022: 235), but maintain that relevance theory undertheorizes the role of pretense: “irony is almost invariably complemented by pretense since in verbal irony we find the speaker’s simulation of a belief or thought” (ibid.: 236).

Scott (2004) addresses the issue how irony can occur in photographs, discussing both multimodal ironies of the verbo-visual variety (e.g., photographs by Margaret Bourne-White and Dorothea Lange) and monomodal visual ironies (e.g., photographs by Elliott Erwitt, Barbara Kruger, Jones Griffiths, and Cindy Sherman). She proposes that “some of the defining properties of irony” can be listed as follows:

- An ideological component, which sets two orders of reality and associated belief systems into conflict with each other.
- A dissembling component, or at least an element of differential awareness, between the ironist-cum-audience and the unwitting victim of irony.

- An incongruity, which alerts the viewer to either the intention or the potential for irony (2004: 35).

Scott finds Sperber and Wilson's approach to irony as a form of "echoic mention" useful, emphasizing that for a purely visual irony to work, the viewer must be able to recognize not just the "echoing" but also the "echoed" element. She proposes

that if a system of beliefs is readily enough available (the very notion of "the usual scheme of things" entails a system of belief), and that if an image can bring to mind this belief system by means of an easily identifiable symbol, then we do not need words in order to access a dominant representation. Once a world view has been summoned, the remainder of the picture must in some way question it in order to achieve ironic effect (2004: 43).

Scott summarizes the essence of what makes the work of the photographers she discusses ironical by pointing out that

they set up a frame of reference, and then subvert it by means of an incongruity. In so doing, they reveal the dominant representation not to be definitive. In all cases, the recognition of a differential awareness between ironist and victim enhances the sense of incongruity (2004: 47).

On the basis of the above sources, let me risk the following mode-independent definition:

Irony holds, or can be construed to hold, when a discourse in any medium presents an evaluation of a state of affairs it purports to represent by explicitly or implicitly echoing a previous, literal discourse of that state of affairs in such a way that the echoic discourse makes transparent a discrepancy between the echoic and the echoed evaluation of the state of affairs at stake.

Figure 5 provides some visual and multimodal examples.



Fig. 5a: Ashtray³



Fig 5b: Ironic traffic sign warning against drinking and driving.



Fig 5c: Plaque commemorating Alois Alzheimer⁴

Figures 5a, 5b, and 5c draw on what Scott calls *icons* (2004: 42), including a visual entity with a ‘coded’ meaning (Forceville 2020a: Chapter 6). Figure 5a presupposes our awareness that an ashtray (the “echoed discourse”) is normally used to tip cigarette ash in, whereas the *no-smoking* pictogram provides the iconic or pictogrammatic ‘echoing’ evaluation that “here it is forbidden to smoke”. While figure 5b might seem to be ironical on a purely verbal level (“Go ahead – drink & drive”), I would argue that the type of ‘traffic sign’ on which the text appears visually reinforces the irony, as its colour reveals it to be an *instruction* sign, not a *forbidding* or *warning* sign (for more discussion, cf. Forceville and Kjeldsen 2018). Similarly, the written-verbal mode alone suffices to make figure 5c (appearing on the English Wikipedia page) ironical. The plaque commemorates Alois Alzheimer, discoverer of the illness mainly responsible for dementia, with pathological forgetfulness as its main symptom. The written text underneath translates as “Alois, we will never forget you.” But as in figure 5b, the design and colour of the ‘echoed’ discourse help identify it, namely as an official commemoration plaque, while the graffiti style of the hand-written comment signals its *unofficial* nature, which adds a visual dimension to the ‘echoing’ comment, for instance as it is likely that the graffiti will at one time be painted over, that is, “forgotten” (for an example of a monomodal musical irony, cf. Forceville 2020a: 235).

³ Source: <https://highjimmie.com/collections/ashtrays/products/at-white-no-smoking> (21.08.2022).

⁴ Source: <https://en.wikipedia.org/wiki/Irony> (21.08.2022).

3.6 Combinations of (visual and multimodal) tropes

The identification and interpretation of verbal tropes is yet further complicated by the fact that two or more tropes can actually occur together. Burgers et al. (2018), analysing a corpus of Dutch newspaper text, chart not only occurrences of metaphor-only, hyperbole-only, and irony-only, but also monitor these tropes' various permutations. While combinations of tropes are less frequent than their isolated occurrence, the authors find a substantial number (although combinations of metaphor *and* hyperbole *and* irony were rare). Clearly, inasmuch as tropes can be reliably distinguished from one another, it makes good sense to broaden analyses like those of Burgers et al. (2018) to the non-verbal and multimodal realm.



Fig. 6.1: Banksy street art: Barcode and Leopard (analysed by Poppi and Kravanja 2019)

Poppi and Kravanja (2019), analysing Banksy's street art, argue that a full interpretation of figure 6.1 requires identifying both a metaphor and an antithesis. The metaphor can be verbalized as *BARCODE IS CAGE*. The authors in addition postulate the antithesis *CAPTIVITY VS. FREEDOM* (2019: 91). I propose that this antithesis could be construed as something like "barcodes facilitate people's freedom to consume while simultaneously constituting a trap from which they want to escape." Moreover, Poppi and Kravanja acknowledge (without further elaboration) that 'irony' often also plays an important role in Banksy's work (2019: 86), as do Peña-Cervel and Ruiz de Mendoza-Ibáñez (2022: 245).



Fig. 6.2a: Damaged brain is clouded sun.

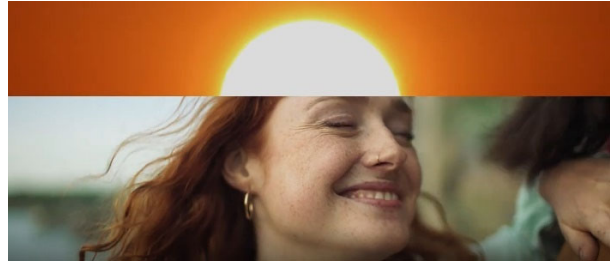


Fig. 6.2b: Healthy brain is shining sun.



Fig. 6.2c: Damaged brain is fading dandelion

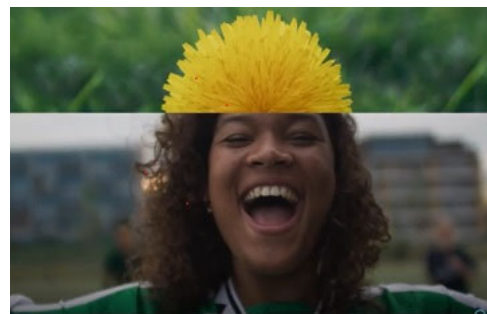


Fig 6.2d: Healthy brain is flowering dandelion

Screenshots from *Hersenstichting* commercial, the Netherlands⁵

Figures 6.2a-6.2d are screenshots from a commercial commissioned by the Dutch *Hersenstichting* ("Brain foundation"), which promotes research to prevent or slow down brain damage. The voice-over text can be translated as follows:

Try to imagine what it is your brains do ...They let you talk, laugh, enjoy ... Try to imagine something happens to your brains. A brain affliction keeps over your life. Try to imagine that everybody has healthy brains. *That* is our goal. Check out [Hersenstichting.nl](https://www.hersenstichting.nl).

Figures 6.2a and 6.2b express (monomodal) visual metaphors that can be verbalized as DAMAGED BRAIN IS CLOUDED SUN and HEALTHY BRAIN IS SHINING SUN, respectively, while figures 6.2c and 6.2d express DAMAGED BRAIN IS FADING DANDELION and HEALTHY BRAIN IS FLOWERING DANDELION, respectively. Routinely, they also feature the part-for-whole metonym CLOSE-UP OF FACE STANDS FOR PERSON. Arguably 6.2a and 6.2c also feature hyperbole in the commercial: the speed with which the clouds darken the sun and the speed with which the

⁵ Source : https://www.youtube.com/watch?v=HMSbmKIDI_A (21.08.2022).

dandelion disperses is much higher than the time it takes for brains to deteriorate. And finally we can also, in combination with the voice-over text, construe an antithesis that might run something like “healthy brains make for happy lives while damaged brains make for unhappy lives”.

Combinations of tropes also occur in print advertising. Given that metaphorical target and source domains are often visually represented via part-for-whole metonymies it is actually likely that most visual and multimodal metaphors automatically involve metonymy (cf. Pérez-Sobrino 2017; Kashanizadeh/Forceville 2020 for discussion of metaphor-metonymy combinations in print advertising).

Similarly, while figure 3a was analysed as an example of antithesis, there surely is also a strong sense that it exemplifies irony. Conversely, figure 5a, analysed as ironical, also suggests “antithesis”. And arguably, figure 4c is not just hyperbolic, but also ironic.

A caveat is in order, though. However important tropes are in persuasion, as theorists and analysts we should not make the mistake to try and squeeze *all* elements that partake in meaning-making in discourse into the mould of one or more tropes. There are many meaning-making elements that simply cannot be accommodated in a catalogue of tropes. It is sensible to try and distinguish between tropes and the many other (types of) meaning-making mechanisms operating in discourse (for examples of this approach cf. Guan/Forceville 2020; Zhang/Forceville 2020).

4. A ‘script’ for developing Multimodal Trope Theory

As suggested above, developing a robust, reliable multimodal trope theory needs to begin by reconsidering the catalogue of ‘classical’ verbal tropes, of which I have discussed only some in this paper. This entails revisiting classic rhetoric (Aristotle, Quintilian, Cicero ...) and to try and extract a supra-modal ‘essence’ from these tropes, rephrasing this essence in terms of criteria in such a way that it can serve as a heuristic irrespective of medium, mode, and genre. Such reformulations will benefit from explicitly using the *target domain* and *source domain* terminology, and specifying how the use of the source transforms the explicit or implicit literal target. It will help to think of test questions to distinguish between different tropes (e.g., when is something a metonym, and

when is it (also) a symbol? How can we differentiate between metaphor, symbolism and allegory? [for a discussion of visual allegory, cf. Cornevin/Forceville 2017]). It makes good sense to first focus on the tropes' verbal manifestations by collecting and analysing a vast number of attested instances of these tropes in order to determine what unites these examples. This should then lead to formulating a supra-modal definition, which can then help find supposed non-verbal and multimodal manifestations.

It is recommendable to *verbalize* all proposed candidates for non-verbal and multimodal tropes, drawing on the terms *target* and *source*, to facilitate checking against the definition of the trope at stake. That said, the analyst should remain aware that *any* verbalization is necessarily no more than a poor approximation of how the trope appears in the original discourse when the trope is non-verbal or multimodal. Moreover, no verbalization is value-free, and different verbalizations may steer different emphases in interpretation.

Assessing whether it is mandatory to analyse a certain phenomenon as exemplifying a certain trope or whether this is optional is both challenging and crucial. In some cases a specific entity only makes sense (i.e., is only relevant) if it is understood as cueing another entity, and thereby constitutes a trope of some sort. In other cases the tropical interpretation is optional, or requires taking into account a broader context than the discourse within which it appears, or is only accessible to interpreters with specific background information.

It is important, moreover, not to look at classical rhetoric for guidelines with too much deference. I think Peña-Cervel and Ruiz-de Mendoza-Ibáñez (2022) make important forays into analysing tropes from an inclusive, cognitive perspective by proposing not only how certain tropes can be clustered hierarchically, but also by proffering cognitive operations and tests to identify specific tropes as well as to distinguish between them.

On the basis of good supra-modal definitions of the various tropes, it becomes possible to address non-verbal manifestations of such tropes. It remains useful to distinguish between monomodal visual (or musical, or sonic, or gestural ...) and multimodal varieties – analyses of the latter requiring expertise in (at least) two modes. It is furthermore fundamental to be optimally open to the affordances and constraints that necessarily characterize each 'mode', as it is highly likely that not all modes display the range of tropes that verbal discourse can. After all, (written and spoken) language has a grammar and a vocabulary,

while other modes at best have structures. For instance, it deserves closer inspection whether what Teng and Sun (2002) discuss as visual “oxymoron” is not the same, after all, as what Tseronis and Forceville (2017b) and Poppi and Kravanja (2019) theorize in terms of “antithesis”. Conversely, it may be the case that there are phenomena in non-verbal and multimodal discourse not appearing in language that nonetheless deserve the label of “trope”. A candidate is Teng and Sun’s (2002) “pictorial grouping”. Similarly, as Wells (1998: 69) points out, one of the most pervasive phenomena in animation film is ‘transformation’: one thing ‘morphs’ into another thing in a way that does not necessarily enable construal as, say, a metaphor or antithesis. Should we, perhaps, promote ‘metamorphosis’ to *trope*-status, on the basis that it is a patterned way to suggest non-literality in animation?

5. By way of conclusion

In this paper I have argued that cognitivist-oriented work on visual and multimodal metaphor and metonymy can serve as a starting point for developing an inclusive Multimodal Trope Theory. While proposals on some multimodal tropes (e.g., antithesis and allegory) had been tentatively addressed in previous cognitivist-oriented research, others (e.g., hyperbole and irony – but also symbolism) are virtually untheorized. To stimulate discussion, examples of some of these latter have been cautiously discussed here. It was pointed out that some examples arguably show traits of two different tropes – something that deserves sustained scrutiny in the examination of other examples in future research.

Crucially, the ambitious project of developing an inclusive Multimodal Trope Theory needs to begin with cognitivist-oriented analyses of *verbal* tropes – which in turn can benefit from classical and modern rhetoric and argumentation theory. Key aspects of the project are examining how specific tropes are both different from, and similar to, each other; which tropes can and which cannot co-occur; and whether it is possible to (hierarchically) cluster various tropes in terms of how they create meaning. In this respect, Peña-Cervel and Ruiz de Mendoza-Ibáñez (2022) have done trail-blazing work by defining the supra-modal essence of a number of tropes on the basis of their *verbal* manifestations. This will, in turn, make it possible to venture further into charting these tropes’ non-verbal and multimodal expressions.

There is a lot of work waiting to be done. In preparing to do this work, let us never forget that models are there to account for and explain data – and not the other way round. Reality has the irritating habit of always turning out to be more complex than the models we build to explain it. Once we discover a new complexity, we thus need to adapt our models, while simultaneously bearing in mind that categorizations are a means for better understanding the world, not a goal in themselves.

That said, the study of multimodal tropes, as a subdiscipline of multimodal discourse in general, is a highly worthwhile pursuit within humanities research. Mass-communication is becoming more, rather than less, multimodal. The project of exploring how tropes that have long been considered exclusively verbal phenomena can function in non-verbal and multimodal discourses will ultimately benefit all scholars studying communication – and may help linguists get rid of the prejudice that communication simply is another word for language.

Acknowledgments. For several examples, and part of the analyses thereof, I am indebted to students in my metaphor course, who found and discussed them in their essays and theses. I am grateful to Denis Jamet and Adeline Terry for challenging me to develop my ideas for a presentation at their conference at Jean Moulin Lyon 3 University, and to invite me to transform this presentation into a paper. I also want to thank two anonymous reviewers for their comments on an earlier draft of this paper. They helped me improve it and correct an irritating error. Of course, I alone remain responsible for its contents.

6. References

Altman, Rick (1999): *Film/Genre*, British Film Institute.

Aristotle (1991 [4th c. BC]): *On Rhetoric* (translated and edited by G.A. Kennedy), Oxford: University Press.

Aristotle (1999 [4th c. BC]): “Poetics” (translated and edited by Stephen Halliwell; revised by Donald A. Russell), in: *Aristotle, Poetics; Longinus, On the Sublime; Demetrius, On Style* (28-141), Loeb Classical Library no. 199, Harvard: University Press.

Barcelona, Antonio (ed.) (2000): *Metaphor and Metonymy at the Crossroads. A Cognitive Perspective*, Berlin/New York: Mouton De Gruyter.

- Bateman, John A./Wildfeuer, Janina/Hiippala, Tuomo (2020): "A question of definitions. Foundations for multimodality – A response to Charles Forceville's review", in: *Visual Communication* 19(2), 317–320.
- Black, Max (1979): "More about metaphor", in: Ortony, Andrew (ed.): *Metaphor and Thought*, Cambridge: University Press, 19-43.
- Burgers, Christian/Van Mulken, Margot/Schellens, Peter Jan (2011): "Finding irony. An introduction of the Verbal Irony Procedure (VIP)", in: *Metaphor and Symbol* 26(3), 186-205.
- Burgers, Christian/Brugman, Britta C./Renardel de Lavalette, Kiki Y./Steen, Gerard J. (2016): "HIP. A method for linguistic hyperbole identification in discourse", in: *Metaphor and Symbol* 31(3), 163-178.
- Burgers, Christian/Renardel de Lavalette, Kiki Y./Steen, Gerard J. (2018): "Metaphor, hyperbole, and irony. Uses in isolation and in combination in written discourse", in: *Journal of Pragmatics* 127, 71-83.
- Busse, Beatrix (2014): "Genre", in: Stockwell, Peter/Whitely, Sara (eds.): *The Cambridge Handbook of Stylistics*, Cambridge: University Press, 103-116.
- Carroll, Noël (1994): "Visual metaphor", in: Hintikka, Jaakko (ed.): *Aspects of Metaphor*, Dordrecht et al.: Kluwer, 189-218.
- Clark, Billy (2013): *Relevance Theory*, Cambridge: University Press.
- Cornevin, Vanessa/Forceville, Charles (2017): "From metaphor to allegory. The Japanese manga *Afuganisu-tan*", in: *Metaphor and the Social World* 7(2), 235-251.
- Dirven, René/Pöring, Ralf (eds.) (2002): *Metaphor and Metonymy in Comparison and Contrast*, Berlin/New York: De Gruyter Mouton.
- Forceville, Charles (1996): *Pictorial Metaphor in Advertising*, London/New York: Routledge.
- Forceville, Charles (2006): "Non-verbal and multimodal metaphor in a cognitivist framework. Agendas for research", in: Kristiansen, Gitte/Achard, Michel/Dirven, René/Ruiz de Mendoza-Ibáñez, Francisco José (eds.): *Cognitive Linguistics. Current Applications and Future Perspectives*, Berlin/New York: De Gruyter Mouton, 379-402.
- Forceville, Charles (2009): "Metonymy in visual and audiovisual discourse", in: Ventola, Eija/Moya Guijarro, Arsenio Jesús (eds.): *The World Told and the World Shown: Issues in Multisemiotics*, New York: Palgrave MacMillan, 56-74.

- Forceville, Charles (2010): "Why and how study metaphor, metonymy, and other tropes in multimodal discourse?", in: Soares da Silva, Augusto/Cândido Martins, José/Magelhães, Luísa/Gonçalves, Miguel (eds.): *Comunicação, Cognição e Media*, Vol. I, Braga: Aletheia/Associação Científica e Cultural, Faculdade de Filosofia, Universidade Católica Portuguesa, 41-60.
- Forceville, Charles (2016): "Visual and multimodal metaphor in film: Charting the field", in: Fahlenbrach, Kathrin (ed.): *Embodied Metaphors in Film, Television, and Video Games*, New York/London: Routledge, 17-32.
- Forceville, Charles (2019): "Developments in multimodal metaphor studies. A response to Górska, Coëgnarts, Porto & Romano, and Muelas-Gil", in: Navarro i Ferrando, Ignasi (ed.): *Current Approaches to Metaphor Analysis in Discourse*, Berlin/Boston: De Gruyter Mouton, 367-378.
- Forceville, Charles (2020a): *Visual and Multimodal Communication. Applying the Relevance Principle*, Oxford University Press.
- Forceville, Charles (2020b): Book review of John A. Bateman/Wildfeuer, Janina/Hiippala, Tuomo, *Multimodality. Foundations, Research and Analysis – A Problem-Oriented Introduction*, De Gruyter Mouton 2017, in: *Journal of Visual Communication* 19(1), 157-160.
- Forceville, Charles (2021): "Multimodality", in: Xu Wen/Taylor, John R. (eds.): *The Routledge Handbook of Cognitive Linguistics*, New York/London: Routledge, 676-687.
- Forceville, Charles/Kjeldsen, Jens E. (2018): "The affordances and constraints of situation and genre. Visual and multimodal rhetoric in unusual traffic signs", in: *International Review of Pragmatics* 10(2), 158-178.
- Forceville, Charles/Paling, Sissy (2021): "The metaphorical representation of DEPRESSION in short, wordless animation films", in: *Visual Communication Journal* 20(1), 100-120 (published ahead of print 21.09.2018).
- Frow, John (2006): *Genre*, New York/London: Routledge.
- Guan, Yue/Forceville, Charles (2020): "Making cross-cultural meaning in five Chinese promotional clips. Metonymies and metaphors", in: *Intercultural Pragmatics* 17(2), 123-149.
- Kashanizadeh, Zahra/Forceville, Charles (2020): "Visual and multimodal interaction of metaphor and metonymy. A study of Iranian and Dutch print advertisements", in: *Cognitive Linguistic Studies* 7(1), 78-110.
- Lakoff, Georges/Johnson, Mark (1980): *Metaphors We Live By*, Chicago: University Press.

- Lankjær, Birger (2016): "Problems of metaphor, film, and visual perception", in: Fahlenbrach, Kathrin (ed.): *Embodied Metaphors in Film, Television, and Video Games*, London/New York: Routledge, 115-128.
- Mittelberg, Irene/Waugh, Linda R. (2009): "Metonymy first, metaphor second. A cognitive-semiotic approach to multimodal figures of thought in co-speech gesture", in: Forceville, Charles/Urios-Aparisi, Eduardo (eds.): *Multimodal Metaphor*, Berlin/Boston: De Gruyter Mouton, 329-356.
- Neale, Steve (2000): *Genre and Hollywood*, London/New York: Routledge.
- Ortony, Andrew (ed.) (1979): *Metaphor and Thought*, Cambridge: University Press.
- Peña-Cervel, María Sandra/Ruiz de Mendoza-Ibáñez, Francisco José (2022): *Figuring out Figuration. A Cognitive Linguistic Account*, Amsterdam/Philadelphia: John Benjamins.
- Pérez-Sobrino, Paula (2017): *Multimodal Metaphor and Metonymy in Advertising*, Amsterdam/Philadelphia: John Benjamins.
- Poppi, Fabio Indio Massimo/Kravanja, Peter (2019): "Actiones secundum fidei. Antithesis and metaphoric conceptualization in Banksy's graffiti art", in: *Metaphor and the Social World* 9(1), 84-107.
- Scott, Biljana (2004): "Picturing irony. The subversive power of photography", in: *Visual Communication* 3(1), 31-59.
- Sperber, Dan/Wilson, Deirdre (2005 [1986]): *Relevance. Communication and Cognition*, Oxford: Blackwell.
- Teng, Norman Y./Sun, Sewen (2002): "Grouping, simile, and oxymoron in pictures. A design-based cognitive approach", in: *Metaphor and Symbol* 17, 295-316.
- Tseronis, Assimakis (2021): "From visual rhetoric to multimodal argumentation. Exploring the rhetorical and argumentative relevance of multimodal figures on the covers of *The Economist*", in: *Visual Communication* 20(3), 374-396.
- Tseronis, Assimakis/Forceville, Charles (eds.) (2017a): *Multimodal Argumentation and Rhetoric in Media Genres*, Amsterdam/Philadelphia: John Benjamins.
- Tseronis, Assimakis/Forceville, Charles (2017b): "The argumentative relevance of visual and multimodal antithesis in Frederick Wiseman's documentaries", in: Tseronis, Assimakis/Forceville, Charles (eds.): *Multimodal Argumentation and Rhetoric in Media Genres*, Amsterdam/Philadelphia: John Benjamins, 165-188.
- Wales, Katie (2001): *A Dictionary of Stylistics*, Harlow: Longman.

- Wells, Paul (1998): *Understanding Animation*, London/New York: Routledge.
- Whittock, Trevor (1990): *Metaphor and Film*, Cambridge: University Press.
- Wilson, Deirdre/Sperber, Dan (2012): *Meaning and Relevance*, Cambridge: University Press.
- Zhang, Cun/Forceville, Charles (2020): "Metaphor and metonymy in Chinese and American political cartoons (2018-2019) about the Sino-US trade conflict", in: *Pragmatics & Cognition* 27(2), 476-501.