Developments in multimodal metaphor studies: A response to Górska, Coëgnarts, Porto & Romano, and Muelas-Gil.

Forceville, C.

Published in: Current Approaches to Metaphor Analysis in Discourse

DOI: 10.1515/9783110629460-017

Citation for published version (APA):
This is a pre-proof version of a chapter that was subsequently published as:


Colour-marked words and expressions appear in the book’s index. If you want to quote from the chapter *verbatim*, please consult the published version (ChF, September 2019).

“Developments in multimodal metaphor studies: a response to Górska, Coëgnarts, Porto & Romano, and Muelas-Gil”

Keywords: conceptual metaphor theory; multimodal metaphor; multimodal metonymy; non-verbal tropes; image schemas;

1 Introduction

It is both encouraging and telling that a volume entitled *Current approaches to metaphor analysis in discourse* has a robust section devoted to “Metaphor analysis in multimodal discourse”. While for more than a decade since Lakoff and Johnson’s (1980) paradigm-changing *Metaphors we live by* the study of conceptual metaphor (and increasingly: conceptual metonymy) was more or less equivalent with the study of its verbal manifestations, nowadays this line of research is complemented by a healthily strong, and growing, body of work analysing visual, gestural, and multimodal expressions of metaphor and metonymy. This broadening of the discipline is excellent news for several reasons. In the first place, *Conceptual Metaphor & Metonymy Theory (CMMT)* has provided tools for the analysis of discourses in media that do not draw (exclusively) on the verbal mode, such as cartoons, comics, films, commercials, and music (e.g., Forceville and Urios-Aparisi 2009), but also aid the development of the quickly growing discipline of “multimodality” (e.g., Kress and Van Leeuwen 2001, 2006; Kress 2010; Jewitt 2014; Bateman 2008, 2014; Bateman et al. 2017; Machin 2014; Klug and Stöckl 2016). As a result, CMMT turns truly
interdisciplinary. But this comes at a cost: seriously becoming an interdisciplinary scholar requires becoming an expert in at least two media. Since hitherto the vast majority of CMMT scholars have been trained as linguists, they need to invest time and energy in learning about another medium or at least another mode. If they do not make this effort, they run the risk of making painful mistakes in their analyses because of insufficient awareness of medium-specific and mode-specific affordances and constraints.

But even CMMT linguists wishing to restrict themselves to language will need to develop some familiarity with visual and multimodal manifestations of metaphor and metonymy, at least if they are interested in the cognition aspect of these tropes. What has hitherto not been sufficiently acknowledged is that the affordances characterizing modes other than language enable them to metaphorize in ways that are difficult, or even impossible, to achieve in verbal form (e.g., Forceville and Paling 2018). Fortunately, the four papers in the “Metaphor analysis in multimodal discourse” section of this book robustly build on this awareness.

2 Response to Górska, Coëgnarts, Porto & Romano, and Muelas-Gil

The cartoons by the Polish artist Janusz Kapusta examined by Elżbieta Górska show how profound ideas can be presented in deceptively simple verbo-pictorial form. Górska persuasively argues that “image schemas” such as UP-DOWN, FORCE, BALANCE etc. (Johnson 1987; Hampe 2005, 2017) are crucial to helping us understand the visual part of the cartoons. Undoubtedly, the salient role of strongly embodied image schemas contributes to the cartoons’ universal appeal. From a multimodal point of view, it would be an interesting test to see how much of the visual information makes sense even without any verbal text at all. Viewers would definitely have to recognize the recurring protagonist as a “blend” between a Buddha and a chess pawn. The Buddha part suggests the creature’s aspiration to, say, achieve wisdom, whereas its pawn part emphasizes its “everyman” character. To understand this blend, one would thus have to be able to both recognize Buddha and pawn (the least valuable and least unique piece in chess) and to know something about their cultural meaning. The cartoon in example 1, incidentally, shows that elements in visuals are by no means always Peircean icons: the interrupted lines, so important in the overall meaning of the cartoon, function in a very different way than the contour lines that make up its “characters”. I find it intriguing that in this cartoon the circle is used as a bounded space for “pain” whereas the rectangle is used for “love” – as the accompanying text suggests. This is perhaps somewhat counter-intuitive, as edgy things are conventionally considered negative and harsh (“takete”),
whereas round things are positive and soft (“maluma” – see Kennedy 1982: 602). Moreover, while Górska interprets the interrupted lines as cueing “the temporaries of the two states” (Górska 2019/this volume: 5), I submit that the interrupted nature of the lines also suggest that love and pain are permeable – and interact, just as their overlapping does.

The cartoons trigger additional meanings that are attributable to their visual part alone. In example 2 the “line” is not just the instrument that helps the character maintain balance; it is also something that the character firmly “holds on” to (via the GRASP schema, as Górska points out) as if it were a crutch. The character in example 3 strikes me as walking on a relatively less-curved – and hence “wise” – path, avoiding the more curved – and thus “more stupid” – paths that are visible. And in example 4 the figure-ground reversal – the protagonist is “in” his suitcase-full-of-worries rather than carrying it – too, is only conveyed visually.

I fully agree with Górska that “the affordances of the visual mode allow for a straightforward realization of a number of image schemas, namely: the BOUNDED SPACE, CYCLE, LINK, NEAR-FAR, and the PART-WHOLE” (Górska 2019/this volume: 5). As always, the precise meaning of any schema will depend on its interaction with other schemas and with information in other modes (if present) as well as on the genre to which the discourse as a whole belongs. I am not sure, though, that the FORCE schema as theorized in Talmy (1988) cannot be easily represented in visuals, as Górska claims. If the cartoonist had chosen to use speed and movement lines (Kennedy 1982; Forceville 2011a) and “squash and stretch” techniques, this schema, too, could have been rendered exclusively in the visual mode. I note in passing that in moving images, optionally supported by sound and music, this can be accomplished even more effortlessly – see Forceville (2017a).

Maarten Coëgnarts is one of a still rare species, a cognitive film scholar who systematically deploys Conceptual Metaphor Theory to discuss the medium in which he is an expert. His central point is that not just creative metaphors (called “image” metaphors by Lakoff and Turner [1989] and “resemblance” metaphors by Grady [1999]) manifest themselves in film, but that structural or “correlation” (Grady 1999) ones do so as well. These latter, he argues, draw heavily on image schemas such as FRONT-BACK, SOURCE-PATH-GOAL, and CENTRE-PERIPHERY. Since live-action film in many respects “copies” whatever it records, it is of course possible to convey image schemas in a film’s pro-filmic mise-en-scène, and could thereby also occur, for instance, in a theatrical performance. But Coëgnart’s goal here (see also Coëgnarts and Kravanja 2012) is to demonstrate that film in addition has various medium-specific opportunities to influence the viewer’s perception of this pro-filmic reality, such as framing, camera movement, lighting, and editing. In this chapter he focuses
on the pertinence of the CONTAINER schema, discussing two of its manifestations in a scene from Francis Ford Coppola’s *The Conversation*. The CONTAINER schema is specifically pertinent to the medium of film, Coëgnarts argues, because every film shot literally *frames* a portion of the filmed scene. The “frame” is thus a container “in” which the contents of what is seen (directly, or indirectly via what a character in the story world sees in a “point-of-view” shot) are made accessible to the film viewer. Necessarily, much of what is portrayed are, in fact, metonyms leading the viewer to construe larger wholes (see Forceville [2009] and Pérez-Sobrino [2017] for other applications of metonymy to the visual and multimodal realm). This framing is not static: different parts of the filmed scene come “into” view by means of camera movements and editing. We are hardly aware of this mechanism as it is not fundamentally different from how human beings perceive (and focus attention on) things in the real world, as Coëgnarts reminds us by his chapter motto that “the most powerful conveyor of meaning is the immediate impact of perceptual form” (Rudolf Arnheim). A second aspect of the CONTAINER schema he discusses is where it partakes in EMOTION IS A FLUID IN A PRESSURIZED CONTAINER, a subtype of the more general EMOTIONS ARE FORCES metaphor (Kövecses 2008), in which a strong emotion (joy, anger, embarrassment) that rises in a person’s body is metaphorically understood as a liquid that begins to boil and thereby puts pressure on the pan or kettle in which it is located. Coëgnart argues that a series of shots in a crowded elevator (itself a “pressurized container”) depicting the character of Harry Caul (Gene Hackman) first in medium close-up, then in close-up, then in medium-close-up again, nudges the viewer to perceive Harry’s state of mind in terms of an increase, followed by a decrease, of his embarrassment and panic.

It is much more difficult to persuade readers that these filmic choices deserve to be understood in terms of metaphors than in (audio)visual media in which reality-as-we-know-it has been explicitly manipulated, as in comics (e.g., Forceville 2005, 2011a) and animation film (e.g., Forceville and Jeulink 2011; Fahlenbrach 2017; Forceville and Paling 2018). Coëgnart’s interpretation of the shot sequence (carefully designed by the filmmaker, after all) is nonetheless highly plausible. It would be interesting to test his claims experimentally. In the above scene, for example, manipulating the depiction of Harry Caul by showing him only in medium-shots, or only in long shots, should lead test subjects to evaluate the character as less embarrassed/panicking than in the original sequence.

Coëgnars’ project (see also Coëgnarts and Kravanja 2014, 2015a) is an important one. On the one hand, his work provides instruments for film analysis that are not part and parcel of cognitivist film scholarship by providing a metaphorical raison d’être for certain shot
sequences. On the other, he demonstrates how conceptual metaphor studies feed into cognition studies. More generally, combining insights from cognitive linguists and cognitive film scholars (e.g., Tan 1996; Bordwell and Thompson 2008; Grodal 2009, Forceville 2011b; Smith 2017) offer promising opportunities for synergy.

In a fine chapter (Roman and Porto 2019/this volume), Romano and Porto investigate how a metaphor can “migrate” from one mode to another mode or combination of modes – which they baptize “transmodal metaphor” (see Forceville [1999] for an early example of such a project). They direct their attention to the Spanish 15M movement that over the past decade or so has voiced and embodied the dissatisfaction of many citizens with the consequences of the financial crisis, and with the neoliberal policies in response to this crisis. Specifically, the authors focus on the Madrilenean Mareas protests, showing how the neutral A MASS OF PEOPLE IS A TIDE metaphor acquired specific, and positive, connotations in the context of these protests. Once the notion of the “tide” was firmly associated with the Mareas protests, Porto and Romano show, it became a kind of topos or meme that could subsequently be used in other, non-verbal and multimodal media as well, such as banners, placards, logos, and posters. Since different groups partaking in the protests had come to be connected with certain colours, visual manifestations of the “tide” could from then on play with this colour symbolism. When the Solfónica choir was founded, the central metaphor began to be expressed in the musical mode as well.

Romano and Porto’s illuminating case study of the Mareas strikes me as an exemplary demonstration of the concept of transmodal metaphor. The chapter has several other praiseworthy characteristics. For one thing, it shows that metaphor theory can be well combined with critical discourse analysis to engage with politically and culturally charged ideological issues – as was perhaps first demonstrated by Charteris-Black (2004) and has more recently been addressed by Musolff (2016). Related to this, it is encouraging that the authors do not only draw on cognitivist linguistics models, but also benefit from semiotics-oriented approaches, notably the work by Kress and Van Leeuwen. Thirdly, they confirm that making meaning is in various senses a dynamic process. Not only can a given metaphor develop within a medium, for instance in language; it is moreover bound to transform and adapt itself to some extent when it migrates to another medium, with its own affordances and constraints, such as visuals, visuals-plus-written texts, or music. Therefore, a given metaphor may in two different media “hide and highlight” (Lakoff and Johnson 1980: 10-13) different aspects of the source domain. Finally, the authors remind us that meaning-making, of course,
goes beyond metaphor. Undoubtedly, we “live by metaphors” – but we live by many other things – metonyms, stories, colour symbolism … – as well.

In her chapter, Muelas-Gil first analyses 28 covers of the English *The Economist* magazine (November 2014 – August 2015) according to the **VISMIP method** (Šorm and Steen 2018), charting how many of them can be said to feature a visual metaphor or metonym. She establishes among other things that 15 of the 28 covers sport a visual metaphor. The author goes on to report an experiment with 17 Spanish students “talking-out-loud” about three figurative *Economist* covers. Although not all participants understood the metaphors and metonyms in all three covers, the author concludes that “most of the subjects, who were not experts on the area whatsoever, understood the message to one extent or another” (Muelas-Gil 2019/this volume: 15).

It is commendable that the author expands the repertoire of genres featuring metaphors and metonyms by focusing on magazine covers. There are, however, several methodological issues that counsel caution vis-à-vis her findings. In the first place, Muelas-Gil refers to the (possible) pertinence of “headlines” to the interpretation of the trope at issue, but it was left to the participants to decide for themselves how important these headlines were. As the author herself acknowledges, this is a weakness in the experimental design, since it is impossible to attest whether or not participants based their interpretations (consciously or subconsciously) only on the visuals if they were exposed to the complete visuals-plus-written text. The lack of attention in the experimental design for the relative role of visuals and text has another unfortunate consequence, as it blurs the distinction between visual and multimodal metaphor. If the verbal information in the headline is indispensable for identifying one of the terms (target or source), this would make a metaphor or metonym multimodal rather than visual according to the definition adopted in Forceville and Urios-Aparisi (2009). Šorm and Steen (2013), on whose approach Muelas-Gil bases herself, do not make the distinction, however. As a result, she is not able to distinguish between (monomodal) pictorial/visual metaphor and (multimodal) verbo-pictorial metaphor. Thus, rather than hypothesizing that “the shorter a headline, the more visual support it will need to be understood” (Muelas-Gil 2019/this volume: 16), my suggestion would be that if the headline contains a metaphorical target or source that is not also rendered visually, it is much more crucial than if the pertinent metaphorical term is also visualized.

A second thorny issue is that the three alternatives the author proposes (metaphor, metonymy, both) suggest a misleadingly easy manner of categorizing. For one thing, it is not advisable to say that something “is” a metaphor/metonymy or not; it is better to say that it is
(not) possible/advisable/imperative to construe a metaphor to make sense of the cover as a whole. Construing, or not construing, a metaphor inevitably depends to a considerable extent on an audience’s ability to recruit appropriate background knowledge – and the Spanish student participants (some of whom even needed to have the English headlines translated for them) are hardly the envisaged readers of *The Economist*. For another, many metaphors are rooted in metonymy or vice versa (see Pérez-Sobrino [2017] for extensive discussion). As a matter of fact, in cover 1, the USB-portal is a metonym for the machine, specifically the computer, while cover 3 presents an image that invites construal not just as metonym but also as the metaphor **EURO IS DISCOBOLUS** – where “being thrown away” is the key feature mapped from source to target. With reference to cover 2, one might query the verbalization of the metaphor identified. In its discussion no mention is made, for instance, of the metaphor **MERKEL IS ROCK**. Surely, the scenario (Musolff 2006, 2016) that is presented here is a variety of the **POLITICS IS A JOURNEY** metaphor, in which if Tsipras “goes ahead,” away from the whirlpool, Greece’s ship-of-state will crash on the Merkel-rock.

3 The next stage

Where should multimodal CMMT scholars go from here? I see various exciting opportunities for novel developments. It is becoming ever more clear that, just as metaphor and metonymy, hyperbole, irony, allegory, antithesis, and probably other tropes, operate first and foremost on the conceptual level – which means that theorization requires analysis of their manifestations both in exclusively verbal and in visual and multimodal discourses. This obliges linguists and rhetoricians to reread the scholars who first defined these tropes – classical sources such as Aristotle, Cicero, and Quintilian – and then refine (and if necessary; redefine) them through the lens of contemporary cognitive scholarship in linguistics, stylistics, film studies, comics scholarship and whatever other disciplines the cognitive approach flourishes in. The reformulations can then constitute the basis for examining the manifestations of the various tropes in different media, taking into account the affordances and constraints of the modes upon which these media draw. This work has, in fact already begun. Cognitivist-oriented proposals have been made for, at least, **visual puns** (Abed 1994), **visual oxymoron** and “**pictorial grouping**” (Teng and Sun 2002), **hyperbole**, **paradox**, and **onomatopoeia** (Pérez-Sobrino 2017), **allegory** (Cornevin and Forceville 2017), and **antithesis** (Tseronis and Forceville 2017). Incidentally, it is important to carefully consider whether/where the list of classic Aristotelian tropes may need to be conflated for their visual or multimodal varieties.
(For instance, is it tenable and desirable to distinguish oxymoron, paradox and antithesis in visuals as they are now commonly distinguished in language?) This line of research can be captured in the slogan “from CMT, via CMMT, to CTT,” the latter acronym standing for “Cognitive Trope Theory”.

Another promising source of new insights is investigating other visual and multimodal genres and subgenres. After all, metaphors (just like any other potentially meaning-generating pattern) may “behave” in slightly or vastly different ways depending on the medium and genre in which they are expressed. Film metaphor, for one, is receiving more and more scholarly attention (e.g., Ortiz 2011, 2015; Winter 2014; Coëgnarts and Kravanja 2015b; Fahlenbrach 2016; 2017). Abdel-Raheem (2019) examines cartoons and op-ed illustrations, bringing to the genre first-hand knowledge of Arabic perspectives, as does Maalej (2015) to university promotion material (for some thoughts on cultural dimensions of visual and multimodal metaphor, see Forceville 2017b). Another intriguing genre is street art. Poppi and Kravanja (2019) focus on Banksy’s public art, whereas Asenjo (2018) works on the famous political wall paintings in Belfast. It is to be noticed, incidentally, that in most of this work the discussion of modes partaking in multimodal metaphor is restricted to the visual and the written-verbal mode. Multimodal metaphor research – and multimodal discourse analysis more generally – including the sonic and musical modes is still rare.

In carrying out this highly complex and demanding research, it is crucial to bear in mind that models for identifying tropes should eventually benefit the analysis of visuals and multimodal text rather than the other way round. Put differently, research should in the last resort help solve puzzles and problems in all kinds of discourse, and models are nothing less but also nothing more than tools to achieve this. We should therefore not hesitate, whenever necessary, to adapt models if they are not, or not sufficiently, capable of performing the job of accounting for new textual data. The idea that analysts should first try to exhaustively describe the visuals and text of a piece of discourse, then signal incongruity, and then identify any metaphors, as Negro et al. (2017) propose, is in my view a misguided strategy. As these authors themselves discovered, even having only two raters describe a given picture rarely leads to the same results. The problem is that a picture can potentially be described in an infinite number of ways. I suggest we start at the other, pragmatic end, namely by attesting to what genre a discourse belongs. If we do this correctly, we know what interpretation strategy is called for, since genres trigger expectations as to what kind of meaning is intended. For instance, a commercial advertiser wants to make a positive claim about a product, service, or brand, while a political cartoonist intends to make a critical, preferably humorous, comment.
on a political situation. As analysts we then, as it were, “backtrack” and observe what visual and verbal information is pertinent to conveying the message. Once we have inventoried this information, and notice that its presentation manifests some sort of incongruity, we can start reflecting whether, and if so, how, it makes sense to label this incongruity “metaphorical” (or metonymical, symbolical, ironical, hyperbolical …), and/or pertains to odd stylistic features, an intertextual reference, or any of a range of other phenomena. In short, any analysis of any discourse – and hence of any element partaking in this discourse – needs to begin by assessing what sort of information its communicator wants to communicate. Therefore the analysis of tropes must be embedded in a theory of communication and cognition. My candidate for such a theory is relevance theory (Sperber and Wilson 1995; Wilson and Sperber 2012; Forceville 2014, in prep.; Forceville and Clark 2014).

References


Asenjo, Roberto. 2018. The influence of culture in the multimodal murals of Northern Ireland. Warsaw Multimodality Workshop and Masterclass, 7–9 June 2018, University of Warsaw, Poland.


Forceville, Charles. in prep. *Analyzing visual and multimodal mass-communication: A pragmatic model* (working title). To be published by Oxford University Press.


